

Yaşam çözümlenmesinde zamana bağlı açıklayıcı değişkenli Cox regresyon modeli

Cox regression model with time dependent covariate in survival analysis

Durdu Sertkaya, Nihal Ata, M. Tekin Sözer

Hacettepe Üniversitesi, Fen Fakültesi, İstatistik Bölümü,
Ankara

Tıpta, salgın hastalıklara ve kronik hastalıklara ilişkin verilerin incelenmesinde ve bu hastalıkları etkileyen faktörlerin saptanmasında yaşam çözümlenmesi Cox regresyon modeli oldukça önemlidir. Ancak, zamanla değişen açıklayıcı değişkenler olduğunda Cox regresyon modeli yerine zamana bağlı açıklayıcı değişkenli Cox regresyon modeli uygun olmaktadır. Bu çalışmada, sabit ve zamana bağlı açıklayıcı değişkenler durumunda Cox regresyon modelleri incelenmiştir. 116 akciğer kanseri hastalarına ait veriler kullanılarak bir uygulama yapılmış ve sonuçlar tartışılmıştır.

Anahtar sözcükler: *Cox regresyon, Yaşam çözümlenmesi, Zamana bağlı açıklayıcı değişken, Akciğer kanseri*

In medical science, in investigating the survival data of epidemic diseases and chronic diseases and determining the factors which affects these diseases, Cox regression model for survival analysis has gained widespread attention. However, using Cox regression model with time-dependent covariates is more suitable than using Cox regression model with fixed covariates in the case of time-dependent covariates. In this study, Cox regression model with fixed covariates and Cox regression model with time-dependent covariates are described. The data of 116 lung cancer patients is used to illustrate these models and the results are discussed.

Key words: *Cox regression, Survival analysis, Time-dependent covariate, Lung cancer*

Tıbbi araştırmalarda klasik çözümlenme yöntemleri yetersiz kalmaktadır. Bunun birinci nedeni, çoğu kez araştırmanın değerlendirilmesinin tüm hastalar ölmeden veya incelenen sonuç ortaya çıkmadan yapılması gerektiği, aksi durumda hangi tedavi yönteminin daha iyi olduğunun ve hastalığı etkileyen faktörlerin saptanmasının yıllar alabileceğidir. İkinci nedeni ise, hastalara uygulanan tedavilerin aynı zamanda başlamamasıdır. Bu tip çalışmalarda yaşam çözümlenmesi yöntemleri daha uygun sonuçlar vermektedir (1).

İnsan ve çevresindeki birimlerin başarısızlık (failure) ya da ölüm zamanının istatistiksel çözümlenmesine ilişkin çalışmalar yaşam tablosu yardımıyla başlamıştır. Bu çalışmalar daha sonra geliştirilerek başarısızlık modeli ya da hazard modeli olarak adlandırılmıştır. Bu modeller, iyi tanımlanmış herhangi bir olayın olma ya da gözlemlenme zamanının çözümlenmesinde kullanılmıştır.

20. yüzyılda başlayan yaşam çözümlenmesi çalışmaları, bu yüzyılın ikinci yarısı boyunca büyük ilerlemeler göstermiştir. Bu alandaki etkili gelişmeler;

- Yaşam fonksiyonunun tahmini için kullanılan “Kaplan-Meier Yöntemi” (2),
- İki yaşam dağılımının karşılaştırılması için kullanılan “Log-rank Test İstatistiği” (3),
- Yaşam süresi üzerinde açıklayıcı değişkenlerin etkilerini ölçebilmek için kullanılan “Cox Regresyon Modeli” (4),

Geliş tarihi: 31.08.2005 • Kabul tarihi: 02.12.2005

İletişim Adresi

Durdu Sertkaya
Hacettepe Üniversitesi Fen Fakültesi
İstatistik Bölümü 06800 Beytepe, Ankara
Tel : (312) 297 79 30
E-posta adresi : durdu@hacettepe.edu.tr

■ Yaşam çözümlenmesi istatistiklerinin küçük ve büyük örneklem özelliklerini ortaya koyabilmek için birleşik bir yapı sağlayan “Sayma Süreci Martingale Kuramı” (5) ile başlamış ve günümüze kadar gelmiştir (6).

Yaşam çözümlenmesi, herbiri için genellikle başarısızlık olarak adlandırılan bir nokta olayı olan bireyler grubu ya da grupları ile ilgilenmektedir. Başarısızlık belirli bir zaman aralığından sonra oluşur ve buna başarısızlık zamanı (failure time) denir. Yaşayan bir organizmanın ya da cansız bir nesnenin belirli bir başlangıç zamanı ile ölümü (başarısızlığı) arasında geçen zamana “yaşam süresi” ya da “başarısızlık süresi” adı verilmektedir ve genellikle T ile gösterilir. Başarısızlık süresine örnek olarak, makine bileşenlerinin yaşam süreleri, ekonomide işsizlik dönemleri, psikolojik bir deneyde deneğin belirlenen görevi tamamlama süresi ve klinik bir deneyde hastaların yaşam süreleri gösterilebilir.

Yaşam çözümlenmesinde temel olan, gözlenen başarısızlık sürelerinin incelenmesi olduğundan, bu değişkenin iyi tanımlanması gerekmektedir. İlgilenilen olaya göre farklı şekillerde ortaya çıkabilen bu değişkenin duyarlı olarak ölçülmesi için,

- başlangıç zamanı her bir birim ya da birey için şüpheye yer vermeyecek şekilde tanımlanmalıdır,
- geçen süreyi ölçmek için bir zaman ölçüğü kabul edilmelidir,
- her bir birim ya da birey için ömrün sona erdiği ya da başarısızlığın meydana geldiği an tamamen net olmalıdır. Diğer bir deyişle, başarısızlığın tam olarak tanımlanması gerekmektedir (1).

Yaşam çözümlenmesinin amaçları; farklı zamanlarda yaşam olasılığı tahminlerinin elde edilmesi, yaşam süresinin dağılımının tahmin edilmesi, farklı hasta gruplarının yaşam süresi dağılımlarının karşılaştırılması olarak sıralanabilir (7).

Yaşam verilerinin çözümlenmesinde karşılaşılan temel güçlük, gözlem altına alınan bazı birim ya da bireylerin başarısızlık zamanlarının gözlenememiş olmasıdır. Örneğin, tıbbi bir çalışmada gözlem altına alınan hastaların bazıları deney sonunda hâlâ yaşamlarını sürdürüyor olabilir. Ayrıca, gözlem altındaki bir birey, bazı nedenlerden dolayı gözlemden çıkabilir. Eğer başarısızlık zamanı, bu gibi nedenlerden dolayı tamamlanamamış ise durdurulmuş (censored) durum söz konusudur.

Birçok durumda incelenen yaşam süresinin başka faktörler tarafından da etkilenebileceği göz önünde bulundurulduğunda, bağımlı değişken olan yaşam süresi üzerinde açıklayıcı değişkenlerin de etkilerinin modellendiği regresyon modelleri yaşam çözümlenmesinde önemli bir yer almaktadır. Örneğin, akciğer kanseri olan hastaların yaşam süreleri üzerine yapılan bir çalışmada hastanın yaşı, hasta-

nın genel fiziksel durumu, tümörün büyüklüğü, hastalığın teşhisinden sonra geçen süre ve diğer ilgili değişkenler açıklayıcı değişkenler olarak alınabilmektedir.

Yaşam çözümlenmesinde en çok kullanılan regresyon modeli Cox regresyon modelidir. 1972 yılında Cox tarafından geliştirilen regresyon modeli ile yaşam çözümlenmesinde önemli adımlar atılmış, Cox’un (7) önerileri, Kalbfleisch ve Prentice’in (8) katkıları ile bugünkü önemini kazanmıştır.

Son yıllarda yapılan çalışmalarda Cox regresyon modeli sabit açıklayıcı değişkenlerin yanı sıra zamana bağlı açıklayıcı değişkenleri de içeren Cox regresyon modeline genişlemiştir.

Bu çalışmanın amacı, Cox regresyon modelinde zamandan bağımsız (sabit) ve zamana bağlı açıklayıcı değişkenlerin kullanımını incelemek, gerçek bir veri kümesine bu modelleri uygulamak ve elde edilen sonuçları yorumlamaktadır.

Cox regresyon modeli

Yaşam süresine ilişkin etkenlerin hazard fonksiyonu üzerindeki etkilerinin çarpımsal olduğu modeller, yaşam süresi verilerinin çözümlenmesinde önemli bir rol oynamaktadır.

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

$$h(t, x) = h_0(t)\psi(x)$$

biçimindedir. $\psi(x)$ değişik biçimlerde ifade edilebilmektedir. Cox’un 1972’de incelediği model,

$$h(t, x) = H_0(t) \exp(\beta'x) \quad (1)$$

biçimindedir. Burada, x açıklayıcı değişkenler vektörü $(x_1, \dots, x_p)'$, β regresyon katsayıları vektörü ve $h_0(t)$ ise $x=0$ olan bir birimin temel hazard fonksiyonu olarak tanımlanmaktadır (4).

Eşitlik 1’de verilen modelde iki bilinmeyen bileşen vardır. Bunlar regresyon parametresi β ve temel hazard fonksiyonu $h_0(t)$ ’dir. Buna eşdeğer olarak temel yaşam fonksiyonu,

$$S_0(t) = \exp\left[-\int_0^t h_0(u)du\right] = [-H_0(t)] \quad (2)$$

biçiminde yazılabilir. Burada $H_0(t)$ temel birikimli hazard fonksiyonudur. x verildiğinde T’nin yaşam fonksiyonu,

$$S(t, x) = \exp\left(\int_0^t h^*(u, x)du\right) = S_0(t) \exp(\beta'x) \quad (3)$$

olmaktadır (9).

Zamana bağlı açıklayıcı değişkenler durumunda Cox regresyon modeli

Eşitlik 1’de verilen modelde temel hazard fonksiyonu zamanın bir fonksiyonunu içerirken, üstel kısımda yer alan

değişkenler zamanın bir fonksiyonu değildir, yani zamandan bağımsızdır. Modelde zamanı içeren değişkenler de olabilir. Bu değişkenler zamana bağlı açıklayıcı değişkenler olarak adlandırılmaktadır. Zamana bağlı değişken, incelenen birey için değeri zamanla değişebilen herhangi bir değişken olarak tanımlanmaktadır. Zamandan bağımsız değişken ise incelenen birey için değeri zaman içinde sabit kalan bir değişkendir. En yaygın tanımlanan zamana bağlı değişken, zamandan bağımsız bir değişken ile zamanın ya da zamanın bir fonksiyonunun çarpımı biçimindedir (9-11).

Zamana bağlı açıklayıcı değişkenin üç türü vardır. Bunlardan biri tanımlanmış (defined) zamana bağlı açıklayıcı değişkendir. Bu, sabit bir açıklayıcı değişken ile zamanın bir fonksiyonunun çarpımı şeklinde elde edilen değişkendir. Zamana bağlı değişkenlerin bir diğeri içsel (internal) değişkendir. İçsel değişkenler için değişkenin değerindeki değişimin nedeni içsel özelliklere, yani bireyin kendi davranışlarına ve özelliklerine dayanmaktadır. Bu gibi değişkenlere örnek olarak, t zamanındaki sigara içme durumu ve t zamanındaki obezite seviyesi verilebilir. İçsel değişkenler için değişkenin değerindeki değişimin nedeni içsel özelliklere, yani bireyin kendi davranışlarına ve özelliklerine dayanmaktadır. Diğer bir zamana bağlı değişken yardımcı (ancillary) değişkendir. Değişkenin değerindeki değişimin nedeni dışsal, çevreye ait özelliklerdir. Özel bir coğrafi alan için t zamanındaki hava kirliliği indeksi dışsal değişkene bir örnek olarak verilebilir (10, 12).

Kısmen içsel, kısmen dışsal olabilecek değişkenler de vardır. Buna örnek olarak ciddi bir kalp rahatsızlığı olan bir hastanın kalp nakli için uygun olup olmaması durumunu düşünebiliriz. Kalp nakli durumuna ilişkin değişken, bireyin ayırt edici özelliklerinden dolayı içsel olarak, tıbbi vericinin elde edilebilirliğinden dolayı yardımcı değişken olarak adlandırılmaktadır (10, 12).

Birçok durumda açıklayıcı değişken verisi uzun dönemde elde edilmektedir. Örneğin kan basıncı, CD4 sayımı, görel ağırlık ve hastalık öyküsü seçilen periyodik zaman noktalarında elde edilebilir. Tedavi ya da diğer etkiler zaman içerisinde değişebilir. Uygun istatistiksel modelde zamanla değişen açıklayıcı değişken bilgisini kullanmak gerekir. Bunu yapmanın yollarından biri de zamana bağlı açıklayıcı değişkenli Cox regresyon modelini kullanmaktır (10).

Zamana bağlı açıklayıcı değişkenler olduğu durumda, Cox regresyon modeli zamandan bağımsız değişkenleri ve zamanın bazı fonksiyonları ile bu değişkenlerin çarpımını içeren bir modele genişlemektedir. x_1, x_2, \dots, x_p zamandan bağımsız değişkenler ve $x_1(t), x_2(t), \dots, x_{p_2}(t)$ zamana bağlı değişkenler olmak üzere açıklayıcı değişkenler,

$$x(t) = (x_1, x_2, \dots, x_{p_1}, x_1(t), x_2(t), \dots, x_{p_2}(t))$$

biçiminde gösterilmektedir. Buna göre β ve δ açıklayıcı değişkenlerin katsayılar vektörü olmak üzere Cox regresyon modeli,

$$h(t, x(t)) = h_0(t) \exp[(\beta'x) + \delta'xg(t)] \quad (4)$$

biçiminde yazılmaktadır. Burada $g(t)$ zamanın bir fonksiyonu olarak tanımlanmaktadır. $g(t)$ 'nin seçimi kullanılan değişkenlerin durumuna göre ve araştırmacının bilgisine göre değişkenlik göstermektedir. Bu fonksiyon genellikle t , $\log t$ ($\ln t$) ya da adım fonksiyonları biçiminde tanımlanmaktadır (10,11).

Cox regresyon modeli, zamanla değişen açıklayıcı değişkenlerle de kullanılabilir. Uygulamada kullanımları sabit (zamandan bağımsız) açıklayıcı değişkenlere göre daha karmaşıktır. Ayrıca hatalı çıkarılma ve modelleme için potansiyel gittikçe artmaktadır (10).

Gereç ve yöntem

Çalışmada, 1994-1998 yılları arasında İbn-i Sina Hastanesi Göğüs Cerrahisi Bölümü'nde tedavi gören 116 akciğer kanseri hastasına ait veriler kullanılmıştır. Hastalar, ameliyat olduktan sonra hastalıklarının ilk nüks etmesine kadar geçen süre boyunca izlenmiştir. Akciğer kanseri hastalarının yaşam sürelerini etkileyen faktörler, sabit ve zamana bağlı açıklayıcı değişkenler durumunda Cox regresyon modelleri kullanılarak belirlenmeye çalışılmıştır.

Bu uygulamada, analiz sonuçlarının tıbbi yorumundan çok, tıbbi bir veri kümesinde sabit ve zamana bağlı açıklayıcı değişkenler durumunda Cox regresyon modellerinin nasıl uygulandığını ve yorumlandığını vurgulamak amaçlanmıştır.

Bulgular

Çalışmada, hastaların ameliyat olduğu tarihten hastalığın ilk nüks etmesine kadar geçen süre (ay olarak) yaşam süresi olarak alınmıştır. Hastalığın nüks etmesi başarısızlık olarak ifade edilmiştir. Hastalığı nüks etmeyen hastalar durdurulmuş olarak tanımlanmıştır. Hastaların izlenme süresi sona erdiğinde 116 hastadan 43'ünde (%37.1) başarısızlık ve 73'ünde (%62.9) durdurma gözlenmiştir. Uygulamada cinsiyet, yaş, sigara tüketimi (paket yıl olarak, ST), tümörün boyutu (mm olarak), invazyon ve patolojik evre (PE) değişkenleri çözümlenmeye alınmıştır. Bu değişkenler, değişkenlerin düzeyleri ve tanımlayıcı değerleri Tablo 1'de verilmiştir.

Cox regresyon çözümlemesinde değişken düzeylerinden biri (genellikle etkenin olmadığı ya da hastalık üzerinde en az etkisinin olduğu düşünülen düzey) referans kategorisi olarak alınmakta ve değişken düzeylerinin yorumlanması buna göre yapılmaktadır. Bu çalışmada, modeldeki değişken β için parametresi ve standart hatası (S.H.), p değeri, hazard oranı ($\exp(\beta)$) ile değişken düzeyleri için hazard oranının alt ve üst sınırları verilmiştir. β parametresinin pozitif değer olması bu

Tablo 1. Kullanılan değişkenler ve düzeyleri

<i>Değişken</i>	<i>Değişken düzeyleri</i> ($\bar{X} \pm$ standart hata)	<i>N (%)</i>	<i>Başarısız olay</i> <i>Sayısı (%)</i>	<i>Durdurulmuş olay</i> <i>Sayısı (%)</i>
Yaş (yıl)	(57.38±0.99)	116	43 (37.1)	73 (62.9)
Cinsiyet	0. Kadın	8 (6.9)	4 (50)	4 (50)
	1. Erkek	108 (93.1)	39 (36.1)	69 (63.9)
Sigara Tüketimi (paket/yıl)	1. ≤ 5	9 (7.7)	3 (33.3)	6 (66.7)
	2. 6-30	36 (31)	11 (30.6)	25 (69.4)
	3. 31-60	56 (48.3)	20 (35.7)	36 (64.3)
	4. ≥ 61	15 (13)	9 (60)	6 (40)
Boyut (mm)	(48.54±2.11)	116	43 (37.1)	73 (62.9)
Patolojik Evre	1. Evre I	61 (52.6)	17 (27.8)	44 (72.2)
	2. Evre II	27 (23.3)	8 (29.6)	19 (70.4)
	3. Evre III+Evre IV	28 (24.1)	18 (64.3)	10 (35.7)
İnvazyon	0. Yok	70 (60.3)	25 (35.7)	45 (64.3)
	1. Var	46 (39.7)	18 (39.1)	28 (60.9)
Süre	(33±1.69)	116	43 (37.1)	73 (62.9)

Tablo 2. Tek açıklayıcı değişken olması durumunda Cox regresyon çözümlemesinin sonuçları

<i>Modelde Yer Alan Değişken</i>	β	<i>S.H.</i>	<i>p-değeri</i>	<i>Exp(β)</i>	<i>Alt sınır – Üst sınır</i>
Yaş	-0.003	0.015	0.8430	0.99	0.97 – 1.03
Cinsiyet	-0.394	0.527	0.4545	0.67	0.24 – 1.89
Sigara Tüketimi			0.1965		
ST(2)	-0.423	0.654	0.5182	0.66	0.18 – 2.36
ST(3)	-0.281	0.622	0.6514	0.76	0.22 – 2.56
ST(4)	0.474	0.668	0.4784	1.61	0.43 – 5.95
Boyut	0.014	0.006	0.0225	1.01	1.00 – 1.03
Patolojik Evre			0.0013		
PE(II)	0.255	0.432	0.5552	1.29	0.55 – 3.01
PE(III+IV)	1.203	0.342	0.0004	3.33	1.70 – 6.51
İnvazyon	0.449	0.314	0.1531	1.57	0.85 – 2.90

düzeyin referans kategorisine göre daha fazla riskli olduğunu, β parametresinin negatif değer olması ise bu düzeyin referans kategorisine göre daha az riskli olduğunu göstermektedir. Hazard oranı olan $\exp(\beta)$ değeri ise önemli bulunan düzeyin, referans kategorisine göre kaç kat (ya da % ne kadar) daha riskli olduğu yorumunu getirmektedir. Önemli bulunan değişkenlerin herbir düzeyine karşılık gelen p değerlerine bakılarak önemli değişken düzeyleri belirlenmektedir.

Herbir değişken için ilk düzeyler referans kategorisi olarak alındığından tablolarda yer almamaktadır. Çalışmadaki açıklayıcı değişkenlerle tek tek Cox regresyon çözümlemesi yapıldığında elde edilen sonuçlar Tablo 2'de verilmiştir.

Tablo 2 incelendiğinde boyut ve patolojik evre değişkenlerinin önemli olduğu (p -değeri <0.05 olduğundan)

%95 güven ile söyleyebilmektedir. Tümörün boyutu arttıkça hastalığın nüks etme riski de artmaktadır. Patolojik evresi III+IV olan bir hastanın patolojik evresi I olan hastaya göre yaklaşık 3.33 kat daha fazla riskli olduğu söylenebilmektedir.

Çalışmadaki altı açıklayıcı değişken birlikte düşünüldüğünde hastalığı etkileyen faktörlerin hangilerinin olduğunu belirlemek için geriye doğru seçim yöntemiyle Cox regresyon çözümlemesi yapıldığında elde edilen modele ait sonuçlar Tablo 3'te verilmiştir.

Tablo 3'te görüldüğü gibi ileriye doğru seçim yöntemi uygulandığında modelde sigara tüketimi, boyut ve patolojik evre değişkenleri kalmaktadır. p değerleri incelendiğinde sigara tüketimi, boyut ve patolojik evre değişkenlerinin ba-

Tablo 3. Tüm açıklayıcı değişkenler durumunda ileriye doğru seçim yöntemi ile Cox regresyon çözümlemesinin sonuçları

Değişken	β	S.H.	p-değeri	Exp(β)	Alt sınır - Üst sınır
Sigara Tüketimi			0.0396		
ST(2)	0.291	0.684	0.6707	1.34	0.35 – 5.11
ST(3)	0.200	0.637	0.7533	1.22	0.35 – 4.25
ST(4)	1.332	0.701	0.0573	3.79	0.96 – 14.96
Boyut	0.013	0.057	0.0266	1.01	1.00 – 1.02
Patolojik Evre			0.0008		
PE(II)	0.193	0.441	0.6606	1.21	0.52 – 2.88
PE(III+IV)	1.322	0.368	0.0003	3.75	1.82 – 7.72

Tablo 4. Boyut*Int Değişkeni Olduğu Durumunda Cox Regresyon Çözümlemesinin Sonuçları

Değişken	β	S.H.	p-değeri	Exp(β)	Alt sınır – Üst sınır
Boyut x Int	0.004	0.002	0.0331	1.01	1.00 – 1.01

Tablo 5. Tüm açıklayıcı değişkenler ve boyut*Int değişkeni olması durumunda ileriye doğru seçim yöntemiyle Cox regresyon çözümlemesinin sonuçları

Değişken	β	S.H.	p-değeri	Exp(β)	Alt sınır - Üst sınır
Sigara Tüketimi			0.0379		
ST(2)	0.284	0.683	0.6779	1.33	0.35 – 5.07
ST(3)	0.202	0.637	0.7513	1.22	0.35 – 4.26
ST(4)	1.339	0.700	0.5580	3.81	0.97 – 15.04
Patolojik Evre			0.0007		
PE(II)	0.203	0.439	0.6440	1.23	0.52 – 2.89
PE(III+IV)	1.329	0.366	0.0003	3.78	1.84 – 7.75
Boyut x Int	0.004	0.002	0.031	1.00	1.00 – 1.01

şarısızlığı etkileyen önemli risk faktörleri olduğu %95 güven düzeyinde söylenebilmektedir ($p < 0.05$ olduğundan). Sigara tüketimi değişkeninin üçüncü düzeyi önemli bulunmuştur ($p \approx 0.05$). Sigara tüketimi ≥ 61 olan hastanın, sigara tüketimi ≤ 5 olan hastaya göre yaklaşık 3.8 kat daha fazla riskli olduğu söylenebilmektedir. Patolojik evre değişkeninin üçüncü düzeyi önemli olarak bulunmuştur ($p < 0.05$ olduğundan). Patolojik evresi III+IV olan hastanın patolojik evresi I olan hastaya göre 3.75 kat daha fazla riskli olduğu görülmektedir. Tümörün boyutu arttıkça hastalığın nüks etme riski de artmaktadır.

Buraya kadar yapılan Cox regresyon çözümlemesinde zamana bağlı açıklayıcı değişken göz ardı edilmiştir. Oysa ki, boyut değişkeni sabit değil, zamanla değişen bir değiş-

kendir. Bu durumda boyut değişkeni, zamanın bir fonksiyonu ile çarpım biçiminde modelde yer alacaktır. Zamanın bir fonksiyonu olarak t, Int ya da araştırmacının belirleyeceği bir fonksiyon kullanılabilir. Çalışmada Int zamanın bir fonksiyonu olarak kullanılmıştır. Buna göre modelde sadece boyut değişkeni varken zamana bağlı açıklayıcı değişkenler durumunda Cox regresyon çözümlemesi Tablo 4'te verilmiştir.

Tablo 4 incelendiğinde, boyut*Int değişkeninin önemli bir faktör olduğu ($p < 0.05$) %0.95 güvenle söylenebilmektedir. Boyut*Int arttıkça nüks etme riski de artmaktadır.

Tüm açıklayıcı değişkenlerin kullanıldığı ve zamanın bir fonksiyonu Int iken geriye doğru seçim yöntemi ile Cox regresyon çözümlemesi sonuçları ise Tablo 5'te verilmiştir.

Tablo 5 incelendiğinde, modelde sigara tüketimi, patolojik evre ve boyutlu değişkenlerinin kaldığı görülmektedir. Sigara tüketimi, patolojik evre ve boyutlu değişkenlerinin başarısızlığı etkileyen önemli risk faktörleri olduğu %95 güven düzeyinde söylenebilmektedir ($p < 0.05$ olduğundan). Zamanın fonksiyonu İnt iken tümörün boyutu arttıkça hastalığın nüks etme riski de artmaktadır. Sigara tüketimi ≥ 61 olan hastanın, sigara tüketimi ≤ 5 olan hastaya göre 3.8 kat daha fazla riskli olduğu, patolojik evresi III+IV olan bir hastanın patolojik evresi I olan hastaya göre 3.78 kat daha fazla riskli olduğu söylenebilmektedir.

Sabit ve zamana bağlı açıklayıcı değişkenler durumunda yapılan geriye doğru seçim yöntemiyle Cox regresyon çözümlemesi sonuçları karşılaştırıldığında (Tablo 3 ve Tablo 5), hazard oranlarında önemli bir farklılık olmasa da boyutlu değişkeninin standart hatasının boyut değişkeninin standart hatasından daha küçük olduğu görülmektedir.

Sonuç

Bu çalışmada, sabit ve zamana bağlı açıklayıcı değişkenli Cox regresyon modelleri incelendikten sonra akciğer kanseri verilerinde bu iki model kullanılarak hastalığın tekrar nüks etmesini etkileyen risk faktörleri belirlenmiştir.

Kaynaklar

1. Cox DR, Oakes D. Analysis of Survival Data, New York: Chapman and Hall; 1984.
2. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations, Journal of the American Statistical Association 1958; 53: 457-481.
3. Mantel N. Evaluation of survival data and two new rank order statistics arising in its consideration, Cancer Chemotherapy Report 1966; 50: 163-170.
4. Cox DR. Regression models and life-tables, Journal of the Royal Statistical Society, Series B, 1972; 34: 187-220.
5. Aalen OO. Statistical inference for a family of counting processes, PH.D. dissertation, University of California, Berkeley, 1975.
6. Fleming TR, Lin DY. Survival analysis in clinical trials: past developments and future directions, Biometrics 2000; 56(4): 971-983.

Uygulamada, öncelikle tüm açıklayıcı değişkenlerin sabit olduğu düşünülerek Cox regresyon çözümlemesi yapılmıştır. Daha sonra, açıklayıcı değişkenlerden tümörün boyutu değişkeninin zaman içerisinde değiştiği düşünülerek zamana bağlı açıklayıcı değişkenli Cox regresyon çözümlemesi yapılmıştır. Zaman içerisinde değişen açıklayıcı değişken olduğunda sabit değişkenli Cox regresyon çözümlemesi yerine zamana bağlı açıklayıcı değişkenli Cox regresyon çözümlemesinin yapılması gerekliliği vurgulanmaya çalışılmıştır. Burada yapılan uygulamada hazard oranlarında önemli bir farklılık olmamakla birlikte zamana bağlı açıklayıcı değişkenin standart hatasında düşüş olduğu görülmüştür. Zamana bağlı olarak değişen bir açıklayıcı değişkeni zaman içerisinde sabit kabul etmek yanlış olmaktadır.

Uygulamalarda sadece tek bir açıklayıcı değişkenin değil de birden fazla açıklayıcı değişkenin zamana bağlı olduğu durumlarda mümkündür. Bu durumda izlenecek yol, tek bir açıklayıcı değişkenin zamana bağlı olması durumunda kullanılan Cox regresyon çözümlemesi ile aynıdır.

Teşekkür

Bu makalenin uygulamasında kullanılan verileri sağlayan Prof. Dr. Hadi Akay'a ve Dr. Ayten Kayı Cangir'e teşekkür ediyoruz.

7. Collett D. Modelling Survival Data in Medical Research, London: Chapman&Hall; 1994.
8. Kalbfleisch JD, Prentice RL. The Statistical Analysis of Failure Time Data, New York :Wiley; 1980.
9. Pettitt AN, Daud IB. Investigating time dependence in Cox's proportional hazards model, Applied Statistics 1990; 39: 313-329.
10. Kleinbaum DG. Survival Analysis: A Self-Learning Text, New York: Springer; 1996.
11. Therneau TM, Grambsch PM. Modelling Survival Data: Extending the Cox Model, New York: Springer; 2000.
12. Fisher LD, Lin DY Time-dependent covariates in the Cox proportional hazards regression model, Annual Review of Public Health 1999; 20: 145-157.