

**T.C.
MİMAR SİNAN ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
İSTATİSTİK ANABİLİM DALI
İSTATİSTİK YÜKSEK LİSANS TEZİ**

COX REGRESYON ANALİZİ VE BİR UYGULAMASI

**Benay BAYÜLKEN YETKİN
DANIŞMAN: Prof.Dr.Gülay KIROĞLU**

İSTANBUL-2006

ÖNSÖZ

Yüksek Lisans eğitimine başlamam konusunda destekleyen ve eğitimim süresince de beni her konuda yönlendiren ve şuanda aramızda olamamasından üzüntü duyduğum Prof. Dr. Kemal YOĞURTÇUGİL'e minnet ve teşekkürlerimi sunmayı bir borç bilirim.

Hem ders hem de tez aşamasında karşılaştığım tüm soru ve sorunlarımda bana sabırla ve inançla yol gösteren danışman hocam Prof. Dr. Gülay KIROĞLU'na ve eğitimim süresince bilgi desteğini aldığım Prof. Dr. Nalan CİNEMRE'ye sonsuz teşekkürlerimi sunuyorum.

Tezimin hazırlanmasında ve verilerin temininde yardımlarını esirgemeyen Sn. Emire BOR'a ve ayrıca tez konuyla ilgili kaynak bulmamda destek veren Yrd.Doç.Dr. Meral YAY'a teşekkür ederim.

Son olarak, bu eğitimi tamamlamam konusunda beni sonuna kadar destekleyen çok değerli aileme ve eşime sonsuz teşekkürler.

Benay BAYÜLKEN YETKİN

İÇİNDEKİLER

ÖZET	III
SUMMARY	III
GİRİŞ	1
BÖLÜM 1: GENEL BİLGİLER	
1.1 Bazı Kavramların Açıklaması	4
BÖLÜM 2: YAŞAM ANALİZ YÖNTEMLERİ	
2.1 Yaşam Tablosu (YT)	6
2.1.1 Yaşam fonksiyonunun hesaplanması	7
2.1.2 Hazard (Ölüm) ve olasılık yoğunluk fonksiyonları	9
2.1.3 Ortanca yaşam süresi	10
2.2 Kaplan–Meier (KM)	10
2.2.1 Yaşam fonksiyonunun hesaplanması	11
2.2.2 Yığılımlı ölüm fonksiyonu (Cumulative Hazard Function)	12
2.2.3 Ortalama yaşam süresi (Mean Survival Time)	13
2.2.4 Çeyrek değerler (Quartiles)	13
2.3 Cox Regresyon	15
2.3.1 Çoklu cox modelinde β katsayılarının hesaplanması	19
2.3.2 β katsayılarının önemliliğinin test edilmesi	21
2.3.3 Zamana bağlı açıklayıcı değişkenler durumunda Cox regresyon modeli	24
BÖLÜM 3: UYGULAMA	27
3.1 Tüm Parametrelerin Ölüm Üzerine Etkileri	29
3.2 Olguların Yaşam Tablosu	30
3.3 Ölüm Durumuna Göre Cox Regresyon Analizi	30
BÖLÜM 4: SONUÇ	35
KAYNAKLAR	36
EKLER	37
EK-1 Hasta Takip Formu	
EK-2 SPSS Analiz Çıktıları	
EK-3 Data Verilerinin CD'si	

ÖZET

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

Çalışma 4 bölümden oluşmaktadır. Bölüm 1'de yaşam çözümlemesi ile ilgili genel bazı kavramlar açıklanmıştır. Yaşam çözümlemesi yöntemlerinden Yaşam tablosu, Kaplan- Meier ve Cox regresyon'nun açıklamalarına ise Bölüm 2'de yer verilmiştir.

Uygulama bölümünde ise çalışmada takip edilen hasta sayısından ve bu hastalardan toplanan verilerin neler olduğundan ve bu verilere uygulanan yaşam çözümlemelerinin neler olduğundan bahsedilmiştir.

Son olarak sonuç bölümünde de hasta verilerine uygulanmış olan Student-t testi, Ki-Kare testi ve Cox regresyon yöntemi ile ölüm üzerine etkilerinden bahsedilmiştir.

Anahtar kelimeler: Cox regresyon, yaşam analizi

SUMMARY

In this study, 481 patients with heart disease who are observed for 5 years are separated into two groups as death and alive are examined according to factors which affect life period (age, smoking, diabetes mellitus, ischemic transiel attack (ita), changing heart vein(radial),left ventricular, left inner dal(LAD), usage of statine, heart weakness (EF>30), chronical obstructive pulmonary disease (COPD), seen death and examination duration) as using Cox regression model.

Affects of cigarette on death are seen impressive while univariate analize and EF>30, statistically significant radial, COPD, age are added to the model as parameters.

After evaluating the Cox regression analysis which is used Backward stepwise method. The model has been seen statistically significant in high level ($p<0.001$) and cigarette usage and heart weakness has inormous impact on mortality rate on heart patients.

Key words: Cox regression, Survival Analysis

GİRİŞ

Tıbbi arařtırmalarda klasik çözümlene yöntemleri yetersiz kalmaktadır. Bunun birinci nedeni, çoęu kez arařtırmanın deęerlendirilmesinin tüm hastalar ölmeyen veya incelenen sonuç ortaya çıkmadan yapılması gerektięi, aksi durumda hangi tedavi yönteminin daha iyi olduęunun ve hastalığı etkileyen faktörlerin saptanmasının yıllar alabileceęidir. İkinci nedeni ise, hastalara uygulanan tedavilerin aynı zamanda başlamamasıdır. Bu tip çalışmalarda yaşam çözümlene yöntemleri daha uygun sonuçlar vermektedir (1).

Yaşam (saękalım) çözümlene özellikle tıbbın çeşitli alanlarında risk etkilerini, prognostik (belirtici) öğeleri ve tedavi başarılarını ortaya koyabilmesi açısından önem taşımaktadır.

Yaşam çözümlene belirli bir hastalığa maruz kalan bir bireyin hastalığın tanısından sonra, uygulanan bir girişimden (tıbbi tedavi, operasyon, kemoterapi vb.) sonra daha ne kadar yaşayabileceğini ya da hastalığın ne kadar sürede tekrarlanabileceğini tahmin etmek tedavi tiplerinin ve dięer faktörlerin yaşam süresine etkilerini incelemek amacıyla geliştirilmiş yöntemler ailesidir (2).

Bu yöntemlerden en etkili gelişme görülenler ise;

- Yaşam fonksiyonunun tahmini için kullanılan “Kaplan Meier Yöntemi” (3).
- İki yaşam dağılımının karşılaştırılması için kullanılan “Log-Rank Test İstatistięi” (4).
- Yaşam süresi üzerinde açıklayıcı deęişkenlerin etkilerini ölçebilmek için kullanılan “Cox Regresyon Modeli” (5).

Yaşam çözümlenmesinde izlenen tüm bireylerin belli bir konuma (ölüm gibi) erişmiş olmaları beklenmez. Araştırma süresi içinde izlenen bireylerden bazılarının izlemden çıkmaları (kaybolmak); ana konu dışındaki nedenlerle ölmeleri (bir kanser araştırmasında kişinin trafik kazasında ölmesi) veya araştırmacı tarafından izlemden çıkarılmaları (ikincil bir hastalığın ortaya çıkarak izlenen özellikleri maskeleyebilmesi) da söz konusu olabilir.

Bütün bu gözlemlere “tamamlanmamış” (censored) gözlemler denir. Censoreal gözlemler araştırmacının konusu doğrultusunda değerlendirilirler.

Bir izlem araştırmasında incelenen bağımlı değişken (yaşam süresi), bir hastalığa yakalanan bireylerin ölüm zamanlarına kadar geçen izlem süreleri (ölüm) ise; açıklayıcı değişkenler ve bu değişken üzerinde etkide bulunan faktör değişkenler (yaş, cinsiyet, tedavi türü v.b.) olur. Tamamlanmamış izlem verilerinde, bağımlı değişken ile bağımsız değişkenler (prognostik factors, covariables) arasındaki neden-sonuç bağıntısını ortaya koymak için yararlanılan regresyon yöntemine Cox regresyon yöntemi denmektedir.

Sonuç değişken ile açıklayıcı değişkenlerin uyması gereken koşullar vardır. Bağımlı değişken normal dağılım göstermeli, bağımsız değişkenler ise ardışık bağımlı olmamalıdır. Yaşamsal verilerde prognostik değişkenler normal dağılım göstermemekte ve birbirleri ile aralarında orantısal (proportional) ilişkiler vardır.

Bu nedenle yaşam çözümlenmesinde en çok kullanılan regresyon modeli Cox regresyon modelidir. 1972 yılında Cox tarafından geliştirilen regresyon modeli ile yaşam çözümlenmesinde önemli adımlar atılmıştır, Cox'un önerileri (6), Kalbfleisch ve Prentice'in (7) katkıları ile bugünkü önemini kazanmıştır.

Son yıllarda yapılan alıřmalarda Cox regresyon modeli sabit aıklayıcı deęiřkenlerin yanısıra zamana baęlı aıklayıcı deęiřkenleri de ieren Cox regresyon modeline geniřlemiřtir.

Bu alıřmanın amacı, Cox regresyon modelinde zamandan baęımsız (sabit) ve zamana baęlı aıklayıcı deęiřkenlerin kullanımını incelemek, gerek bir veri kmesine bu modelleri uygulamak ve elde edilen sonuları yorumlamaktır.

BÖLÜM 1

GENEL BİLGİLER

1.1 Bazı Kavramların Açıklanması

Yaşam Analizinde geçen bazı kavram ve gösterimler aşağıda kısaca açıklanmıştır.

- Yaşam Süresi (İzlem süresi, Survival time): Bir bireyin belirli girişime ya da etkene maruz kaldıktan sonra iyileşmesine, hastalığının tekrarlamasına ya da ölümüne kadar geçen süreye Yaşam Süresi (izlem süresi) adı verilmekte ve t ile gösterilmektedir.
- Yaşam Fonksiyonu (Survival function, Survival curve): Yaşam sürelerinin olasılık dağılımına, Yaşam dağılımı fonksiyonu ya da Yaşam eğrisi adı verilmektedir. Fonksiyon, Yaşamsal verilerin genel eğilimini matematiksel bir modelle ifade eder. Yaşam fonksiyonu bir olasılıktır ve $Y(x)$ ile gösterilmektedir.
- Yaşam Yoğunluk Fonksiyonu: Sağ olan bir kişinin t zamanı için daha ne kadar yaşayabileceğini ortaya koyan olasılıklardır. Bu olasılıklar, belirli t zamanları için birlikte hesaplanmakta ve p ile gösterilmektedir.
- Anlık Ölüm Olasılığı (Hazard function, hazard rate): Sağ olan bir kişinin, belirli bir zamanda (anda) ölüm olasılığı, taşıdığı ölüm riskidir ve λ ile gösterilir.
- Ölüm Yoğunluk Fonksiyonu: Sağ olan bir kişinin t zamanındaki ölüm riskini ortaya koyan olasılıklardır. Bu olasılıklar, belirli t zamanları için birlikte hesaplanmaktadır ve f_i ile gösterilir.
- Yığılımlı Yaşam Fonksiyonu (Yığılımlı yaşam olasılığı, cumulative hazard function): T zamanı içinde belirli bir t zamanı (anı) için

hesaplanmış olan ölüm olasılıklarının yığılımlı fonksiyonudur ve $\Lambda(t)$ ile gösterilir.

BÖLÜM 2

YAŞAM ANALİZ YÖNTEMLERİ

Yaşam Analizi, veri yapılarına ve test edilecek hipotezlere bağlı olarak üç başlık altında incelenir. Bu yöntemler aşağıdaki isimlerle anılırlar.

- 1- Yaşam Tablosu Yöntemi (YT, Actuarial Life Table Method, Cutler-Ederer Method)
- 2- Kaplan-Meier Yöntemi (KM, Product Limit Method)
- 3- Cox Regresyon Yöntemi (Cox Orantısal Riskler Regresyon Yöntemi, Cox Proportional Hazards Regression, Cox Regresyon)

Bu yöntemler kısaca birer örnek ile açıklanmıştır.

2.1 Yaşam Tablosu Yöntemi

Yaşam Tablosu (YT) yöntemi, yaşam süresi verilerini eşit zaman aralıklarına göre frekans tablosuna dönüştürerek analiz eden ve her bir zaman aralığında yaşam fonksiyonlarını hesaplamayı amaçlayan bir yöntemdir.

YT yöntemi, ölüm düzeylerinin ölçülmesi ve belirli bir yılda doğan kuşağın (kohort) herhangi bir yaşta, beklenen yaşam süresini tahmin etmek için geliştirilmiştir. Sonraları uygulama alanı genişleyen yaşam tablosu yöntemi, nüfus yapısı ve özellikleri, 0 yaşta beklenen yaşam ümidi, sağlıklı ve hastalıklı yaşam süresi, tedaviden sonra hastanın kaç yıl yaşayacağı gibi konularda tahmin ve projeksiyon aracı olarak kullanılmaya başlanmıştır.

YT yöntemi, yaşam sürelerinin k sayıda ($k > 6$) sayıda eşit aralıklı sınıfa ayrıldığı, eşit (tekrarlanan) zaman sürelerinin fazla olduğu, birim

sayısının 100'ün üzerinde ($n > 100$) olduğu veri setlerinde X etkenine maruz kalan birimlerin yaşam ve ölüm olasılıklarını, ortalama yaşam süresini hesaplamayı sağlayan bir yöntemidir.

YT yönteminde yaşam ve ölümlle ilgili fonksiyonların, parametre tahminlerinin hesaplanması aşağıdaki gibi yapılmaktadır.

2.1.1 Yaşam fonksiyonunun hesaplanması

YT yönteminde her sınıftaki birimlerin eşit ölüm riskine sahip olduğu varsayılır. i . aralıkta ölen hastalar dışındaki hastalar diğer $(i+1)$ nci sınıfa eşit ölüm riski / yaşam olasılığı ile geçerler.

i 'nci aralıktaki hasta sayısı, $r_i, r_i = n_i - 0.5 * c_i$ şeklinde hesaplanır.

Burada;

$n_i = i$ 'nci aralıkta yaşayan hasta sayısını (takipten çıkan, kayıp hastalar dahil) belirtir.

YT yönteminde her aralıktaki hastaların buldukları aralığa eşit olarak dağıldığı (Uniformly distributed) varsayılır. Bu varsayıma göre, yaşayan kişilerin sınıf değerinde (sınıf orta noktasında) riske maruz kaldıkları varsayılır.

i 'nci aralıktaki kişilerin taşıdığı ölüm riski:

$$q_i ; q_i = d_i / r_i \quad (2.1)$$

şeklinde hesaplanır.

Burada d_i , i 'nci aralıkta ölen kişilerin sayısıdır.

$$\text{Yaşam olasılığı: } p_i, p_i = 1 - q_i \quad (2.2)$$

şeklinde hesaplanır.

$$i\text{'nci aralıktaki yığılımlı yaşam olasılığı: } Y_i \text{ ise } Y_i = p_i * Y_{i-1} \quad (2.3)$$

şeklinde hesaplanır.

i 'nci aralıkta yaşayanların $(i+1)$. Aralığa canlı olarak geçeceği varsayılır ve izlem periyodunda araştırmaya dahil olacakları varsayılır. Burada $Y_1 = 1$ olarak alınır.

Y_i 'nin standart hatası ise;

$$\text{s.e.}(Y_i) = \sqrt{\sum_{j=1}^i \frac{q_j}{r_j p_j}} \quad (2.4)$$

şeklinde hesaplanır.

Ölüm olasılığı yoğunluk fonksiyonu;

$$f_i = \frac{Y_i - Y_{i+1}}{h_i} \quad (2.5)$$

şeklinde hesaplanır.

Burada f_i , i 'nci aralığın sınıf orta noktasındaki ölüm olasılığını vermektedir. h_i , sınıf aralığıdır.

Ölüm olasılığının standart hatası aşağıdaki gibi hesaplanır.

$$SH(f_i) = \frac{Y_i q_i}{h_i} \sqrt{\sum_{j=1}^{i-1} \frac{q_j}{(r_j p_j)} + \frac{p_i}{(r_i q_i)}} \quad (2.6)$$

2.1.2 Hazard (Ölüm) ve olasılık yoğunluk fonksiyonları

Sağ olan bir kişinin belirli bir zamanda ölme riskini, ölüm oranını veren fonksiyonlardır.

YT yönteminde Hazard fonksiyonu aşağıdaki gibi hesaplanır.

$$\lambda_i = \frac{2q_i}{h_i(1+p_i)} \quad (2.7)$$

Burada h_i sınıf aralığı, q_i , ölüm olasılığı ve p_i yaşam olasılığıdır.

λ 'nin standart hatası ($SH(\lambda)$) aşağıdaki gibi hesaplanır.

$$SH(\lambda_i) = \lambda_i \sqrt{\frac{1 - (h_i \lambda_i / 2)^2}{r_i q_i}} \quad (2.8)$$

Burada r_i , i 'nci aralıktaki hasta sayısıdır.

Ölüm ve yoğunluk fonksiyonları, verileri açıklamak için seçilecek modelleri belirleme imkanı veren ve yaşamsal veriler için alternatif açıklamalar yapmayı kolaylaştıran fonksiyonlardır.

2.1.3. Ortanca yaşam süresi

İzlem sürelerini içeren veri setlerinde değişim genişliği büyük olabilir ve uç değerler bulunabilir. Bu nedenle “ortalama yaşam süresi” yerine “ortanca yaşam süresi” hesaplanır. Ortanca yaşam süresi aşağıdaki gibi hesaplanır.

$$\theta_{50} = t_i + \frac{Y_{i-1} - 0.5}{f_i} \quad (2.9)$$

Burada, t_i , i 'nci aralığın yığılımlı yaşama olasılığıdır ve 0.5'ten küçük olan önceki sınıfın sınıf başlangıç değeridir.

Y_{i-1} : t_i ile belirlenmiş aralıktan önceki yığılımlı yaşam olasılığıdır ve 0.5'ten büyüktür.

2.2 Kaplan –Meier (KM) Yöntemi

Kaplan-Meier (KM) Yöntemi (Product Limit Method), yaşam sürelerine ilişkin, verileri zaman aralıklarına bölmeden yaşam ve ölüm fonksiyonlarını hesaplamayı sağlayan bir yöntemdir. İzlem sürelerinin tekrarlandığı birimlerin bulunmadığı ve izlem sürelerinin birbirlerini büyüklük sırasına göre izledikleri durumlarda YT yöntemine tercih edilen bir yöntemdir.

Yaşam Tablosu (YT) ve KM yöntemleri yaşam ve ölüm fonksiyonlarını hesaplamada benzerlikler ve farklılıklar içermektedir. Bu farklılıklar aşağıdaki gibi özetlenebilir:

1. KM yönteminde az sayıda bireyle çalışılabilir. YT yönteminde aralıklara düşen birim sayısının azalması tahminleri etkilemektedir. Çok sayıda birey olduğunda YT ve KM yöntemi benzer sonuçlar vermektedir. YT yöntemi çok sayıda izlenen birim bulunduğu takdirde tercih edilmelidir. KM yönteminde, izleme süresini belirli zaman dilimlerine (gruplara) ayırmaya gerek duyulmamaktadır.

2. KM yönteminde tekrarlı ölçüm zamanlarına ilişkin olasılıklar hesaplanmaz. İzlem zamanlarının küçükten büyüğe doğru dizildiği serilerde ($t_i < t_{i+1} < \dots < t_{i+k}$) tekrarlı ölçüm çok az ise KM, YT'ye tercih edilir.

3. KM yönteminde kayıplar, eksik veriler dikkate alınmaz, sadece ölümler üzerinden yaşam olasılığı hesaplanır. Yaşam olasılığı ise ölüm olayının gerçekleştiği ana ilişkin olarak hesaplanır. YT yönteminde ise sınıf orta değerine göre tahmin yapılır.

4. KM yönteminde kesin ölüm tarihi kullanıldığı için, nokta yaşam olasılığı bulunur. YT yöntemi ise yaklaşık bir olasılık verir. Çünkü izlem süreleri k gruba (sınıf) ayrılarak incelenir ve tahminler sınıf değerine göre hesaplanır.

2.2.1 Yaşam fonksiyonunun hesaplanması

N birimin gözlenen yaşam süreleri $t_1 < t_2 < t_3 < \dots < t_N$ biçiminde sıralanmış olsun. N bireyin yaşam fonksiyonu;

$$Y(t) = \prod_{t_i < t} \frac{N - i - 1 \delta_i}{N - i + 1} \quad (2.10)$$

şeklinde hesaplanır. Bu formülde;

$$\delta_i = \begin{cases} 1 & \text{eğer } t_i \text{ ölüm zamanı ise} \\ \text{eğer } t_i \text{ canlı izleme zamanı ise (kaybolan, tamamlanma mı, censored)} & \end{cases}$$

şeklinde belirlenir.

Y (t)'nin standart hatası;

$$s.e.Y(+)=Y(t)\sqrt{\sum \frac{\delta_i}{(N-i)(N-i+1)}} \quad (2.11)$$

biçiminde hesaplanır.

Y (t), her cevap zamanı için değişebilen aşamalı bir fonksiyondur. t zamanı değiştikçe Y (t) fonksiyonu da kademeli olarak değişim gösterir. Benzer izlem zamanına sahip gözlemler (tied observations) bulunması durumunda, hesaplanan değer, dizideki canlı gözlemi ifade eden sıradaki değerden bir önceki ölümü belirten birime ilişkin olarak hesaplanır.

2.2.2 Yığılımlı ölüm fonksiyonu (Cumulative Hazard Function)

Yığılımlı ölüm fonksiyonu ($\hat{\Lambda}(t)$), tahmin edilen P (t) yaşam fonksiyonunun eksi işaretli tabii logaritması olarak hesaplanır ve Peterson tahmincisi olarak bilinir.

$$\hat{\Lambda}(t)=\ln P(t) \quad (2.12)$$

biçiminde hesaplanır.

2.2.3 Ortalama yaşam süresi (Mean Survival Time)

İzlenen n birimin ortalama yaşam süresini belirtir. Örneğin izlenen hastalardan ölenlerin sayısı D tane ve bunların ölüm zamanları da $t_1 < t_2 < t_3 < \dots < t_D$ olarak sıraya dizilmiş olsun.

Bu dizide ortalama yaşam süresi tahmini (μ);

$$\mu = t_1 + \sum_{i=1}^{D-1} Y(t_i)(t_{i+1} - t_i) \quad (2.13)$$

şeklinde belirlenir.

Yukarıdaki $Y(t_i)$, t_i 'deki tüm verilere dayalı olarak hesaplanan Kaplan-Meier yaşam olasılığıdır.

μ 'nün varyansı $\text{Var}(\mu)$;

$$\text{Var}(\mu) = \left[\sum_j \frac{s_j^2}{(n-j)(n-j+1)} \right] \left[\frac{D}{D-1} \right] \quad (2.14)$$

formülü ile elde edilir. Burada s_j ;

$$s_j = \sum_{i=j}^{D-1} Y(t_i)(t_{i+1} - t_i) \quad (2.15)$$

biçiminde hesaplanan değerdir.

2.2.4 Çeyrek değerler (Quartiles)

KM yöntemine göre yapılan hesaplamalarda, $t(q)$ q , çeyrek değerlerin tahminini belirtir. $q=0.5$ medyan değeri, $q=0.25$ 1.çeyrek değeri $q=0.75$ ise 3.çeyrek değeri ifade eder. t_i zamanındaki yaşam fonksiyonu tahmini $p(t_i)$ olarak alındığında çeyrek değerler;

$$\begin{aligned} t(q) &= \inf\{t_i \mid Y(t_i) \leq q\} \quad \text{ve} \\ u(q) &= \sup\{t_i \mid Y(t_i) \geq q\} \end{aligned} \quad (2.16)$$

alınarak, $t(q)$ 'deki olasılık yoğunluk fonksiyonu tahmini hesaplanabilir.

$$\text{Eğer, } u(q + 0.05) = t(q - 0.05) \text{ ise,} \quad (2.17)$$

Olasılık yoğunluk fonksiyonu

$$f[t(q)] = \frac{Y[u(q + 0.05)] - Y[t(q - 0.05)]}{t(q - 0.05) - u(q + 0.05)} \quad (2.18)$$

olarak, varyansı ise;

$$\text{Var}[t(q)] = \left[\frac{1}{f[t(q)]} \right] \text{var}[Y[t(q)]] \quad (2.19)$$

olarak hesaplanır.

Bazı durumlarda aynı hastalığa yakalanan bireyler iki yada daha fazla gruba ayrılarak her gruba farklı bir tedavi yöntemi uygulanabilir (örneğin; farklı ilaç, farklı ameliyat v.b). Farklı yöntemlerle tedavi edilen hastalar için birden fazla yaşam fonksiyonu hesaplanabilir. Bu yaşam fonksiyonlarının birbirleriyle olan farklılıkları test edilebilir. Farklı gruplara göre elde edilen yaşam fonksiyonlarının karşılaştırılmasını yapmak için yaygın olarak kullanılan dört test yöntemi vardır.

Bunlar; Mantel-Cox Testi, Log-Rank Testi, Breslow Wilcoxon Testi ve Tarone-Ware Testi olarak sayılır.

2.3 Cox Regresyon Yöntemi

Bir izlem araştırmasında incelenen bağımlı değişken (Yaşam Süresi), bir hastalığa yakalanan bireylerin ölüm zamanlarına kadar geçen izlem süreleri (ölüm) ise; açıklayıcı değişkenler, bu değişken üzerinde etkide bulunan faktör değişkenler (Yaş, cins, tedavi türü v.b) olur. Tamamlanmamış izlem verilerinde, bağımlı değişken ile bağımsız değişkenler (prognostik factors, covariates) arasındaki neden-sonuç bağıntısını ortaya koymak için yararlanılan regresyon yöntemine Cox regresyon yöntemi adı verilir.

Regresyon analizinde sonuç değişken ile açıklayıcı değişkenler arasındaki neden-sonuç ilişkisi kurulurken bu değişkenlerin uyması gereken koşullar vardır (bağımlı değişken normal dağılmalıdır, bağımsız değişkenler ardışık bağımlı olmamalıdır v.b.). Yaşamsal verilerde prognostik değişkenler normal dağılım göstermemekte, birbirleri ile korelasyon göstermekte ve aralarında orantısal (proportional) ilişkiler bulunmaktadır. Bu nedenle yaşamsal verilerin neden sonuç ilişkilerini analizde çoklu regresyon analizi

uygulanamamaktadır. Yaşamsal verilerin nedensellik analizlerinde Cox regresyon yönteminden yararlanılır.

Cox'un (1972) Orantısal Ölüm Riski Bağintı Modeli (Proportional Hazard Regression Model) ölüm hızlarının bu değişken değerlerinin loglinear fonksiyonu olduğu varsayımından yola çıkarak geliştirilmiştir. Yaşamsal verilerde prognostik değişkenler arasında orantısal bir etki söz konusudur.

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 ölçeği 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 (6).

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 hala 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.

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

Yaşam çözümlenmesinde 1972 yılında Cox tarafından geliştirilen regresyon modeli ile yaşam çözümlenmesinde önemli adımlar atılmış, Cox'un (5) önerileri, Kalbfleisch ve Prentice'in (6) katkıları ile bugünkü önemini kazanmıştır.

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

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

Eğer Z_i kadar prognostik değişken, yaşam süresinin loglinear bir fonksiyonu olarak, loglinear modellerle incelemeye alınabiliyorsa, Cox'un Propotional Hazard Regression yönteminden yararlanarak her bir prognostik değişkenin yaşam süresi üzerine etkisi incelenebilir.

Z_i prognostik değişkenler vektörü z ve yaşam süresi t olsun. Bir bireyin Z_i ortak değişkenlerine göre ölüm fonksiyonu $h(t; z)$ olarak alınabilir. Buna göre orantısal ölüm modeli (Proportional hazard model, bir başka anlatımla z açıklayıcı değişkenlere göre regresyon modeli);

$$h(t; z) = h_0(t) \exp(\beta'z) \quad (2.20)$$

olarak yazılır.

Bu modelde; β , regresyon katsayıları vektörü; $h_0(t)$, $z=0$ olduğunda temel (baseline) ölüm fonksiyonudur.

Cox regresyon modelinde iki temel varsayım vardır. Bunlar;

1. Prognostik değişkenlerin ölüm fonksiyonu üzerine etkileri loglineer'dir.
2. Prognostik değişkenlerin loglinear fonksiyonu ile ölüm fonksiyonu arasında çarpımsal bir ilişki vardır.

Bu varsayımlara göre, farklı prognostik değişken setlerine sahip olan iki birimin ölüm fonksiyonları oranı zamana bağlı değildir, ölüm riskleri orantısaldır.

Yaşamsal verilerde neden-sonuç ilişkilerinin ortaya konmasında Cox tarafından önerilen $h(t; z) = h_0(t) \exp(\beta'z)$ regresyon modeli kullanılır.

z ortak değişkenler matrisinin tek ya da çok değişkenli olmasına göre regresyon katsayıları aşağıdaki modellere göre tahmin edilir.

Tek değişkenli Cox regresyon modeli;

$$h(t) = [h_0(t)] x e^{\beta X_1} \quad (2.21)$$

biçiminde yazılır.

Çoklu Cox regresyon modeli,

$$h(t) = [h_0(t)] e^{(\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p)} \quad (2.22)$$

biçiminde yazılır.

Bu eşitlikte X_1, X_2, \dots, X_p ortak değişkenlerdir. Ortak değişkenler, yaşam süresine etkide bulunan yaş, kan basıncı, sıcaklık gibi Sürekli değişkenler ya da hastalık evresi, histolojik evreler gibi kategorik değişkenler olabilir. Eğer ortak değişkenler setinde kategorik değişkenler varsa bunların orijinal kategorileri göz önüne alınıp, yeni değişken veri setleri oluşturularak transform edilmesi gerekir. Kategorik

değişkeni ifade eden yeni transform değişkenin kategori sayısı, orijinal kategori sayısından az olması gerekir.

2.3.1 Çoklu Cox modelinde β katsayılarının hesaplanması

Gözlenen n tane yaşam süresi arasından k tanesi sıralanmış olarak ($t_1 < t_2 < \dots < t_k$), ölüm sonucu olan verileri gösterebilir. Bir R_i setinde, t_i zamanında değerleri saptanan z_i ortak değişken vektörü belirlenmiş olsun. Bir prognostik değişkenin, yaşam süresi üzerine etkide bulunan tüm değişkenler dikkate alınarak belirlenecek genel risk içindeki oranı riskler oranı biçiminde belirlenir.

$$\exp(\beta'z_i) / \sum_{j \in R_i} \exp(\beta'z_j) \quad (2.23)$$

k farklı ölüm zamanlarının bu oranlarla çarpımı, kısmi benzerlik fonksiyonunu verir. Regresyon katsayıları bu kısmi benzerlik fonksiyonu yardımı ile tahmin edilirler.

Kısmi Benzerlik Fonksiyonu $L(\beta)$;

$$L(\beta) = \prod_{i=1}^k \exp(\beta'z_i) / \sum_{j \in R} \exp(\beta'z_j) \quad (2.24)$$

biçiminde hesaplanır.

β katsayılarının en büyük benzerlik tahminleri, logaritmik benzerlik fonksiyonunu en büyükleyerek aşağıdaki gibi hesaplanır.

$$\ln L(\beta) = \sum_{i=1}^k \left\{ \sum_{j \in D_i} \ln r(\beta_j z_j(t_i)) - m_i \ln \left[\sum_{l \in R_i} r(\beta z_l(t_i)) \right] \right\} \quad (2.25)$$

Burada;

D_i : t_i zamanında ölen kişilerin seti

M_i : t_i zamanındaki cevapların sayısı

R_i : t_i zamanında canlı olan kişilerin setini belirtmektedir.

$Z_i(t_i)$: t_i zamanında i.birey için p-ortakdeğişken vektörü

$$r(\beta, z_i(t_i)) : \exp(\beta, z_i(t_i)) \quad (2.26)$$

β katsayılarının hesaplanmasında Newton-Raphson algoritması kullanılır ve ardışık tekrarlanan çözümlenmeler ile (iterative parameter estimation) β 'nin tahminleri yapılır.

Veri setinde benzer süre gözlemleri olduğunda $L(\beta)$ 'nin maximizasyonu Breslow (1974) tarafından ileri sürülen yaklaşım,

$$L(\beta) = \left\{ \prod_{i=1}^k \exp(\beta' z_i) / \left[\sum_{j \in R_i} \exp(\beta' z_j) \right]^{m_i} \right\} \quad (2.27)$$

ile hesaplanır.

2.3.2 β katsayılarının önemliliğinin test edilmesi

β katsayılarının önemliliği için $H_0:\beta=0$ hipotezi test edilir. Bu amaçla üç test yöntemi ileri sürülmüştür. Bunlar;

- Wald testi
- Benzerlik Oranı (Likelihood, LR) testi
- Score testidir.

Wald testi

Wald testi en büyük benzerlik tahminlerinin (MLE) normal dağıldığı varsayımına dayanır. Regresyon katsayısının, standart hatasına oranı;

$$z = (\beta / SH_{\beta}) \quad (2.28)$$

Wald istatistiği olarak adlandırılır. Bu durumda Wald istatistiği Standart Normal Dağılım (SND) gösterir ve SND'nin kritik değerleri ile karşılaştırılarak önemliliği belirlenir. Aynı zamanda Wald istatistiği,

$$w = z^2 = (\beta / SH_{\beta})^2 \quad (2.29)$$

olarak da kullanılabilir. Bu durumda ise Wald test istatistiği 1 serbestlik dereceli kare dağılımı gösterir ve 1 serbestlik dereceli χ^2 dağılımının kritik değerleri ile karşılaştırılarak önemliliği belirlenir.

Benzerlik Oranı Testi (LRATIO)

LRATIO testi Wald testinden daha genel bir testtir. Kategorik değişkenlerin iki ve daha fazla sınıfı bulunduğu ve Cox modeline aynı anda birçok değişkenin alındığı durumlarda tercih edilir.

Cox regresyonda regresyon katsayılarının önemliliği $H_0: \beta_i = 0$ $i=1,2,\dots,p$ varsayımı test edilerek belirlenir. Modelin en büyük benzerlik istatistiği I_0 değeri ile modelde v tane birbirinden farklı ve Yaşam Süresi üzerine önemli etkide bulunan başka bir modelin en büyük benzerlik istatistiği I_v belirlenerek LRATIO test istatistiği hesaplanır.

Modelde, v farklı regresyon katsayısı olduğunda bu modelin maximum benzerliği I_v ile gösterilir. Her bir modelin $\beta=0$ durumuna göre benzerlik oranı (LR) test edilir.

$$\begin{aligned} LR &= -2 \log (I_0 / I_v) \\ LR &= -2 (L_0 / L_v) \end{aligned} \quad (2.30)$$

biçiminde hesaplanır.

Burada, $L_v = \log I_v$ ve $L_0 = \log I_0$ 'dir.

LR istatistiği v serbestlik dereceli ($sd=v$) χ^2 dağılımı gösterir. Önemliliğinin belirlenmesinde $sd=v$ serbestlik dereceli χ^2 dağılımının kritik değerlerinden yararlanılır. (SPSS, 1998; Parmar & Machin, 1995)

Score (Skor) Testi

Skor testi S ile gösterilir. Logaritmik benzerlik istatistiklerinden yararlanılarak hesaplanır. İkili değer alabilen (binary) değişkenin P gözlenme oranına göre r kez gözlenme olasılığından yararlanılarak logaritmik benzerlik istatistiği;

$$L = \log I = r \log \pi + (n - r) \log(1 - \pi) \quad (2.31)$$

şeklinde hesaplanır.

Skor istatistiği, pay'a L 'nin birinci türevinin karesi, payda'ya ise L 'nin ikinci türevi alınarak belirlenir. S istatistiği payın paydaya oranı olarak bulunan bir değerdir. Modeldeki değişkenler Sürekli olduğunda ya da birden fazla değişken bulunduğu Skor testi Mantel-Haenszel logrank testinin bir genellemesidir.

Skor test istatistiği, sd tahmin edilen parametre sayısı olmak üzere sd serbestlik dereceli χ^2 dağılımı gösterir. Önemliliği sd serbestlik dereceli dağılımının kritik değerleri ile karşılaştırılarak yapılır.

2.3.3 Zamana bağlı açıklayıcı değişkenler durumunda Cox regresyon modeli

Eşitlik (2.20) 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 deęişebilen herhangi bir deęişken olarak tanımlanmaktadır. Zamandan bağımsız deęişken ise incelenen birey için deęeri zaman içinde sabit kalan bir deęişkendir. En yaygın tanımlanan zamana baęlı deęişken, zamandan bağımsız bir deęişken ile zamanın ya da zamanın bir fonksiyonunun çarpımı biçimindedir (8,10).

Zamana baęlı açıklayıcı deęişkenin üç türü vardır. Bunlardan biri tanımlanmış (defined) zamana baęlı açıklayıcı deęişkendir. Bu, sabit bir açıklayıcı deęişken ile zamanın bir fonksiyonunun çarpımı şeklinde elde edilen deęişkendir. Zamana baęlı deęişkenlerin bir dięeri içsel (internal) deęişkendir. İçsel deęişkenler için deęişkenin deęerindeki deęişimin nedeni içsel özelliklere, yani bireyin kendi davranışlarına ve özelliklerine dayanmaktadır. Bu gibi deęişkenlere örnek olarak, t zamanındaki sigara içme durumu ve t zamanındaki obezite seviyesi verilebilir. İçsel deęişkenler için deęişkenin deęerindeki deęişimin nedeni içsel özelliklere, yani bireyin kendi davranışlarına ve özelliklerine dayanmaktadır. Dięer bir zamana baęlı deęişken yardımcı (ancillary) deęişkendir. Deęişkenin deęerindeki deęişimin nedeni dışsal, çevreye ait özelliklerdir. Özel bir coęrafi alan için t zamanındaki hava kirlilięi indeksi dışsal deęişkene bir örnek olarak verilebilir (9,11).

Kısmen içsel, kısmen dışsal olabilecek deęişkenler de vardır. Buna örnek olarak ciddi bir kalp rahatsızlıęı olan bir hastanın kalp nakli için uygun olup olmaması durumunu düşünebiliriz. Kalp nakli durumuna ilişkin deęişken, bireyin ayırt edici özelliklerinden dolayı içsel olarak, tıbbi vericinin elde edilebilirliğinden dolayı yardımcı deęişken olarak adlandırılmaktadır (9,11).

Bir çok durumda açıklayıcı deęişken verisi uzun dönemde elde edilmektedir. Örneęin kan basıncı, CD4 sayımı, görelilię ve hastalık öyküsü seçilen periyodik zaman noktalarında elde edilebilir. Tedavi ya da dięer etkiler zaman içerisinde deęişebilir. Uygun

istatistiksel modelde zamanla deęişen açıklayıcı deęişken bilgisini kullanmak gerekir. Bunu yapmanın yollarından biri de zamana baęlı açıklayıcı deęişkenli Cox regresyon modelini kullanmaktır (9).

Zamana baęlı açıklayıcı deęişkenler olduęu durumda, cox regresyon modeli zamandan baęımsız deęişkenleri ve zamanın bazı fonksiyonları ile bu deęişkenlerin çarpımını içeren bir modele genişlemektedir. x_1, x_2, \dots, x_p zamandan baęımsız deęişkenler ve $x_1(t), x_2(t), \dots, x_{p2}(t)$ zamana baęlı deęişkenler olmak üzere açıklayıcı deęişkenler,

$$x(t) = (x_1, x_2, \dots, x_{p1}, x_1(t), x_2(t), \dots, x_{p2}(t)) \quad (2.31)$$

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

$$h(t, x(t)) = h_0(t) \exp[(\beta' x) + \delta' xg(t)] \quad (2.32)$$

biçiminde yazılmaktadır. Burada $g(t)$ zamanın bir fonksiyonu olarak tanımlanmaktadır. $g(t)$ 'nin seçimi kullanılan deęişkenlerin durumuna göre ve arařtırmacının bilgisine göre deęişkenlik göstermektedir. Bu fonksiyon genellikle $t, \log t$ ($\ln t$) ya da adım fonksiyonları biçiminde tanımlanmaktadır (9,10).

Cox regresyon modeli, zamanla deęişen açıklayıcı deęişkenlerle de kullanılabilir. Uygulamada kullanımları sabit (zamandan baęımsız) açıklayıcı deęişkenlere göre daha karmaşıktır. Ayrıca hatalı çıkarsama ve modelleme için potansiyel gittikçe artmaktadır (9).

BÖLÜM 3 UYGULAMA

Çalışmada 2001-2005 yılları arasında Marmara Üniversitesi Hastanesi Kalp Damar Cerrahisinde tedavi gören 481 kalp hastasına ait veriler kullanılmıştır. Hastalar ameliyat edildikten sonra 5 yıl boyunca takipleri yapılmıştır. Kalp ameliyatı geçirmiş hastalarının yaşam süreleri ve yaşam sürelerini etkileyen faktörler Cox regresyon modeli kullanılarak belirlenmeye çalışılmıştır.

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

Çalışmada; olguların yaşları, iskemik transiyel atak durumu (ita), kalbi besleyen damarlardan radial, sol ventrikül damar (sv), sol iner dal (LAD), ilaçlardan statin kullanımı, ejeksiyon fraksiyonu yani kalp yetersizliği (EF>30); shunt takılması, kronik obstrüktif pulmoner disease (kronik kalp yetmezliği); diyabet (şeker hastalığı), sigara kullanımı, ölüm görülmesi ve takip süreleri parametreleri alınarak çalışma verileri oluşturulmuştur.

Ölüm üzerine parametrelerin etkileri değerlendirildi. Takip süresi olayımızda etken bir faktör olduğundan yaşam analizi ve Cox regresyon analizi uygulaması yapılmıştır.

Çalışmada elde edilen bulgular değerlendirilirken, istatistiksel analizler için SPSS for Windows 10.0 programı kullanılmıştır. Çalışma verileri değerlendirilirken tanımlayıcı istatistiksel metodların (Ortalama, Standart sapma) yanısıra niceliksel verilerin karşılaştırılmasında student t test kullanılmıştır. Niteliksel verilerin karşılaştırılmasında ise Ki-Kare testi uygulanmıştır. Sonuçlar üzerine değişkenlerin etkileri ise

Cox regression analizi ile değerlendirilmiştir. Sonuçlar % 95'lik güven aralığında, anlamlılık $p < 0,05$ düzeyinde incelenmiştir.

Çalışmada ele alınan 481 olgunun yaşları 28 ile 88 arasında değişmekte olup ortalama yaş $67,27 \pm 10,69$ 'dur. Öncelikle tüm parametrelerin ölüm üzerine etkilerini univariate analiz yöntemleri ile değerlendirdik.

3.1 Tüm Parametrelerin Ölüm Üzerine Etkileri

Tablo 1: Ölüm durumuna göre değerlendirmeler

	Ölüm		Test değ,p
	Var(n=24)	Yok (n=457)	
Yaş	70,20±10,37	67,11±10,69	t:1,382; p:0,168
Takip zamanı	25,50±14,12	38,41±12,74	t:4,812; p:0,001**
ITA	22 (% 91,7)	435 (%95,2)	χ^2: 0,596; p:0,440
Radial	21 (%87,5)	323 (%70,7)	χ^2: 3,168; p:0,075
Sol ventrikül	18 (%75,0)	378 (%82,7)	χ^2: 0,932; p:0,334
LAD	22 (%91,7)	437 (%95,6)	χ^2: 0,818; p:0,366
Statin	16 (%66,7)	293 (%64,1)	χ^2: 0,065; p:0,799
Diyabet	7 (%29,2)	118 (% 25,8)	χ^2: 0,133; p:0,716
Sigara	4 (%16,7)	16 (%3,5)	χ^2: 9,918; p:0,002**
EF>30	13 (%54,2)	66 (%14,4)	χ^2: 26,214; p:0,0001**
COPD	5 (%20,8)	42 (% 9,2)	χ^2: 3,506; p:0,061

t: student t testi

χ^2 : Ki-kare testi

**p<0,01

Ölüm üzerine yaş parametresinin, İskemik transiyel atak geçirme oranı istatistiksel olarak anlamlı bir etkisi görülmemiştir ($p > 0,05$). Radial damar oranı ölen olgularda yüksek olmakla beraber anlamlı değildir ($p > 0,05$). LAD, Statin kullanımı ve diyabet oranı ölüm üzerine anlamlı bir etki göstermemiştir ($p > 0,05$). COPD varlığı ölen olgularda daha yüksek görülmesine rağmen istatistiksel olarak anlamlı bulunmamıştır ($p > 0,05$).

Ölen olguların takip süresi elbette ki anlamlı düzeyde düşük bulunmuştur ($p<0,01$). Sigara kullanan olguların ölüm oranları anlamlı düzeyde yüksektir ($p<0,01$). EF>30 üzeri olanların ölüm oranı anlamlı düzeyde yüksektir ($p<0,01$). Ölüm durumuna göre değerlendirmeler tablo 1'de sunulmuştur.

4.2 Olguların Yaşam Tablosu

Çalışmaya alınan 481 olgunun; en son 54.ayda mortality gözlenmekte, sağkalım oranı % 88 ve standart hatası % 0,4'dür; bu grupta 457 (%95) deneğin yaşamakta ve 24 denekte ölüm gözlenmekte olup ortalama sağkalım süresi 55,76 ay dır.

Tablo 2: Olguların Yaşam Tablosu

	Zamanın başında sağ olan hasta sayısı	Zaman içerisinde ölen hasta sayısı	Kümülatif sağkalım oranı	Kümülatif sağkalım oranının SE
0 AY	481	2	0,99	0,002
6 AY	479	2	0,99	0,004
12 AY	470	2	0,98	0,005
18 AY	448	6	0,97	0,007
24 AY	380	3	0,96	0,008
30 AY	338	3	0,95	0,010
36 AY	281	4	0,94	0,012
42 AY	230	0	0,94	0,012
48 AY	162	1	0,93	0,015
54 AY	37	1	0,88	0,049

4.3 Ölüm Durumuna Göre Cox Regresyon Analizi

Bu çalışmada, modeldeki değişken β parametresi ve standart hatası (S.H.), p değeri, hazard oranı ($\exp(\beta)$) ile düzeyleri için hazard oranının alt ve üst sınırları verilmiştir. β parametresinin pozitif değer olması bu düzeyin referans kaegorisine göre daha fazla riskli olduğu, β parametresinin negatif değer olması ise bu düzeyin referans kategorisine göre daha az riskli olduğunu göstermektedir. Hazart oranı olan $\exp(\beta)$ değeri ise önemli bulunan düzeyin, referans kategorisine göre kaç kat (yada % ne kadar) daha riskli olduğu yorumunu getirmektedir. Önemli bulunan değişkenlerin her bir düzeyine karşılık gelen p değerlerine bakılarak önemli değişken düzeyleri belirlenmektedir.

Regresyona Univariate analizde ölüm üzerine etkilerini önemli bulduğumuz sigara kullanımı, EF>30 olması alınmış bunun yanında anlamlılık düzeyine yakın bulunan radial, COPD ve yaş parametresi de analize dahil edilmiştir.

Çalışmadaki değişkenlerle tek tek Cox regresyon çözümlemesi yapıldığında elde edilen sonuçlar Tablo 3'de gösterilmiştir.

Tablo 3: Ölüm Durumuna göre Cox Regression Sonuçları

Case Processing Summary		N	Percent
Cases available in analysis	Event ^a	24	5,0%
	Censored	457	95,0%
	Total	481	100,0%
Cases dropped	Cases with missing values	0	,0%
	Cases with non-positive time	0	,0%
	Censored cases before the earliest event in a stratum	0	,0%
	Total	0	,0%
	Total	481	100,0%

a. Dependent Variable: TAKIP.SÜ

Tablo 3.1**Categorical Variable Codings a,b,c**

	Frequency	(1)
RADIAL ,00	137	-,500
1,00	344	,500
SIGARA ,00	461	-,500
1,00	20	,500
EF.30 ,00	402	-,500
1,00	79	,500
COPD ,00	434	-,500
1,00	47	,500

- a. Simple Parameter Coding
b. Category variable: RADIAL
c. Category variable: SIGARA
d. Category variable: EF.30
e. Category variable: COPD

Tablo 3.2**Omnibus Tests of Model Coefficients**

Step	-2 Log Likelihood	Overall (score)		
		Chi-square	df	Sig.
1 ^a	238,233	46,098	5	,000
2 ^b	239,462	43,760	4	,000
3 ^c	241,170	42,860	3	,000

- a. Variable(s) Entered at Step Number 1: YAŞ
RADIAL SIGARA EF.30 COPD
b. Variable Removed at Step Number 2: COPD
c. Variable Removed at Step Number 3: YAŞ

Tablo 3.3**Omnibus Tests of Model Coefficients**

Step	Change From Previous Step		
	Chi-square	df	Sig.
1 ^a	35,522	5	,000
2 ^b	1,228	1	,268
3 ^c	1,708	1	,191

a. Variable(s) Entered at Step Number 1: YAŞ
RADIAL SIGARA EF.30 COPD

b. Variable Removed at Step Number 2: COPD

c. Variable Removed at Step Number 3: YAŞ

Tablo 3.4.**Omnibus Tests of Model Coefficients**

Step	Change From Previous Block		
	Chi-square	df	Sig.
1 ^a	35,522	5	,000
2 ^b	34,293	4	,000
3 ^c	32,585	3	,000

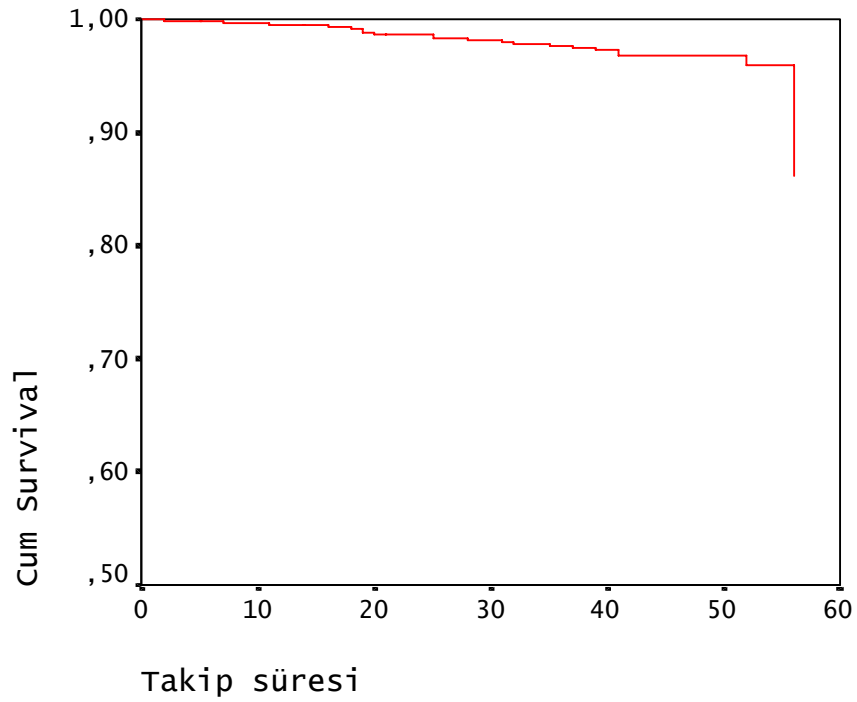
a. Variable(s) Entered at Step Number 1: YAŞ
RADIAL SIGARA EF.30 COPD

b. Variable Removed at Step Number 2: COPD

c. Variable Removed at Step Number 3: YAŞ

Tablo 3.5**Variables in the Equation**

		B	SE	wald	df	Sig.	Exp (B)	95% CI for Exp(B)	
								Lower	Upper
Step 1	YAŞ	,024	,021	1,379	1	,240	1,025	,984	1,067
	RADIAL	1,083	,638	2,881	1	,090	2,954	,846	10,32
	SIGARA	2,300	,630	13,339	1	,000	9,977	2,903	34,28
	EF.30	2,151	,467	21,248	1	,000	8,593	3,443	21,45
Step 2	COPD	,601	,513	1,375	1	,241	1,824	,668	4,985
	YAŞ	,026	,021	1,618	1	,203	1,027	,986	1,069
	RADIAL	1,128	,638	3,132	1	,077	3,091	,886	10,78
	SIGARA	2,374	,624	14,491	1	,000	10,74	3,164	36,48
Step 3	EF.30	2,183	,465	22,079	1	,000	8,873	3,570	22,06
	RADIAL	1,003	,625	2,575	1	,109	2,727	,801	9,289
	SIGARA	2,363	,623	14,376	1	,000	10,62	3,132	36,04
	EF.30	2,208	,465	22,556	1	,000	9,093	3,657	22,61



Şekil 1: Cox regresyon analizi grafiği

Yapılan Cox regresyon analizinde Backward Stepwise (Conditional LR) yöntemi kullanıldı. Cox regresyon analizi ile değerlendirdiğimizde; modelin ileri düzeyde anlamlı ($p < 0.001$) bulunduğu, analizin 3.step sonunda sonlandığı ve modelde sigara, EF>30 ve radial parametrelerinin ölüm üzerine etkilerinin anlamlı olduğu görülmektedir. Radial parametresi modelde yerini almıştır, hastanın radial olması ölümü 2.72 kat arttırıcı etkiye sahiptir, modelde istatistiksel olarak önemli bulunmamasına bizim için önemli bir parametre olarak yerini almaktadır. Olguların sigara kullanmasının ölüm üzerine etkisi çok önemli olarak bulunmuş olup ölümü 10,62 kat arttırıcı etkiye sahiptir. EF>30 parametresinin ölüm üzerine etkisi yine çok önemli bulunmuş olup yani kalp yetersizliğinin ölümü 9,09 kat arttırıcı etkiye sahip olduğu görülmüştür.

BÖLÜM 5

SONUÇ

Bu çalışmada, ölüm üzerine tüm açıklayıcı değişkenlerin önce univariate analizi yapılmış daha sonra sağkalım analizi uygulanmıştır. Sonunda ise tüm açıklayıcı değişkenlerin sabit olduğu düşünülerek Cox regresyon çözümlenmesi yapılmıştır.

Çalışmaya alınan 481 olguda en son 54.ayda ölüm gözlenmekte, sağkalım oranı % 88 ve standart hatası % 0,4'dür. 457 (%95) deneğin yaşamakta ve 24 denekte ise ölüm gözlenmekte olup ortalama sağkalım süresi 55,76 ay dır.

Cox regresyon analizine örnek olarak tıbbi çalışma sonuçlarında da Backward Stepwise yöntemi kullanıldığı ve modelin ileri düzeyde anlamlı ($p < 0.001$) bulunduğu görmekteyiz. Analizin 3.adım sonunda sonlandığı ölüm üzerine sigara, EF>30 ve radial parametrelerinin etkilerinin anlamlı olduğu görülmektedir. Radial damar ile ilgili girişimin ölümü 2.72 kat arttırıcı; sigara kullanmasının ölümü 10,62 kat arttırıcı ve kalp yetersizliğinin ölümü 9,09 kat arttırıcı etkiye sahip olduğu görülmüştür.

KAYNAKLAR

1. Cox DR, Oakes D. Analysis of Survival Data, New York: Chapman and Hall. 1984
2. Özdemir K., SPSS ile Biyoistatistik, Eskişehir 5. Baskı 2003
3. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. Journal of the American Statistical Association 1958; 53 457-481.
4. Mantel N. Evaluation of survival data and two new rank order statistics arising in its
5. Cox DR, Regression models and life-tables, Journal of the Royal Statistical Society, Series B, 1972; 34: 187-220
6. Colett D. Modelling Survival Data in Medical Research, London: Chapman & Hall 1994.
7. Kalbfleisch JD, Prentice RL. The Statistical Analysis of Failure Time Data, New York: Wiley; 1980.
8. Pettitt AN, Daud IB. Investigating time dependence in Cox's proportional hazards model, Applied Statistics 1990; 39: 313-329.
9. Kleinbaum DG. Survival Analysis: A Self-Learning Text, New York: Springer; 1996.
10. Therneau TM, Grambsch PM. Modelling Survival Data: Extending the Cox Model, New York: Springer; 2000.
11. Fisher LD, Lin DY Time-dependent covariates in the Cox proportional hazards regression model, Annual Review of Public Health 1999; 20: 145-157

EKLER**EK-1****HASTA TAKİP FORMU**

Hasta protokolü :
Tarih :
Hasta yaşı :
Sigara kullanımı :
Diyabet :
Yapılan operasyon :
ITA :
Radial :
SV :
LAD :
Statin :
EF>30 :
Shunt :
KOAHA :
Ex :

EK-2 : SPSS Çıktıları

	ÖLÜM	N	Mean	Std. Deviation
YAŞ	1,00	24	70,2083	10,3755
	,00	457	67,1182	10,6959
TAKIP	1,00	24	25,5000	14,1298
	,00	457	38,4136	12,7444

t-test for Equality of Means			
	t	df	Sig. (2-tailed)
YAŞ	1,382	479	,168
TAKIP	-4,812	479	,000

		ÖLÜM		Total
		,00	1,00	
TAKIP	,00	22	2	24
		4,8%	8,3%	5,0%
	1,00	435	22	457
		95,2%	91,7%	95,0%
Total		457	24	481
		100,0%	100,0%	100,0%

	value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	,596	1	,440

		ÖLÜM		Total
		,00	1,00	
RADIAL	,00	134	3	137
		29,3%	12,5%	28,5%
	1,00	323	21	344
		70,7%	87,5%	71,5%
Total		457	24	481
		100,0%	100,0%	100,0%

	value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	3,168	1	,075

		ÖLÜM		Total
		,00	1,00	
SV	,00	79	6	85
		17,3%	25,0%	17,7%
	1,00	378	18	396
		82,7%	75,0%	82,3%
Total		457	24	481
		100,0%	100,0%	100,0%

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	,932	1	,334

	ÖLÜM		Total
	,00	1,00	
LAD ,00	20 4,4%	2 8,3%	22 4,6%
1,00	437 95,6%	22 91,7%	459 95,4%
Total	457 100,0%	24 100,0%	481 100,0%

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	,818	1	,366

	ÖLÜM		Total
	,00	1,00	
STATIN ,00	164 35,9%	8 33,3%	172 35,8%
1,00	293 64,1%	16 66,7%	309 64,2%
Total	457 100,0%	24 100,0%	481 100,0%

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	,065	1	,799

	ÖLÜM		Total
	,00	1,00	
DIYABET ,00	339 74,2%	17 70,8%	356 74,0%
1,00	118 25,8%	7 29,2%	125 26,0%
Total	457 100,0%	24 100,0%	481 100,0%

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	,133	1	,716

	ÖLÜM		Total
	,00	1,00	
SIGARA ,00	441 96,5%	20 83,3%	461 95,8%
1,00	16 3,5%	4 16,7%	20 4,2%
Total	457 100,0%	24 100,0%	481 100,0%

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	9,918	1	,002

	ÖLÜM		Total
	,00	1,00	
EF.30 ,00	391 85,6%	11 45,8%	402 83,6%
1,00	66 14,4%	13 54,2%	79 16,4%
Total	457 100,0%	24 100,0%	481 100,0%

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	26,214	1	,000

	ÖLÜM		Total
	,00	1,00	
COPD ,00	415 90,8%	19 79,2%	434 90,2%
1,00	42 9,2%	5 20,8%	47 9,8%
Total	457 100,0%	24 100,0%	481 100,0%

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	3,506	1	,061

	ÖLÜM		Total
	,00	1,00	
STUNT ,00	105 23,0%	10 41,7%	115 23,9%
1,00	352 77,0%	14 58,3%	366 76,1%
Total	457 100,0%	24 100,0%	481 100,0%

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	4,379	1	,036

Survival

This subfile contains: 481 observations

Life Table

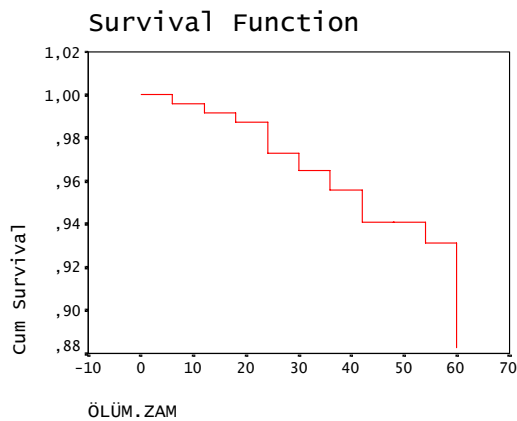
Survival variable ÖLÜM.ZAM

Cumul Intrvl Propn Start Surv Time End	Number Entrng Proba- this bility Intrvl Densty	Number wdrawn During Hazard Intrvl Rate	Number Exposd to Risk	Number of Termnl Events	Propn Termi- nating	Propn Sur- viving	at
,0	481,0	,0	481,0	2,0	,0042	,9958	
,9958	,0007	,0007					
6,0	479,0	7,0	475,5	2,0	,0042	,9958	
,9917	,0007	,0007					
12,0	470,0	20,0	460,0	2,0	,0043	,9957	
,9873	,0007	,0007					
18,0	448,0	62,0	417,0	6,0	,0144	,9856	
,9731	,0024	,0024					
24,0	380,0	39,0	360,5	3,0	,0083	,9917	
,9650	,0013	,0014					
30,0	338,0	54,0	311,0	3,0	,0096	,9904	
,9557	,0016	,0016					
36,0	281,0	47,0	257,5	4,0	,0155	,9845	
,9409	,0025	,0026					
42,0	230,0	68,0	196,0	,0	,0000	1,0000	
,9409	,0000	,0000					
48,0	162,0	124,0	100,0	1,0	,0100	,9900	
,9315	,0016	,0017					
54,0	37,0	36,0	19,0	1,0	,0526	,9474	
,8824	,0082	,0090					

The median survival time for these data is 54,00+

Intrvl Start Time	SE of Cumul Sur- viving	SE of Proba- bility Densty	SE of Hazard Rate
,0	,0029	,0005	,0005
6,0	,0042	,0005	,0005
12,0	,0051	,0005	,0005
18,0	,0077	,0010	,0010
24,0	,0089	,0008	,0008
30,0	,0103	,0009	,0009
36,0	,0126	,0012	,0013
42,0	,0126	,0000	,0000
48,0	,0156	,0016	,0017

54,0 ,0499 ,0080 ,0090



Cox Regression

Case Processing Summary

		N	Percent
Cases available in analysis	Event ^a	24	5,0%
	Censored	457	95,0%
	Total	481	100,0%

a. Dependent Variable: TAKIP.SÜ

Categorical Variable Codings^{a,b,c}

	Frequency	(1)
RADIAL ,00	137	-,500
1,00	344	,500
SIGARA ,00	461	-,500
1,00	20	,500
EF.30 ,00	402	-,500
1,00	79	,500
COPD ,00	434	-,500
1,00	47	,500

a. Simple Parameter Coding

b. Category variable: RADIAL

c. Category variable: SIGARA

d. Category variable: EF.30

e. Category variable: COPD

Block1: Method = Backward Stepwise (Conditional LR)

Omnibus Tests of Model Coefficients

Step	-2 Log Likelihood	Overall (score)		
		Chi-square	df	Sig.
1 ^a	238,233	46,098	5	,000
2 ^b	239,462	43,760	4	,000
3 ^c	241,170	42,860	3	,000

a. Variable(s) Entered at Step Number 1: YAŞ
RADIAL SIGARA EF.30 COPD

b. Variable Removed at Step Number 2: COPD

c. Variable Removed at Step Number 3: YAŞ

Omnibus Tests of Model Coefficients

Step	Change From Previous Step		
	Chi-square	df	Sig.
1 ^a	35,522	5	,000
2 ^b	1,228	1	,268
3 ^c	1,708	1	,191

- a. Variable(s) Entered at Step Number 1: YAŞ
RADIAL SIGARA EF.30 COPD
- b. Variable Removed at Step Number 2: COPD
- c. Variable Removed at Step Number 3: YAŞ

Omnibus Tests of Model Coefficients

Step	Change From Previous Block		
	Chi-square	df	Sig.
1 ^a	35,522	5	,000
2 ^b	34,293	4	,000
3 ^c	32,585	3	,000

- a. Variable(s) Entered at Step Number 1: YAŞ
RADIAL SIGARA EF.30 COPD
- b. Variable Removed at Step Number 2: COPD
- c. Variable Removed at Step Number 3: YAŞ

Variables in the Equation

	B	SE	wald	df	Sig.	Exp (B)	95% CI for Exp(B)		
							Lower	Upper	
Step 1	YAŞ	,024	,021	1,379	1	,240	1,025	,984	1,067
	RADIAL	1,083	,638	2,881	1	,090	2,954	,846	10,32
	SIGARA	2,300	,630	13,339	1	,000	9,977	2,903	34,28
	EF.30	2,151	,467	21,248	1	,000	8,593	3,443	21,45
	COPD	,601	,513	1,375	1	,241	1,824	,668	4,985
Step 2	YAŞ	,026	,021	1,618	1	,203	1,027	,986	1,069
	RADIAL	1,128	,638	3,132	1	,077	3,091	,886	10,78
	SIGARA	2,374	,624	14,491	1	,000	10,74	3,164	36,48
	EF.30	2,183	,465	22,079	1	,000	8,873	3,570	22,06
Step 3	RADIAL	1,003	,625	2,575	1	,109	2,727	,801	9,289
	SIGARA	2,363	,623	14,376	1	,000	10,62	3,132	36,04
	EF.30	2,208	,465	22,556	1	,000	9,093	3,657	22,61

