

T.C.

MARMARA ÜNİVERSİTESİ
FEN BİLİMLER ENSTİTÜSÜ

**ÇİLİK DURUMUNDA COX ORANSAL HAZARD
MODELİNİN FARKLI TAHMİN YÖNTEMLERİNİN
İNCELENMESİ**

DENİZ NAN

(141203120050001)

DOKTORA TEZİ

MATEMATİK ANABİLİM DALI

UYGULAMALI MATEMATİK PROGRAMI

DANIŞMAN

Prof. Dr. Müjgan TEZ

STANBUL 2010



Bu çal, mada, büyük katk,lar, ve eme i olan, beni her konuda destekleyen sayg,de er dan, man,m Say,n Prof. Dr. Müjgan TEZøe, çal, mam boyunca bilgileri ile beni ayd,nlatan Say,n Prof. Dr. Ayd,n ERAR ve Say,n Yrd. Doç. Dr. Bahar SENNARO LUøna, Mathematica program, konusunda benden bilgilerini esirgemeyen Say,n Prof Dr. U ur Yah iøye ve çal, mam boyunca beni destekleyen aileme te ekkürlerimi sunar,m.

Kas,m, 2010

Deniz NAN

Ç NDEK LER

SAYFA

ÖNSÖZİ	í í	.i
Ç NDEK LERİ	í í	.ii
ÖZETİ	í í v
ABSTRACT	í í	.í í ...vi
YEN L K BEYANI	í í	..í í í í í í í .í .. vii
SEMBOLLERİ	í í	..í í í .í ...viii
KISALTMALARI	í í	.í í ..ix
EK LLERİ	í í í í í í í í ..í í í í í í í í í í í í í í í í í í	.í í x
TABLolar	í í í í í í í í í í í í ..í í í í í í í í í í í í í í í	.. xi
BÖLÜM I. G R VE AMAÇ	1
BÖLÜM II. YA AM SÜRDÜRME ZAMANI ANAL Z N N KAPSAMI VE TEMEL KAVRAMLARI	4
II.1 YA AM SÜRDÜRME ANAL Z N N KAPSAMI ve ÇALI MA ALANI		4
II.2 YA AM SÜRDÜRME ANAL Z N N TEMEL KAVRAMLARI	5
II.2.1 Ya am sürdürme zaman,	5
II.3. YA AM SÜRDÜRME ZAMANI ANAL Z NDE KULLANILAN TEMEL FONKS YONLAR	11
II.3.1 Ya am sürdürme fonksiyonu	11

.....nu	14
II.4. YA AM SÜRDÜRME VE HAZARD FONKS YONUNUN TAHM N	
ED LMES	17
II.4.1. Ya am tablosu tahmini	17
II.4.2. Kaplan-Meier tahmini.....	19
BÖLÜM III. COX ORANSAL HAZARD REGRESYON MODEL	
VE KARAKTER ST KLER	21
III.1 COX ORANSAL HAZARD REGRESYON MODEL N N YAPISI.....	22
III.2 COX ORANSAL HAZARD MODEL Ç N KISM ENÇOK	
OLAB L RL K TAHM N ED C S	23
III.2.1 Newton-Raphson Yöntemi.....	27
III.3 PARAMETRELERE L K N H POTEZ TESTLER VE GÜVEN	
ARALIKLARININ BEL RLENMES	28
III.3.1 Wald Test statisti i.....	29
III.3.2 Olabilirlik Oran Test statisti i	29
III.4 HAZARD ORANININ HESAPLANMASI VE YORUMU.....	30
III.5 AÇIKLAYICI DE KEN ETK LER N ÇEREN YA AM	
SÜRDÜRME FONKS YONU Ç N KAPLAN-MEIER TAHM N	32
III.6 ORANTILI HAZAR VARSAYIMININ NCELENMES	33
III.6.1 Grafiksel Yöntemler.....	33
III.6.2 istatistiksel Yöntemler	38
BÖLÜM IV. PARAMETR K YA AM SÜRDÜRME MODELLER	
.....	41
IV.1 ÜSSEL MODEL.....	43
IV.2 WE BULL MODEL	46

BÖLÜM V. Ç L K DURUMUNDA COX ORANSAL HAZARD MODEL Ç N FARKLI TAHM N YÖNTEMLER	51
V.1. COX ORANSAL HAZARD REGRESYON MODEL Ç N R DGE TAHM N ED C S	51
V.2. COX ORANSAL HAZARD REGRESYON MODEL Ç N L U-T P TAHM N ED C	54
V.3 COX ORANSAL HAZARD REGRESYON MODEL Ç N TEMEL B LE ENLER TAHM N ED C S	62
V.4 COX ORANSAL HAZARD REGRESYON MODEL Ç N L U- T P TAHM N ED C LE TEMEL B LE ENLER TAHM N ED C S N N B R KOMB NASYONU.....	65
BÖLÜM VI. BENZET M ÇALI MASI.....	71
BÖLÜM VII. SONUÇLAR VE TARTI MA	76
BÖLÜM VIII. SON DE ERLEND RMELER VE ÖNER LER.....	78
KAYNAKLAR.....	80
EKLER.....	83
ÖZGEÇM	84

Ç L K DURUMUNDA COX ORANSAL HAZARD MODEL Ç N FARKLI TAHM N ED C LER N NCELENMES

Bu çal, mada, aç,klay,c, de i kenler aras, iç ili ki olmas, durumunda Cox oransal hazard regresyon model parametrelerinin kestirimi için farklı yakla ,mlar incelenmektedir.

Genel olarak Cox oransal hazard regresyon modelinin parametrelerinin tahmininde kısmi çok olasılık tahmin edicisi kullanılmaktadır. Fakat iç ili kinin varlığı, çok olasılık tahminlerini ciddi şekilde etkileyebilir. Parametre tahminleri çok büyük varyanslara sahip olacaklarından gerçek değerlerden çok uzak olabilirler.

2007 yılında Ridge regresyon yaklaşımını, Cox oransal hazard regresyon modeline genellemlerdir. Fakat özellikle yüksek iç ili ki durumunda bu yaklaşım sorunun çözümünde yetersiz kalmaktadır.

Bu çal, mada Cox oransal hazard regresyon modeline; Liu-tipi tahmin edici, temel bileşenler tahmin edicisi ve bu iki tahmin edicinin bir kombinasyonu olan yeni bir tahmin edici üretilmiş ve bu tahmin ediciler Ridge tahmin edicisi ile ortalama hata kareler ölçütüne göre karşılaştırılmışlardır. Son olarak tahmin edicilerin performansları, benzetim çalışmalarını ile ayrıntılı olarak incelenmiştir.

A STUDY ABOUT DIFFERENT KIND OF ESTIMATORS FOR COX PROPORTIONAL HAZARD MODEL IN PRESENCE OF COLLINEARITY

This paper considers new approaches to estimate Cox proportional hazard regression model parameters in presence of collinearity between covariates.

Usually partial maximum likelihood estimator is used to estimate Cox proportional hazard regression model parameters. But when there exist collinearity partial maximum likelihood estimates can be effected seriously . Parameter estimates have large variances so they may be far from true values.

In 2007 ridge regression approach was generalized to the Cox proportional hazard regression model. But especially when there exist severe collinearity this approach may not fully addresses the collinearity problem.

In this study we developed; Liu-type estimator, principal component estimator and a new estimator, which combine these two estimators, for Cox proportional hazard regression model and compared with ridge regression estimator in terms of mean squared error (MSE). Finally we evaluated their performances through simulation studies.

Ç L K DURUMUNDA COX ORANSAL HAZARD MODEL Ç N FARKLI TAHM N ED C LER N NCELENMES

Bu çal, mada, iç ili ki durumunda Cox Oransal Hazard Regresyon modeline uyarlanm, olan ridge regresyon parametre tahmin yöntemine alternatif olarak, modele; Liu-tipi tahmin yöntemi, temel bile enler tahmin yöntemi ve bu iki tahmin yönteminin bir kombinasyonu olan yeni bir tahmin yöntemi uyarlanm, t,r.

Yap,lan kuramsal ispatlarla, Cox oransal hazard modele uyarlanm, olan, Liu-tipi tahmin yöntemi ve Liu-tipi ile temel bile enler tahmin yöntemlerinin bir kombinasyonu olarak üretilmi olan yeni tahmin yönteminin, ridge tahmin yönteminden daha küçük hata kareler ortalamas, verdi i özel ko ullar belirlenmi tir. Ayr,ca bu yeni yöntemler birbirleri ile hata kareler ortalamas, ölçütü bak,m,ndan k,yaslanm, t,r.

Yap,lan benzetim çal, malar, ile önerilen yeni yöntemlerin performanslar, daha ayr,nt,l, bir ekilde incelenmi ve özellikle yüksek iç ili ki durumunda, Cox oransal hazard model parametrelerinin tahmininde, k,smi ençok olabilirlik ve ridge tahmin yöntemleri yerine kullan,lmalar, önerilmi tir.

- T : Ya am sürdürme süresi
- t : Takip süresi
- $F(t)$: Da ,l,m fonksiyonu
- $f(t)$: Olas,l,k yo unluk fonksiyonu
- $S(t)$: Ya am sürdürme fonksiyonu
- $S_0(t)$: Temel ya am sürdürme fonksiyonu
- $S(t, X)$: X aç,klay,c, de i kenlerine sahip birimin ya am sürdürme fonksiyonu
- $h(t)$: Hazard fonksiyonu
- $h_0(t)$: Temel hazard fonksiyonu
- $h(t, X)$: X aç,klay,c, de i kenlerine sahip birimin hazard fonksiyonu
- $L(\beta)$: β en çok olabilirlik fonksiyonu
- $R(t_j)$: t_j an,ndaki risk kümesi
- δ_j : j. veriye ait durdurma de i keni
- $I(\beta)$: Kestirilm bilgi matrisi
- $V(\beta)$: β parametresinin varyans,
- $S(\beta)$: β parametresinin standart hatas,
- γ : H,zland,rma faktörü
- t_m : Ya am sürdürme süresi medyan,
- κ : Kötü ko ulluluk katsay,s,
- β_{EK}^{\prime} : Lineer model için en küçük kareler tahmin edicisi
- β_{PML}^{\prime} : Cox oransal hazard modeli için k,smi en çok olabilirlik tahmin edicisi
- β_R^{\prime} : Lineer model için ridge tahmin edicisi

1 modeli için ridge tahmin edicisi

$\beta_{TB}^{\prime\prime}$: Lineer model için temel bile enler tahmin edicisi

β^{TB} : Cox oransal hazard modeli için temel bile enler tahmin edicisi

$\beta_{kd}^{\prime\prime}$: Lineer model için Liu-tipi tahmin edici

β^{kd} : Cox oransal hazard modeli için Liu-tipi tahmin edici ($\beta^{\prime\prime} = \beta_{PML}^{\prime\prime}$ için)

β^{kdR} : Cox oransal hazard modeli için Liu-tipi tahmin edici ($\beta^{\prime\prime} = \beta^{R}$ için)

$\beta_r^{\prime\prime}(k)$: Lineer model için (r-k) s,n,f, tahmin edici

$\beta_r^{\prime\prime}(d)$: Lineer model için (r-d) s,n,f, tahmin edici

$\beta_r^{\prime\prime}(d,k)$: Lineer model için (r-d,k) s,n,f, tahmin edici

$\beta^r(d,k)$: Cox oransal hazard modeli için (r-d,k) s,n,f, tahmin edici

d_{opt} : Optimum d değeri



*Your complimentary
use period has ended.
Thank you for using
PDF Complete.*

[Click Here to upgrade to
Unlimited Pages and Expanded Features](#)

HKO : Hata kareler ortalamas,

HO : Hazard oran,

Kov : Kovaryans

KLM : Kandaki lökosit miktar,

VK : Vücut kitle indeksi

SAYFA NO

ekil II.1 Durdurmaya ili kin grafiksel örnek	7
ekil II.2 Ya am sürdürme fonksiyonu	12
ekil II.3 Ya am sürdürme sürelerine ili kin basamak fonksiyonu.....	13
ekil II.4 Artan hazard fonksiyonu	15
ekil II.5 Tek tepeli hazard fonksiyonu	15
ekil III.1 Bireylere ili kin log-log e rileri.....	36
ekil V.1 Teorem(5.1)ın ispat,na yönelik grafiksel örnek	59
ekil V.2 Teorem(5.2)ın ispat,na yönelik grafiksel örnek	62



Your complimentary
use period has ended.
Thank you for using
PDF Complete.

[Click Here to upgrade to
Unlimited Pages and Expanded Features](#)

SAYFA NO

Tablo III.1 Örnek veri.....	25
Tablo III.2 Bireylere ili kin hazardlar	26
Tablo VI.1 Benzetim çal, mas, sonuçlar,	75

G R VE AMAÇ

Hastal,k ve ölüm gibi insan ya am,n, do rudan etkileyen olaylar, belirli çevre ve çal, ma ko ullar, alt,nda bir ürünün ar,zalanmas,, bir irketin iflas etmesi ya da bo anma, grev gibi sosyal ba ar,s,zl,k durumu olarak tan,mlanabilen durumlar, ara t,rmac,lar için her zaman ilgi çekici olmu tur. Yap,lan birçok ara t,rmada birer ba ar,s,zl,k durumu olarak varsay,lan bu olaylar,n gerçekle mesi üzerinde etkili olan faktörlerin belirlenmesine ve ba ar,s,zl,k zamanlar,n,n analiz edilebilmesine çal, ,lm, t,r. Önceleri sadece insan ya am, ve özellikle ölüm zaman,n,n istatistiksel çözümlenmesine ili kin olarak ya am tablolar, yard,m,yla yap,lan çal, malar, daha sonra herhangi bir olay,n meydana gelmesi zaman,n,n çözümlenmesi amac,yla olu turulan modellerle geli tirilmi tir.

Ba ar,s,zl,k zaman, analizi ilk olarak 17. yüzy,lda kullan,lmaya ba lanm, t,r. 1687-1691 y,llar, aras,nda Edmund Halley ilk ya am tablosunu tasarlam, t,r. Halley'in tasarlad, , ya am tablosu günümüzde demografi ve aktüerya çal, malar,nda kullan,lan ya am tablolar, ile benzerlik göstermektedir.

20. yüzy,lda kinci Dünya Sava , s,ras,nda özellikle askeri teçhizatlar,n güvenilirli i ve ya am sürdürme süreleri üzerine ara t,rmalar h,zlanm, t,r. Sava sonras,nda da özellikle elektronik endüstrisi alan,nda ya am sürdürme analizlerinin önemi artm, t,r. 20. yüzy,l,n ikinci yar,s,ndan sonra ba ta t,p olmak üzere fen ve sosyal bilimlerde yap,lan ara t,rmalarda ya am sürdürme analizi daha çok kullan,lr hale gelmi tir.

Zaman, da ,l,ma ili kin hiç bir varsay,m gerektirmeyen parametrik olmayan yöntemler kullan,larak yap,labilir. Ancak gerek insan ya am,nda, gerekse endüstriyel ve sosyal ya amda görülen ba ar,s,zl,klar üzerinde zaman kadar, olay,n gözlem birimlerine ait özellikler de etkilidir. Bu nedenle söz konusu de i kenlerin geli tirilen modellere eklenmesiyle ba ar,s,zl,k zaman, analizi daha kapsaml, olarak incelenmeye ba lanm, t,r.

Ba ar,s,zl,k zamanlar,n,n ba ,ml, de i ken ve bu zaman de i kenini üzerinde etkili oldu u dü ünülen di er de i kenlerin de ba ,ms,z de i ken olarak yer ald, , ba ar,s,zl,k zaman, modellerinde, ba ,ms,z de i kenler hazard fonksiyonu üzerinde (oransal hazard model) ya da ya am sürdürme zaman, üzerinde (h,zland,r,lm, ba ar,s,zl,k zaman, modeli) çarp,msal etkiye sahiptirler. Ba ar,s,zl,k zaman, analizinde kullan,lan bu modeller yard,m,yla, durdurulmu gözlem olarak adland,r,lan ve herhangi bir ekilde çal, ma kapsam,ndan ç,kar,lm, ya da çal, man,n sona ermesine ra men hala ba ar,s,zl, a u ramam, olan veriler de analize dahil edilir. Elde edilen tüm verilerin kullan,labilmesi, ba ar,s,zl,k zaman, analizini di er yöntemlerden ay,ran en büyük özelliktir.

Ba ar,s,zl,k zaman, analizinin uygulamaya yönelik çal, malar,nda en fazla kullan,lan modellerden birisi, da ,l,msal hiçbir varsay,m gerektirmeyen ve yar, parametrik bir model olan Cox oransal hazard regresyon modelidir. Cox taraf,ndan 1972de geli tirilen model, ba ta kronik hastal,k çal, malar, olmak üzere t,bbi alanda, demografik ve endüstriyel ya am,n test edilmesi çal, malar,nda uzun süre kullan,lm, ve son y,llarda da ekonomik çal, malarda artarak kullan,lmaya ba lanm, t,r [1].

Ba ar,s,zl,k zaman, analizinde, de i kenler aras, iç ili ki problemi s,kl,kla rastlan,lan bir problemdir. ç ili ki durumunda, literatürde genel olarak kullan,lan

olabilirlik tahmin yönteminin kullanılması, dengersiz tahminde etkinliği yitirileceği gibi, önemli bir de etkisini gizleyebilir ya da oldu undan fazlaym, gibi gösterebilir. Birbiri ile ilgili olan de etkilerin gruplandırılması, iç ilgili düzeyini dürebilir fakat genellikle etkinlik kaybına sebep olur [2]. Bunun yanında iç ilgili sebepler olan de etkilerin analizden çıkarılması, daha hızlı tahminler elde edilmesini sağlar.

Son yıllarda bu problemin çözümüne yönelik Cox oransal hazard regresyon modeline ridge tahmin edicisi uygulanmıştır. Küçük çok olabilirlik tahmin edicisine göre ridge tahmin edicisi genellikle daha küçük hata kareler ortalamasına sahiptir. Ayrıca küçük çok olabilirlik tahmin edicisi sadece asimptotik olarak yansız oldu undan, ridge tahmin edicisinin kullanılmadan kaynaklanan yanlış çok önemsenmeyecek boyutlardadır. Fakat özellikle çok yüksek iç ilgili kinin söz konusu olduğu durumlarda ridge tahmin edicisi kötü kulluluk problemini çözmekte yetersiz kalmaktadır.

Bu çalışmanın amacı, Cox oransal hazard regresyon modelinin parametrelerinin tahmini için alternatif tahmin yaklaşımlar, üretmek ve özellikle yüksek iç ilgili kinin varlığında bu yaklaşımların performanslarını incelemektir. Geliştirilen teoremler ile tahmin edicilerin hata kareler ortalaması (HKO) ölçütü açısından birbirlerine göre durumları ortaya konmuştur. Son olarak yapılan benzetim çalışmaları ile tahmin edicilerin performansları daha ayrıntılı bir şekilde incelenmiştir.

YA AM SÜRDÜRME ZAMANI ANALİZİNİN KAPSAMI VE TEMEL KAVRAMLARI

II.1 YA AM SÜRDÜRME ANALİZİNİN KAPSAMI ve ÇALIŞMA ALANI

Ya am sürdürme analizi, sonuç de i keninin ilgili olay meydana gelene kadar geçen süre oldu u veri analizi için kullan,lan istatistiksel yöntemler toplulu udur. Bu yöntemler do rultusunda kurulan modeller yard,m, ile ya am sürdürme zaman,n, etkileyen de i kenler belirlenebilece i gibi, ölüm ya da ba ar,s,zl,k gibi aç,k ve kesin olarak tan,mlanm, herhangi bir olay,n meydana gelmesi zaman, da analiz edilebilir.

Ya am sürdürme analizi, hastalara uygulanan tedavi biçimlerinin ba ar,s,n,n belirlenmesi, farklı tedavi yöntemlerinin ba ar,s,n,n k,yaslanmas, gibi t,p ara t,rmalar,nda, herhangi bir sürecin güvenilirlik uygulamalar,n,n test edilmesi, elektronik parçalar,n garanti sürelerinin belirlenmesi gibi endüstriyel alanda kullan,labilece i gibi demografik bir çal, mada ikametgah de i ikli i süresi, evli çiftlerin evli kalma süresi, ekonomide i sızlık dönemleri, i çilerin grevde bulunma süreleri, irketlerin ya da bankalar,n iflas süreleri gibi ölçülebilir bir sürece ili kin verilere de uygulanabilir.

Bu tip veriler, lojistik regresyon ya da ikili yan,t de i kenleri için kullan,lan di er yöntemlerle de incelenebilir. Ancak bu yöntemlerin bilgi kayb,, hesaplamada k,s,tlama ve yanl,l,k gibi bir tak,m olumsuz sonuçlar, görülebilir. Oysa ya am sürdürme analizi, hem bir olay,n meydana gelmesi durumunun, hem de olay,n meydana gelme süresinin incelenebilmesi gibi avantajlara sahiptir. Bu analizde

edilen bilgiler kullanılır ve çalınan madaki bütün birimler gözlemlenir. Bu süreçte analiz katkısında bulunurlar.

Veri analizinde kullanılan standart istatistiksel ölçümler, yaşam sürdürme verilerine uygulanamaz. Bunun birinci sebebi, yaşam süresi verilerinin genellikle simetrik dağılımı olmamasıdır. Benzer birim ya da bireylerden oluşan bir grubun yaşam sürelerinden oluşan histogram, sağa eğimli olacak ve sonuç olarak, bu tip verilerin normal dağılımı varsaymak uygun olmayacaktır. İkinci sebep, incelenen baz, birim ya da bireylerin yaşam sürdürme zamanlarının net olarak gözlenememesi olmasıdır. Durdurulmuş gözlemler olarak adlandırılan bu gözlemler, daha çok yaşam sürdürme zamanı takip süresinden uzun olan birim ya da bireylerden oluşur. Analize dahil edilen bu gözlemlere ilişkin durdurma olayı, sonraki bölümlerde incelenecektir.

II.2 YAŞAM SÜRDÜRME ANALİZİNİN TEMEL KAVRAMLARI

II.2.1 Yaşam sürdürme zamanı,

Yaşam sürdürme zamanı, analizinde, bir birim için, belirli bir başlangıç zamanından, birimin bozulmasına kadar geçen süreye, o birim için ömür ya da yaşam sürdürme zamanı denilir. Bu tanımda birim ile kastedilen canlı bir varlık (insan, hayvan, bitki) olabileceği gibi herhangi bir nesne (elektronik eşya, araba vs.) ya da sosyal bir kurum (aile, banka, şirket vs.) olabilir. Bozulma ise; ölüm, arıza, boşanma, iflas, suç işleme gibi pek çok farklı olay olabilir.

Yaşam sürdürme zamanı, analizinde, belirli bir takip süresi içerisinde izlenen grup ya da gruplara ilişkin veriler ile ilgilenilir. Yaşam sürdürme sürelerinin tam olarak belirlenebilmesi için çalınan madaki her bir birim için kesin başlangıç ve bozulma zamanı tanımlanmalıdır. Aksi takdirde elde edilen gözlemler durdurulmuş gözlemler olacaktır. Başlangıç zamanı, genellikle her bir birim için aynı takvim

Örne in birçok t,bbi ara t,rmada
bir eyer, be n bir zaman periyodunda ayr, ayr, ba lang,ç zamanlar,na sahiptir ve bir
hastan,n ya am sürdürme zaman,, çal, maya girdi i tarihten itibaren ölçülür. Bu
kat,l,m, özel bir durumun te hisi, bir tedavi eklinin ba lang,c, ya da baz, olumsuz
durumlar,n olu umu ile ba layabilir.

Ya am sürdürme sürelerinin belirlenmesinde bir zaman ölçe i kullan,lmal,d,r.
Ço u çal, mada zaman,n ölçülmesi için ölçek olarak gerçek zaman (gün, ay, y,l)
kullan,lmas,na ra men örne in ta ,t ömrü için araban,n kilometre ile ölçülen ya ,
gibi farkl, ölçekler de vard,r.

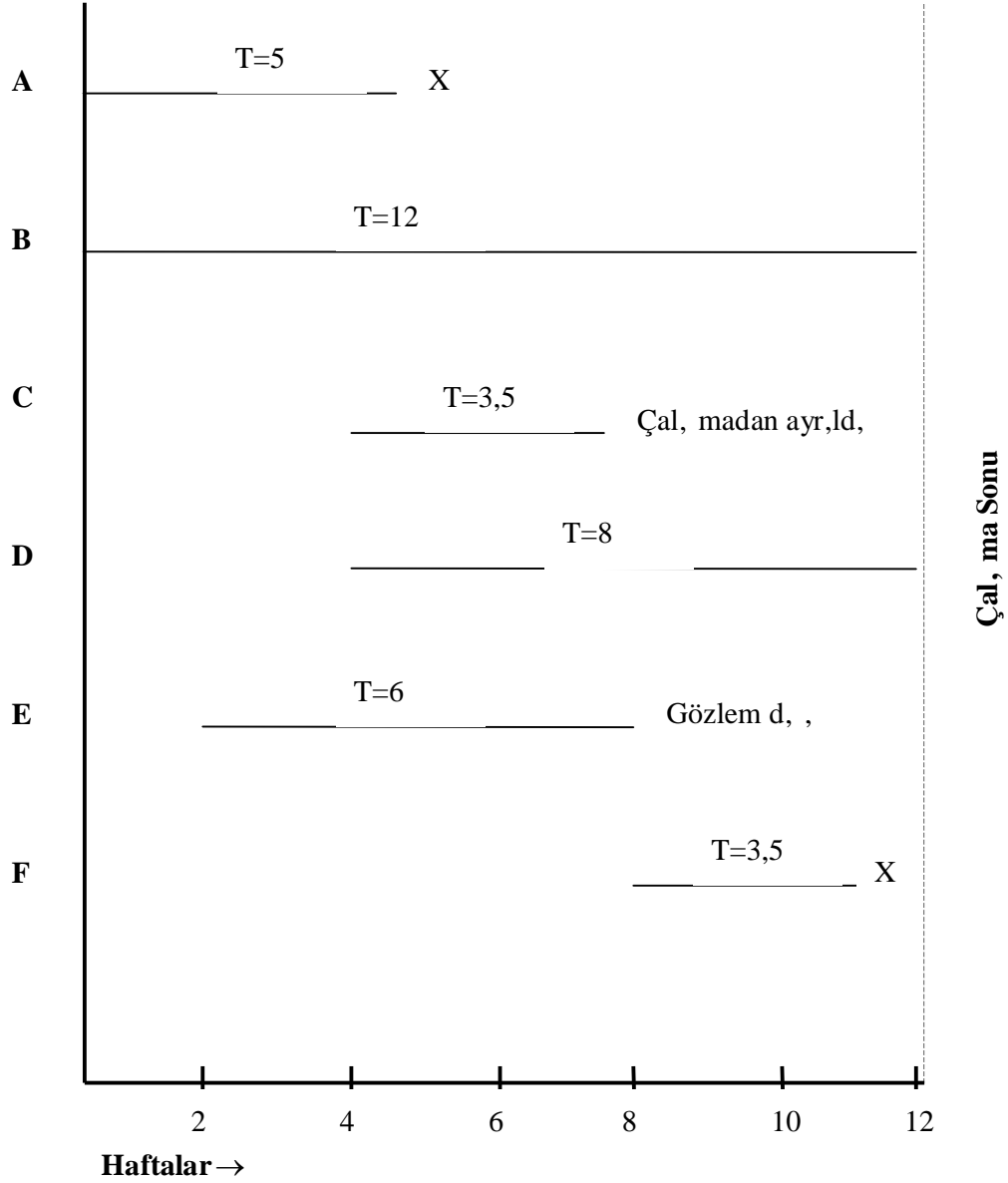
Bozulma yani ilgilenilen olay net bir ekilde tan,m lanmal,d,r. Örne in,
akci er kanseri ile ilgili yap,lan bir çal, mada ilgilenilen olay hastan,n akci er
kanserinden dolay, ölmesidir. E er hasta bu takip süreci içersinde akci er
kanserinden ba ,ms,z olarak beyin kanamas, sebebiyle öldüyse bozulma yani
ilgilenilen olay gerçekleşme mi tir. Dolay,s, ile bu hastadan elde edilecek olan
gözlem durdurulmu gözlem olacakt,r. Bu uygulamada bozulma tan,m, net bir
ekilde tan,ml,d,r ve keyfilikten uzakt,r. Fakat baz, uygulamalarda bozulma
tan,m,nda çok az miktarda keyfilik olabilir. Örne in bir irketin insan kaynaklar,
müdürü çal, anlar,n ba ar,s,zl,k durumlar, ile ilgileniyor olabilir. Bu gibi bir durum
da bozulma tan,m,, herhangi bir nicel yolla ölçülen performans,n, kabul edilebilir bir
düzeyin alt,na dü mesi olarak tan,m lan,r. Bu durumda ba ar,s,zl, ,n tan,m,nda bir
miktar keyfilik olabilecektir.

II.2.2.Durdurma

Ya am sürdürme analizinde s,kl,kla rastlanan bir problem durdurulmu
verilerdir. Bu tip veride, birimin ya am sürdürme zaman,na ili kin bir tak,m
bilgilere sahip olmam,za ra men ya am sürdürme zaman, net olarak bilinmez.

emek için a a ,daki gibi grafiklenmi olan 12 haftal,k

çal, may, meclerenn.



ekil II.1 Durdurmaya ili kin grafiksel örnek

mi çal, man,n ba lang,c,nda çal, maya kat,lm, ve 5. haftada olay, ya am, t,r. Bu birim için ya am sürdürme zaman, $T=5$ dir ve veri durdurulmu veri de ildir.

B birimi de çal, man,n ba lang,c,nda çal, maya kat,lm, fakat çal, man,n sonuna kadar olay, ya amam, t,r. Bu durumda bu birim için gerçek ya am sürdürme zaman,n,n 12 haftadan büyük oldu unu bilmemize ra men tam olarak kaç hafta oldu unu bilemeyiz. Bu veri durdurulmu veridir ve ya am sürdürme zaman, $T=12$ al,n,r.

C birimi çal, maya 4. haftada kat,lm, ve 3,5 hafta sonra olay, ya amadan çal, madan çekilmi tir. Bu birim için de gerçek ya am sürdürme zaman,n,n 3,5 haftadan fazla oldu unu bilmemize ra men tam olarak kaç hafta oldu unu bilemeyiz. Bu veride durdurulmu tur ve $T=3,5$ al,n,r

D birimi çal, maya 4. haftada ba lam, ve çal, man,n sonuna kadar olay, ya amam, t,r. Bu veride durdurulmu tur ve $T=8$ dir.

E birimi çal, maya 2. haftada kat,lm, ve 8. haftada olay, ilgilenilen neden den farklı, bir nedenle ya am, t,r. Bu veride durdurulmu tur ve $T=6$ d,r.

F birimi çal, maya 8. haftada girmi , 3,5 hafta sonra olay, ya am, t,r. Bu veri durdurulmu de ildir ve $T=3,5$ dir [3].

ncelenen olay,n yap,s, ve elde edilen verilerin türüne göre birçok durdurma tipinden bahsedilebilir. Analizde kullan,lan temel durdurma çe itleri:

- Sa dan durdurma
 - a) I. tip durdurma b) II. tip durdurma
- Soldan durdurma
- Çifte durdurma
- Aral,k durdurmas,

olarak s,n,fland,r,labilir.

durma

çok kar ,la ,lan tiptir. Bu tip gözlemler de bozulma gerçeyle memektir. Bu tip durdurmada gerçek ya am sürdürme süresi (T) takip süresi (t) den daha fazladır. Örneğin bir kanser hastasının gerçek ya am sürdürme süresi (T) olarak alındığında üç farklı durum söz konusudur.

- Hasta takip süresi boyunca ölmemiştir. Takip süresi çalınma süresine eşittir ve $T > t$ dir. (Çalınma bitmemiştir fakat bozulma gerçeyle memiştir.)
- Hasta takip süresi içerisinde gözlem deneysel olarak kalabilir. Burada da $T > t$ dir. (Hasta çalınmadan çekilebilir.)
- Ölüm kanser deneysel olarak farklı bir olaydan dolayı gerçeyle memi olabilir.

Hastanın ya ama devam etmesi anlaşılmamıştır, tır.

a) I. Tip durdurma

Sadandan durdurmanın özel bir tipidir. T_p ve endüstride ya am sürdürme zamanı deneylerinde sıklıkla karşılanmaktadır. I. tip durdurma, deneyin sabit bir tamamlanma zamanında sonlandırılması ile oluşur.

b) II. Tip durdurma

Bu tip durdurmada deney, önceden belirlenmiş sayıda olay meydana geldiğinde sona erdirilir.

II.2.2.2 Soldan durdurma

Soldan durdurma ile tanımlanan olayın başlangıç zamanının kesin olarak bilinmediği durumlarda karşılanır. Örneğin, belirli bir kanserin incelendiği ilk tümörün alınmasından sonra ikinci tümörün oluşması süresinin takip edildiği bir araştırmada, hastalar üç ayda bir kontrol edilsin. İlk kontrol esnasında tümör tespit

kesin süre bilinmemektedir. Bu tip verilere soldan durdurulmuş veri denir. Soldan durdurulmuş veride $T < t$ dir. Soldan durdurulmuş veri ile, sağdan durdurulmuş veriye oranla daha az karışık olur.

II.2.2.3. Çifte durdurma

Soldan ve sağdan durdurulmuş gözlemlerin bir arada bulunduğu veri tipidir. Örneğin bir çalınan mada, belirli bir sayıdaki farelerin bulunduğu ortama bulaşmış, bir hastalıklı olan başka bir fare konulsun ve farelerin hastalığa yakalanma süreleri araştırılsın. Bu amaçla bir ay süre ile fareler iki günde bir kontrol edilsin. Kontrol esnaslında hastalıklı kapılmış olan fareler tespit edilecektir fakat bunların hastalıklı kaç saat içerisinde kaptıkları belirlenemeyecektir. Bunun yanında takip süresi boyunca hastalıklı kapmayan ya da takip süresi içerisinde farklı sebeplerden ölen fareler de olacaktır. Bu tip verilere çifte durdurulmuş veri denir.

II.2.2.4. Aralık durdurması,

Bu tip durdurmada beklenen olay iki zaman noktası arasında bir yerde gerçekleşmektedir. Ancak gerçekleştirmenin hakkındaki kesin bir bilgi yoktur. Örneğin üç ayda bir kontrol edilmek üzere takibe alınan bir hastanın birinci ve ikinci kontrolünde hastalıklı belirtilerine rastlanmamış, fakat üçüncü kontrolünde hastalıklı tekrarlamış olduğu görülsün. Hastalıklı gerçek tekrarlaması süresinin, ikinci kontrolle üçüncü kontrol arasında geçen üç aylık süre içerisinde olduğu bilinir. Fakat hastalıklı olarak ne zaman ortaya çıktığı hakkında kesin bir bilgi yoktur. Dolayısıyla ile gözlenen tekrarlaması süresinin aralık durdurulmuş olduğu söylenebilir. Aralık durdurması, soldan durdurmaya benzetmekle birlikte, incelenen olayın

II.3. YA AM SÜRDÜRME ZAMANI ANAL Z NDE KULLANILAN TEMEL FONKS YONLAR

Ba lang,ç zaman, iyi tan,mıanm, belirli bir gözlemlene sürecinde, her birimin bir ya am sürdürme zaman,na sahip oldu u homojen (aç,klay,c, de i kenleri olmayan) bir kütle içinden seçilen birimlerin T ile gösterilen, negatif olmayan ve ço unlukla sürekli oldu u varsay,lan ba ar,s,zl,k zamanlar,n,n olas,l,k da ,l,m,, genellikle ve aksi belirtilmedikçe $[0, \hat{O})$ aral, ,nda, ya am sürdürme fonksiyonu ve hazard fonksiyonu olmak üzere iki ekilde tan,mıan,rlar.

Veriye uygun olan ya am sürdürme da ,l,m,n,n bilinmedi i durumlarda bu fonksiyonlar parametrik yöntemlerle belirlenemez. Bu fonksiyonlar,n tahmin edilmesinde ençok kullan,lan yöntemler, öya am tablosu tahmin yöntemiö ve öKaplan-Meier tahmin yöntemiö dir. Ya am sürdürme zamanlar,n,n da ,l,m, ile ilgili olarak özel da ,l,msal varsay,mılar gerektirmeden, parametrik olmayan ya da da ,l,mıan ba ,ms,z olarak adland,r,lan bu yöntemler, oransal frekans tablolar,n,n, olas,l,k noktalamalar,n,n ve deneysel da ,l,m fonksiyonunu kullan,ımas,n, içerirler.

II.3.1 Ya am sürdürme fonksiyonu

Ya am sürdürme fonksiyonu $S(t)$ ile ifade edilir ve bir birimin belirli bir t an,ndan daha uzun süre ya am sürdürmesi olas,l, ,n, verir.

Bir zaman noktas,na ili kin anl,k bir ölçüm olan ve bir birimin t an,nda ko ulsuz ba ar,s,z olma yo unlu unu gösteren $T_{\text{öin}}$ olas,l,k yo unluk fonksiyonu $f(t)$ ile ve da ,l,m fonksiyonu $F(t)$ ile gösterilsin.

$$(u)du \quad (2.1)$$

dur.

$S(t)$ fonksiyonunu, olasılıklı yoğunluk fonksiyonu $f(t)$ ve dağılım fonksiyonu $F(t)$ nin yardımıyla, T nin sürekli ya da kesikli değişken olmasına göre iki şekilde yazabiliriz. T sürekli değişken ise;

$$S(t) = P(T > t) = \int_t^{\infty} f(t)dt = 1 - F(t) \quad (0 < t < \infty) \quad (2.2)$$

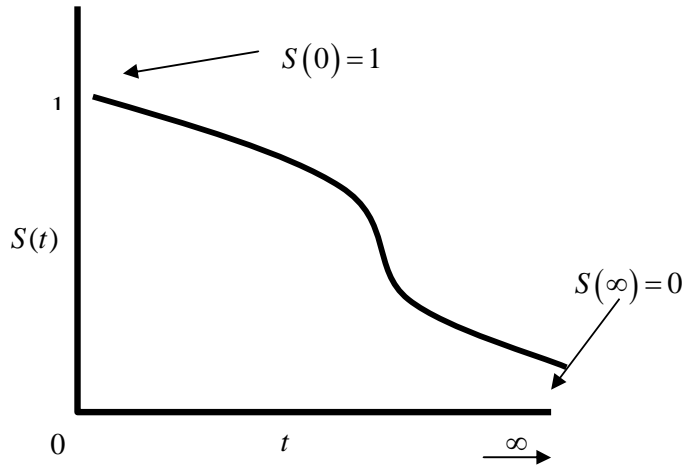
T kesikli değişken ise; $f_j = P(T = t_j)$ olmak üzere,

$$S(t) = P(T \geq t) = \sum_{t_j > t} f_j = 1 - F(t) \quad (j=1,2,\dots) \quad (2.3)$$

olarak tanımlanır.

Yaşam sürdürme fonksiyonu, yaşam sürdürme analizinin temellerindedir. Çünkü farklı değerleri için yaşam olasılıklarını elde etmek, bize veri ile ilgili çok önemli ipuçları verir.

Kuramsal olarak $t=0$ ile \hat{O} arasında değer aldığında, bu fonksiyonun grafiği aşağıdaki gibi düzgün bir eğri olmalıdır.



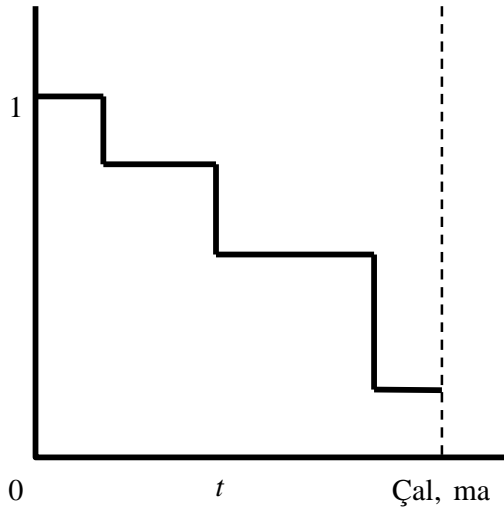
ekil II.2 Yaşam sürdürme fonksiyonu

ne fonksiyonlar, a a ,daki özellikleri sa lar:

ksiyonlard,r. t artarken $S(t)$ azal,r.

- $t=0$ için $S(t)=S(0)=1$ dir. Çal, ma ba larken hiç bir birim olay, ya amam, t,r. Bu yüzden birimlerin $t=0$ an,ndan daha uzun süre ya amas, olas,l, , $S(0)=P(T>0)=1$ dir.
- $t=\hat{O}$ için $S(t)=S(\hat{O})=0$ d,r. Kuramsal olarak çal, ma periyodu limitsiz ise hiçbir birim sa kalmayaca ,ndan $t=\hat{O}$ için ya am sürdürme fonksiyonunun de eri 0 olacakt,r.

Bu özellikler ya am sürdürme e rilerinin kuramsal özellikleridir. Pratikte ya am sürdürme fonksiyonlar, için, düzgün e riler yerine basamak fonksiyonu grafikleri kullan,l,r. Bunun d ,nda çal, ma süreleri s,n,r,l, oldu undan ve bu süre içersinde olay, ya amam, birimler olabilece inden dolayı,, çal, ma sonunda $S(t)$ s,f,rılanmak zorunda de ildir.



ekil II.3 Ya am sürdürme sürelerine ili kin basamak fonksiyonu

iyonu

Özünde t- an,na kadar ya am, olan bir birimin t- an,ndaki anl,k bozulma riskini verir. h(t) ile gösterilir ve matematiksel olarak öyle ifade edilebilir [3]:

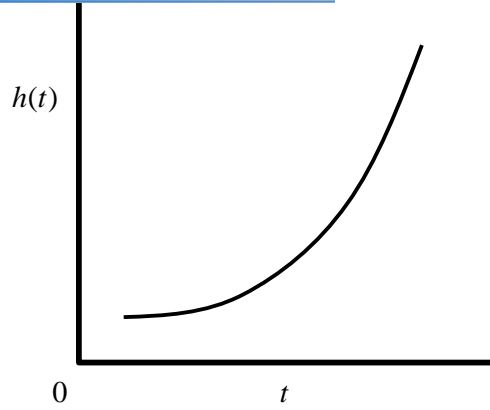
$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T \leq t + \Delta t | T \geq t)}{\Delta t} \quad (2.4)$$

Burada \hat{t} ifadesi çok küçük bir zaman aral, n, göstermektedir. Hazard fonksiyonunun matematiksel ifadesinden de anla ,laca , üzere, hazard bir olas,l,k de il bir h,zd,r ve 0 ile \hat{O} aras,nda de er alabilir.

Hazard fonksiyonu, bozulmam, olma ile ilgili bilgi veren ya am sürdürme fonksiyonunun tersine bozulma ile ilgili bilgi vermektedir. Bu anlamda hazard fonksiyonu, ya am sürdürme fonksiyonunun tersi gibi dü ünülebilir.

Ya am sürdürme fonksiyonunun tersine, hazard fonksiyonu 1'den ba lamak zorunda de ildir. Herhangi bir noktada ba layabilir ve zaman üzerinde artan ya da azalan olarak istedi i gibi hareket edebilir. Hiç bir zaman negatif de er almaz ve bir üst s,n,ra sahip de ildir.

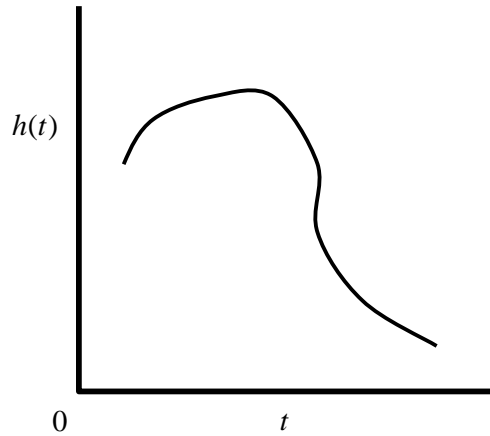
Hazard fonksiyonu çok farklı grafik tipleri sergiler. Bu anlamda veriye ait hazard grafi i tipinin bilinmesi, o veriye uygun olabilecek model formunun belirlenmesine yardımcı olur. Örne in akci er kanseri ile ilgili çal, ma yapan bir ara t,rmac, bilir ki normal artlar alt,nda zaman, n ilerlemesi ile hastalar, n ölüm potansiyeli artan bir ivmeyle artar. Buradan yola ç,karak diyebiliriz ki tahmini olarak bu tip bir veri için hazard grafi i a a ,daki gibi bir e ridir:



ekil II.4. Artan hazard fonksiyonu

Bu tip hazard grafi ine sahip bir veri için weibull da ,1,m,n,n uygun oldu u bilinmektedir. Böylelikle veri için uygun olan model belirlenebilir.

Ba ka bir örnek verilirse; hamile kad,nlar için, dü ük yapma riski ilk üç ay,n sonuna kadar artar ve üç aydan sonra giderek azal,r. Bu bilgiye dayanarak tahmini olarak denilebilir ki, bu veri için hazard grafi i a a ,daki gibidir.



ekil II.5. Tek tepeli hazard fonksiyonu

bir hazard fonksiyonuna sahip olan bir veri için lognormal da ,l,m,n uygun oldu u olmamaktadır. Bunlar gibi birçok farklı grafik tipi için uygun olabilecek da ,l,m bilgisi literatür de vardır. Böylelikle hazard grafiğinin tipi bize veriye uygun olabilecek olan yaşam sürdürme da ,l,m, konusunda çok önemli ipuçları verir.

Eğer veriye uygun olan da ,l,m formu biliniyor ise $S(t)$ ve $h(t)$ fonksiyonları parametrik olarak belirlenebilir.

$f(t)$, veriye uygun olan da ,l,m,n olasılık yoğunluk fonksiyonu ve $F(t)$, karışık gelen da ,l,m fonksiyonu olmak üzere,

$$h(t) = \frac{f(t)}{S(t)} \quad (2.5)$$

ve $S(t)$ ile $h(t)$ kendi aralarında,

$$S(t) = \exp \left[- \int_0^t h(u) du \right] \quad (2.6)$$

$$h(t) = - \left[\frac{dS(t)/dt}{S(t)} \right] \quad (2.7)$$

eklinde ilişkilidirler [3].

ÖRNEK VE HAZARD FONKSİYONUNUN

Ayrıca, bu bölümde, da, bu bilgilerin var olmadığı, durumlarda ya da am sürdürme ve hazard fonksiyonları, tahmininde kullanılan özyaşam tablosu yöntemi ve Kaplan-Meier yöntemiyle incelenecektir.

II.4.1. Yaşam tablosu tahmini

Yaşam tablosu, yaşam sürdürme zaman verilerinin tanımlanmasında en çok kullanılan ve en eski yöntemlerden birisidir. Bu yöntem ilk olarak 17. yüzyılda kullanılmaya başlanmış ve Edmund Halley (1686) tarafından geliştirilmiştir. Önceleri demografiler ve sigortacılar tarafından nüfusun ölümlülük yapısını incelemeye yönelik çalışmalarında neslin yaşam tabloları olarak kullanılmış ve daha sonraları, biyolojik zaman analizinin çeşitli alanlardaki uygulamalarında bu tablolardan geniş ölçüde yararlanılmıştır.

Yaşam tabloları, yaşam sürdürme zamanları, çalışmaları, genellikle zaman uygun aralıklar içinde bölünmesi ile oluşturulur ve her bir zaman aralığı için, aralığın başlangıcında ya da bir birimlerin sayısı, aralık içerisinde bozulan birimlerin sayısı ve izi kaybedilen ya da çalışmadan çekilen birimlerin sayısı hesaplanabilir. Çalışmadan çekilen ya da çeşitli nedenlerle gözlenemeyen birimler durdurulmuş gözlemlerdir. Belirlenen zaman aralığının sonunda hala yaşamakta olduğu varsayılırsa, bu birimler de, aralık içinde risk altındadır. Yaşam tabloları, kullanılarak yaşam sürdürme fonksiyonunun tahmin edilmesinde; her bir zaman aralığı içindeki yaşam sürdürme zaman olasılığı; herhangi bir durdurulmuş gözlemin, önce durdurulmuş olmasaydı, aralık sona ermeden başlangıçta olabileceği ve ancak aralık yarısı için risk altında sayılacağı varsayımına dayanarak hesaplanır. Kaplan-Meier yönteminden farklı olarak, her bir aralık içinde bozulan ya da durdurulan birimlerin sayısı, bilinmesi koşuluyla, her bir birim için kesin yaşam sürdürme ya da durdurma zamanı,

lar, yöntemiyle ya am sürdürme fonksiyonu tahmin edilmektedir.

Ya am tablosu uygun aralıklarla gruplandırılm, ya am sürdürme zaman, verilerinin, ya am sürdürme durumları, bir özetidir. N sayı, gözlemin $t_0 = 0$ ve $t_k = T$ (gözlenen ya am sürdürme zamanları, üst sınırları), $t_{k+1} = \infty$ olmak üzere $I_j = [t_{j-1}, t_j)$, $j=1,2,\dots,(k+1)$ aralıklarında gruplandırılmış, varsayalım. I_j zaman aralığında gözlenen başarısızlık zamanları sayıları D_j , gözlenen durdurulmuş başarısızlık sayıları C_j ve j-inci aralıkta meydana gelen ve bu nedenle bozulma riski altında olan birim sayıları da N_j ile gösterilirse; durdurulmuş başarısızlık zamanları j-inci aralık boyunca hep aynı şekilde olacaktır varsayalım, altındaki söz konusu aralıkta risk altında olan ortalama birim sayısı;

$$N'_j = N_j - C_j / 2 \quad (2.8)$$

olarak hesaplanır. Bu durumda I_j aralığında meydana gelen sürdürme ü bilinen bir birimin, bu aralıkta bozulma olasılığı;

$$\hat{q}_j = D_j / N'_j \quad (2.9)$$

olarak tahmin edilebilir ve standart ya am tablosu tahmini olarak adlandırılır.

Her bir aralık için olasılık başarı, \hat{q}_j ve buna bağlı olarak olasılık sürdürme başarı, $\hat{p}_j = 1 - \hat{q}_j$ bulunduğundan sonra, $t \in I_j$ için tahmini ya am sürdürme başarı, aşağıdaki gibi hesaplanmaktadır:

$$\hat{S}(t) = \prod_{i=1}^j \left(\frac{N_i - D_i}{N'_i} \right), \quad j=1,2,\dots,k+1 \quad (2.10)$$

oyunca bozulma h, z, n, n sabit oldu u ve bu aral, kta

ortalama hayatta kalma süresinin $(N_j - D_j / 2) \Gamma_j$ oldu u varsay, ls, n. Burada Γ_j , j-
inci zaman aral, , n, n uzunlu udur. Bu durumda j-inci zaman aral, , nda hazard
fonksiyonunun ya am tablosu tahmini, $t_j \leq t \leq t_{j+1}$, $j=1, 2, \dots, k+1$ için

$$\hat{h}(t) = \frac{D_j}{(N_j - D_j / 2) \Gamma_j} \quad (2.11)$$

olarak gösterilir [5].

Ya am tablosu yönteminin veri say, s, n, n oldukça yüksek oldu u çal, ma
örnekleri için ($n \geq 100$) kullan, lmas, önerilmektedir. Grafik gösterimde de
ara t, rmac, n, n seçti i dönemlere göre, dönem sonu ya am sürdürme olas, l, klar,
belirtilir. Ba lang, ç an, ya am sürdürme olas, l, , "1" (% 100) dir. Dönem sonu
olas, l, k de er noktalar, birle tirilerek çizim yap, l, r ve dönem içinde bozulmalar
olmas, nedeni ile dönem içi ya am sürdürme olas, l, klar, n, n azalarak sürmesi
gösterilmi olur.

II.4.2. Kaplan-Meier tahmini

Ya am sürdürme fonksiyonunun Kaplan-Meier tahmini, 1900dü y, llar, n
ba lar, ndan bu yana kullan, lmaktad, r. Kaplan ve Meier (1958) taraf, ndan tahminin
ç, kar, m, en çok benzerlik yöntemi kullan, larak yap, lm, ve daha sonralar, bu
yöntemin özellikleri Breslow ve Crowley (1974) ve Meier (1975) taraf, ndan ayr, nt, l,
olarak incelenmi tir [6-8].

Bu yöntemde hayatta kalma süresinin tahmini, hayat tablosu ile benzerdir.
Fakat Kaplan-Meier yönteminde hesaplamalar frekans da , l, m tablosu üzerinden
de il, tek tek veriler üzerinden yap, l, r. Kaplan-Meier tahmininin bu özelli i az say, da
birey içeren örnekler için kullan, lmas, na izin verir.

lemde örö tane bozulma gözlemlenir ve t_1, t_2, \dots, t_r der s,ral,
in.

- $n_i : [t_i, t_{i+1})$ aral, ,n,n ba ,nda risk alt,nda bulunan birimlerin say,s,
- $d_i : t_i$ an,nda gerçekleşen bozulmalar,n say,s,
- $c_i : [t_i, t_{i+1})$ aral, ,nda bulunan durdurulmuş verilerin say,s,
- $t_{r+1} = \infty$

olmak üzere, ya am sürdürme fonksiyonunun Kaplan-Meier tahmini

$$\tilde{S}(t) = \prod_{t_i \leq t} \frac{n_i - d_i}{n_i}, \quad i=1, \dots, r \quad (2.12)$$

olarak tanımlanır [9]. Burada;

$$n_i = n_{i-1} - (d_{i-1} + c_{i-1}), \quad i=1, \dots, r \quad (2.13)$$

dir . Benzer biçimde hazard fonksiyonu, $t_i \leq t \leq t_{i+1}$ için;

$$\tilde{h}(t) = \frac{d_i}{n_i \tau_i} \quad (2.14)$$

olarak tahmin edilebilir. Burada $\tau_i = t_{i+1} - t_i$ dir. Son bozulma zaman,nda ba layan aral,k aç,k oldu u için, bu aral,ktaki hazard fonksiyonunun tahmin edilmesinde (2.14)ü kullanmak olanaks,zd,r.

Ya am tablosu yönteminden farklı olarak, Kaplan-Meier yönteminde bütün bozulma zamanlar, tek tek değerlendirilir. Bir önceki bozulma an,ndan sonra, yeni bozulma zaman,na dek, ba ka bozulma olmamas, nedeni ile olas,l,k de i mez kalmaktad,r ve grafiksel gösterimde, ya am sürdürme olas,l,klar,, basamak fonksiyonu biçiminde belirmektedir.

COX ORANSAL HAZARD REGRESYON MODEL VE KARAKTERİSTİKLERİ

Yaşam sürdürme analizinin kullanıldığı, çalışmalarda, ilgilenilen olaya bağlı olarak, her birey ya da birim için ilave bilgiler de kaydedilir. Örneğin iki tedavi eklinin karşılaştırıldığı, tıbbi araştırmalarda, hastanın yaşı, ve cinsiyeti gibi demografik değişkenler, kandaki beyaz hücre sayısı, ve nabız gibi fizyolojik değişkenler dışında, hastanın yaşam tarzına bağlı, yeme-içme alışkanlıkları, sigara kullanması gibi ek bilgiler de çalışmaya dahil edilir. Bu değişkenler hastanın yaşam süresinde ve tedavi eklinde etkili olabilirler. Yaşam süresini etkilediği düşünülen bu tip değişkenlere örneğin, yaş, cinsiyet, deyişkenler denir [10].

Yaşam sürdürme zamanı analizine, örneğin, yaş, cinsiyet, deyişkenlerin de dahil edilmesi sonucu bu değişkenlerin bozulma riski üzerindeki etkileri incelenebilir, ayrıca birim için yaşam sürdürme fonksiyonu tahminlerine de ulaşılabilir. Böylece birimin tahmini yaşam sürdürme zamanı konusunda çıkarımlar yapılabilir.

Orantılı hazard modellerinin en popüler olanı, 1972 yılında Cox tarafından önerilen öCox Orantılı Hazard Regresyon Modeli'dir. Cox regresyon modelinde yaşam sürdürme zamanı için belirli bir olasılık dağılımı yoktur. Bu nedenle Cox model yapı, parametrik bir modeldir. Modelin en çok kullanılan modellerden biri olması, sebeplerinden biri de bu özelliğidir.

Modelin en önemli varsayımı, farklı birimlerin hazardları, oranlı, zamandan bağımsız olmasıdır. Eğer bu oran zamanla değişiyorsa, oransal hazard varsayımı sağlanmaz.

GENEL HAZARD REGRESYON MODELİNİN YAPISI

Model olarak aşağıdaki gibi verilir [3]:

$$h(t, X) = h_0(t) \cdot e^{\sum_{i=1}^p \beta_i X_i} \quad (3.1)$$

Burada (X_1, X_2, \dots, X_p) açıklayıcı değişkenlerdir. Model, t-anında belirli açıklayıcı değişkenlere sahip olan bir birimin hazard oranı verir.

Model iki kısımdan oluşur: ilk kısım $h_0(t)$, temel hazard fonksiyonu; ikinci kısım ise $\beta_i X_i$ değişkenlerinin üssel fonksiyonudur. Bu modelin, oransal hazard varsayımı ile ilgili olan en önemli özelliklerinden biri, tıya bağlı, temel hazard fonksiyonunun, X açıklayıcı değişkenlerinden ve aynı şekilde X bağımlı üssel ifadenin tıdan tamamen bağımsız olmasıdır. X açıklayıcı değişkenleri zaman bağımsız değişkenler olarak adlandırılır. X açıklayıcı değişkenlerinin zaman bağımlı olacağı durumlar da söz konusu olabilir. Fakat bu durumda oransal hazard varsayımı sağlanmayacağından dolayı, veriye klasik Cox model formu uygulanamaz. Bu tip veri kümeleri için geliştirilmiş Cox model formu uygundur.

Cox modelde $h_0(t)$ fonksiyonu belirlenmemiş olması rağmen regresyon katsayıları, hazard oranları ve yaşam sürdürme etkileri etkili bir şekilde elde edilebilmektedir. Ayrıca bu elde edilen değişkenler, uygun dağılımın bilinmesi durumunda elde edilecek olan değişkenlere oldukça yakındır.

Ayrıca Cox modelin çok kullanılması sağlayan bir başka özelliği de, yaşam sürdürme zamanları ile ilgili açıklayıcı bilgi var olduğunda ve durdurulmuş veriler söz konusu ise lojistik modele göre üstünlüktür. Lojistik model, yantı değişkeni olarak sadece $\{0,1\}$ kullanılır, durdurulmuş verileri yok sayarken; Cox model yaşam sürdürme zamanları, yantı değişkeni olarak kullanılır ve durdurulmuş verileri de analize dahil eder.

Cox regresyon parametrelerinin tahmininde k,smi ençok olabilirlik yöntemi kullan,labilir. Genel olarak ençok olabilirlik fonksiyonu yan,t de i keninin da ,l,m,n, kullan,r. Fakat daha önce de belirtti imiz gibi Cox modelde yan,t de i keninin da ,l,m, bilinmemektedir. Bu sebeple Cox model için tam bir ençok olabilirlik fonksiyonu olu turulamamaktad,r. Cox model için k,smi ençok olabilirlik fonksiyonu, ortak olas,l,k yo unluk fonksiyonu yerine, olaylar,n gözlemlenme s,ralar,n, kullan,r.

K,smi ençok olabilirlik fonksiyonu, yaln,zca durdurulmam, gözlemler için olas,l,k de erini dikkate al,rken, durdurulmu verileri belirgin olarak hesaba katmaz. Bu sebeple k,smi ençok olabilirlik tahmin edicisi olarak adland,r,l,r.

Genel olarak k,smi ençok olabilirlik fonksiyonu, ayr, ayr, olas,l,klar,n çarp,m, olarak yaz,labilir. r- tane ayr, ölüm zaman,na sahip n tane birey olsun ve bu bireylerin n-r tanesi sa dan durdurulmu olsun. Bu durumda t_j , j. s,radaki ya am sürdürme zaman, olmak üzere, r tane ya am sürdürme zaman,, t_1, t_2, \dots, t_r ve t_j an,nda risk alt,nda bulunan bireylerin kümesi, $R(t_j)$ olarak gösterilsin. Risk kümesi olarak adland,r,lan $R(t_j)$, en az t_j an,na kadar ya am, olan birimlerin grubudur.

Genel olarak k,smi ençok olabilirlik fonksiyonu

$$l_j(\beta) = \frac{h_0(t_j) \exp(\beta^T X_j)}{\sum_{l \in R(t_j)} h_0(t_j) \cdot \exp(\beta^T X_l)} = \frac{\exp(\beta^T X_j)}{\sum_{l \in R(t_j)} \exp(\beta^T X_l)} \quad (3.2)$$

eklindedir.

,c, de i ken vektörü X_j olan bir bireyin, t_j anındaki ölüm riskidir. Payda ise, t_j anında ölüm riski ile kar , kar ,ya olan bütün bireylerin ölüm risklerinin toplamıdır. Ayrıca $l_j(\beta)$, t_j anına kadar yaşam, olan bir birimin t_j anında bozulması olasılığıdır. Sonuçta r tane bozulma zamanı üzerinden olasılıkların çarpımı olan, orantılı, hazard regresyon modeli için kısmi çok olasılabilirlik fonksiyonu,

$$l(\beta) = \prod_{j=1}^r \frac{\exp(\beta^T X_j)}{\sum_{l \in R(t_j)} \exp(\beta^T X_l)} \quad (3.3)$$

biçiminde yazılır. (3.3) de verilen olasılabilirlik fonksiyonuna durdurulmuş birimler katılmazlar. Ancak herhangi bir durdurma zamanından hemen önceki zamanda risk kümesinde bulunurlar. Her bozulma zamanında risk kümesi yeniden belirlenmiştir için olasılabilirlik fonksiyonu sadece bozulma zamanları, sıralanması, na bağlıdır. Eğer olasılabilirlik fonksiyonuna gözlemlenmiş bireylerin yanı sıra durdurulmuş bireyler de katılırsa olasılabilirlik fonksiyonu de ierek durdurmaya uygun hale gelir.

Cox oransal hazard modeli için kısmi çok olasılabilirlik fonksiyonu;

$$\delta_j = \begin{cases} 0, & \text{durdurma var ise} \\ 1, & \text{durdurma yok ise} \end{cases}$$

durdurma de i keni olmak üzere,

$$\left[\frac{\beta^T X_j}{\sum_{l \in R(t_j)} \exp(\beta^T X_l)} \right]^{\delta_j} \quad (3.4)$$

biçiminde gösterilir [3].

ÖRNEK

Tablo III.1 Örnek veri

Birey	Hastalanma Zaman,		Sigara içme durumu (X_1)
A	4	1	1
B	5	1	1
C	7	0	0
D	9	1	1

Tabloda, sigara içen bireylerin herhangi bir hastalığa yakalanma sürelerini veren, tek aç,klay,c, de i keni sigara içme durumu olan, bir veri kümesi verilmiş olsun. durdurma de i keni olmak üzere, veri kümesinden anlamaktadı, r ki, A ki isi t=4 an,nda hastalanmış, sigara içen; B ki isi t=5 an,nda hastalanmış, sigara içen; C ki isi t=7 an,nda çalış, ma d, , kalm, sigara içmeyen; D ki isi t=9 an,nda hastalanmış, sigara içen bir bireydir.

Bu veri kümesi için Cox oransal hazard modeli,

$$h(t, x_1) = h_0(t)e^{\beta_1 x_1}$$

eklindedir. A, B, C, D bireyleri için bireysel hazardlar,

ili kin hazardlar

	rd
A	$h_0(t)e^{\beta_1}$
B	$h_0(t)e^{\beta_1}$
C	$h_0(t)e^0$
D	$h_0(t)e^{\beta_1}$

olacaktır.

A ki isinin olay, ya ad, , anda tüm bireyler risk kümesindedir ve

$$l_1(\beta) = \left[\frac{h_0(t)e^{\beta_1}}{h_0(t)e^{\beta_1} + h_0(t)e^{\beta_1} + h_0(t)e^0 + h_0(t)e^{\beta_1}} \right]$$

eklinde yazılır. B ki isinin olay, ya ad, , anda A ki isi hariç herkes risk kümesindedir ve

$$l_2(\beta) = \left[\frac{h_0(t)e^{\beta_1}}{h_0(t)e^{\beta_1} + h_0(t)e^0 + h_0(t)e^{\beta_1}} \right]$$

dir. C ki isi durdurulmuş gözlem oldu undan kısmi en çok olabilirlik fonksiyonuna katılmaz. Son olarak D ki isinin olay, ya ad, , anda risk kümesinde kendisinden başka hiçbir birey kalmamıştır. Dolayısıyla ile

$$l_3(\beta) = \left[\frac{h_0(t)e^{\beta_1}}{h_0(t)e^{\beta_1}} \right]$$

dir.

ik fonksiyonu bu veri kümesi için üç terimden oluşur ve

$$l(\beta) = l_1(\beta) \times l_2(\beta) \times l_3(\beta)$$

eklenmiştir.

Yani,

$$l(\beta) = \left[\frac{e^{\beta_1}}{e^{\beta_1} + e^{\beta_1} + e^0 + e^{\beta_1}} \right] \times \left[\frac{e^{\beta_1}}{e^{\beta_1} + e^0 + e^{\beta_1}} \right] \times \left[\frac{e^{\beta_1}}{e^{\beta_1}} \right]$$

eklinde yazılabilir.

$l(\beta)$ fonksiyonunun logaritması, en büyük yapılarak β katsayıları tahminleri elde edilir. β katsayıları tahmin edilmesinde Newton-Raphson yöntemi kullanılmaktadır.

III.2.1 Newton-Raphson Yöntemi

Durdurulmuş amaç sürdürme verisinde β katsayıları tahminleri için çok olasılık fonksiyonunun maksimumunun elde edilmesinde Newton-Raphson yöntemi kullanılmaktadır.

$U(\beta)$, çok olasılık fonksiyonunun logaritması β göre birinci türevlerinden oluşan $p \times 1$ tipinde vektör olsun. Bu değer etkin skorlar vektörü olarak bilinmektedir.

$I(\beta)$, çok olasılık fonksiyonunun logaritması β göre ikinci türevlerinin negatif iaretlilerinden oluşan $p \times p$ tipinde bir matristir ve

(3.5)

$p = p_{PML}$

eklindedir ve $I(\beta)$ matrisi ö tahmin edilen bilgi matrisiö olarak bilinmektedir [11].

Newton-Raphson yöntemi ad,msal bir yöntemdir. Bu yöntem $(s+1)$. iterasyonda iken β parametre vektörünün tahmini $\hat{\beta}_{s+1}$ dir. $U(\beta_s)$, etkin skorlar,n vektörü ve $I^{-1}(\hat{\beta}_s)$ bilgi matrisinin tersi olarak ele al,nd, ,nda, $s=1,2,\dots$ ve m ad,m say,s, olmak üzere;

$$\hat{\beta}_{s+1} = \hat{\beta}_s + m \left[I^{-1}(\hat{\beta}_s) U(\hat{\beta}_s) \right] \quad (3.6)$$

olacakt,r.

Log-olabilirlik fonksiyonundaki de i im yeterince küçük oldu unda ya da katsay, tahminlerinin de erlerindeki görelî de i im yeterince küçük oldu unda hesaplama süreci sona erdirilir [10].

III.3 PARAMETRELERE L K N H POTEZ TESTLER VE GÜVEN ARALIKLARININ BEL RLENMES

β_i parametresi için, $i = 1, 2, \dots, p$ olmak üzere

$$H_0 : \beta_i = 0$$

$$H_1 : \beta_i \neq 0$$

hipotezi genel olarak α ,daki istatistikler kullan,larak test edilir.

istatisti i

lar,n,n istatistiksel anlaml,l, ,n, test etmede kullan,lan

Wald testi, ençok olabilirlik tahmin edicilerinin normal da ,ld, , varsay,m,na

dayan,r. $\hat{S}(\hat{\beta}_i)$, $\hat{\beta}_i$ nin tahmin edilen standart sapmas, olmak üzere;

$$\frac{\hat{\beta}_i}{\hat{S}(\hat{\beta}_i)} \in Z \quad (3.7)$$

dir.

Test istatisti i nin standart normal da ,ld, , varsay,l,r. $V(\hat{\beta})$, $\hat{\beta}$ n,n varyans kovaryans matrisi olmak üzere, $\hat{\beta}' [V(\hat{\beta})]^{-1} \hat{\beta}$ ifadesi p serbestlik derecesi ile asimptotik ki-kare da ,l,m,na sahip olacakt,r. Dolay,s,yla Wald test istatisti i

$$\chi_w^2 = \hat{\beta}' [V(\hat{\beta})]^{-1} \hat{\beta} \quad (3.8)$$

eklinde de ifade edilebilir.

III.3.2 Olabilirlik Oran Test istatisti i

Modeldeki β katsay,lar,n,n istatistiksel anlaml,l, ,n, test etmede kullan,lan bir ba ka test de olabilirlik oran test istatisti i dir. Herhangi bir β_i parametresinin anlaml,l, ,n,n testi için olabilirlik oran test istatisti i elde edilirken; indirgenmi yani X_i yi içermeyen modelin ençok olabilirlik fonksiyonunun logaritmas,n,n -2 kat,ndan, tam modelin yani X_i yi de içeren modelin ençok olabilirlik fonksiyonunun logaritmas,n,n -2 kat, ç,kar,l,r. Elde edilen test istatisti i nin p serbestlik dereceli ki-kare da ,l,m,ndan oldu u kabul edilir

$$-2 \left[\ln L_{indirgenmi} - \ln L_{tam} \right] \sim \chi^2 \quad (3.9)$$

Genel olarak Wald testi ile olabilirlik oran test istatisti i parametrelerin anlamlı, için aynı sonuçlar, vermez. Bu gibi bir durumda olabilirlik oran test istatisti i daha iyi istatistiksel özelliklere sahip olması, nedeniyle tercih edilmelidir.

Bir β_i parametresi için $1-\alpha$ anlamlı, düzeyinde güven aralığı,

$$\hat{\beta}_i \pm z_{\alpha/2} \hat{S}(\hat{\beta}_i) \quad (3.10)$$

dir [9].

III.4 HAZARD ORANININ HESAPLANMASI VE YORUMU

Genel olarak hazard oranı, iki farklı birimin hazardları, birbirlerine oran, olarak tanımlanır. Yaşam sürdürme analizinde iki farklı grubu karşılaştırmak için hazard oranları kullanılabilir. X_1, X_2, \dots, X_p ve $X_1^*, X_2^*, \dots, X_p^*$ iki ayrı gruba ait açıklayıcı değişkenler olmak üzere, hazard oranı ($H\ddot{O}$);

$$H\ddot{O} = \frac{h(t, X^*)}{h(t, X)} = \frac{\hat{h}_0(t) e^{\sum_{i=1}^p \hat{\beta}_i X_i^*}}{\hat{h}_0(t) e^{\sum_{i=1}^p \hat{\beta}_i X_i}} = \theta \quad (3.11)$$

eklinde tanımlanır. Burada θ , sabit bir değerdir.

Kanser tedavisi gören bir grup hasta üzerinde A ve B gibi iki farklı ilaç etkisi ara t r l yor olsun. Kullanılan ilaç tipi d , nda, hastaların ö kanındaki lökosit miktarı (KLM) kaydedilmiş olsun.

$X^* = (X_1^*, X_2^*)$, $X_1^* = 1$ olmak üzere A ilacını kullanan hasta grubunu,

$X = (X_1, X_2)$, $X_1 = 0$ olmak üzere B ilacını kullanan hasta grubunu temsil etsin.

İkinci açıklama, de i kenler KLM 'yi gösterebilir. **ilaç tipi de i ken** için tahmin edilen hazard oranı,

$$HO = \frac{h(t, X^*)}{h(t, X)} = \frac{\dot{h}_0(t) e^{\sum_{i=1}^2 \beta_i^* X_i^*}}{\dot{h}_0(t) e^{\sum_{i=1}^2 \beta_i^* X_i}} = e^{\beta_1^*(1-0) + \beta_2^*(KLM - KLM)} = e^{\beta_1^*}$$

bulunur.

Genel iki seviyeli bir de i ken için hazard oranı, o de i kene ait olan katsayıların üstelidir. Bu oran, 1 ile kodlanmış olan etkenin 0 ile kodlanmış olan etkene göre bozulma üzerinde kaç kat daha etkili olduğunu göstermektedir.

Eğer de i kenimiz sürekli bir de i kense bu durumda uygun şekilde gruplandırılması gereklidir. Eğer ikiden fazla seviye söz konusu ise her bir seviye için hazard oranı hesaplanırken, diğer seviyeler birleştirilip tek bir seviyeyi gibi düşünülmelidir.

Ayrıca genel olarak anlamlı olan de i kenlerin hazard oranlarına bakıldığında en büyük hazard oranına sahip olan de i ken bozulma üzerinde en etkili ve s ras yla daha küçük hazard oranlarına sahip olan de i kenlerin bozulma üzerinde daha az etkili oldukları söylenilebilir.

DE KEN ETK LER N ÇEREN YA AM NU Ç N KAPLAN-MEIER TAHM N

Ya am sürdürme verisi için uygun da ,l,m,n bilinmemesi durumunda, Kaplan-Meier yönteminden yararlan,larak ya am sürdürme tahminleri elde edilebilir .

Ya am sürdürme verisi için Cox model kullan,ld, ,nda, ya am sürdürme e rileri aç,klay,c, de i kenlerin etkisini de içerecek ekilde geli tirilebilir ve ayn, Kaplan-Meier e rileri gibi bu e rilerde basamak fonksiyonlar, ekinde çizilebilir.

Kaplan-Meier e rileri olu turulurken, sadece bozulma zamanlar, kullan,larak elde edilen temel ya am sürdürme fonksiyonunun ($S(t)$), farklı t de erleri için alaca , de erler, parametrik olmayan yöntemlerle tahmin ediliyordu. (2.6) da genel olarak ya am sürdürme fonksiyonu,

$$S(t) = \exp \left[- \int_0^t h(u) du \right]$$

olarak verilmi ti. Buradan yola ç,karak Cox modele ait ya am sürdürme fonksiyonu olan $S(t, X)$ için,

$$S(t, X) = \exp \left[- \int_0^t h(u) e^{\sum_{i=1}^p \beta_i X_i} du \right] = \exp \left[e^{\sum_{i=1}^p \beta_i X_i} \left(- \int_0^t h(u) du \right) \right]$$

$$= \left[\exp \left(- \int_0^t h(u) du \right) \right]^{e^{\sum_{i=1}^p \beta_i X_i}}$$

$$S(t, X) = \left[S(t) \exp \sum_{i=1}^p \beta_i X_i \right] \quad (3.12)$$

elde edilir.

olmadı, bu nedenle $S(t)$ parametrik olarak bilinmemektedir.

Bu durumda, $S(t, X)$ için Kaplan-Meier tahmini,

$$\tilde{S}(t, X) = \left[\tilde{S}(t) \exp \sum_{i=1}^p \beta_i X_i \right] \quad (3.13)$$

şeklinde yazılabilir. Burada $\tilde{S}(t)$, $S(t)$ için Kaplan-Meier tahminidir.

III.6 ORANTILI HAZAR VARSAYIMININ İNCELENMESİ

Cox regresyon modelinin en önemli varsayımı, olan orantılı, hazard varsayımı, herhangi bir birime ili kin hazard fonksiyonunun di er herhangi bir birimin hazard fonksiyonu ile oranının zamandan ba msız bir sabit olmasıdır. Yani, bir açıklama, de i kenin iki farklı seviyesinin bozulma riski üzerindeki etkisini kapsayan hazard oranının, zaman içerisinde sabit kalmasıdır.

Bu varsayımın geçerliliği için orantılılık testi yapılması önemlidir. Genel olarak orantılı, hazard varsayımının kontrol etmek için grafiksel ve istatistiksel yöntemler kullanılır.

III.6.1 Grafiksel Yöntemler

Orantılı, hazard varsayımının testi edilmesinde kullanılan iki grafiksel yaklaşım aşağıda anlatılmaktadır.

III.6.1.1 Log-log Grafikleri

Orantılı, hazard varsayımının kontrolünde en sık kullanılan grafiksel yöntem, incelenen de i kenlerin farklı seviyelerine ili kin log-log yaşam sürdürme eğrilerinin

paralel ise oran, r, t, k varsa, m, sa lan, yor demektir.

Log-log ya am sürdürme e risi, tahmin edilen ya am sürdürme fonksiyonunun iki kere logaritmas, n, n al, nmas, yla olu turulur. Matematiksel olarak log-log e risi,

$$-\ln(-\ln \tilde{S}(t)) \quad (3.14)$$

eklinde ifade edilmektedir.

Tahmini ya am sürdürme fonksiyonu $\tilde{S}(t)$, daima (0,1) aral, nda de erler alaca ndan, logaritmas, negatif bir say, d, r. Bu yüzden ilk logaritma al, nd, ktan sonra negatifli in yok edilmesi için ifade (-) ile çarp, larak pozitif hale getirilir. $-\ln(-\ln \tilde{S})$ sonucun da elde edilen de er pozitif veya negatif olabilir. Ya am sürdürme e rileri basamak fonksiyonu olarak çizildi inden log-log e rileri de basamak fonksiyonu olarak çizilmektedir.

Ya am sürdürme e risinin y eksenindeki de erleri (0,1) aral, ndayken, $-\ln(-\ln \tilde{S})$ e risinde $(-\infty, +\infty)$ aral, ndad, r. Orant, l, hazard modeli ya am sürdürme fonksiyonu

$$S(t, X) = S_0(t) \exp \left[\sum_{i=1}^p \beta_i X_i \right] \quad (3.15)$$

eklinde ifade edilir. Ya am sürdürme fonksiyonunun iki kez logaritmas, al, nd, nda,

$$\left[\sum_{i=1}^p \beta_i X_i \right] \cdot \ln S_0(t)$$

$$\ln[-\ln S(t, X)] = \ln \left\{ -\exp \left[\sum_{i=1}^p \beta_i X_i \right] \cdot \ln S_0(t) \right\}$$

$$\ln[-\ln S(t, X)] = \ln \left\{ \exp \left[\sum_{i=1}^p \beta_i X_i \right] \right\} + \ln[-\ln S_0(t)]$$

$$-\ln[-\ln S(t, X)] = -\sum_{i=1}^p \beta_i X_i - \ln[-\ln S_0(t)] \quad (3.16)$$

ifadesi elde edilir [3]. ki farklı seviyesi olan bir aç, klay, c, de i kenin oransal hazard varsay, m, n, sa lay, p sa lamad, , n, kontrol etmek istensin. $X_1 = (x_{11}, x_{12}, \dots, x_{1p})$ bu de i kenin birinci seviyesine ait bir birimin aç, klay, c, de i ken vektörü, $X_2 = (x_{21}, x_{22}, \dots, x_{2p})$ de bu de i kenin ikinci seviyesine ait bir birimin aç, klay, c, de i ken vektörünü gösteriyor olsun. Bu iki birime ait log-log e rileri,

$$-\ln[-\ln S(t, X_1)] = -\sum_{j=1}^p \beta_j X_{1j} - \ln[-\ln S_0(t)]$$

$$-\ln[-\ln S(t, X_2)] = -\sum_{j=1}^p \beta_j X_{2j} - \ln[-\ln S_0(t)]$$

eklinde ifade edilir.

kinici log-log e risi birinciden ç, kart, ld, , nda

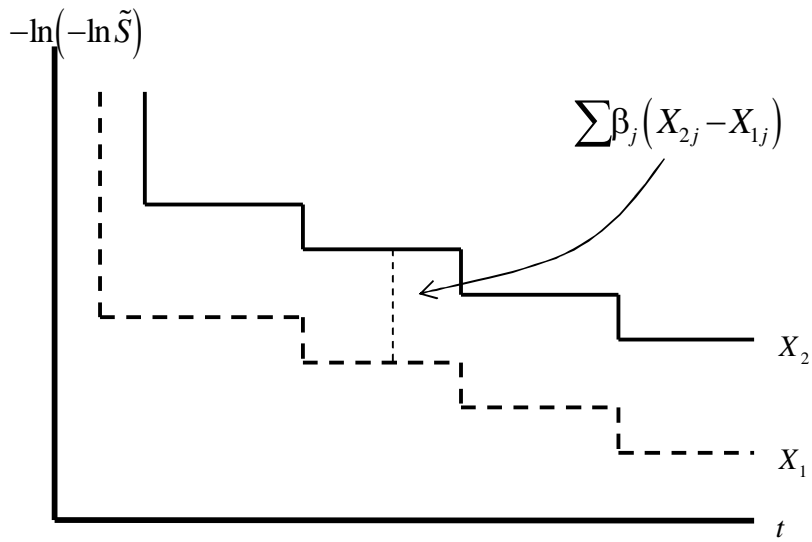
$$-\ln[-\ln S(t, X_1)] - \{-\ln[-\ln S(t, X_2)]\} = \sum_{j=1}^p \beta_j [X_{2j} - X_{1j}] \quad (3.17)$$

e itli i elde edilir. (3.14) yeniden düzenlendi inde

$$= -\ln[-\ln S(t, X_2)] + \sum_{j=1}^p \beta_j [X_{2j} - X_{1j}] \quad (3.18)$$

olur.

Bu iki birey için log-log ya am e rileri ayn, grafikte çizildi inde, e rilerin paralellik gösterip göstermedi i anla ,labilir.



ekil III.I Bireylere ili kin log-log e rileri

Uygulamada elde edilen e riler kesi miyor ve kabaca paralel gözüküyorsa varsay,m,n sa land, , kabul edilebilir. Bu karar ara t,rmac,n,n inisiyatifinde oldu undan yöntemin güvenilirli i sars,lmaktad,r. Ayr,ca sürekli de i kenlerin s,n,fland,rma ekli farkl, sonuçlar elde edilmesine sebep olabilir. S,n,f say,s,n,n mümkün oldu unca az olmas,na ve s,n,flar,n seçiminde anlaml,l, a dikkat edilmelidir. Bu kararlar da yine ara t,rmac,ya ba l, oldu undan grafiksel yöntemin uygulan, ,nda ki inin seçimlerinin ön planda oldu u görülür.

Beklenen Yaşam Etkileri Yaklaşım,

Özellikle, de erlendirmekte kullanılan grafiksel yöntemlerden biri tahmin edilen etkilerin gözlenen etkilerle karşılaştırılmasıdır.

Gözlenen ve beklenen yaşam etkilerini oluştururken gözlenen ve beklenen yaşam olasılıkları kullanılır. Bu etkileri oluşturmak için öncelikle veri ilgilenilen de ikenin seviyelerine göre ayrıştırılır. Daha sonra her seviye için ayrı ayrı Kaplan-Meier etkileri çizilerek gözlenen etkiler elde edilir. Aynı şekilde ilgilenilen de ikenin her seviyesi için beklenen etkiler, (3.13)de verilen açıklayıcı de iken etkilerini içeren yaşam sürdürme fonksiyonu tahminleri kullanılarak çizilir.

Gözlenen ve beklenen değerler karşılaştırıldığında, her bir seviye için ayrı ayrı oluşturulan gözlenen ve beklenen etkiler aynı grafik üzerine çizilir. İlgilenilen de ikenin her bir düzeyi için, beklenen ve gözlenen etkiler birbirine yakınsa, birbirisiyle örtüşüyorsa, orantılı hazard varsayımına karar verilir.

Eğer sürekli bir de ikenin orantılı hazard varsayımına sahip sağlamada, aynı kontrol etmek istiyorsak, öncelikle ilgilenilen de iken uygun şekilde seviyelere ayrılır. Gözlenen etkilerin elde edilmesi için aynı şekilde her seviye için Kaplan-Meier etkileri oluşturulur.

Beklenen etkilerin çizimi için iki ayrı yöntem söz konusudur. Yöntemlerden ilki, k tane seviyesi bulunan bir de iken için $k-1$ tane kukla de iken içeren Cox oransal hazard regresyon modelinin kullanılmasıdır. Model parametreleri belirlendikten sonra, her bir seviye için, açıklayıcı de iken etkilerini içeren yaşam sürdürme fonksiyonu tahminleri kullanılarak beklenen etkiler oluşturulur. İkinci yöntemde ise önce Cox oransal hazard modeli kullanılarak ilgilenilen sürekli de ikene ilişkin parametre tahmini elde edilir. Daha sonra her bir seviye için açıklayıcı de iken etkilerini içeren yaşam sürdürme fonksiyonu tahminleri, o seviyeye ait ortalamalar kullanılarak elde edilir.

Yaklaşım, m, n, da oldu u gibi, Bu yaklaşım, m, n, da da, gözlenen ve beklenen c, t, n, e, n, i, n, yaklaşıp olmad, , karar, ara t, r, m, a, c, n, n, de erlendirmesidir. Bu da yöntemin zay, f, yan, d, r. Fakat çizilen grafiklerde ciddi bir farklılık varsa orant, l, hazard varsay, m, n, n, sa lanmad, , na karar verilebilir.

III.6.2 istatistiksel Yöntemler

Orant, l, hazard varsay, m, n, de erlendirmede kullanılan istatistiksel yöntemler uyum iyili i yaklaşım, ve zamana ba l, de i kenlerin kullan, l, mas, d, r. Bu yöntemlerde test istatisti i kullan, l, aca, n, dan elde edilen sonuçlar grafiksel yöntemlere göre daha objektif olacaktır, r.

III.6.2.1 Uyum iyili i Yaklaşım,

Uyum iyili i testi, orant, l, hazard varsay, m, n, de erlendirmek amacıyla, de i ken için p de eri ve istatistiksel güvenilirli i yüksek bir test istatisti i sa lar. Bu nedenle grafiksel yöntemlere göre daha kesin ve objektif sonuç verir.

Literatürde uyum iyili i yaklaşım, için geli tirilmi bir çok farklı test bulunmaktadır, r. Bu bölümde Harrel ve Lee tarafından 1986 y, l, nda geli tirilmi olan yaklaşım k, s, a, c, a, anlat, l, aca, t, r, [12].

Testi uygulamak için ilk adım her de i ken için Schoenfeld art, klar, n, n, olu turulmas, d, r. Örne in vücut kitle indeksi (VK) aç, k, l, a, y, c, de i keni için Schoenfeld art, klar, n, hesaplamak istiyelim. Herhangi bir t_j an, nda ölmü birey için Schoenfeld art, , o birey için gözlenen VK de erinden, hala risk kümesinde bulunan bireylerin a, r, l, k, l, a, n, d, r, l, m, VK de erleri ç, k, a, r, t, l, a, r, a, k, elde edilir. Burada kullanılan a, r, l, k, l, a, r, her bireye ili kin hazard fonksiyonlar, n, n, (2.14) de verilen Kaplan-Meier tahminleridir.

Bozulma zamanlar, n , s,ralayan bir s,ralama de i kenini
oldu tartılır. Örne ğin ilk bozulma birim için s,ralama de i keninin alaca ğı, de er 1
olacaktır ve bozulma s,ras,na göre bu de i ken 2,3ı gibi de erler olacaktır.

Son olarak, ilk ve ikinci ad,mda olu turulan de i kenler aras,ndaki
korelasyon test edilir. Bu test için yokluk hipotezi Schoenfeld art,klar, ile s,ralama
de i kenini aras,ndaki korelasyonun s,f,r olmas,d,r. Yokluk hipotezinin ret edilmesi,
orant,l, hazard varsay,m,n,n sa lanmadı ğı, n, gösterir.

Uyum iyili ği testleri avantajlarla birlikte dezavantajlara da sahiptir. Uyum
iyili ği testi, genel anlamda orant,l, hazard varsay,m,ndan sapmalar, tespit etmek için
yap,lm, bir kontrol mekanizmas,d,r. Ancak bu test, orant,l, hazard varsay,m,nda
anl,k bir sapmay, tespit edemeyebilir. Grafikselle yöntemler ise bu sapmalar,n
belirlenmesine olanak tan,r. Sonuç olarak orant,l, hazard varsay,m,n,
de erlendirmede hem grafikselle yakla mlar,n hem de istatistiksel testlerin birlikte
kullan,lm, önerilir.

III.6.2.2 Zamana Ba lı, De i kenlerin Kullan,lm,ı,

Orant,l, hazard varsay,m,n, de erlendirmede zamana ba lı, de i kenleri
kullanmak için, Cox model, zamandan ba ms,z de i ken ve herhangi bir zaman
fonksiyonunun çarp,m,ndan olu an terimi içerecek ekilde geni letilir. Bu çarp,m
de i keninin katsay,s, anlams,z ç,karsa orant,l, hazard varsay,m,n,n zamandan
ba ms,z de i ken için sa landı ğı, söylenir.

t birim için t zaman,ndaki aç,klay,c, de i ken de eri X_i olmak üzere zamana
ba lı, de i kenleri içerecek ekilde geni letilmi Cox modeli,

$$\exp \left[\sum_{i=1}^p \beta_i X_i + \sum_{i=1}^p \Psi_i X_i \cdot g_i(t) \right] \quad (3.19)$$

biçiminde elde edilir. Modelde yer alan $g_i(t)$ fonksiyonu bir zaman fonksiyonudur ve bu fonksiyon için aşağıdaki gibi fonksiyonlar kullanılabilir:

$$g(t) = t \quad (3.20)$$

$$g(t) = \log t \quad \text{ya da} \quad (3.21)$$

$$g(t) = \begin{cases} 1 & t \geq t_0 \\ 0 & t < t_0 \end{cases} \quad (3.22)$$

X_i de i keninin orantılı olup olmadığını test edilmesi için (3.19) e ilgili inde verilen Ψ_i katsayılarının anlamlılığını test edilir.

$$H_0 : \Psi_i = 0$$

$$H_1 : \Psi_i \neq 0$$

Burada H_0 hipotezi reddedilemez ise X_i de i keninin orantılı, hazard varsayımına, sağladığına karar verilir. Eğer H_0 hipotezi reddedilir ve $\Psi_i < 0$ ise hazard oran, zamana bağlı olarak azalır, $\Psi_i > 0$ ise hazard oran, zamana bağlı olarak artar.

Zamana bağlı de i kenler için geneltilmi Cox orantılı, hazard regresyon modelinin testi için kullanılan istatistik; (3.9) da verilen 1 serbestlik dereceli olabilirlik oran test istatistiki dir.

Zamana bağlı de i kenler orantılı, hazard varsayımına, sağlanmadığına kontrolünde kullanıldığında, gibi, orantılı, hazard varsayımına, sağlanmadığında her bir zaman aralığında de i en hazard oranının belirlenmesinde de kullanılabilir [13].

PARAMETRİK YA AM SÜRDÜRME MODELLERİ

Bu bölümde, parametrik ya am sürdürme modelleri ve bu modellere ait varsayımlar incelenecektir. Özel olarak Weibull, log-normal, bozulma zaman, modeli varsayımlar, ve buna karşılık oransal hazard modeli varsayımlar, irdelenecek, üssel, Weibull ve log-lojistik da, içeren parametrik ya am sürdürme modelleri ile ilgili örnekler verilecektir.

Parametrik ya am sürdürme modelleri, açıklayıcı, de iken olan zaman, da, net olarak bilindiği modellerdir. Weibull, üssel, log-lojistik, log-normal ve genelleştirilmiş gamma da, ya am sürdürme zaman, için en çok kullanılan da,lardır. Bu bölümde bu da,lardan üssel, Weibull ve log-lojistik da, üzerinde durulacaktır.

Parametrik modellerin aksine Cox oransal hazard regresyon modeli yarı parametrik bir modeldir. Modelde regresyon parametreleri bilinmesine karşın, açıklayıcı, de ikenin da, bilinmemekte bu sebeple hazard fonksiyonu parametrik olarak belirlenememektedir. Dolayısıyla, bu model da, varsayımlar gerektirmemektedir. Bu da modelin popüler olmasının başlıca sebebidir.

Kaplan-Meier eğrileri, da, olmayan deneysel bir takım yöntemler kullanılarak elde edilmekte ve teorideki ya am sürdürme fonksiyonlarından farklı olarak basamak fonksiyonu olarak oluşturulmaktadır. Parametrik modellerde ise da, bilgisi var olduğundan, kuramsal ya am sürdürme eğrileri ile daha tutarlı olan tahmini ya am sürdürme eğrileri elde edilir. Eğer ara tirmacı, zaman de ikeni için uygun olan da, net olarak belirleyebilirse, hazard ve ya am sürdürme fonksiyonunu parametrik olarak kolaylıkla belirleyebilir. Bu durumda, Cox modelle

ve tahmin de erleri elde edilebilir. Bu kullan,m kolayl ,
ve tutar,r,k parametrik model yakla ,m,n, cazip hale getiren özelliklerdir.

Ya am sürdürme modelleri oransal hazard ve h,zland,r,lm, bozulma zaman, modelleri olarak iki gruba ayr,l,r. Bunlardan ilki olan oransal hazard modellerinde, aç,klay,c, de i kenlerin hazard fonksiyonu üzerinde çarp,msal etkisi oldu u varsay,m,na kar ,n, h,zland,r,lm, ya am sürdürme modellerinde, aç,klay,c, de i kenlerin, ya am sürdürme zaman, üzerinde çarp,msal etkisi oldu u varsay,m, söz konusudur.

Örne in, köpeklerin insanlara göre yedi kat daha h,zl, ya land,klar, bilinmektedir. Yani, on ya ,ndaki bir köpe in ya lanma düzeyi, yetmi ya ,ndaki bir insan ile ayn,d,r. H,zland,r,lm, bozulma modellerine göre, bir köpe in on y,l ya am sürdürme olas,l, ,, bir insan,n 70 y,l ya am sürdürme olas,l, ,na e ittir. Daha genel olarak S_K ve S_I s,ras,yla, köpeklere ve insanlara ait ya am sürdürme fonksiyonlar,n, göstermek üzere, $S_K(t) = S_I(7t)$ dir.

H,zland,r,lm, bozulma modelleri, ya am sürdürme zaman, üzerindeki bu daralmay,, aç,klay,c, de i kenlerin bir fonksiyonu olarak ifade eder.

Genel olarak h,zland,r,lm, bozulma zaman, varsay,m,,

$$S_2(t) = S_1(\gamma t) \quad t \geq 0 \quad (4.1)$$

eklinde ifade edilir. Bu e itlikte γ , h,zland,r,ma faktörü olarak adland,r,lan bir sabittir. Regresyon çal, malar,nda herhangi bir de i kene ait h,zland,r,ma faktörü, $\exp(\alpha)$ ekinde yaz,labilir. Burada α , ilgilenilen de i kene ili kin parametre tahminidir. H,zland,r,lm, bozulma zaman, varsay,m, rastgele de i kenler cinsinden,

$$T_1 = \gamma T_2 \quad (4.2)$$

Oransal hazard modellerinde, hazard oran, λ , de i kenlerin bozulma riski üzerindeki etkilerini belirlerken, $h(t)$, bozulma modellerinde, $h(t)$ faktörü, λ , de i kenlerin ya am sürdürme süresi üzerindeki etkilerini belirler.

Bu bölümde inceleyece imiz üssel model ve iki weibull model hem oransal hazard varsay,mlar,n,, hemde $h(t)$, bozulma zaman, modeli varsay,mlar,n,, sa layan modellerdir.

IV.1 ÜSSEL MODEL

Üssel da λ , ba ar,s,zl,k zaman, modeli olarak kullan,lan ilk da λ ,md,r. T,bbi ara t,rmalarda ve mühendislikte güvenilirlik çal, malar,nda geni ölçüde kullan,lr.

Üssel da λ ,mda, ara t,rma ba lad,ktan sonra herhangi bir zamanda ba ar,s,zl,k riski ayn,d,r ve λ öya e ittir. Bu durumda, λ parametrelili ve $1/\lambda$ ortalamal, üssel da λ ,m için hazard fonksiyonu zamana göre sabit olup;

$$h(t) = \lambda \quad 0 \leq t < \infty \quad (4.3)$$

olarak gösterilir ve λ parametresi, pozitif bir katsay,d,r. Üssel da λ ,m,n ya am sürdürme fonksiyonu (2.6) e itli inden yararlan,larak,

$$S(t) = \exp\left(-\int_0^t \lambda du\right) \quad 0 \leq t < \infty \quad (4.4)$$

$$0 \leq t < \infty \quad (4.5)$$

olarak gösterilir. Bu durumda (2.5) e itli inden üssel da ,l,m için Təmin olas,l,k yo unluk fonksiyonu;

$$f(t) = \lambda \exp(-\lambda t) \quad 0 \leq t < \infty \quad (4.6)$$

olarak elde edilir.

Parametrik ya am sürdürme modellerinde, tipik olarak, ölçek parametresi olarak adlandırılan λ , yeniden parametrelendirilir. Kolay anla ,lmas, aç,s,ndan, bu bölümde anlataca ,m,z tüm modelleri, cinsiyet gibi tek bir aç,klay,c, de i ken içeren bir örnek üzerinde aç,klayal,m. Cinsiyet de i keni için öö de eri kad,nlar,, ölöde eri erkekleri temsil etsin ve λ , $\exp(\beta_0 + \beta_1 \cdot \text{Cinsiyet})$ ekinde yeniden parametrelendirilsin. Bu durumda;

$$H(t, X) = \exp(\beta_0 + \beta_1 \cdot \text{cinsiyet}) \quad (4.7)$$

olacaktır. Bu modelde erkek bireyler için hazard $\exp(\beta_0 + \beta_1)$ ve kad,n bireyler için hazard $\exp(\beta_0)$ ekinde dir. Kad,n ve erkek gruplar,n, kar ,la t,ran hazard oran,;

$$HO (\text{cinsiyet} = 1' e \text{ kar} , \text{cinsiyet} = 2) = \frac{\exp(\beta_0 + \beta_1)}{\exp(\beta_0)} = \exp(\beta_1) \quad (4.8)$$

olarak elde edilir.

an üssel model, parametrik oransal hazard regresyon modelidir. Açıklayıcı, de i kenler hazard fonksiyonu üzerinde, çarpımsal bir etkiye sahiptir ve de i kenlere ait hazard oranlar, sabit olarak belirlenebilir.

Üssel dağılıma ait, $S(t) = \exp(-\lambda t)$, yaşam sürdürme fonksiyonu, t için çözümlenirse;

$$t = [-\ln(S(t))] \cdot \frac{1}{\lambda} \quad (4.9)$$

olarak bulunur. Burada $\frac{1}{\lambda} = \exp(\alpha_0 + \alpha_1 \cdot \text{cinsiyet})$ şeklinde ya da buna denk olarak $\lambda = \exp(-\alpha_0 - \alpha_1 \cdot \text{cinsiyet})$ şeklinde yeniden parametrelendirilmesi sonucu ;

$$t = [-\ln(S(t))] \exp(\alpha_0 + \alpha_1 \cdot \text{cinsiyet}) \quad (4.10)$$

modeli elde edilir. Görüldüğü üzere bu modelde, yaşam sürdürme fonksiyonunun belirli bir de eri için, açıklayıcı, de i kenler yaşam sürdürme süresi üzerinde çarpımsal bir etkiye sahiptir. Örneğin, $S(t)=0.5$ için yaşam sürdürme süresi medyan,;

$$t_m = [-\ln(0.5)] \exp(\alpha_0 + \alpha_1 \cdot \text{cinsiyet}) \quad (4.11)$$

olarak bulunur.

Herhangi bir sabit $S(t)=q$ olasılık için hızlandırma faktörü γ ;

$$\gamma = \frac{[-\ln(q)] \exp(\alpha_0 + \alpha_1)}{[-\ln(q)] \exp(\alpha_0)} = \exp(\alpha_1) \quad (4.12)$$

Model özel olarak; $\gamma = 1/HO$ dır.

Üssel oransal ve üssel h, zland, r, lm, bozulma zaman, modelleri, her ne kadar farklı, varsay, mlara dayansalar da asl, nda aralar, ndaki tek farklı, l, k parametrelendirilme biçimleridir. ki model için de ya am sürdürme fonksiyonu, hazard fonksiyonu ve ya am sürdürme süresi medyan, ayn, d, r.

IV.2 WE BULL MODEL

Son yıllarda, en çok ya am sürdürme zaman, da , l, mlar, ndan olan weibull da , l, m,, özellikleri nedeniyle özellikle endüstriyel çal, malarda tercih edilmektedir. Üssel da , l, m, n uygulanabilirli i sabit hazard h, z, na sahip oldu u varsay, m, ndan dolayı, s, n, rl, olmas, na ra men, weibull da , l, mlar, ailesi, hem artan hem de azalan hazard h, z, n, içerd i için uygulanabilirlik aç, s, ndan daha esnektir.

Biz bu bölümde iki parametrel i weibull da , l, m, ile ilgilenece iz. ki parametrel i weibull da , l, m, için ya am sürdürme fonksiyonu, olas, l, k yo unluk fonksiyonu ve hazard fonksiyonu s, ras, yla a a , daki gibi verilir:

$$S(t) = \exp(-\lambda t^p) \quad (4.13)$$

$$f(t) = \lambda p(t)^{p-1} \exp(-\lambda t^p) \quad (4.14)$$

$$h(t) = \lambda p(t)^{p-1} \quad (4.15)$$

Hazard fonksiyonunun ekli, p parametresine ba l, d, r. Bu nedenle p , biçim parametresi olarak ve λ , üssel da , l, mda oldu u gibi ölçek parametresi olarak adland, r, l, r. ki parametrel i weibull da , l, m, n, n hazard fonksiyonu, $p < 1$ için

monoton artan ve $p = 1$ için sabit fonksiyondur. Burada $p = 1$ için iki parametrelili Weibull da λ, m , üssel da λ, m dönüşümlü olarak ifade edilebilir.

k parametrelili Weibull da λ, m, n önemli bir özelliği, da λ, m, n ya am sürdürme fonksiyonu olan $S(t)$ için $\ln[-\ln(S(t))]$ fonksiyonu ile $\ln(t)$ fonksiyonu arasında lineer bir ilişki bulunmasıdır. Bu şekilde var olan verinin iki parametrelili Weibull da λ, m, n uygunluğu, $\tilde{S}(t)$ ya am sürdürme fonksiyonunun Kaplan-Meier tahmini olmak üzere, $\ln[-\ln(\tilde{S}(t))]$ değerlerine karşı $\ln(t)$ değerleri grafiklenerek kontrol edilebilir. Eğer bu plotlama lineer bir ilişkiyi ortaya getiriyorsa, t değeri için iki parametrelili Weibull da λ, m, n uygun olduğu düşünülebilir.

k parametrelili Weibull da λ, m, n hazard fonksiyonu $h(t) = \lambda p(\lambda t)^{p-1}$ olmak üzere, Weibull oransal hazard modeli λ için $\lambda = \exp(\beta_0 + \beta_1 \cdot \text{cinsiyet})$ şeklinde yeniden parametrelendirilmesi ile elde edilebilir. Cinsiyet değeri için hazard oranı;

$$HO = \frac{\exp(\beta_0 + \beta_1) p t^{p-1}}{\exp(\beta_0) p t^{p-1}} = \exp(\beta_1) \quad (4.16)$$

şeklinde elde edilir. Bu değer, hazard fonksiyonunun eğilimi belirleyen, eğilim parametresi p için, cinsiyet değeri için her iki seviyesi için aynı olması durumunda hesaplanmaz, t.r. Aksi durumda, t değerleri sadeleştirilmez, hazard oranı zamanla değişken olarak elde edilir. Bu da hazard varsayımını sağlamadığını gösterir.

Weibull da λ, m, n için $S(t) = \exp(-\lambda t^p)$, ya am sürdürme fonksiyonu t için çözümlenirse;

$$t = [-\ln(S(t))]^{1/p} \frac{1}{\lambda^{1/p}} \quad (4.17)$$

$= \exp(\alpha_0 + \alpha_1 \cdot \text{cinsiyet})$ parametrelendirilmesi yapılacak

olursa;

$$t = [-\ln S(t)]^{1/p} \exp(\alpha_0 + \alpha_1 \cdot \text{cinsiyet}) \quad (4.18)$$

Weibull h, zland, r, lm, bozulma zaman, modeli elde edilir. Bu model kullanılarak ya am sürdürme zaman, medyan,;

$$t_m = [-\ln(0.5)]^{1/p} \exp(\alpha_0 + \alpha_1 \cdot \text{cinsiyet}) \quad (4.19)$$

olarak bulunur. Herhangi bir sabit $S(t)=q$ olasılığ, için cinsiyet de i kenine ait h, zland, rma faktörü γ ;

$$\gamma = \frac{[-\ln(q)]^{1/p} \exp(\alpha_0 + \alpha_1)}{[-\ln(q)]^{1/p} \exp(\alpha_0)} = \exp(\alpha_1) \quad (4.20)$$

eklinde hesaplanır. Oransal hazard modelde olduğu gibi bu de erin elde edilmesi de i kenin her iki seviyesi için p de erinin aynı olmas, durumunda mümkündür. Aksi takdirde h, zland, rma faktörü, q de erine ba lı olarak bulunur. Bu durumda, h, zland, r, lm, bozulma zaman, modeli varsayımlar, sağlanmaz.

Weibull model için e er oransal hazard varsayımı gerçekleşiyorsa, h, zland, r, lm, bozulma modeli varsayımı, tersine h, zland, r, lm, bozulma zaman, modeli varsayımı, gerçekleşiyorsa, oransal hazard varsayımı, gerçekleşir. Bu özellik yalnızca weibull modele özgü bir özelliktir.

K MODEL

[Click Here to upgrade to
Unlimited Pages and Expanded Features](#)

hazard fonksiyonunun tekdüze olmadı, , durumlarda kullanılır. Örneğin; bir organ naklinden sonra, vücut yeni organa uyum sağlarken ilk on gün içinde hastanın ölme riski ve böylece hazard fonksiyonu artmakta, hastanın iyileşmesi ile birlikte zamanla azalmaktadır. Böyle bir araştırmaya ilişkin verilerin analizinde log-lojistik dağılımı kullanılması uygundur.

Log-lojistik modele ait yaşam sürdürme fonksiyonu, olasılık yoğunluk fonksiyonu ve hazard fonksiyonu sırasıyla aşağıdaki gibi verilir:

$$S(t) = \frac{1}{1 + \lambda t^p} \quad (4.21)$$

$$f(t) = \frac{\lambda p t^{p-1}}{(1 + \lambda t^p)^2} \quad (4.22)$$

$$h(t) = \frac{\lambda p t^{p-1}}{1 + \lambda t^p} \quad (4.23)$$

Log-lojistik dağılımı, hastalandırılmı, bozulma zamanı modeli için uygun bir dağılım olması nedeniyle, oransal hazard modeli için uygun değildir. Bu dağılım için oransal hazard varsayımı sağlanmaz.

Log-lojistik dağılımı için, $S(t) = \frac{1}{1 + \lambda t^p}$, yaşam sürdürme fonksiyonu t için çözümlenirse;

$$t = \left[\frac{1}{S(t)} - 1 \right]^{1/p} \frac{1}{\lambda^{1/p}} \quad (4.24)$$

bulunur. Bu eşitlikte $\frac{1}{\lambda^{1/p}} = \exp(\alpha_0 + \alpha_1 \cdot \text{cinsiyet})$ parametrelendirilmesi yapılacaktır; olursa;

log-lojistik hızlandırılmış bozulma zaman modeli elde edilir. Bu model kullanılarak yaşıyla am südürme zamanı, medyanı;

$$t_m = \exp(\alpha_0 + \alpha_1 \cdot \text{cinsiyet}) \quad (4.25)$$

olarak bulunur. Herhangi bir sabit $S(t)=q$ olasılığı için cinsiyet de ikenine ait hızlandırılma faktörü γ ;

$$\gamma = \frac{[q^{-1} - 1]^{1/p} \exp(\alpha_0 + \alpha_1)}{[q^{-1} - 1]^{1/p} \exp(\alpha_0)} = \exp(\alpha_1) \quad (4.26)$$

eklinde hesaplanır.

Ç İL K DURUMUNDA COX ORANSAL HAZARD MODELİN FARKLI TAHMİN YÖNTEMLERİ

Bu bölümde, iç ili ki durumunda Cox oransal hazard regresyon modeli parametrelerinin tahmini için alternatif tahmin yöntemleri incelenmiştir.

Genel olarak Cox model parametrelerinin tahmininde kullanılan, kısmi en çok olabilirlik yöntemi, iç ili ki durumundan ciddi şekilde etkilenebilir. iç ili ki durumunda, belirli varyanslara sahip ve dolayısıyla diğerlerinden çok uzak kalan tahmin değerleri elde edilebilir.

Bu problemi çözmeye yönelik, 2007 yılında, ridge regresyon yaklaşım, Cox oransal hazard regresyon modeline uyarlanmıştır. Bu bölümde ilk olarak bu yaklaşım, daha bahsedilecek daha sonra da incelenilen diğer yöntemler anlatılacaktır.

V.1. COX ORANSAL HAZARD REGRESYON MODELİNİN RIDGE TAHMİNİ

ç ili ki durumunda, ridge regresyon yöntemi, lineer modellerde sıklıkla kullanılan bir yöntemdir ve aşağıdaki şekilde tanımlanır.

$$\hat{\beta}_R = (X'X + kI)^{-1}(X'X)\hat{\beta}_{EK} \quad (5.1)$$

Burada $\hat{\beta}_{EK}$, β parametresinin en küçük kareler tahmin edicisi ve $X'X$ matrisi de lineer modele ait olan bilgi matrisidir ve net olarak belirlidir. Daha önceki bölümlerde

sal hazard regresyon modeli;

$$h(t, X) = h_0(t) \cdot e^{\sum_{i=1}^p \beta_i X_i}$$

formundadır ve lineer modelin tersine Cox model için bilgi matrisi net olarak bilinmez. Cox model için bilgi matrisini tahmin etmeye yönelik bir takım yaklaşımlar mevcuttur [14].

$l(\beta)$; Cox modelin bilinmeyen β parametresine ilişkin kısmi ençok olabilirlik fonksiyonu olmak üzere, β parametresinin kısmi ençok olabilirlik tahmin edicisi $\hat{\beta}_{PML}$ için kovaryans matrisi [11];

$$K\ddot{ö}v(\hat{\beta}_{PML}) = \left[-\frac{\partial^2 l(\beta)}{\partial \beta \partial \beta'} \Big|_{\beta = \hat{\beta}_{PML}} \right]^{-1} \quad (5.2)$$

eklenmiştir ve (3.5) de verildiği üzere;

$$I(\beta) = -\frac{\partial^2 l(\beta)}{\partial \beta \partial \beta'} \Big|_{\beta = \hat{\beta}_{PML}} = \ddot{D}' \ddot{D}$$

matrisi tahmin edilen bilgi matrisi olarak bilinen $p \times p$ boyutlu bir matristir. Bu tahmin edilen bilgi matrisinden yararlanarak, Cox model için ridge regresyon tahmin edicisi;

$$\hat{\beta}^R = (\ddot{D}' \ddot{D} + kI)^{-1} \ddot{D}' \ddot{D} \hat{\beta}_{PML} \quad (5.3)$$

(5.3) ve (5.2) bir arada dü ünülürse, ridge tahmin edicisi için kovaryans matrisi,

$$Kov(\beta^R) = (D'D + kI)^{-1} D'D (D'D + kI)^{-1} \quad (5.4)$$

eklinindedir.

Cox modele ridge regresyon yakla ,m,, 2007 y,l,nda Xue ve arkada lar, taraf,ndan yap,lm, t,r. Bir benzetim çal,mas, olan bu çal,mada modelin performans, farklı durumlar için incelenmi ve iç ili kinin oldu u durumlarda genel olarak ridge regresyon yakla ,m,n,n, HKO ölçütü aç,s,ndan daha iyi sonuçlar verdi i sonucuna var,lm, t,r.

ç ili ki durumunda, bilgi matrisi kötü kulludur. Ridge regresyon yöntemiyle amaçlanan, bilgi matrisinin kö egen elemanlar,na ökö gibi küçük bir sabit ekleyerek, kötü kulluluk problemini yok etmek ve a a ,daki gibi tan,mlanan ko ul say,s, K øy, istenilen seviyeye çekmektir.

$$K = (en büyük özde er / en küçük özde er)^{1/2} \quad (5.5)$$

Uygulamada, ridge regresyonun ökö parametresi oldukça küçük seçilir. Oysa kolayl,kla gösterilebilir ki, $D'D + kI$ n,n ko ul say,s,, k n,n azalan bir fonksiyonudur. Bu yüzden, k ne kadar büyük seçilirse, $D'D + kI$ n,n ko ul say,s, da o kadar küçük olacakt,r. Özellikle çok yüksek iç ili ki durumunda, $D'D$ matrisi çok kötü kulludur ve uygulamada seçilen küçük k de eri kötü kulluluk probleminin çözümünde yetersiz kal,r. Di er taraftan, k n,n büyük seçilmesi regresyonun yanl,l, ,n, artt,r,r.

Bu problem göz önünde bulundurularak, lineer model için Liu-tipi tahmin edici geli tirilmi tir. Bölümde V.2. de, lineer model için geli tirilmi olan bu tahmin

regresyon modeline uyarlan, , ve uyarlanan bu tahmin edicinin performans,na m kın geli tirilmi teoremler verilecektir.

V.2. COX ORANSAL HAZARD REGRESYON MODEL Ç N L U- T P TAHM N ED C

Liu-tipi tahmin edici ilk olarak, lineer modellerde, ridge regresyon parametresi k ,n regresyona katt, , yanl,l, , azaltmak amac,yla tasarlanm, t,r ve a a ,daki ekilde tan,m,l,d,r;

$$\hat{\beta}_{kd} = (X'X + kI)^{-1}(X'X - dI)\hat{\beta} \quad ; \quad k > 0, \quad -\infty < d < \infty \quad (5.6)$$

Burada $\hat{\beta}$, β n,n herhangi bir tahmin edicisi $k > 0$ ve $-\infty < d < \infty$ belirlenmesi gereken sabitlerdir.

Liu-tipi tahmin yönteminde ilk olarak k parametresi kullan,larak bilgi matrisinin ko ul say,s, küçültülür daha sonra ikinci bir ödö parametresi yard,m,yla, k ,n kullan,m,n,n regresyona katt, , yanl,l,k azalt,l,r ve istatistiksel özellikler iyile tirilir. Bu özellikleri bak,m,ndan, bu tahmin edici, kötü ko ulluluk probleminin çözümünde oldukça etkilidir ve ayn, zamanda uygun d parametresinin seçilmesi durumunda, ridge tahmin edicisinden daha küçük HKOøya sahiptir [16].

Lineer modeldekine benzer olarak, Cox oransal hazard regresyon modeli için Liu-tipi tahmin edici;

$$\hat{\beta}^{kd} = (\hat{D}'\hat{D} + kI)^{-1}(\hat{D}'\hat{D} - dI)\hat{\beta} \quad , \quad k > 0, \quad -\infty < d < \infty \quad (5.7)$$

er modelde oldu u gibi burada da $\hat{\beta}$, Cox modelin bilinmeyen β parametresinin herhangi bir tahmin edicisi, $k > 0$ ve $-\infty < d < \infty$ belirlenmesi gereken sabitlerdir. Bu çal, mada $\hat{\beta}_{\text{ML}}$, $\hat{\beta}_{PML}$ ve $\hat{\beta}^R$ olmas, durumlar, ile ilgilenilmi tir.

lk olarak bilgi matrisi $D'D$ kanonik formda yaz,ls,n. θ , sütunlar,, $D'D$ matrisinin özvektörlerinden olu an ortogonal matris olmak üzere, $Z = D\theta$ ve $\hat{\alpha}_{PML} = \theta' \hat{\beta}_{PML}$ olarak tan,mlans,n. Bu durumda,

$$Z'Z = \theta' D' D \theta = \Lambda = \text{kö eger}(\lambda_1, \dots, \lambda_p) \quad (5.8)$$

olacaktır. Burada $\lambda_1 \geq \dots \geq \lambda_p > 0$, $D'D$ matrisinin sıral, özde erleridir. (5.3) de verilen $\hat{\beta}^R$ öye benzer ekilde, kanonik form için ridge tahmin edicisi, $\hat{\alpha}^R$;

$$\hat{\alpha}^R = (\Lambda + kI)^{-1} \Lambda \hat{\alpha}_{PML} \quad (5.9)$$

ve kanonik form için Liu-tipi tahmin edici $\hat{\alpha}^{kd}$, (5.7) de verilen $\hat{\beta}^{kd}$ öye benzer olarak

$$\hat{\alpha}^{kd} = (\Lambda + kI)^{-1} (\Lambda - dI) \hat{\alpha} \quad (5.10)$$

eklinedir.

$\hat{\alpha}^R$ ve $\hat{\alpha}^{kd}$ tahmin edicilerinden herhangi biri ve $\hat{\beta}_a$

karşılaştıkları gelen tahmin edici olmak üzere $\hat{\beta}_a$,daki ilikler geçerlidir.

$$\hat{\beta}_a = \theta \hat{\alpha}_a \quad (5.11)$$

$$HKO(\hat{\beta}_a) = HKO(\hat{\alpha}_a) \quad (5.12)$$

Eklerde bu iliklerin bazıları, n, n ispat, verilmiştir. İmdi sırasıyla $\hat{\beta} = \hat{\beta}_{PML}$

ve $\hat{\beta} = \hat{\beta}^R$ durumları için geliştirilmiş olan iki teorem verilecektir.

TEOREM 5.1

$E(\hat{\alpha}_{PML}) = (c_1, \dots, c_p)$ olmak üzere, herhangi bir $k > 0$ ve $\hat{\beta} = \hat{\beta}_{PML}$ için;

$$d = \frac{\sum_{i=1}^p [(\lambda_i c_i^2 - \alpha_i c_i (\lambda_i + k) + 1) / (\lambda_i + k)^2]}{\sum_{i=1}^p (\lambda_i c_i^2 + 1) / [\lambda_i (\lambda_i + k)^2]}$$

olarak seçilirse $HKO(\hat{\beta}^{kd}) \leq HKO(\hat{\beta}^R)$ dir. Etilik $d=0$ durumunda geçerlidir.

SPAT

(5.12)den $HKO(\hat{\beta}^{kd}) \leq HKO(\hat{\beta}^R)$ e itisizliğin ispat edilebilmesi için

$HKO(\hat{\alpha}^{kd}) \leq HKO(\hat{\alpha}^R)$ e itisizliğin ispat edilmesi yeterlidir. İlk olarak $\hat{\alpha}^{kd}$

tahmin edicisinin beklenen değeri

$$\begin{aligned} E(\hat{\alpha}^{kd}) &= E[(\Lambda + kI)^{-1}(\Lambda - dI)\hat{\alpha}_{PML}] \\ &= (\Lambda + kI)^{-1}(\Lambda - dI)E(\hat{\alpha}_{PML}) \end{aligned} \quad (5.13)$$

eninin beklenen de eri olan $E(\hat{\alpha}_{PML})$, px1 tipinde bir vektördür ve bu vektörün bile enleri,

$$E(\hat{\alpha}_{PML}) = (c_1, \dots, c_p) \quad (5.14)$$

olmak üzere;

$$E(\hat{\alpha}^{kd}) = \left(\frac{\lambda_1 - d}{\lambda_1 + k} c_1, \dots, \frac{\lambda_p - d}{\lambda_p + k} c_p \right) \quad (5.15)$$

olarak elde edilir.

$\hat{\alpha}^{kd}$ tahmin edicisinin kovaryans matrisi;

$$\begin{aligned} K\ddot{O}v(\hat{\alpha}^{kd}) &= K\ddot{O}v[(\Lambda + kI)^{-1}(\Lambda - dI)\hat{\alpha}_{PML}] \\ &= (\Lambda - dI)'(\Lambda + kI)^{-1'} K\ddot{O}v(\hat{\alpha}_{PML})(\Lambda + kI)^{-1}(\Lambda - dI) \\ &= (\Lambda - dI)'(\Lambda + kI)^{-1'}(\Lambda)^{-1}(\Lambda + kI)^{-1}(\Lambda - dI) \end{aligned} \quad (5.16)$$

eklinde hesaplan,r.

(5.15) ve (5.16) dan yararlan,arak $\hat{\alpha}^{kd}$ tahmin edicisinin hata kareler ortalamas,;

$$HK\ddot{O}(\hat{\alpha}^{kd}) = \left\| E(\hat{\alpha}^{kd}) - \alpha \right\|^2 + zK\ddot{O}v(\hat{\alpha}^{kd}) \quad (5.17)$$

olmak üzere,

$$HK\ddot{O}(\hat{\alpha}^{kd}) = \sum_{i=1}^p \left(\frac{(\lambda_i - d)c_i}{(\lambda_i + k)} - \alpha_i \right)^2 + \sum_{i=1}^p \frac{(\lambda_i - d)^2}{\lambda_i(\lambda_i + k)^2} = g(d) \quad (5.18)$$

olarak bulunur.

üzere $H\check{K}O(\alpha^{kd})$, d 'nin kuadratik bir fonksiyonudur

ve bu fonksiyonu minimum yapan d de eri (5.19) da verildi i gibi belirlenir;

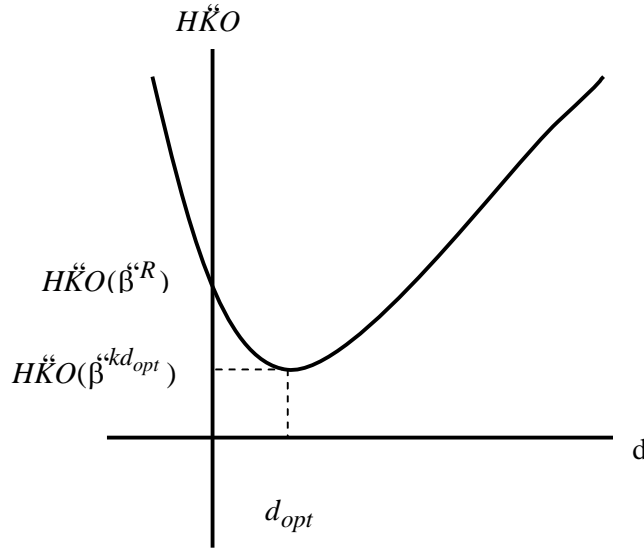
$$d_{opt} = \frac{\sum_{i=1}^p [(\lambda_i c_i^2 - \alpha_i c_i (\lambda_i + k) + 1) / (\lambda_i + k)^2]}{\sum_{i=1}^p (\lambda_i c_i^2 + 1) / [\lambda_i (\lambda_i + k)^2]} \quad (5.19)$$

$g(0) = H\check{K}O(\alpha^R)$ oldu u dü ünülürse, d parametresinin teorem 5.1 de

verildi i gibi seçilmesi durumunda, $H\check{K}O(\alpha^{kd}) \leq H\check{K}O(\alpha^R)$ ve buna denk olarak

$H\check{K}O(\beta^{kd}) \leq H\check{K}O(\beta^R)$ dir ve böylece ispat tamamlanm olur. spat,n daha rahat

anla ,labilmesi için a a ,daki grafiksel örnek verilmi tir.



ekil V.1 Teorem(5.1)'nin ispat,na yönelik grafiksel örnek

β^{kR} durumu için geli tirilmi tir. Anlat,m kolayl ,

aç,s,ndan bu durum için elde edilen Liu-tipi tahmin ediciyi β^{kdR} olarak gösterelim.

TEOREM 5.2

Herhangi bir $k > 0$ ve $\beta^{\ast} = \beta^{R}$ için;

$$d = \frac{\sum_{i=1}^p [\lambda_i^2 + \lambda_i^3 c_i^2 - \alpha_i \lambda_i c_i (\lambda_i + k)^2] / (\lambda_i + k)^4}{\sum_{i=1}^p [\lambda_i + \lambda_i^2 c_i^2] / (\lambda_i + k)^4}$$

olarak seçilirse $H\check{K}O(\beta^{kdR}) \leq H\check{K}O(\beta^{R})$ ødir. E itlik $d=k$ durumunda geçerlidir.

SPAT

Ayn, ekilde (5.12)øden diyebiliriz ki $H\check{K}O(\beta^{kdR}) \leq H\check{K}O(\beta^{R})$ e itsizli inin

ispat edilebilmesi için $H\check{K}O(\alpha^{kdR}) \leq H\check{K}O(\alpha^{R})$ e itsizli inin ispat edilmesi yeterlidir.

lk olarak α^{kdR} tahmin edicisinin beklenen de eri,

$$\begin{aligned} E(\alpha^{kdR}) &= E[(\Lambda + kI)^{-1}(\Lambda - dI)(\Lambda + kI)^{-1} \Lambda \alpha_{PML}^{\ast}] \\ &= (\Lambda + kI)^{-1}(\Lambda - dI)(\Lambda + kI)^{-1} \Lambda E(\alpha_{PML}^{\ast}) \end{aligned} \quad (5.20)$$

olarak hesaplan,r. (5.14) den;

$$E(\alpha^{kdR}) = \left(\frac{(\lambda_1 - d)\lambda_1}{(\lambda_1 + k)^2} c_1, \dots, \frac{(\lambda_p - d)\lambda_p}{(\lambda_p + k)^2} c_p \right) \quad (5.21)$$

eklinde elde edilir.

α^{kdR} -nin kovaryans matrisi;

$$\begin{aligned}
 & (\Lambda + kI)^{-1}(\Lambda - dI)(\Lambda + kI)^{-1} \Lambda \hat{\alpha}_{PML} \\
 & = (\Lambda + kI)^{-1}(\Lambda - dI)(\Lambda + kI)^{-1}(\Lambda)C\ddot{ö}v(\hat{\alpha}_{PML}) (\Lambda) \\
 & \quad (\Lambda + kI)^{-1}(\Lambda - dI)(\Lambda + kI)^{-1}
 \end{aligned}$$

$$\begin{aligned}
 K\ddot{ö}v(\hat{\alpha}^{kdR}) & = (\Lambda + kI)^{-1}(\Lambda - dI)(\Lambda + kI)^{-1}(\Lambda)(\Lambda + kI) \\
 & \quad (\Lambda - dI)(\Lambda + kI)^{-1}
 \end{aligned} \tag{5.22}$$

eklinde hesaplan,r.

(5.21) ve (5.22) dan yararlan,larak $\hat{\alpha}^{kdR}$ tahmin edicisinin hata kareler ortalamas,

$$HK\ddot{O}(\hat{\alpha}^{kdR}) = \left\| E(\hat{\alpha}^{kdR}) - \alpha \right\|^2 + zK\ddot{ö}v(\hat{\alpha}^{kdR}) \tag{5.23}$$

olmak üzere;

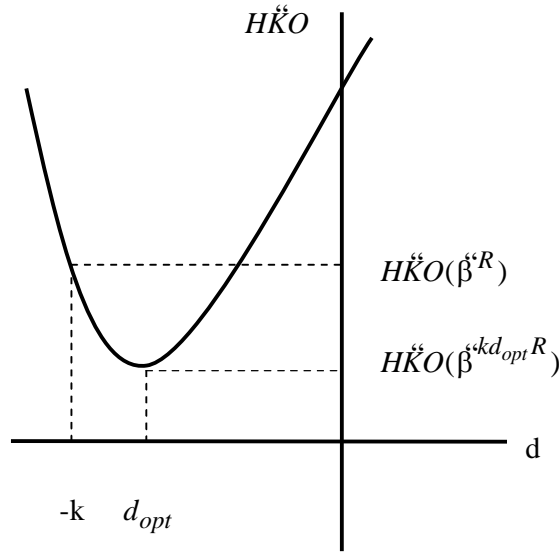
$$HK\ddot{O}(\hat{\alpha}^{kdR}) = \sum_{i=1}^p \frac{(\lambda_i - d)^2 \lambda_i}{(\lambda_i + k)^4} + \sum_{i=1}^p \left(\frac{(\lambda_i - d) \lambda_i c_i}{(\lambda_i + k)^2} - \alpha_i \right)^2 = h(d) \tag{5.24}$$

olarak bulunur.

(5.24)den görüldü ü üzere $HK\ddot{O}(\hat{\alpha}^{kdR})$, dğnin kuadratik bir fonksiyonudur ve bu fonksiyonu minumum yapan d de eri rahat,kla (5.25) da verildi i gibi belirlelenebilir;

$$d_{opt} = \frac{[c_i^2 - \alpha_i \lambda_i c_i (\lambda_i + k)^2] / (\lambda_i + k)^4}{\sum_{i=1}^p [\lambda_i + \lambda_i^2 c_i^2] / (\lambda_i + k)^4} \quad (5.25)$$

$h(-k) = H\ddot{K}O(\alpha^R)$ oldu u dü ünülürse, d parametresinin teorem 5.2 de verildi i gibi seçilmesi durumunda, $H\ddot{K}O(\alpha^{kdR}) \leq H\ddot{K}O(\alpha^R)$ ve buna denk olarak $H\ddot{K}O(\beta^{kdR}) \leq H\ddot{K}O(\beta^R)$ dir. spatin daha rahat anlaşılması için aşağıdaki grafiksel örnek verilmiştir.



ekil V.2 Teorem(5.2)ün ispatına yönelik grafiksel örnek

Lineer model için temel bile enler tahmin edicisi a a ,daki ekilde
tan,mlanm, t,r;

$$\beta_{TB}^{\prime} = T_r (T_r' X' X T_r)^{-1} T_r' X' X \beta_{EK}^{\prime} \quad (5.26)$$

Burada, T_r ; sütunlar, $X'X$ matrisinin özvektörlerini içeren ortogonal matris ve $T_r' X' X T_r = \Delta = kö\ egen(\lambda_1, \lambda_2, \dots, \lambda_p)$, $X'X$ ın s,ral, özde erlerini içeren kö egen matris olmak üzere, T_r ; T ın p-r sütununun silinmesi ile elde edilen matristir. Bu silinen p-r sütun, $X'X$ ın s,f,ra çok yak,n olan özde erlerine kar ,l,k gelen özvektörlerdir. Bu durumda $T_r' X' X T_r = \Delta_r = diag(\lambda_1, \dots, \lambda_r)$ olacakt,r.

Temel bile enler tahmin edicisinin kullan,m amac,, iç ili ki probleminden kaynaklanan, çok küçük özde erlerin sebep oldu u, i irilmi varyanslar, kabul edilebilir seviyelere indirgemektir. (5.26)dan anla ,laca , , üzere, r=p olmas, durumunda, temel bile enler tahmin edicisi en küçük kareler tahmin edicisine dönü ür.

(5.26)ya benzer ekilde, Cox oransal hazard regresyon modeli için temel bile enler tahmin edicisi;

$$\beta_{TB}^{\prime} = \theta_r (\theta_r' D' D \theta_r)^{-1} \theta_r' D' D \beta_{PML}^{\prime} \quad (5.27)$$

ada θ_r ; θ matrisinin p-r sütununun silinmesi ile elde edilen ortogonal matristir. Silinen p-r sütun, (5.8) de verilen Λ matrisinin çok küçük özdeğerlerine karşılık gelen özvektörlerdir.

Bu durumda,

$$\theta_r' \ddot{D}' \ddot{D} \theta_r = \Lambda_r = \text{kö eger}(\lambda_1, \lambda_2, \dots, \lambda_r) \quad (5.28)$$

olacaktır.

θ ortogonal bir matris oldu undan, (5.11)den $\alpha'_{PML} = \theta \beta'_{PML}$ dir ve kanonik form için temel bileşenler tahmin edicisi;

$$\alpha'^{TB} = B \alpha'_{PML} \quad (5.29)$$

olarak verilir. Burada $B = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ tipinde pxp boyutlu bir matristir [18]. (5.29)den

yararlanılarak β'^{TB} tahmin edicisi, (5.7)eye denk olarak a a ,daki gibi de hesaplanabilir;

$$\beta'^{TB} = \theta \alpha'^{TB} \quad (5.30)$$

A a ,da gösterildiği üzere $H\hat{K}O(\beta'^{TB}) = H\hat{K}O(\alpha'^{TB})$ dir.

$$\begin{aligned}
 & \left[\text{Köv}(\hat{\beta}^{TB}) \right] + \left\| E(\hat{\beta}^{TB}) - \beta \right\|^2 \\
 &= z \left[\text{Köv}(\theta \hat{\alpha}^{TB}) \right] + \left\| E(\theta \hat{\alpha}^{TB}) - \theta \alpha \right\|^2 \\
 &= z \left[\theta^T \text{Köv}(\hat{\alpha}^{TB}) \theta \right] + \left\| E(\hat{\alpha}^{TB}) - \alpha \right\|^2 \theta \theta^T \\
 &= z \left[\theta \theta^T \text{Köv}(\hat{\alpha}^{TB}) \right] + \left\| E(\hat{\alpha}^{TB}) - \alpha \right\|^2 \\
 &= z \left[\text{Köv}(\hat{\alpha}^{TB}) \right] + \left\| E(\hat{\alpha}^{TB}) - \alpha \right\|^2 = H\hat{K}O(\hat{\alpha}^{TB})
 \end{aligned} \tag{5.31}$$

Bu durumda, hesapsal kolaylık açısından, $H\hat{K}O(\hat{\beta}^{TM})$ yerine $H\hat{K}O(\hat{\alpha}^{TM})$ nin hesaplanması tercih edilir.

İlk olarak $\text{Köv}(\hat{\alpha}^{TM})$;

$$\begin{aligned}
 \text{Köv}(\hat{\alpha}^{TB}) &= B' \text{Köv}(\hat{\alpha}_{PML}) B \\
 &= B'(\Lambda)^{-1} B \\
 \text{Köv}(\hat{\alpha}^{TB}) &= (\Delta_r)^{-1}
 \end{aligned} \tag{5.32}$$

eklinde hesaplanır ve $\hat{\alpha}^{TB}$ tahmin edicisinin beklenen değeri, (5.14) göz önünde bulundurularak;

$$\begin{aligned}
 E(\hat{\alpha}^{TB}) &= E(B\hat{\alpha}_{PML}) \\
 &= B.E(\hat{\alpha}_{PML}) = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_r \\ 0 \\ \vdots \\ 0 \end{bmatrix}
 \end{aligned}$$

(5.33)

olarak bulunur.

olarak, $HKO(\hat{\alpha}^{TB})$ a a ,daki gibi hesaplan,r.

$$HKO(\hat{\alpha}^{TB}) = z \left[K\ddot{v}(\hat{\alpha}^{TB}) \right] + \left\| E(\hat{\alpha}^{TB}) - \alpha \right\|^2$$

$$HKO(\hat{\alpha}^{TB}) = \sum_{i=1}^r \frac{1}{\lambda_i} + \sum_{i=1}^r (c_i - \alpha_i)^2 + \sum_{i=r+1}^p \alpha_i^2 \quad (5.34)$$

(5.31)den yararlan,larak, $HKO(\hat{\beta}^{TB})$;

$$HKO(\hat{\beta}^{TB}) = \sum_{i=1}^r \frac{1}{\lambda_i} + \sum_{i=1}^r (c_i - \alpha_i)^2 + \sum_{i=r+1}^p \alpha_i^2 \quad (5.35)$$

olarak elde edilir.

V.4 COX ORANSAL HAZARD REGRESYON MODEL Ç N L U- T P TAHM N ED C LE TEMEL B LE ENLER TAHM N ED C S N N B R KOMB NASYONU

Lineer regresyon modeli için, 1984 de Baye ve Parker taraf,ndan, en küçük kareler, ridge ve temel bile enler tahmin edicisinin bir kombinasyonu olan

$$\hat{\beta}_r(k) = T_r (T_r' X X T_r + k I_r)^{-1} T_r' X X \hat{\beta}_{EK} \quad (5.36)$$

r-k s,n,f, tahmin edici geli tirilmi tir [19]. Nomura ve Ohkubo 1985teki bir çal, malar,nda bu yeni tahmin ediciyi hata kareler ortalamas, aç,s,ndan, enküçük kareler ve ridge tahmin edicisi ile kar ,la t,rm, lard,r [20].

ve Sakall,o lu en küçük kareler, temel bile enler ve Liu tahmin edicisinin bir kombinasyonu olan

$$\beta_r^{\epsilon}(d) = T_r (T_r' X X T_r + I_r)^{-1} (T_r' X' y + d T_r' \beta_{TB}^{\epsilon}) \quad (5.37)$$

tahmin edicisini tanımlayanlar [17]. Özkale ve Kaçaranlar 2007'de bu tahmin ediciyi matris hata kareler ortalaması, ölçütü açısından en küçük kareler, temel bile enler ve Liu tahmin edicisi ile karşılaştırmışlardır [21].

Bu kombinasyon ilk olarak lineer model için en küçük kareler, temel bile enler ve Liu-tipi tahmin edicinin bir kombinasyonu ya da bir anlamda Baye ve Parker'ın r-k s.n.f. tahmin edicisi ile Kaçaranlar ve Sakall,o lu'nun tanımladığı r-d s.n.f. tahmin edicinin bir kombinasyonu olan;

$$\beta_r^{\epsilon}(k, d) = T_r (T_r' X X T_r + k I_r)^{-1} (T_r' X X T_r - d I_r) T_r' \beta_{TB}^{\epsilon} \quad (5.38)$$

tahmin edicisi tanımlanmıştır. Bu yeni tanımlanan tahmin edici, r-(k,d) s.n.f. tahmin edici olarak adlandırılacak olursa; k=0 için, r-d s.n.f. tahmin ediciye, d=0 için r-k s.n.f. tahmin ediciye dönüşür. Bu anlamda (5.36) ve (5.37) ile verilen her iki tahmin ediciyi de kapsar.

Bu çalışmada bu yeni tahmin edici, Cox oransal hazard regresyon modeline uygulanacak ve hata kareler ortalaması, ölçütü bakımından performans, incelenecektir. Tahmin edicinin lineer modeldeki performansı, ilerleyen çalışmalarda incelenilmesi düşünülen açık bir konudur.

(5.38)'dekine benzer olarak Cox model için r-(k,d) s.n.f. tahmin edici;

$$(\Lambda_r + kI_r)^{-1}(\theta_r' \hat{D}' \hat{D} \theta_r - dI_r) \theta_r' \hat{\beta}^{*TB} \quad (5.39)$$

eklinde tanımlanabilir. İmdi bu tahmin edicinin $H\hat{K}O$ özünü hesaplayalım.

Kanonik form için r - (k,d) s,n,f, tahmin edici a a ,daki gibi verilebilir.

$$\hat{\alpha}^{*r}(k, d) = B\hat{\alpha}^{*kd} \quad (5.40)$$

Burada B daha öncede bahsedildiği gibi $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ tipinde $p \times p$ boyutlu bir

matristir. (5.40)dan yararlanılarak $\hat{\beta}^{*r}(k, d)$ tahmin edicisi, (5.39)den denk olarak a a ,daki gibi de hesaplanabilir;

$$\hat{\beta}^{*r}(k, d) = \theta \hat{\alpha}^{*r}(k, d) \quad (5.41)$$

(5.31)deki ispata benzer şekilde $H\hat{K}O(\hat{\beta}^{*r}(k, d)) = H\hat{K}O(\hat{\alpha}^{*r}(k, d))$ olduğu kolayca gösterilebilir. Daha öncede bahsedildiği üzere, hesaplamaları kolaylaştırmak bakımından, $H\hat{K}O(\hat{\beta}^{*r}(k, d))$ yerine $H\hat{K}O(\hat{\alpha}^{*r}(k, d))$ için hesaplanmas, tercih edilir.

İlk olarak $K\hat{O}v(\hat{\alpha}^{*r}(k, d))$, (5.16)den yararlanılarak;

$$\begin{aligned} K\hat{O}v(\hat{\alpha}^{*r}(k, d)) &= B'K\hat{O}v(\hat{\alpha}^{*kd})B \\ &= B'(\Lambda - dI)'(\Lambda + kI)^{-1}(\Lambda)^{-1}(\Lambda + kI)^{-1}(\Lambda - dI)B \end{aligned} \quad (5.42)$$

eklinde hesaplanabilir ve $\hat{\alpha}^{*r}(k, d)$ tahmin edicisinin beklenen değeri, (5.15) göz önünde bulundurularak,

$$= \left(\frac{\lambda_1 - d}{\lambda_1 + k} c_1, \dots, \frac{\lambda_r - d}{\lambda_r + k} c_r, 0, 0, \dots, 0 \right) \quad (5.43)$$

olarak bulunur.

(5.42) ve (5.43) kullanılarak, $H\ddot{K}O(\hat{\alpha}^r(k, d))$ aşağıdaki gibi hesaplanabilir.

$$H\ddot{K}O(\hat{\alpha}^r(k, d)) = z \left[K\ddot{O}v(\hat{\alpha}^r(k, d)) \right] + \left\| E(\hat{\alpha}^r(k, d)) - \alpha \right\|^2$$

$$H\ddot{K}O(\hat{\alpha}^r(k, d)) = \sum_{i=1}^r \frac{(\lambda_i - d)^2}{\lambda_i(\lambda_i + k)^2} + \sum_{i=1}^r \left(\frac{\lambda_i - d}{\lambda_i + k} c_i - \alpha_i \right)^2 + \sum_{i=r+1}^p \alpha_i^2 \quad (5.44)$$

$H\ddot{K}O(\hat{\beta}^{sr}(k, d)) = H\ddot{K}O(\hat{\alpha}^r(k, d))$ olduğundan;

$$H\ddot{K}O(\hat{\beta}^{sr}(k, d)) = \sum_{i=1}^r \frac{(\lambda_i - d)^2}{\lambda_i(\lambda_i + k)^2} + \sum_{i=1}^r \left(\frac{\lambda_i - d}{\lambda_i + k} c_i - \alpha_i \right)^2 + \sum_{i=r+1}^p \alpha_i^2 \quad (5.45)$$

Bu k,s,mda (V.3) ve (V.4)de tanımlan tahmin ediciler için geliştirilmiş iki teorem verilecektir.

TEOREM V.3

Herhangi bir $k > 0$ için,

$$d = \frac{\sum_{i=1}^r [\lambda_i c_i^2 - \alpha_i c_i (\lambda_i + k) + 1] / (\lambda_i + k)^2}{\sum_{i=1}^r (\lambda_i c_i^2 + 1) / [\lambda_i (\lambda_i + k)^2]}$$

olarak seçilirse $H\ddot{K}O[\hat{\beta}^{sr}(k, d)] \leq H\ddot{K}O(\hat{\beta}^{sTB})$ dir. Etilik $d = k$ durumunda geçerlidir.

için $\beta^{r}(k, d)$ nin HKO ösu a a ,da verildi i gibi dain

kuadratik bir fonksiyonudur.

$$HKO(\beta^{r}(k, d)) = \sum_{i=1}^r \frac{(\lambda_i - d)^2}{\lambda_i(\lambda_i + k)^2} + \sum_{i=1}^r \left(\frac{\lambda_i - d}{\lambda_i + k} c_i - \alpha_i \right)^2 + \sum_{i=r+1}^p \alpha_i^2 = f(d)$$

Bu fonksiyonu minimum yapacak olan d de eri a a ,daki gibi hesaplanabilir.

$$d_{opt} = \frac{\sum_{i=1}^r [(\lambda_i c_i^2 - \alpha_i c_i (\lambda_i + k) + 1) / (\lambda_i + k)^2]}{\sum_{i=1}^r (\lambda_i c_i^2 + 1) / [\lambda_i (\lambda_i + k)^2]}$$

$f(-k) = HKO(\beta^{TB})$ oldu u dü ünülürse, d_{opt} de erinin -k dan farklı olarak belirlendi i durumlarda, $f(d_{opt}) < f(-k)$ olacaktır. Dolay,s, ile d_{opt} de erinin teoremd e verildi i gibi belirlenmesi durumunda $HKO[\beta^{r}(k, d)] < HKO(\beta^{TB})$ olacaktır. E itlik $d_{opt} = -k$ olmas, durumunda geçerlidir.

TEOREM V.4

Herhangi bir $k > 0$ için $\lambda_1 \geq \dots \geq \lambda_p > 0$ olmak üzere;

$$d < \min[\lambda_p, \min(\lambda_i - \frac{2(\lambda_i + k)\lambda_i \alpha_i}{1 + c_i^2 \lambda_i})] \quad i = 1, \dots, p$$

veya

$$d > \max[\lambda_1, \max(\lambda_i - \frac{2(\lambda_i + k)\lambda_i \alpha_i}{1 + c_i^2 \lambda_i})] \quad i = 1, \dots, p$$

seçilmesi durumunda, $HKO(\beta^{r}(k, d)) < HKO(\beta^{kd})$ dır.

lk olarak, $H\check{K}O(\beta^r(k, d)) - H\check{K}O(\beta^{r-1}(k, d)) > 0$ e itsizli i d için çözülsün.

(5.45) den

$$\begin{aligned} H\check{K}O(\beta^r(k, d)) - H\check{K}O(\beta^{r-1}(k, d)) &= \frac{(\lambda_r - d)^2}{\lambda_r(\lambda_r + k)^2} + \left(\frac{\lambda_r - d}{\lambda_r + k}c_r - \alpha_r\right)^2 - \alpha_r^2 \\ &= (\lambda_r - d)^2(1 + \lambda_r c_r^2) - 2(\lambda_r - d)(\lambda_r + k)\lambda_r c_r \alpha_r > 0 \end{aligned}$$

$d < \lambda_r$ için

$$\begin{aligned} (\lambda_r - d)(1 + \lambda_r c_r^2) &> 2(\lambda_r + k)\lambda_r c_r \alpha_r \\ d &< \lambda_r - \frac{2(\lambda_r + k)\lambda_r c_r \alpha_r}{(1 + \lambda_r c_r^2)} \\ \Rightarrow d &< \min\left[\lambda_r, \lambda_r - \frac{2(\lambda_r + k)\lambda_r c_r \alpha_r}{(1 + \lambda_r c_r^2)}\right] \text{ için e itsizlik} \end{aligned}$$

sa lan,r.

$d > \lambda_r$ için

$$\begin{aligned} (\lambda_r - d)(1 + \lambda_r c_r^2) &< 2(\lambda_r + k)\lambda_r c_r \alpha_r \\ d &> \lambda_r - \frac{2(\lambda_r + k)\lambda_r c_r \alpha_r}{(1 + \lambda_r c_r^2)} \\ \Rightarrow d &> \max\left[\lambda_r, \lambda_r - \frac{2(\lambda_r + k)\lambda_r c_r \alpha_r}{(1 + \lambda_r c_r^2)}\right] \text{ için e itsizlik} \end{aligned}$$

sa lan,r.

$i, r = 1, \dots, p$ ve $\lambda_1 \geq \dots \geq \lambda_p > 0$ olmak üzere genellenecek olunursa, d'nin;

$$d < \min\left[\lambda_p, \min\left(\lambda_i - \frac{2(\lambda_i + k)\lambda_i c_i \alpha_i}{(1 + \lambda_i c_i^2)}\right)\right] \text{ veya } d > \max\left[\lambda_1, \max\left(\lambda_i - \frac{2(\lambda_i + k)\lambda_i c_i \alpha_i}{(1 + \lambda_i c_i^2)}\right)\right]$$

eklinde seçilmesi durumunda, r nin her de eri için,

$$H\check{K}O(\beta^r(k, d)) - H\check{K}O(\beta^{r-1}(k, d)) > 0 \text{ olacakt,r.}$$

Bu durumda,

$$H\check{K}O[\beta^p(k, d)] > H\check{K}O[\beta^{p-1}(k, d)] > \dots > H\check{K}O[\beta^r(k, d)]$$

$$\Rightarrow H\check{K}O[\beta^{kd}] > H\check{K}O[\beta^r(k, d)]$$

bulunur.

BENZETİM ÇALIŞMASI

Bu bölümde, Cox model için önerilmiş olan yeni tahmin yöntemlerinin performansları, karşılaştırmak amacıyla, Monte Carlo benzetimleri düzenlenmiştir. Bu yeni yöntemlerin performansları, farklı iç ili ki düzeylerinde ve farklı örneklem büyüklüklerinde değerlendirilmiştir. Çalışma için gerekli olan kod Wolfram Mathematica 7 programında yazılmıştır.

Benzetim çalışmasında, açıklanmış dört ve gerçek parametre değerleri, (0.4, 0.2, -0.5, -0.2) olarak seçilmiştir.

Örneklem büyüklüğünün küçük olduğu durumlarda iç ili ki parametre tahminleri üzerinde daha kötü etkilere sahip olduğundan, $n=20$ ve $n=30$ olacak şekilde küçük örneklem seçilmiştir.

Yapam sürdürme zamanları üretildiğinde, ölçek parametresi $p=0.2$ olan dört parametrelili Weibull model kullanılmıştır. Durdurma değeri, Bernoulli dağılımından, durdurma olasılığı, $q=0.3$ olacak şekilde üretilmiştir.

Açıklanmış ölçeklerin üretilmesinde Box-Müller kullanılmıştır. W_i değerleri, standart normal rasgele değişkenler olmak üzere, X_j açıklanmış ölçekleri;

$$\begin{aligned} \varphi_{w_i,5} & \quad i=1,\dots,n \quad j=1, 2 \\ \rho_{w_i,5} & \quad i=1,\dots,n \quad j=2, 3 \end{aligned} \quad (6.1)$$

eklinde üretilmi tir. φ ve ρ , sırasıyla iki erli korelasyonlar, φ^2 ve ρ^2 olacak ekilde ayarlanm, t,r [22].

Çal, mada;

$$\varphi = 0,20 \text{ ve } \rho = 0,80$$

$$\varphi = 0,20 \text{ ve } \rho = 0,90$$

$$\varphi = 0,20 \text{ ve } \rho = 0,95$$

gibi üç farklı korelasyon durumu, $n=20$ ve $n=30$ örneklem büyüklükleri için incelenmi tir.

V. bölümde, önerilen yeni tahmin yöntemleri için farklı d_{opt} de erleri belirlenmi tir. Benzetim çal, mas,nda, yöntemlerin tamam,n,n HKO ölçütü aç,s,ndan birbirleri ile k,yaslanabilmesi için, bu d_{opt} de erlerinden yalnızca, teoremV.1 ile verilen

$$d_{opt} = \frac{\sum_{i=1}^p [(\lambda_i c_i^2 - \alpha_i c_i (\lambda_i + k) + 1) / (\lambda_i + k)^2]}{\sum_{i=1}^p (\lambda_i c_i^2 + 1) / [\lambda_i (\lambda_i + k)^2]}$$

de eri ile ilgilenilmi tir.

metrelerin gerçek de eri ve beklenen de eri belirli

olmad, ,ndan, d_{opt} de erinin hesaplanmas,nda, α ve $E(\hat{\alpha}_{PML}) = (c_1, \dots, c_p)$

yerine $\hat{\alpha}_{PML}$ kullan,lm,as, önerilebilir. Benzetim çal, mas,nda bu e kilde elde edilen;

$$d_{opt} = \frac{\sum_{i=1}^p [(1 - k(\hat{\alpha}_{PML(i)}))^2 / (\lambda_i + k)^2]}{\sum_{i=1}^p (\lambda_i(\hat{\alpha}_{PML(i)})^2 + 1) / [\lambda_i(\lambda_i + k)^2]} \quad (6.2)$$

de eri kullan,lm, t,r ve böylece önerilen de er test edilmi tir.

Hesapsal zorluklar nedeniyle benzetimler 100 defa tekrarlanm, t,r. Her deneyde a a ,daki tahminler kaydedilmi tir.

- $\hat{\beta}_{PML}$: K,smi en çok olabilirlik kestiricisi
- $\hat{\beta}^{R}$: Ridge kestiricisi, $k = \frac{1}{\hat{\beta}'_{PML}\hat{\beta}_{PML}}$
- $\hat{\beta}^{kd}$: Liu-tipi kestirici, $k = \frac{1}{\hat{\beta}'_{PML}\hat{\beta}_{PML}}$ ve $\hat{\beta} = \hat{\beta}_{PML}$
- $\hat{\beta}^{kdr}$: Liu-tipi kestirici, $k = \frac{1}{\hat{\beta}'_{PML}\hat{\beta}_{PML}}$ ve $\hat{\beta} = \hat{\beta}^R$
- $\hat{\beta}^{TB}$: Temel bile enler kestiricisi
- $\hat{\beta}^r(k, d)$: Temel bile enler kestiricisi ile Liu-tipi kestiricinin kombinasyonu

Her durum için 100 benzetim tamamland,ktan sonra yukar,daki tahmin ediciler için hata kareler ortalamas, ve yan de erleri hesaplanm, t,r.

Elde edilen sonuçlar Tablo VI.1'de görölmektedir.

n çal, mas, sonuçlar,

		n=20		n=30	
		HKO	Yan	HKO	Yan
$\varphi = 0.20$ $\rho = 0.80$	$\hat{\beta}_{PML}^c$	6.55	0.91	$\hat{\beta}_{PML}^c$	2.00 0.58
	$\hat{\beta}^{R}$	4.56	0.76	$\hat{\beta}^{R}$	1.48 0.48
	$\hat{\beta}^{kd}$	2.26	0.53	$\hat{\beta}^{kd}$	0.78 0.36
	$\hat{\beta}^{kdR}$	1.63	0.44	$\hat{\beta}^{kdR}$	0.61 0.31
	$\hat{\beta}^{TB}$	0.49	0.22	$\hat{\beta}^{TB}$	0.43 0.28
	$\hat{\beta}^r(k,d)$	0.38	0.24	$\hat{\beta}_r^c(k,d)$	0.31 0.22
$\varphi = 0.20$ $\rho = 0.90$	$\hat{\beta}_{PML}^c$	8.71	0.82	$\hat{\beta}_{PML}^c$	2.79 0.45
	$\hat{\beta}^{R}$	5.16	0.64	$\hat{\beta}^{R}$	1.70 0.36
	$\hat{\beta}^{kd}$	3.42	0.58	$\hat{\beta}^{kd}$	1.03 0.36
	$\hat{\beta}^{kdR}$	2.33	0.50	$\hat{\beta}^{kdR}$	0.76 0.31
	$\hat{\beta}^{TB}$	0.55	0.20	$\hat{\beta}^{TB}$	0.37 0.24
	$\hat{\beta}^r(k,d)$	0.46	0.20	$\hat{\beta}_r^c(k,d)$	0.30 0.22
$\varphi = 0.20$ $\rho = 0.95$	$\hat{\beta}_{PML}^c$	31.55	0.86	$\hat{\beta}_{PML}^c$	12.08 0.47
	$\hat{\beta}^{R}$	16.40	0.74	$\hat{\beta}^{R}$	6.59 0.43
	$\hat{\beta}^{kd}$	10.75	0.84	$\hat{\beta}^{kd}$	4.47 0.58
	$\hat{\beta}^{kdR}$	7.03	0.70	$\hat{\beta}^{kdR}$	3.03 0.50
	$\hat{\beta}^{TB}$	0.58	0.22	$\hat{\beta}^{TB}$	0.44 0.28
	$\hat{\beta}_r^c(k,d)$	0.54	0.22	$\hat{\beta}_r^c(k,d)$	0.40 0.26

Tablo VI.10den, önerilen tüm tahmin edicilerin tahmin edilen hata kareler

ortalamas, de erlerinin, k,smi ençok olabilirlik ve ridge tahmin edicilerine göre

gözlemlenmektedir. Genel olarak tahmin edicilerin

yanlıklarında da bir azalma söz konusudur.

$n=20$ ve $\varphi = 0.20$, $\rho=0.95$ durumunda, $\hat{\beta}^{kd}$ önin yanlı, ridge tahmin edicisinin yanlına göre biraz daha büyüktür. Fakat hata kareler ortalamasındaki dü ü ün yannda, yanlıktaki bu art, önemsiz denecek kadar küçüktür. Aynı ekilde $n=30$ ve $\varphi = 0.20$, $\rho=0.95$ durumunda da, $\hat{\beta}^{kd}$ ve $\hat{\beta}^{kdR}$ önin yanlıklarında hafif bir art, olmasına rağmen, hata kareler ortalamalarında ciddi bir dü ü gözlenmektedir.

Çalışmada uygun optimum d değerleri kullanılmamasına rağmen, $\hat{\beta}^{TB}$ ve $\hat{\beta}_r^{(k,d)}$ tahmin edicilerinin her durumda en küçük hata kareler ortalaması ve yanlılık değerlerine sahip oldukları görülmektedir. Ayrıca uygulamada kullanılması önerilen optimum d değerinin uygunluğu, benzetim çalışması ile test edilmiş ve uygun olmadığı bir durum gözlemlenmemiştir.

SONUÇLAR VE TARTI MA

Cox oransal hazard regresyon modelinin uygulandığı, çalışmalarda parametre tahminleri için genel olarak kısımlı çok olasılırlık tahmin yöntemi kullanılmaktadır. Fakat bu tip çalışmalar, açıklayıcı, değişkenler arasında iç içeliği olması durumu sıklıkla rastlanabilen olağan bir durumdur. Bu durum göz ardı edilerek model parametrelerinin tahmini için kısımlı çok olasılırlık tahmin yönteminin kullanılması, irilimsi varyanslara sahip, sıklıkla tahmin değerleri elde etmemizi sağlayacaktır.

2007 yılında yapılan bir çalışmada, iç içeliği durumunda kullanılmak üzere Cox oransal hazard regresyon modeline ridge regresyon tahmin yöntemi uyarlanmıştır ve bu tahmin yönteminin performansını bir benzetim çalışması ile incelenmiştir. Yapılan benzetim çalışması sonucu, ridge regresyon tahmin yönteminin, kısımlı çok olasılırlık tahmin yöntemine göre genellikle daha küçük hata kareleri ortalamasına sahip olduğu gözlemlenmiştir. Fakat bu çalışma herhangi bir kuramsal ispat yapılmamıştır.

Liu-tipi tahmin yöntemi, 2003 yılında Kejan Liu tarafından lineer modeller için geliştirilmiş olan bir tahmin yöntemidir. Bu çalışmada ilk olarak Liu-tipi tahmin yönteminin iki farklı tipi, yarı parametrik bir model olan Cox oransal hazard regresyon modeline uyarlanmıştır ve bu yeni tahmin yönteminin her iki tipinin de,



Your complimentary
use period has ended.
Thank you for using
PDF Complete.

[Click Here to upgrade to
Unlimited Pages and Expanded Features](#)

re daha küçük hata kareler ortalamas,na sahip oldu u

kuramsal olarak ispatlanm, t,r.

Benzer olarak temel bile enler tahmin yöntemi, lineer modellerde, iç ili ki
durumunda kullan,lan tahmin yöntemlerindendir. Çal, mada ikinci olarak temel
bile enler tahmin yöntemi, Cox oransal hazard regresyon modeline uyarlanm, ve
daha sonra bu yöntem Liu-tipi tahmin yöntemi ile birle tirilerek, Cox model için yeni
bir tahmin yöntemi geli tirilmi tir. Yap,lan kuramsal ispatlarla yöntemin temel
bile enler tahmin yöntemine ve baz, ko ullar alt,nda Liu-tipi tahmin yöntemine,
dolay,s, ile ridge tahmin yöntemine göre daha küçük hata kareler ortalamas,na sahip
oldu u gösterilmi tir.

Son olarak yap,lan benzetim çal, mas, ile geli tirilmi olan yeni tahmin
yöntemlerinin performanslar, daha ayr,nt,l, ekilde incelenmi ve uygulamada da
oldukça iyi sonuçlar elde edildi i görülmü tür.

SON DE ERLEND RMELER VE ÖNER LER

Birçok kullan,m alan,n,n yan, s,ra özellikle t,pta, salg,n hastal,klar,n ve kronik hastal,klara ili kin verilerin incelenmesinde ve bu hastal,klar, etkileyen de i kenlerin saptanmas,nda Cox oransal hazard regresyon modeli en çok kullan,lan modellerdendir.

Modelde aç,klay,c, de i kenlere ait parametrelerin de erleri, de i kenlerin etkinli ini belirlemektedir. Bu nedenle parametre tahminlerinin minimum hata ile yap,lmas,, yap,lan incelemeler için oldukça önemlidir.

Cox oransal hazard regresyon modeline son olarak adapte edilen ridge tahmin yöntemine kar ,n, bu çal, mada modele, daha küçük hata pay,na sahip tahmin yöntemleri uyarlanm, ve böylece özellikle iç ili ki durumunda daha sa l,kl, tahmin de erleri elde edilebilmesine olanak sa lanm, t,r.

Ya am sürdürme analizi için bugüne kadar yap,lm, makaleler yard,m,yla çe itli algoritmalar geli tirilmi ve bunlar paket programlara aktar,lm, t,r. Bu paket programlar,n baz,lar,nda Cox oransal hazard regresyon model formu da mevcuttur. Fakat Cox oransal hazard regresyon modelini içeren paket programlar,n hepsi, model parametrelerinin tahmini için klasik yöntem olan k,smi ençok olabilirlik yöntemini kullanmaktad,r. Bu paket programlar,n hiç birinde Cox oransal hazard regresyon



PDF
Complete

*Your complimentary
use period has ended.
Thank you for using
PDF Complete.*

[Click Here to upgrade to
Unlimited Pages and Expanded Features](#)

nda kullan,labilecek olan farkl, tahmin yöntemlerine yer

verilmemektedir. Bu yeni tahmin yöntemleri ile ilgili olarak geli tirilen

algoritmalar,n, mevcut paket programlara aktar,lmas, ile ara t,rmac,lara uygulamada

kolayl,k sa lanabilece i dü ünülmektedir.

- [1] Cox, D.R.: "Regression Models and Life Tables (with Discussion)", *Journal of the Royal Statistical Society*, B34 ,(1972), 187-220
- [2] Lagakos S. W.: "Effects of Mismodelling and Mismeasuring Explanatory Variables on Tests of Their Association with a Response Variable", *Stat Med*, 7, (1988), 257-274
- [3] Kleinbaum, D.G.; Klein. M.,: "Survival Analysis", *Springer*, Second Edition, Gail, M., Krickeberg, K., Samet, J., Tsiatis, A., Wong, W.Editors.; New York, USA, (2005), 1-167
- [4] Lawless, J.F.: "Statistical Models and Methods for Lifetime Data", *John Willey&Sons*, New York, USA, (1982), 13.
- [5] Bülbul, E.S.: "statistiksel Ba ar,s,zl,k Zaman, Modelleri ve Finansal Analizlerde Uygulaması", *Doktora Tezi*, Marmara Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul, Türkiye, (1999), 34-35.
- [6] Kaplan, E.L.; Meier, P.: "Non Parametric Estimation From Incomplete Observations", *Journal of the American Statistical Association*, 53, (1958), 457-481
- [7] Breslow, N.; Crowley, J.: "A Large Sample Study of Life Table and Product Limit Estimates Under Random Censorship", *The Annals of Statistics*, 2, 3, (1974), 437-453
- [8] Meier, P.: "Perspectives in Probability and Statistics", *Academic Press*, Gani, J. Editors.; New York, USA, (1975), 67-87
- [9] Hosmer, D.W.; Lemeshow, S.: "Applied Survival Analysis, Regression Modelling of the Time To Event Data", *John Willey&Sons*, New York, USA, (1999), 93-105, 354-357.
- [10] Collet ,D.: "Modeling Survival Data in Medical Research", *Chapman&Hall*,

- [11] Lee, E.T., Wang, J.W.: "Statistical Methods for Survival Data Analysis", *John Wiley & Sons*, Balding, D.J.; Bloomfield, P.; Cressie, N.A.C.; Fisher, N.I.; Johnstone, I.M.; Kadane, J.B.; Ryan, L.M.; Scott, D.W.; Smith, A.F.M.; Teugels, J.L.: Canada, (2003), 298-376.
- [12] Harrel, F. and Lee, K.: "Verifying assumptions of the proportional hazards model", *Proceedings of the Eleventh Annual SAS Users Group International*, (1986), 823-828.
- [13] Abrahamowicz, M.; Mackenzie, T. And Esdaile J.M.: "Time Dependent Hazard Ratio: Modeling and Hypothesis with Application in Lupus Nephritis", *Journal of the American Statistical Association*, 436, (1996), 1432-1439.
- [14] Lustbader, E.D.: "Relative Risk Regression Diagnosis", *Modern Statistical Methods in Chronic Disease Epidemiology*, Moolgavkar, S.H., Prentice, R.L., SIAM, Philadelphia, (1986).
- [15] Xue, X.; Kim, M.Y.; Shore, R.E.: "Cox Regression Analysis In Presence of Collinearity: An Application to Assessment of Health Risk Associated With Occupational Radiation Exposure", *Life Time Data Analysis*, 13, 3, (2007), 333-350.
- [16] Liu, K.: "Using Liu Type Estimator to Combat Collinearity", *Communications in Statistics*, 32, 5, (2003), 1009-1020.
- [17] Kaçaranlar, S.; Sakallıoğlu, S.: "Combining the Liu Estimator and the Principal Component Regression Estimator", *Communications in Statistics-Theory and Methods*, 30, (2001), 2699-2705 .
- [18] Montgomery, D.C.; Peck, A.E.; Vining, G.G.: "Introduction to the Linear Regression Analysis", *John Wiley & Sons*, Bloomfield, P.; Cressie, N.A.C.; Fisher, N.I.; Johnstone, I.M.; Kadane, J.B.; Ryan, L.M.; Scott, D.W.; Silverman, B.W.;



PDF Complete

Your complimentary use period has ended.
Thank you for using PDF Complete.

[Click Here to upgrade to Unlimited Pages and Expanded Features](#)

; Barnett, V., Bradley, R.A.; Hunter, J.S.; Kendall, D.G.:
Canada, (2001), 525-582.

[19] Baye, M.R.; Darell, F.P.: "Combining Ridge and Principal Component Regression: A Money Demand Illustration." *Communications in Statistics-Theory and Methods*, 13, 2, (1984), 197-205.

[20] Nomura, M.; Ohkubo, T.: "A Note on Combining Ridge and Principal Component Regression." *Communications in Statistics-Theory and Methods*, 14, 10, (1985), 2489-2493.

[21] Özkale, M.R.; Kaçaranlar, S. "Superiority of the r-d Class Estimator Over Some Estimators by the Mean Square Error Matrix Criterion" *Statistics & Probability Letters*, 77, (2007), 438-446.

[22] Gibbons, D.G., "A Simulation Study of Some Ridge Estimators" *Journal of the American Statistical Association*, 76, 373, (1981), 131-139.

SPAT I:

$$\begin{aligned}
 H\ddot{K}O(\hat{\alpha}_{PML}) &= zK\ddot{\sigma}v(\hat{\alpha}_{PML}) + \|E(\hat{\alpha}_{PML}) - \alpha\|^2 \\
 &= z(\ddot{Z}'\ddot{Z})^{-1} + \|E(\theta \hat{\beta}_{PML}) - \theta \beta\|^2 \\
 &= z(\ddot{Z}'\ddot{Z})^{-1} + \|\theta'[E(\hat{\beta}_{PML}) - \beta]\|^2 \\
 &= z(\ddot{D}\theta\theta'\ddot{D}')^{-1} + \|E(\hat{\beta}_{PML}) - \beta\|^2 \theta\theta' \\
 &= z(\ddot{D}\ddot{D}')^{-1} + \|E(\hat{\beta}_{PML}) - \beta\|^2 = H\ddot{K}O(\hat{\beta}_{PML}) \\
 &\Rightarrow H\ddot{K}O(\hat{\alpha}_{PML}) = H\ddot{K}O(\hat{\beta}_{PML})
 \end{aligned}$$

SPAT II:

$$\begin{aligned}
 H\ddot{K}O(\hat{\alpha}^{kd}) &= H\ddot{K}O(\theta \hat{\beta}^{kd}) = zK\ddot{\sigma}v(\theta \hat{\beta}^{kd}) + \|E(\theta \hat{\beta}^{kd}) - \theta \beta\|^2 \\
 &= z[\theta'K\ddot{\sigma}v(\hat{\beta}^{kd})\theta] + \|\theta'[E(\hat{\beta}^{kd}) - \beta]\|^2 \\
 &= zK\ddot{\sigma}v(\hat{\beta}^{kd}) + \|E(\hat{\beta}^{kd}) - \beta\|^2 \\
 &\Rightarrow H\ddot{K}O(\hat{\alpha}^{kd}) = H\ddot{K}\hat{O}(\hat{\beta}^{kd})
 \end{aligned}$$



Your complimentary
use period has ended.
Thank you for using
PDF Complete.

[Click Here to upgrade to
Unlimited Pages and Expanded Features](#)

Deniz nan, 1980 y,l,nda Almanya-n,n Scheim ehrinde do du. lk ve orta ö renimi Yenibosna lkö retim Okulunda ve lise ö renimini Çemberlita K,z Lisesinde tamamlad,. 1998 y,l,nda Marmara Üniversitesi Fen-Edebiyat Fakültesi Matematik Bölümüne girdi. 2002 y,l,nda mezun oldu. Ayn, y,l Marmara Üniversitesi Fen Bilimleri Enstitüsünde Matematik Ana Bilim Dal, Uygulamal, Matematik Program,nda yüksek lisans yapmaya hak kazand,. 2005 y,l,nda mezun olduktan sonra yine ayn, enstitünün ayn, program,nda doktora program,na kabul edildi. 2003 y,l,nda Fen-Edebiyat Fakültesinde ara t,rma görevlisi olarak görev yapmaya ba lad,. Halen bu görevi devam etmektedir.

T.C.
MARMARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

KABUL ve ONAY BELGESİ

Deniz İnan.'ın *İç İlişki Durumunda Cox Oransal Hazard Modeli için Farklı Tahmin Edicilerin İncelenmesi* başlıklı Lisansüstü tez çalışması, M.Ü. Fen Bilimleri Enstitüsü Yönetim Kurulu'nun 29.12.2010 tarih ve 2010/24-05 sayılı kararı ile oluşturulan jüri tarafından Matematik Anabilim Dalı Uygulamalı Matematik Programında DOKTORA Tezi olarak Kabul edilmiştir.

Danışman : Prof. Dr. Müjgan TEZ (MÜ)
1. Üye : Prof. Dr. Aydın ERAR (MSGÜ)
2. Üye : Yar. Doç. Dr. Bahar Sennaroğlu (MÜ)
3. Üye : Prof. Dr. Gülay Kiroğlu (MSGÜ)
4. Üye : Doç. Dr. Nural Bekiroğlu (MÜ)

Tezin Savunulduğu Tarih : 13. 12. 2010

ONAY

M.Ü. Fen Bilimleri Enstitüsü Yönetim Kurulu'nun ~~27.12.2010~~ tarih ve ~~2010/26-02~~ sayılı kararı ile Deniz İnan'ın Matematik Anabilim Dalı Uygulamalı Matematik Programında Doktora (Dr, PhD.) derecesi alması onanmıştır.

Fen Bilimleri Enstitüsü Müdürü

Prof.Dr.Meral Ünal

