

YAŞAM ÇÖZÜMLEMESİNDE AYKIRI DEĞERLER

OUTLIERS IN SURVIVAL ANALYSIS

NURAY TUNCER

PROF. DR. DURDU KARASOY

Tez Danışmanı

Hacettepe Üniversitesi

Lisansüstü Eğitim-Öğretim Yönetmeliğinin

İstatistik Anabilim Dalı İçin Öngördüğü

YÜKSEK LİSANS TEZİ olarak hazırlanmıştır.

2014

NURAY TUNCER' in hazırladığı “**Yaşam Çözümlemesinde Aykırı Değerler**” adlı çalışma aşağıdaki jüri üyeleri tarafından **İSTATİSTİK ANABİLİM DALI'** nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Doç. Dr. Mehtap AKÇİL OK

Başkan

.....

Prof. Dr. Durdu KARASOY

Danışman

.....

Yard. Doç. Dr. Nihal ATA TUTKUN

Üye

.....

Bu tez Hacettepe Üniversitesi Fen Bilimleri Enstitüsü tarafından **YÜKSEK LİSANS TEZİ** olarak onaylanmıştır.

PROF.DR. FATMA SEVİN DÜZ

Fen Bilimleri Enstitüsü Müdürü

ETİK

Hacettepe Üniversitesi Fen Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmada,

- tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- atıfta bulunduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- ve bu tezin herhangi bir bölümünü bu üniversitede veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.

04/06/2014

Nuray TUNCER

ÖZET

YAŞAM ÇÖZÜMLEMESİNDE AYKIRI DEĞERLER

Nuray TUNCER

Yüksek Lisans, İstatistik Bölümü

Tez Danışmanı: Prof. Dr. Durdu KARASOY

Mayıs 2014, 84 sayfa

Yaşam çözümülemesi, tanımlanan herhangi bir olayın ortaya çıkmasına kadar geçen sürenin incelenmesinde kullanılan istatistiksel yöntemler bütünüdür. Aykırı değerler, modele iyi uyum sağlamayan değerlerdir. Yaşam çözümülemesinde aykırı değerler klasik regresyonda kullanılan yöntemlerden farklı yöntemler kullanılarak hesaplanmaktadır. Çünkü yaşam çözümülemesinde veriler durdurulmuş gözlemler içerir ve durdurulmuş gözlemlerin varlığı "beklenen eksi gözlenen" değerini hesaplamayı zorlaştırır.

Yaşam çözümülemesinde aykırı değer belirleme yöntemleri artıklara ve artıkların analizine dayanmaktadır. Yaşam çözümülemesinde kullanılan başlıca artık türleri Cox-Snell, Martingale, Schoenfeld, Sapma, Log-odds ve Normal sapma artıklarıdır. Etkili gözlemleri belirlemek için kullanılan yöntemler ise DFBETA, LMAX ve Olabilirlik Değişim değerleridir.

Çalışma boyunca incelenen artık türleri mide kanseri ile ilgili verilere uygulanmış ve aykırı değerler belirlenmiştir. Belirlenen aykırı değerler çıkarılarak model yeniden kurulmuştur. Aykırı değerler çıkarıldığında sonuçların daha iyi olduğu görülmüştür.

Anahtar Kelimeler: Yaşam çözümülemesi, yaşam modelleri, aykırı değerler, etkili gözlemler, artıklar

ABSTRACT

OUTLIERS IN SURVIVAL ANALYSIS

Nuray TUNCER

Master of Science, Department of Statistics

Supervisor: Prof. Dr. Durdu KARASOY

May 2014, 84 pages

Survival analysis is a collection of statistical methods for analyzing data where the outcome variable is the time until the occurrence of an event of interest. Outliers are the individuals whose survival time is not well fitted by the model. Outliers in survival analysis are calculated differently from classical regression analysis. Because in survival analysis data contain censored observations and this makes it difficult to calculate the difference "observed minus predicted" values.

In survival analysis outlier detection methods are commonly carried out based on residuals and residual analysis. In survival analysis there are different types of residuals that are Cox-Snell, Martingale, Schoenfeld, Deviance, Log-odds and Normal deviance residuals. There are methods which are DFBETA, LMAX and Likelihood Displacement values for detecting influential observations.

The residuals are analyzed during the study which is applied on a stomach cancer data set and the outliers are detected. After omitting these outliers, the model is set up again and results were found better.

Key Words: Survival analysis, survival models, outliers, influential observations, residuals

TEŐEKKÜR

Bu tezin oluŐmasında en büyük desteęi saęlayan, tez alıŐmamın her aŐamasında bana yol gÖsteren ve yakın ilgisiyle beni teŐvik eden danıŐmanım Sayın Prof. Dr. Durdu KARASOY'a, sabırları ve destekleri iin alıŐma arkadaŐlarıma ve aileme itenlikle teŐekkür ederim.

İÇİNDEKİLER

Sayfa

ÖZET	i
ABSTRACT	ii
TEŞEKKÜR	iii
İÇİNDEKİLER.....	iv
ÇİZELGELER.....	vi
ŞEKİLLER	vii
SİMGELER VE KISALTMALAR	ix
1. GİRİŞ.....	1
2. GENEL BİLGİLER	3
2.1. Yaşam Çözümlemesi	3
2.1.1. Durdurma	5
2.1.2. Yaşam Çözümlemesinde Kullanılan Fonksiyonlar.....	6
2.2. Cox Orantılı Tehlikeler Modeli	8
2.3. Orantılı Tehlikeler Varsayımının İncelenmesi.....	9
2.4. Parametrik Dağılımlar.....	12
2.4.1. Üstel Dağılım	13
2.4.2. Weibull Dağılımı	14
2.4.3. Log-Normal Dağılım	15
2.4.4. Log-Lojistik Dağılım	16
2.4.5. Gamma ve Genelleştirilmiş Gamma Dağılımı.....	17
2.5. Aykırı Değerler	18
3. YAŞAM ÇÖZÜMLEMESİNDE AYKIRI DEĞERLER.....	20
3.1. Cox-Snell Artıkları	22
3.2. Martingale Artıkları	23
3.3. Sapma Artıkları.....	25
3.4. Schoenfeld Artıkları	26
3.5. Log-Odds Artıkları	29
3.6. Normal Sapma Artıkları	30
3.7. Etkili Gözlemlerin Belirlenmesi	32
4. UYGULAMA	35
5. SONUÇLAR	64

KAYNAKLAR.....	66
EKLER	69
ÖZGEÇMİŞ	84

ÇİZELGELER

	<u>Sayfa</u>
Çizelge 4.1. Açıklayıcı değişkenler ve düzeyleri.....	36
Çizelge 4.2. Zamana bağlı değişkenlerin anlamlılığı için model sonuçları.....	37
Çizelge 4.3. Yaşam süresi rankının Schoenfeld artıkları ile ilişkisi	39
Çizelge 4.4. Cox orantılı tehlikeler modelinin sonuçları	39
Çizelge 4.5. Modellerin karşılaştırılması	43
Çizelge 4.6. Log-odds ve normal sapma artık değerleri	46
Çizelge 4.7. Aykırı değerler çıkarıldıktan sonra Cox orantılı tehlikeler modeli sonuçları.....	56
Çizelge 4.8. Log-normal regresyon modelinin sonuçları.....	57
Çizelge 4.9. Log-normal regresyon modelinde log-odds ve normal sapma artık değerleri	61
Çizelge 4.10. Aykırı değerler çıkarıldıktan sonra log-normal regresyon modeli sonuçları.....	62
Çizelge 4.11. Log-normal regresyon modeli için adımsal seçim yöntemi sonuçları	63

ŞEKİLLER

Sayfa

Şekil 4.1. Yaş değişkeni için ölçeklendirilmiş Schoenfeld artığı grafiği	37
Şekil 4.2. Diğer açıklayıcı değişkenler için ölçeklendirilmiş Schoenfeld artığı grafikleri.....	38
Şekil 4.3. Kaplan-Meier yaşam eğrisi	40
Şekil 4.4. Birikimli tehlike eğrisi	40
Şekil 4.5. Cox orantılı tehlikeler modeli için Cox-Snell artığı grafiği.....	41
Şekil 4.6. a-e. Parametrik modeller için Cox-Snell artığı grafikleri	42
Şekil 4.7. Yaş değişkeni için Martingale artığı grafiği	43
Şekil 4.8. Doğrusal tahmine karşı Martingale artığı grafiği	44
Şekil 4.9. Yaş değişkeni için sapma artığı grafiği	45
Şekil 4.10. Doğrusal tahmine karşı sapma artığı grafiği	45
Şekil 4.11. Yaş değişkeni için DFBETA grafiği	47
Şekil 4.12. Cinsiyet değişkeni için DFBETA grafiği.....	48
Şekil 4.13. Kilo Kaybı (Semptom) değişkeni için DFBETA grafiği.....	48
Şekil 4.14. Anemi (Hb) değişkeni için DFBETA grafiği	49
Şekil 4.15. Diseksiyon değişkeni için DFBETA grafiği	49
Şekil 4.16. Tümörün midedeki lokalizasyonu değişkeninin orta üçlük düzeyi için DFBETA grafiği	50
Şekil 4.17. Tümörün midedeki lokalizasyonu değişkeninin alt üçlük düzeyi için DFBETA grafiği	51
Şekil 4.18. Tümörün midedeki lokalizasyonu değişkeninin tüm mide düzeyi için DFBETA grafiği	51
Şekil 4.19. Evre değişkeninin evre 3 düzeyi için DFBETA grafiği.....	52
Şekil 4.20. Evre değişkeninin evre 4 düzeyi için DFBETA grafiği.....	53
Şekil 4.21. Kemoterapi değişkeni için DFBETA grafiği.....	53
Şekil 4.22. Log-olabilirlik değişim grafiği.....	54
Şekil 4.23. Lmax istatistiği grafiği	55
Şekil 4.24. Log-normal regresyon modelinde yaş değişkeni için Martingale artığı grafiği	58
Şekil 4.25. Log-normal regresyon modelinde doğrusal tahine karşı Martingale artığı grafiği	58

Şekil 4.26. Log-normal regresyon modelinde yaş değişkeni için sapma artığı grafiği	59
Şekil 4.27. Log-normal regresyon modelinde doğrusal tahmine karşı sapma artığı grafiği	60

SİMGELER VE KISALTMALAR

AIC	Akaike Bilgi Kriteri
ÇEÖ	Çok Erken Ölen
ÇUY	Çok Uzun Yaşayan

1. GİRİŞ

Eldeki veri kümesine bir model uyarlandıktan sonra, uygulanan modelin varsayımlarının sağlanıp sağlanmadığının kontrol edilmesi gereklidir. Modelin kontrolü için tanı yöntemleri, modelleme sürecinin önemli bir kısmıdır [1]. Bu süreçlerin çoğu artıkların analizine dayanmaktadır. Yaşam çözümlemesinde özellikle Cox orantılı tehlikeler modelinde değişik amaçlarla kullanılabilecek değişik artık türleri vardır [2].

Yaşam çözümlemesinde en yaygın kullanılan artık türü Cox ve Snell (1968) tarafından önerilen Cox-Snell artıklarıdır [3]. Bu artık türü modelin uygunluğunun kontrolü için kullanılır. Diğer bir artık türü olan Martingale artıkları Barlow ve Prentice (1988) tarafından önerilmiştir ve Cox-Snell artıklarının doğrusal dönüşümüdür. Martingale artıkları açıklayıcı değişkenlerin modele uyumunu belirlemede kullanılır [4]. Benzer şekilde sapma artıkları da Martingale artıklarının dönüşümüdür. Aykırı değerler genellikle Therneau v.d. (1990) tarafından önerilen sapma artıkları ile görüntülenir [5]. Fleming ve Harrington (1991) sapma artıklarının referans örnekleme dağılımına sahip olmadığına ve veri kümesinde durdurulmuş gözlemler olmadığına bile standart normal dağılım yaklaşımının tatmin edici olmadığına dikkat çekmiştir [6]. Nardi ve Schemper (1999) bu problemin üstesinden gelmek için yeni artık türleri önermiştir. Bu yeni artık türlerinin, aykırı değerleri belirlemede daha doğru bir sonuç verdiği iddia edilmektedir. Bu yeni artıklar log-odds ve normal sapma artıklarıdır [7]. Diğer bir artık türü ise Cox orantılı tehlikeler modelinde orantılılık varsayımının testinde yaygın olarak kullanılan ve Schoenfeld (1982) tarafından önerilen Schoenfeld artıklarıdır [8].

Çalışmanın ikinci bölümünde yaşam çözümlemesi hakkında genel bilgiler verilmiş, yaşam çözümlemesinde durdurma kavramından bahsedilmiştir. Yaşam çözümlemesinde kullanılan fonksiyonlara ve birbiriyle ilişkilerine, yaşam çözümlemesinde en çok kullanılan model olan Cox orantılı tehlikeler modeline ve temel varsayımı olan orantılılık kavramı hakkında bilgilere yer verilmiştir. Ayrıca bu bölümde yaşam çözümlemesinde kullanılan parametrik dağılımlar ve aykırı değer kavramı ile ilgili genel bilgiler de yer almaktadır.

Üçüncü bölümde ise yaşam çözümlemesinde aykırı değerler hakkında ve aykırı değerleri belirlerken kullanılan yöntemler hakkında bilgiler yer almaktadır. Yaşam çözümlemesinde aykırı değerleri belirlerken kullanılan başlıca artık türlerinden ve etkili gözlemleri belirme yöntemlerinden bahsedilmiştir.

Dördüncü bölümde ise çalışma boyunca değinilen yöntemler mide kanseriyle ilgili gerçek verilere uygulanmış, aykırı değerler belirlenmiş ve gerekli analizler yapılarak sonuçlar yorumlanmıştır.

Bu çalışmanın amacı, yaşam çözümlemesinde karşılaşılan aykırı değerleri tanımlama yöntemlerini ve uygulama alanlarını incelemektir. Bu amaçla literatürde yer alan bu artık türleri incelenmiş ve gerçek bir veri kümesi üzerinde uygulanmıştır.

2. GENEL BİLGİLER

2.1. Yaşam Çözümlemesi

Yaşam çözümlemesi, ilgilenilen herhangi bir olayın ortaya çıkmasına kadar geçen sürenin incelenmesinde kullanılan çözümleme yöntemleri topluluğudur. Yaşam çözümlemesinin kendine özgü bir terminolojisi vardır. Geçen süre; araştırmanın başından ilgilenilen olay gerçekleşinceye kadar geçen yılları, ayları, haftaları, günleri ya da olay gerçekleştiği anda birimin yaşını ifade eder. Olay ise ölüm, hastalığa yakalanma, hastalığın kötüye gitmesi, iyileşme, işe dönme ya da birimin başına gelebilecek ilgilenilen herhangi bir olayı ifade eder. Olumlu (iyileşme, hastaneden taburcu olma gibi), olumsuz (ölüm, hastalığın kötüye gitmesi gibi) ya da birimin başına gelen ilgilenilen herhangi bir durum olayın konusunu oluşturabilir [9].

Yaşayan bir organizmanın ya da cansız bir nesnenin belirli bir başlangıç zamanı ile başarısızlığı arasında geçen zamana “yaşam süresi” ya da “başarısızlık süresi” adı verilmektedir. Her bir birime ait yaşam süresi, tanım gereği sürekli ve pozitif bir değere sahiptir [10].

Herhangi bir birim için başarısızlık ancak bir kez oluşabilir. Başarısızlık süresine örnek olarak, makine bileşenlerinin yaşam süreleri, işçilerin grev süreleri ya da ekonomide işsizlik dönemleri, psikolojik bir deneyde deneğin belirlenen görevi tamamlama süresi ve klinik bir deneyde hastaların yaşam süreleri gösterilebilir. Yaşam çözümlemesinde temel olan, gözlenen başarısızlık sürelerinin incelenmesi olduğundan, bu değişkenin iyi tanımlanması gerekmektedir. İlgilenilen olaya göre farklı şekillerde ortaya çıkabilen bu değişkenin duyarlı olarak ölçülmesi için,

- Başlangıç zamanı her bir birim için şüpheye yer vermeyecek şekilde tanımlanmalıdır.
- Geçen süreyi ölçmek için bir zaman ölçeği kabul edilmelidir.
- Her bir birim için başarısızlığın meydana geldiği an tamamen net olmalıdır [11].

Yaşam çözümlemesi verilerinin kaynağı genel olarak sağlık bilimleri olmakla beraber nüfus bilimi, mühendislik, ekonomi ve sosyal bilimler gibi diğer alanlar da olabilir.

Sosyolojide olay geçmişi çözümlenmesi (event history analysis), mühendislikte başarısızlık zamanı çözümlenmesi (failure time analysis), ekonomide süre çözümlenmesi (duration analysis) ya da geçiş çözümlenmesi (transition analysis) ve sağlık alanında yaşam çözümlenmesi (survival analysis) olarak adlandırılır [12].

Yaşam sürelerinin çözümlenmesinde kullanılan bazı yöntemler oldukça eskidir ama 1970'lerde başlayan bu alan, yöntem, teori ve uygulama açısından hızlıca ilerlemiştir.

Yaşam çözümlenmesinin amaçları,

- farklı zamanlarda yaşam olasılığı tahminlerinin elde edilmesi,
- yaşam süresinin dağılımının tahmin edilmesi,
- farklı hasta gruplarının yaşam süresi dağılımlarının karşılaştırılması,
- yaşam süresini etkileyen faktörlerin belirlenmesi

olarak sıralanabilir.

Yaşam çözümlenmesinde ilgilenilen problemin çözümüne ilişkin farklı yaklaşımlar;

- başarısızlık sürelerinin bağımlı değişken olarak kabul edildiği ve bu değişkeni açıkladığı düşünülen diğer değişkenlerin de çözümlenmede yer aldığı modellerin kullanıldığı yaklaşım,

Bu yaklaşım için iki model düşünülebilir:

- Bağımsız değişkenler olarak da nitelendirilen açıklayıcı değişkenlerin tehlike fonksiyonu üzerinde çarpımsal bir etkiye sahip olduğu orantılı tehlikeler modelleri,
- Açıklayıcı değişkenlerin yaşam süresinin logaritması üzerinde toplamsal ya da yaşam süresi üzerinde çarpımsal bir etkiye sahip olduğu log-doğrusal modeller.
- çeşitli parametrik yaşam dağılımlarını kullanarak tahminlerde bulunmak ve hipotez testlerini hazırlamak,
- herhangi bir dağılım varsayımına dayanmayan parametrik olmayan süreçleri kullanarak tahminlerde bulunmak

olarak özetlenebilir [1].

Yaşam çözümlemesinin iki önemli avantajı; yaşam süresinin dağılımı hakkında bilgi gerektirmemesi ve durdurma zamanına kadar tüm bilgileri kullanması ve hiçbir bilgiyi göz ardı etmemesidir [13].

2.1.1. Durdurma

Yaşam çözümlemesinde kullanılan modellerin diğer istatistiksel modellerden temel farkı zaman değişkeninin yapısından dolayı durdurulmuş (censored) gözlem içeren veri kümesi için tasarlanmış istatistiksel yöntemler bütünü olmasıdır.

Durdurma genel olarak;

1. Çalışmanın bitiş noktasına kadar olayın gözlenememesi (administrative censoring)
2. Çalışma bitmeden birimle ilgili bilgi alınamaması (lost to follow-up)
3. Başka bir olayla (başka nedenden ölüm, ilaç reaksiyonu gibi) karşılaşma (withdrawing)

nedenleri ile meydana gelir.

Bu gibi nedenlerle birimler daha fazla gözlemlenemez ve bu birimler "durdurulmuş gözlem" olarak ifade edilir [13].

Durdurulmuş birimler için başarısızlık gerçekleşmediğinden bu birimler için başarısızlık süresi yerine durdurma süreleri söz konusu olur [14].

Bu gibi durumlarda hiçbir araştırmacı bu birimlerin tamamını kayıp veri (missing data) olarak çalışmanın dışında bırakmak istemez, çünkü bunların çoğu araştırma sonucunu etkileyebilir. Birimin yaşam süresi ile ilgili bilgi, en son görüldüğü anda elde edilen bilgi olur ve durdurulmuş gözlem olarak araştırmaya katılır [1].

Durdurma soldan, sağdan ve aralıklı olmak üzere üç çeşittir. Ancak, en yaygın kullanılanı sağdan durdurmadır (right censoring). Örneğin, hastaların yirmi hafta boyunca gözlemlendiği bir çalışmada, bir hasta çalışma süresi boyunca ilgilenilen olayı yaşamamışsa bu sağdan durdurmadır. Bu kişinin yaşam süresi en az çalışma süresi kadar olmalıdır. Bu kişinin yaşam süresi durdurulmuş olarak tanımlanır [15].

Yaşam süresi her birim için farklı olabilir. Birimler farklı zamanlarda araştırmaya katılabilir. Tanımlanan olay farklı zamanlarda ortaya çıkabilir.

2.1.2. Yaşam Çözümlemesinde Kullanılan Fonksiyonlar

Yaşam süresi verileri başarısızlık, ölüm, tepki, iyileşme, verilen hastalığın gelişimi, tahliye ya da boşanma gibi belirli olayların zamanı ile ilgilidir. Bu zamanlar rasgeledir ve bir dağılım oluşturur. Yaşam sürelerinin dağılımları genel olarak üç fonksiyonla gösterilir. Bunlar; yaşam fonksiyonu, olasılık yoğunluk fonksiyonu ve tehlike fonksiyonudur. Bu fonksiyonlar matematiksel olarak birbirinden elde edilebilir [16].

Olasılık Yoğunluk Fonksiyonu

Birimin geçen yaşam süresi t , T değişkeninin bir değeri olarak gösterilebilir ve bu değer negatif olamaz. T rastlantı değişkeni $f(t)$ olasılık dağılımına sahiptir. $f(t)$ olasılık yoğunluk fonksiyonu

$$f(t) = \lim_{\delta t \rightarrow 0} \frac{P(t \leq T \leq t + \delta t)}{\delta t}$$

şeklindedir. T 'nin dağılım fonksiyonu,

$$F(t) = P(T < t) = \int_0^t f(u) du$$

biçimindedir ve yaşam süresinin t değerinden küçük olması olasılığını gösterir [1].

Yaşam Fonksiyonu

Yaşam fonksiyonu $S(t)$, yaşam süresinin t 'den daha büyük olması olasılığını verir ve

$$S(t) = P(T \geq t) = 1 - F(t)$$

biçimindedir [1].

Yaşam fonksiyonu monoton azalan soldan sürekli bir fonksiyondur ve

$$t = 0 \text{ iken; } S(t) = S(0) = 1$$

$$t = \infty \text{ iken; } S(t) = S(\infty) = 0$$

olur [15].

Tehlike Fonksiyonu

Tehlike fonksiyonu, t zamanında yaşadığı bilinen bir birimin t ile $t+\delta t$ zaman aralığında başarısız olma riskinin bir tanımıdır. Anlık ölüm hızı olarak da tanımlanır.

Tehlike fonksiyonu,

$$h(t) = \lim_{\delta t \rightarrow 0} \left\{ \frac{P(t \leq T < t + \delta t / T \geq t)}{\delta t} \right\}$$

biçimindedir. $h(t)$, 0 ile ∞ arasında değer alır.

Bu tanımdan yaşam ve tehlike fonksiyonları arasında bazı ilişkiler görülebilir.

$F(t)$, T ' nin dağılım fonksiyonu olmak üzere,

$$h(t) = \lim_{\delta t \rightarrow 0} \left\{ \frac{F(t + \delta t) - F(t)}{\delta t} \right\} \frac{1}{S(t)}$$

olur ve buradan

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

elde edilir.

Tehlike fonksiyonu,

$$h(t) = -\frac{d}{dt} \{\log S(t)\}$$

biçiminde de yazılabilir.

$$H(t) = \int_0^t h(u) d_u$$

iken

$$S(t) = \exp\{-H(t)\}$$

biçiminde olur.

$H(t)$, birikimli tehlike fonksiyonu olarak adlandırılır. Yaşam fonksiyonundan,

$$H(t) = -\log S(t)$$

ile elde edilebilir [1].

Birikimli tehlike fonksiyonu, artan, sağdan sürekli ve $\lim_{t \rightarrow \infty} H(t) = \infty$ olan bir fonksiyondur [11].

2.2. Cox Orantılı Tehlikeler Modeli

Yaşam çözümlemesinde en çok kullanılan model Cox orantılı tehlikeler (Cox proportional hazards) modelidir. 1972 yılında Cox tarafından geliştirilen regresyon modeli ile yaşam çözümlemesinde önemli adımlar atılmış, Cox (1972)' un önerileri, Kalbfleisch ve Prentice (1980) 'in katkıları ile bugünkü önemini kazanmıştır [17].

Cox orantılı tehlikeler modeli; bir birimin yaşam süresi ile birden fazla açıklayıcı değişken arasındaki ilişkiyi ortaya çıkaran istatistiksel bir yöntemdir. Cox orantılı tehlikeler modelinin amacı, yaşam verilerinin genel durumunu yansıtacak bir model oluşturmaktır. Bu yolla, yaşam süresi üzerinde etkili olduğu düşünülen değişkenlerin etkileri eşanlı olarak açıklanabilmektedir [18].

Cox orantılı tehlikeler modeli, Cox modeli veya Cox regresyon modeli (Cox regression model) olmak üzere farklı şekillerde adlandırılabilen, dağılım bilgisi gerektirmeyen bir modeldir. Bu modelde, yaşam süresi ve bu süre üzerinde etkili olarak görülen açıklayıcı değişkenler yer alır. Açıklayıcı değişkenler, modeli toplamsal değil, çarpımsal olarak etkiler [19].

Cox orantılı tehlikeler modeli,

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

biçimindedir. Burada, $\boldsymbol{\beta}$ regresyon katsayıları vektörü ve $h_0(t)$ ise açıklayıcı değişkene sahip olmayan ($\mathbf{x}=\mathbf{0}$ olan) bir birimin temel tehlike fonksiyonu olarak tanımlanmaktadır [20].

Cox regresyon modelindeki katsayılar, kısmi olabilirlik fonksiyonu ile tahmin edilir. Birbirinden farklı başarısızlık süreleri sayısı k olmak üzere, kısmi olabilirlik fonksiyonu $L(\boldsymbol{\beta})$ aşağıdaki gibi hesaplanır:

$$L(\boldsymbol{\beta}) = \prod_{i=1}^k \exp(\boldsymbol{\beta}' x_i) / \sum_{j \in R_j} \exp(\boldsymbol{\beta}' x_j)$$

Burada x_i , i. sıralı başarısızlık zamanı t_i 'de başarısız olan birimler için açıklayıcı değişkenler vektörüdür. Paydadaki toplam, t_i zamanda riskte olan birimler üzerinden $\exp(\beta'x)$ değerlerinin toplamıdır.

Regresyon katsayılarının en çok olabilirlik tahminleri, logaritmik olabilirlik fonksiyonunu en büyükleyerek $\ln L(\beta)$ ile hesaplanır. Newton-Raphson algoritması kullanılarak iteratif çözümlenmeler yapılır ve bu yöntemle β katsayıları tahmin edilir.

Cox regresyon modelinde kullanılan tehlike oranı, farklılığın, durumlar arasında kaç kat olduğunu belirtmektedir. Tehlike oranı $\exp(\beta)$ şeklinde ifade edilir. Tehlike oranı 1 ise durumlar arasında bir farklılık olmadığı anlamına gelir. 1'den büyük olan tehlike oranı riskin arttığına, 1'den küçük olan tehlike oranı ise riskin azaldığına işaret eder [21].

Parametrik modellerin gerektirdiği varsayımların sağlanmadığı durumlarda Cox orantılı tehlike modeli, parametrik analizlerden daha etkilidir.

Cox orantılı tehlikeler modelinde;

- açıklayıcı değişkenlerin tehlike fonksiyonu üzerindeki etkileri log-doğrusaldır.
- açıklayıcı değişkenlerin log-doğrusal fonksiyonu ile tehlike fonksiyonu arasındaki ilişki çarpımsaldır.

Bu model için varsayım, tehlike oranının zamana göre değişmemesi, yani sabit olmasıdır. Tehlike oranı ile ilgili bu varsayım orantılı tehlikeler varsayımı olarak bilinir [18].

2.3. Orantılı Tehlikeler Varsayımının İncelenmesi

Cox orantılı tehlikeler modelinin temel varsayımı orantılı tehlikelerdir. Orantılı tehlike varsayımı, tehlike oranının zamana karşı sabit olması ya da bir birimin tehlikesinin diğer birimin tehlikesine orantılı olması anlamına gelmektedir [22].

$\mathbf{X}^* = (x_1^*, x_2^*, \dots, x_p^*)$ ve $\mathbf{X} = (x_1, x_2, \dots, x_p)$ iki birime ait açıklayıcı değişkenler vektörü olmak üzere tehlike oranı,

$$H\hat{O} = \frac{\hat{h}(t, \mathbf{x}^*)}{\hat{h}(t, \mathbf{x})} = \frac{\exp(\sum_{j=1}^p \hat{\beta}_j \mathbf{x}_j^*) \hat{h}_0(t)}{\exp(\sum_{j=1}^p \hat{\beta}_j \mathbf{x}_j) \hat{h}_0(t)} = \exp \left[\sum_{j=1}^p \hat{\beta}_j (\mathbf{x}_j^* - \mathbf{x}_j) \right]$$

biçimindedir. Üstteki eşitlikte görüldüğü gibi tehlike oranı t 'yi içermez. Bir başka deyişle, model uydurulduğunda \mathbf{x}^* ve \mathbf{x} için değerler belirlendiğinde, tehlike oranı tahmini için üstel ifadenin değeri sabittir, zamana bağlı değildir. Bu sabit $\hat{\theta}$ ile gösterilirse, tehlike oranı

$$\hat{\theta} = \frac{\hat{h}(t, \mathbf{x}^*)}{\hat{h}(t, \mathbf{x})}$$

biçiminde yazılabilir. Bu orantılı tehlike varsayımını gösteren matematiksel bir ifadedir. Orantılı tehlike varsayımı,

$$\hat{h}(t, \mathbf{x}^*) = \hat{\theta} \hat{h}(t, \mathbf{x})$$

biçiminde de ifade edilebilir. Burada $\hat{\theta}$, orantılılık sabiti (proportionality constant) olarak adlandırılır ve zamandan bağımsızdır [1, 20].

Orantılı tehlikeler varsayımını incelemek için grafiksel ya da sayısal yöntemler kullanılmaktadır. Orantılı tehlike varsayımının incelenmesinde en çok kullanılan yöntemler, log(-log) yaşam eğrileri, gözlenen ve beklenen yaşam eğrileri, Arjas grafikleri, modele zamana bağlı değişkenlerin eklenmesi, Schoenfeld artıkları ile yaşam süresinin rankı arasındaki korelasyon testi biçiminde sıralanabilmektedir [20].

En yaygın kullanılan grafiksel yöntem olan log(-log) yaşam eğrileri, açıklayıcı değişkenlerin farklı düzeyleri üzerinden $-\ln(-\ln)$ yaşam eğrilerinin tahmininin karşılaştırılmasını içermektedir. Elde edilen paralel eğriler orantılı tehlikeler varsayımının sağlandığını göstermektedir.

Orantılı tehlikeler varsayımını değerlendirmek için gözlenen yaşam olasılıkları ile beklenen yaşam olasılıklarının çizimlerinin kullanımı uyum iyiliği testi yaklaşımının grafiksel karşılığıdır. Beklenen yaşam eğrilerini elde etmek için, değerlendirilen tahmin ediciyi içeren Cox orantılı tehlikeler modeli uydurulur. Yaşam eğrisinin tahmini, tahmin edicinin her bir sınıfı için değerleri formülde yerine yerleştirilerek

elde edilir. Tahmin edicinin her bir sınıfı için gözlenen ve beklenen yaşam eğrileri birbirine yakınsa, orantılı tehlikeler varsayımının sağlandığına karar verilir. Bir ya da daha çok sınıf için bu eğriler farklılık gösterirse, orantılı tehlikeler varsayımının bozulduğuna karar verilir.

Arjas grafikleri ise orantılı tehlikeler modelinde uyum iyiliğinin test edilmesi için kullanılan grafiksel bir yöntemdir. Alternatif modellerin tahminine gereksinim duymadığından ve sadece parametre tahmininde kullanılan kısmi olasılık ifadesindeki niceliklere benzer nicelikleri içerdiğinden kullanılması kolay bir yöntemdir. Grafiksel yöntemler içinde en uygun sonucu veren yöntemdir [23].

Zamandan bağımsız değişkenin orantılı tehlike varsayımını sağlayıp sağlamadığını incelemek için zamana bağlı değişkenler de kullanılır. Bu durumda model, zamandan bağımsız değişkenleri ve zamanın fonksiyonlarını kapsayan çarpım terimlerini içeren Cox orantılı tehlikeler modeli olmaktadır. Bu genişletilmiş Cox orantılı tehlikeler modeli x açıklayıcı değişkeni için

$$h(t, x) = h_0(t) \exp[\beta x + \delta(xg(t))]$$

biçimindedir.

$g(t)$ fonksiyonu için farklı seçenekler vardır. Bunlar;

$$g(t) = t$$

$$g(t) = \log t$$

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

biçiminde olabilir [15].

Genişletilmiş Cox orantılı tehlikeler modelinde, çarpım terimlerinin önemliliği için test yapılarak orantılı tehlikeler varsayımı değerlendirilir. Yokluk hipotezi, $H_0 : \delta = 0$ biçimindedir. Yokluk hipotezi doğru ise, model tek bir değişken içeren orantılı tehlikeler modeline indirgenir [23].

Orantılı tehlikeler varsayımı Schoenfeld artıkları ile de kontrol edilebilir. Schoenfeld artıkları açıklayıcı değişkenin gerçek değeri ile ağırlıklı risk skorlarının ortalaması arasındaki farktır. Schoenfeld artıkları zamana karşı çizilir ve orantılı tehlikeler varsayımının geçerliliğini kontrol amacıyla kullanılır. Eğer artıklar yatay bir doğru

etrafında (sıfır merkezli) tesadüfi olarak yer alıyorsa, orantılı tehlikeler varsayımı sağlanır. Orantılı tehlikeler varsayımı sağlanmazsa grafik bir trend gösterir [24].

Bu test istatistiği, orantılı tehlikeler varsayımının incelenmesi için kullanılan grafiksel yöntemlere göre daha nesnel bir kriter sağlamaktadır. Grafiksel yöntemler ise daha öznedir [23].

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

Cox orantılı tehlikeler modelinde orantılı tehlike varsayımı sağlanmıyorsa yaşam verisinin modellenebilmesi için birçok yaklaşım önerilmektedir. Bu yaklaşımlar aşağıda verilmiştir:

- Orantısızlığa neden olan değişkenlerle tabakalandırma yapmak (Tabakalandırılmış Cox regresyon modeli)
- Orantısızlığı zamana bağlı açıklayıcı değişkenlerle modellemek (Genişletilmiş Cox regresyon modeli)
- Farklı modeller kullanmak (Ağırlıklandırılmış Cox regresyon modeli, hızlandırılmış başarısızlık zamanı modeli ya da toplamsal tehlike modeli) [22].

2.4. Parametrik Dağılımlar

Yaşam çözümlemesinde, elde edilen veriye hangi dağılımın uygulanacağı çok önemlidir. Bu nedenle model seçimi yapılırken ilk olarak modelin veriye uygunluğu için gerekli test yapılmaktadır. Model veriye uygunsa parametrik test varsayımlarının sağlanıp sağlanmadığına bakılarak modelin geçerliliği ile ilgili gerekli analizler yapılmaktadır.

Yaşam çözümlemesi verilerinde yaygın olarak kullanılan parametrik dağılımlar;

- Üstel dağılım
- Weibull dağılımı
- Log-normal dağılım
- Log-lojistik dağılım
- Gamma dağılımı

olarak ifade edilebilir.

2.4.1. Üstel Dağılım

Üstel dağılım için “Tamamen Rastgele Başarısızlık Modeli” ifadesi de kullanılmaktadır. Birçok yaşam çözümlemesi verisi üstel dağılım tarafından yeterli derecede tanımlanmasa da, bazı genel durumlarda anlamayı kolay hale getirmektedir [25].

Sabit tehlike fonksiyonuna sahip üstel dağılım, yaşam çözümlemesi ve güvenilirlik teorisinde oldukça yaygın uygulama alanı bulmuştur. Böylece üstel dağılımın tehlike fonksiyonu,

$$h(t; \beta) = \beta^{-1}, t \geq 0$$

biçimindedir. Buradan yaşam fonksiyonu,

$$S(t; \beta) = \exp\left(-\int_0^t h(u) d_u\right) = \exp(-t / \beta)$$

biçiminde yazılabilir. Dağılım fonksiyonu ise,

$$F(t; \beta) = 1 - \exp(-t / \beta), t \geq 0$$

biçimindedir. Böylece β ortalamalı bir üstel dağılımın olasılık yoğunluk fonksiyonu,

$$f(t; \beta) = \beta^{-1} \exp(-t / \beta), t \geq 0$$

biçimindedir.

Bu fonksiyondaki β parametresi ölçek parametresidir ve $\beta > 0$ 'dır. T , β ölçek parametrelili bir üstel dağılımlı rasgele değişken ise $T \sim E(\beta)$ ile gösterilir. Böyle bir rasgele değişkenin ortalaması,

$$E(T) = \mu = \beta$$

ortancası, $Med(T) = \beta \log 2$ ve standart sapması ise $\sigma = \beta$ 'dir.

İntegralenmiş tehlike fonksiyonu ise,

$$H(t; \beta) = \beta^{-1} \int_0^t d_u = \beta^{-1} t$$

eşitliği ile verilir [1].

2.4.2. Weibull Dağılımı

Weibull dağılımı 1960 ve 1970'li yıllarda çok yaygın olarak kullanılmaya başlanmıştır. Bu model daha esnek bir dağılım sınıfı vermekte ve özel bir durum olarak üstel dağılımı içine almaktadır.

Weibull dağılımı üstel dağılımın genelleştirilmiş halidir. Sabit tehlike oranına sahip olmaması Üstel dağılımdan tek farkıdır. Sabit tehlike oranına sahip olmadığından genelleştirici uygulamaları Üstel dağılımdan daha fazladır. Bu dağılım daha çok güvenilirlik çalışmalarında ve hastalık ölümlerinde kullanılmaktadır [25].

Weibull dağılımı yaşam çözümlemesi modellerinde en yaygın kullanılan dağılımdır. Risk ve sigortacılık gibi alanlarda sıkça kullanılmaktadır. Başarısızlığın oluşmasına kadar geçen süreyi ya da başarısızlıktan sonra ikinci bir başarısızlığın oluşmasına kadar geçen süreyi modellemede kullanılan bir dağılımdır [26].

Weibull dağılımı iki parametrelidir ve $W(\alpha, \beta)$ ile gösterilir. Bu dağılım Johnson v.d. (1970) ve Barnett (1982) gibi birçok araştırmacı tarafından da çalışılmıştır [25]. Tehlike fonksiyonu,

$$h(t; \alpha, \beta) = \alpha \beta^{-1} (\beta^{-1} t)^{\alpha-1}$$

ile verilir. Weibull dağılımı için yaşam fonksiyonu,

$$S(t; \alpha, \beta) = \exp\left[-(\beta^{-1} t)^\alpha\right]$$

ve dağılım fonksiyonu,

$$F(t; \alpha, \beta) = 1 - \exp\left[-(t/\beta)^\alpha\right], t \geq 0$$

biçiminde verilir.

Burada β ölçek ve α ise biçim parametresidir. $\alpha = 1$ olduğu zaman Weibull dağılımı üstel dağılım olur. Buradan Weibull dağılımının olasılık yoğunluk fonksiyonu,

$$f(t; \alpha, \beta) = \alpha \beta^{-1} (\beta^{-1} t)^{\alpha-1} \exp\left[-(t/\beta)^\alpha\right], t \geq 0$$

ile ifade edilir. Weibull dağılımının beklenen değeri,

$$E(T) = \beta \Gamma_1$$

ile verilir. Ortancası, $Med(T) = \beta(\log 2)^{1/\alpha}$ 'dır. Burada Γ_k gamma fonksiyonudur ve

$$\Gamma_k = \Gamma\left(1 + \frac{k}{\alpha}\right) \text{ 'dır. Varyansı,$$

$$V(T) = \beta^2(\Gamma_2 - \Gamma_1^2)$$

ile verilir. İntegrallenmiş tehlike fonksiyonu ise,

$$H(t; \alpha, \beta) = (\beta^{-1}t)^\alpha$$

ile verilir [27].

2.4.3. Log-Normal Dağılım

Yaşam çözümlenmesinde, elde edilen veriye hangi dağılım şeklinin uygulanacağı konusu ayrı bir önem taşımaktadır. Bazı durumlarda belli bir dağılımı seçmek için bir neden olmaktadır. Örneğin, geçmiş deneyimler modelin, daha önce incelenen benzer yığınları oldukça iyi bir şekilde tanımladığını göstermektedir ya da geçerliliğini sağlayan başarısızlık sürecine ait bir bilgi elde edilebilmektedir. Bazen model, var olan veriye iyi uyum sağlasa bile tehlike fonksiyonunun formundan dolayı uygun olmayabilir. Log-normal dağılım böyle bir dağılımdır ve azalan tehlike fonksiyonunun sahip olduğu forma göre uygun olmaktan uzaklaşabilmektedir. Benzer olasılık yoğunluk fonksiyonlarına sahip iki dağılım çok farklı tehlike fonksiyonlarına sahip olabilmektedir. Bu durumda tehlike fonksiyonunun yapısına ilişkin özel bir bilgi ya da varsayım, dağılımlarından biri lehine olmak üzere karar verilmesi sağlanabilmektedir [25].

Eğer $\log T$, μ ortalamalı ve σ^2 varyansı ile normal dağılıma sahip ise, T rasgele değişkeninin μ ve σ parametrelili log-normal dağılıma sahip olduğu söylenir.

Log-normal dağılımın tehlike fonksiyonu,

$$h(t; \mu, \sigma) = f(t)S^{-1}(t)$$

ile verilir. Log-normal dağılım için yaşam fonksiyonu,

$$S(t; \mu, \sigma) = 1 - \Phi\left[\frac{\ln t - \mu}{\sigma}\right]$$

ile verilir. Burada Φ , normal birikimli dağılım fonksiyonudur.

Eğer T rasgele değişkeni log-normal dağılıma sahip ise $T \sim LN(\mu, \sigma)$ ile gösterilir.

Böylece $T \sim LN(\mu, \sigma)$ 'nin dağılım fonksiyonu,

$$F(t; \mu, \sigma) = \Phi\left(\frac{\ln t - \mu}{\sigma}\right), t > 0$$

ile verilir.

Buradan olasılık yoğunluk fonksiyonu ise,

$$f(t; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} t^{-1} \exp\left[-(\log t - \mu)^2 / 2\sigma^2\right], t > 0$$

ile verilir [27].

2.4.4. Log-Lojistik Dağılım

Log-lojistik dağılım, Weibull dağılımının alternatifi olmaktadır. Eğer $\log(T)$ lojistik dağılıma sahipse T yaşam süresi Log-lojistik dağılmaktadır. Bu dağılım iki parametreye sahiptir. Bunlar α ve λ 'dir. Bu dağılım başta artan daha sonra azalan bir şekle sahip olmaktadır. Log-normal dağılım ile karşılaştırıldığında benzer yaşam fonksiyonuna ve tehlike oranına sahip olmaktadır [25].

Weibull tehlike fonksiyonunun kısıtlı olduğu durumlarda Log-lojistik dağılım yararlı olabilir. Bu dağılımın tehlike fonksiyonu,

$$h(t; \alpha, \lambda) = \alpha \lambda t^{\alpha-1} (1 + \lambda t^\alpha)^{-1}, t \geq 0$$

ile verilir. Eğer $\alpha \leq 1$ ise tehlike fonksiyonu monoton olarak azalır. Fakat $\alpha > 1$ ise tehlike fonksiyonu monoton olarak artar ve yaşam fonksiyonu,

$$S(t; \alpha, \lambda) = [1 + \lambda t^\alpha]^{-1}, t \geq 0$$

ile verilir. Log-lojistik dağılımın dağılım fonksiyonu,

$$F(t; \alpha, \lambda) = \lambda t^\alpha [1 + \lambda t^\alpha]^{-1}, t \geq 0$$

biçimindedir. Buradan olasılık yoğunluk fonksiyonu,

$$f(t; \alpha, \lambda) = \alpha \lambda t^{\alpha-1} [1 + \lambda t^\alpha]^{-2}, t \geq 0$$

denklemleri ile verilir. Böyle bir olasılık yoğunluk fonksiyonuna sahip T rasgele değişkenine λ ve α parametrelili Log-Lojistik dağılıma sahiptir denir. Bu dağılımın ortalaması ve varyansı sırası ile

$$E(T) = \frac{1}{\lambda^{1/\alpha}} \left[\Gamma\left(1 + \frac{1}{\alpha}\right) \Gamma\left(1 - \frac{1}{\alpha}\right) \right]$$

$$V(T) = \frac{1}{\lambda^{2/\alpha}} \left\{ \left[\Gamma\left(1 + \frac{2}{\alpha}\right) \Gamma\left(1 - \frac{2}{\alpha}\right) \right] - [E(T)]^2 \right\}, \alpha > 2$$

ile verilir [27].

2.4.5. Gamma ve Genelleştirilmiş Gamma Dağılımı

Gamma dağılımı Ki-kare ve Üstel dağılımı içermektedir. Brown ve Flood (1947) tarafından bir kafeteryadaki bardak kullanım zamanı için tanımlanmıştır. Birnbaum ve Saunders (1958) ise materyallerin yaşam uzunluğu için bir istatistiksel model olarak tanımlamışlardır. İzleyen dönemlerde de yaşam çözümlemesi problemlerinde sık bir şekilde kullanılmaya başlanmıştır [25].

Gamma dağılımı λ ve γ olmak üzere iki parametreye sahiptir.

$0 < \gamma < 1$ iken, grafik azalış yönündedir, yani zaman 0'dan sonsuza artarken tehlike oranı monoton bir şekilde sonsuzdan λ' ya doğru azalmaktadır. $\gamma > 1$ ise bir artma söz konusudur, yani tehlike oranı zaman artarken λ' ya göre artmaktadır. $\gamma = 1$ ise tehlike oranı λ' ya eşit olmaktadır.

Gamma dağılımının olasılık yoğunluk fonksiyonu,

$$f(t) = \frac{\lambda}{\Gamma(\gamma)} (\lambda t)^{\gamma-1} e^{-\lambda t} \quad t > 0, \lambda > 0, \gamma > 0$$

biçimindedir. Burada γ şekil, λ ölçüm parametresidir. Yani γ 'daki değişim grafiğin şeklini değiştirirken, λ 'daki değişim ölçümünü değiştirmektedir.

$\gamma > 1$ iken grafik tek bir tepe noktasına sahipken, bu tepe noktası $t = (\gamma - 1)/\lambda$ biçimindedir.

Gamma dağılımı için yaşam fonksiyonu,

$$S(t) = e^{-t} \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!},$$

Tehlike fonksiyonu ise,

$$h(t) = \frac{\lambda(\lambda t)^{n-1}}{(n-1)! \sum_{k=0}^{n-1} (1/k!)(\lambda t)^k}$$

biçimindedir.

Genelleştirilmiş Gamma dağılımının λ, α ve γ olmak üzere üç parametresi vardır.

Genelleştirilmiş Gamma dağılımının olasılık yoğunluk fonksiyonu,

$$f(t) = \frac{\alpha \lambda^{\alpha \gamma}}{\Gamma(\gamma)} t^{\alpha \gamma - 1} \exp[-(\lambda t)^\alpha] \quad t > 0, \gamma > 0, \lambda > 0, \alpha > 0$$

biçimindedir.

Üstel, Weibull, log-normal ve Gamma dağılımları Genelleştirilmiş Gamma dağılımının özel halleridirler. Buna göre;

$\alpha = \gamma = 1$ ise Üstel dağılım,

$\gamma = 1$ ise Weibull dağılımı,

$\gamma \rightarrow \infty$ ise Log-normal dağılım,

$\alpha = 1$ ise Gamma dağılımı elde edilir [16].

2.5. Aykırı Değerler

Eldeki veri kümesine bir model uyarlandıktan sonra, uygulanan modelin varsayımlarının sağlanıp sağlanmadığının kontrol edilmesi gereklidir. Modelin kontrolü için tanı yöntemleri, modelleme sürecinin önemli bir kısmıdır [1].

Aykırı değerler, veri kümesindeki diğer verilerle karşılaştırıldıklarında çok büyük ya da çok küçük oldukları gözlemlenen değerlerdir. Analiz sonuçları üzerinde hiçbir etkiye sahip olmayabilecekleri gibi çok büyük bir etki de yaratabilirler. Bir veri kümesinde aykırı değerlerin belirlenmesi istatistiksel analizler açısından önemlidir. Aykırı değerler hatalı veriler olmamaları durumunda bize bir süreci nasıl iyileştirebileceğimiz konusunda ipuçları verir [28].

Artıklar en yaygın kullanılan tanı yöntemleridir. Eğer uygulanan model uygunsa artıklar için çizilen grafiklerde beklenilmeyen örüntüler oluşmaz. Doğrusal regresyondaki artık değerleri en basit haliyle tahmin edilen değer ile gerçek değer

arasındaki farktan hesaplanır. Bu durumda grafiklerin sıfır etrafında rasgele saçılımlı olması beklenir [29].

Regresyon verilerinde aykırı değerler literatürde tartışılan bir problemdir. Beckman ve Cook (1983) aykırı değeri, hedef dağılımı gerçekleştirilmeyen gözlemler olarak tanımlarken, Barnett ve Lewis (1984) ise veri kümesindeki diğer gözlemlerden tutarsız bir gözlem olarak tanımlamışlardır. Hawkins (1980)'e göre ise, bir aykırı değer, diğer gözlemlerden çok fazla sapan, dolayısıyla kuşku uyandıracak şekilde farklı bir mekanizmadan türetilmiş bir gözlemdir. Aykırı değerlerin varlığı çalışmanın her alanındaki uygun model parametre tahminlerini değiştirir. Belsley v.d. (1980), Barnett ve Lewis (1984) ve Montgomery ve Pack (1992) de doğrusal regresyon problemlerinde aykırı değer bulmayı tartışmışlardır [30].

Aykırı değerler değişik nedenlerle ortaya çıkabilir. Örneğin, kaba (gross) hatalardan kaynaklanabilir. Bu hatalar kopyalama ya da veri girişi hataları, hatalı ondalık noktalama, çalışmadan elde edilen ölçümleri hatalı ölçeklendirme, farklı anlamdaki iki veriyi karıştırma, farklı kitleden hatalı alınan bir gözlem, geçici etkiler ya da donanım yetersizliği gibi hatalar olabilir. Bu tür hatalar, iyi koşullar altında özel ilgiyle gözlemlenirken yüksek nitelikli verilerde nadiren rastlanır. Hampel v.d. (1986) verilerin tipik olarak %1-10 oranında bu tür hataya sahip olduklarını belirtmişlerdir [30].

Aykırı değerler genellikle üç sebeple ortaya çıkmaktadırlar:

- Veri kümesinin yapısındaki doğal farklılıklardan: Veri kümesindeki gözlemlerin doğal farklılıklarından kaynaklanır ve bu tür durumlar kontrol edilemez.
- Ölçüm yanlışlıkları: Ölçme aracının yetersizliğinden ya da yanlış okuma veya kaydetmeden kaynaklı durumlardır. Bu tür durumların kontrolü mümkündür.
- Uygulama hatası: Verinin yanlış kopyalanmasından kaynaklı durumlardır. Veri kümesinin kitleyi iyi temsil etmeyen yanlış örneklemelerinden seçilmesi durumudur [31].

3. YAŞAM ÇÖZÜMLEMESİNDE AYKIRI DEĞERLER

Yaşam çözümlemesinde verinin kendine has özelliklerinden dolayı birçok yazar aykırı değere özel bir anlam vermeyi denemiştir. Collet (1994) yaşam çözümlemesinde aykırı değerlerin, son derece uzun yaşam süresine sahip birimler olduğunu belirtmiş ama açıklayıcı değişkenlerin erken başarısız olmuş ya da tam tersi birimler olmasını tavsiye etmiştir [1]. Therneau v.d. (1990) ve Nardi ve Schemper (1999) aykırı değerleri çok erken ölen ya da çok uzun yaşayan birimlerle ilişkilendirmişlerdir [5, 7]. Nardi ve Schemper (1999) daha sonradan yaşam çözümlemesinde aykırı değerleri "çok uzun yaşayan" ya da "çok erken ölen" birimler olarak belirtmişlerdir [7].

Yaşam çözümlemesi çalışmalarında toplanmış veriler aykırı değerler içerebilir. Aykırı değerler genellikle modele iyi uyum sağlamayan değerler olarak tanımlanır. Bu aykırı değerler, model tahminine göre "çok geç" ya da "çok erken" başarısız olan birimler olmasına göre büyük pozitif ya da negatif artık değerleri alabilir. Birimler uzun yaşam süresine sahip olabilir ama açıklayıcı değişkenlerin değerleri birimin daha erken ölmesi gerektiğini gösterebilir [30].

Bazı durumlarda, veri kümesinde yapılan dikkatli kontrol belirgin özellikleri saptamaya yardımcı olur. Örneğin, bazen uzun ya da kısa yaşam sürelerine sahip birimler dikkat çeker. Bu durum durdurma ile daha karmaşıklaşır ki durdurulmuş yaşam sürelerinin ortaya çıkışı modelin yeterliliğini test etmeyi en basit durumda olsa bile zorlaştırır. Bu yüzden verinin görsel testi modelin yeterliliğinin testi için daha biçimsel uygulamalarla desteklenmelidir.

Yaşam çözümlemesindeki bağımlı değişkenler çalışmadaki birimlerin yaşam süresi ve durumu hakkında bilgi içerdiğinden, aykırı değerlerin çözümlemeleri doğrusal regresyondakinden biraz farklıdır ve modellerin yeterliliğinin incelenmesinde kullanılan yöntemler doğrusal regresyon modellemesinde kullanılan uyum yöntemlerinden daha karmaşıktır [1].

Yaşam çözümlemesinde aykırı değerler, modeldeki parametre tahminlerini etkileyebilir, tehlike oranını, seçilen modeli değiştirebilir ve modele dayanan tahminleri etkileyebilir. Bu aykırı değerler etkili gözlemler olarak tanımlanır. Etkili gözlemlere genelde uzun yaşam sürelerinde rastlanılmaktadır. Cox orantılı tehlikeler modelinde, etkili gözlemler parametre tahminlerini etkileyebilirler ve

sonuçta tehlike oranını değiştirebilirler. Kısacası, yaşam çözümlemesinde veri kümesindeki etkili gözlemlerin ve aykırı değerlerin ortaya çıkışının araştırılması oldukça önemlidir [30].

Aykırı değerleri belirlemenin yanı sıra etkili gözlemlerin belirlenmesi için de bir prosedürün geliştirilmesine gerek vardır. Regresyon problemlerinde, Belsley ve arkadaşları (1980), Hadi (1992) ve Imon (2005) doğrusal regresyon modelinde leave-one-methodunu kullanarak gözlemlerin etkilerini başarılı bir şekilde tanımlamıştır. Diğer yandan, benzer yöntemler yaşam çözümlemesi problemlerindeki etkili gözlemleri belirlemek için de önerilmektedir. Standart tam delta-beta prosedürünün yanı sıra, Storer ve Crowley (1985) genelleştirilmiş yaklaşımı (augmented approach (AUG)) önerirken, Cain ve Lange (1984) ve Reid ve Crépeau (1985) etki fonksiyonunu (influence function (IF)) bu problemi incelemek için kullanmışlardır. Wang v.d. (2006) tüm yaklaşımları karşılaştırmış ve AUG yaklaşımının daha üstün performansa sahip olduğunu göstermiştir [30].

Regresyon modellerindeki etkili gözlemleri belirlemek için Atkinson and Riani (2000) ileriye doğru araştırma yöntemi (forward search method) olarak adlandırılan alternatif bir yöntem önermiştir. Bu yaklaşım aykırı değer içermeyen ilk altküme ile başlar. Etkili gözlemler, bu alt kümeye değişkenler eklendikçe ilgilenilen parametrede oluşan değişikliklere bakılarak belirlenir. Bu yaklaşım yaşam çözümlemesinde etkili gözlemleri belirlemek için de kullanılmıştır. Burada ilk altkümenin hangisi olacağı ve değişkenlerin bu altkümeye giriş sırası her gözlemin artık karelerinin sıralanması ile belirlenmiştir. Gözlemlerin, altkümeye girişte hesaplanan istatistikler üzerindeki etkileri devamlılık eğrileri ile gözlemlenebilir. İlgilenilen istatistikte büyük etkiye sahip olan gözlemleri belirlemek için bir basit istatistik önerilmiştir [30].

Birçok model kontrol yöntemleri artıklara ve artıkların analizine dayanmaktadır. Uygulanan model yeterli olduğunda, her değişken için çalışmada hesaplanabilecek değerler ve en azından yaklaşık olarak bilinen davranış özellikleri vardır [1].

Yaşam çözümlemesinde, başlıca artık türleri;

- a) Cox-Snell artıkları
- b) Martingale artıkları
- c) Sapma artıkları
- d) Schoenfeld artıkları
- e) Log-odds artıkları
- f) Normal sapma artıkları

olarak ifade edilebilir.

Aşağıda artık türlerinin anlatılmasında kullanılacak bazı eşitlikler verilmiştir:

$t_1 < t_2 < \dots < t_D$ başarısızlık sürelerini ve d_i , t_i zamanındaki başarısızlık sayılarını gösterebilir. Birikimli temel tehlike oranı tahmin edicisi;

$$\widehat{H}_0(t) = \sum_{t_i \leq t} \frac{d_i}{\sum_{j \in R(t_i)} \exp(\widehat{\beta}' Z_j(s))} \quad (3.1)$$

biçimindedir. Burada $\widehat{\beta}$, tahmini regresyon katsayılarını, Z_j 'ler de açıklayıcı değişkenleri göstermektedir.

$$H_0(t) = \int_0^t h_0(u) du$$

olarak verilir ve bu gözlemlenen başarısızlık sürelerinde sıçramalar ile bir adım fonksiyonudur. Burada $h_0(u)$ temel tehlike fonksiyonunu göstermektedir.

3.1. Cox-Snell Artıkları

Cox-Snell artıkları, yaşam verilerinin çözümlemesinde en yaygın kullanılan artık türüdür ve Cox ve Snell tarafından önerilmiştir [3]. Bu artıklar, modelin uygunluğunu değerlendirmek için kullanılabilir.

i. gözlem için Cox-Snell artığı,

$$r_{c_i} = \widehat{H}_0(t_i) \exp(\widehat{\beta}' Z_i) \quad (3.2)$$

biçimindedir. Burada $\widehat{H}_0(t)$, Eşitlik (3.1)'de verildiği gibidir.

Eğer seçilen model veriye uygunsa ve $\hat{\beta}$ değerleri β değerlerine yakınsa o zaman r_{C_i} 'ler birim üstel dağılımlı durdurulmuş bir örneklem gibi dağılır [32].

Bu artıklar negatif değildir ve grafiklerinin yorumlanması zordur.

Cox-Snell artıkları modele uyumu araştırmak için çok kullanışlıdır. Ancak, tahmin edilen birikimli tehlike grafiği doğrusal olmadığında bulunan modelden sapma türlerini belirleyememektedir [33].

Cox-Snell artıkları, doğrusal regresyon analizinde kullanılan artıklardan biraz farklı özelliklere sahiptir. Bu özellikler;

- sıfır etrafında simetrik dağılmaz,
- sıfır ile sonsuz arasında değer alır yani negatif olamaz,
- uygun model uydurulduğunda Cox-Snell artıklarının üstel dağılımlı olduğu varsayıldığından oldukça çarpık bir dağılımı vardır ve i. artığın ortalaması ve varyansı 1 olur,

biçiminde verilebilir.

Diğer bir nokta ise eğer en büyük yaşam süresi durdurulmamış ise r_{C_i} bu gözlem için tanımsızdır [1].

$\hat{H}_{r_{C_i}}(r_{C_i})$ 'nin r_{C_i} 'ye karşı grafiği bir güven bandı içermelidir. Böylece önem belirlenebilir. Ama r_{C_i} 'ler tam olarak bir dağılımın durdurulmuş bir örneklemini değildir. Yani bu grafik genellikle sadece kaba bir tanı olarak kullanılır [33].

3.2. Martingale Artıkları

Martingale artıkları, Barlow ve Prentice tarafından önerilmiştir [29]. Zamana bağlı olmayan açıklayıcı değişkenlerle Cox modeli için, i. gözlemin t_i zamanındaki değeri ve δ_i olay durumu ($\delta_i = 0$ iken t_i durdurma süresi ve $\delta_i = 1$ iken t_i olay süresi) ise Martingale artığı,

$$r_{m_i} = \delta_i - \hat{H}_0(t_i) \exp(\hat{\beta}' Z_i) = \delta_i - r_{C_i}$$

(3.3)

biçimindedir.

Eşitlik (3.3)'te görüldüğü gibi Martingale artıkları Cox-Snell artıklarının doğrusal dönüşümüdür.

Martingale artıkları ve dönüşümleri model belirlemede kullanılabilir. Açıklayıcı değişkenlere karşı çizilen Martingale artıkları grafiği modele dahil edilen açıklayıcı değişkenlerin fonksiyonel formunu belirlemede kullanılır. Modelin uygunluğu için değişkenlerin herhangi bir dönüşüme ihtiyacı olup olmadığını gösterir [34]. En uygun biçimi belirlemek için model kurulduktan sonra Martingale artıklarına karşı değişkenlerin istenilen dönüşümlerinin grafiği çizdirilir. Eğer dönüşüm uygun ise grafikteki eğri yaklaşık olarak doğrusal olur. Ayrıca aykırı değerleri belirlemede de kullanılır [35].

Martingale artıkları, her birim için gözlenen ve beklenen başarısızlık sayıları arasındaki farkla ilgilenir. Çünkü hiç bir olayda bir başarısızlıktan fazla başarısızlık olamaz varsayılır. Eşitlik (3.3)'ün ikinci faktörü $(0, t_i)$ aralığının üstündeki birim için başarısızlıkların birikimli tehlikelerinin tahminidir [33].

Martingale artıkları $\sum_{i=1}^n r_{m_i} = 0$ özelliğine sahiptir [32].

Büyük n için, r_{m_i} 'ler sıfır ortalamalı kitleden ilişkisiz örneklerdir. Ama sıfır etrafında simetrik değillerdir [33].

Martingale artıkları $-\infty$ ve 1 arasında değer alır ve durdurulmuş gözlemler $(\delta_i = 0)$ için artıklar negatif değer alır [1].

Martingale artıkları bir değerinin yakınlarında değer alıyorsa beklenenden daha kısa yaşam süresi, büyük negatif bir değer alıyorsa beklenenden daha uzun yaşam süresi olduğu anlamına gelir.

Doğrusal regresyon analizindeki artıklarla benzer özellikleri vardır. Fakat Martingale artıkları sıfır etrafında simetrik dağılmaz, çarpıktır [1].

Martingale artıkları genellikle konuya özeldir, çalışmanın sonunda hesaplanır ve tehlikenin logaritmasında açıklayıcı değişkenlerin doğrusal etkileri varsayımının test edilmesinde yararlıdır [36].

Doğrusal olmama Cox orantılı tehlikeler modellerinde karşılaşılabilecek olası bir problemdir. Martingale artıkları eğriselliği saptamak için açıklayıcı değişkenlere

karşı çizilebilir. Ayrıca doğrusal modellerde bileşen-artı-artık (ya da kısmi artık) grafikleri oluşturmak için de kullanılabilir [37].

3.3. Sapma Artıkları

Sapma artıkları Therneau v.d. tarafından önerilmiştir [5]. Sapma artıkları olan r_{d_i} 'ler Martingale artıklarından dönüştürülmüştür ve

$$r_{d_i} = \text{sign}(r_{m_i}) \sqrt{2[-r_{m_i} - \delta_i \log(\delta_i - r_{m_i})]} \quad (3.4)$$

biçimindedir.

Eşitlik (3.4)'te görüldüğü gibi, simetrik bir dağılım elde etmek için Martingale artıklarının dönüştürümüyle elde edilir. Sonuç olarak sapma artıkları sıfır etrafında simetrik dağılır ve yaklaşık olarak 1 standart sapmaya sahiptir [32, 38].

$\text{sign}(r_{m_i})$ Martingale artıklarının işaretini göstermektedir. Bu nedenle sapma artıkları Martingale artıkları ile aynı işarete sahiptir [30].

Karekök alma işlemi büyük negatif Martingale artıklarıyla ilişkilendirirken logaritma fonksiyonu Martingale artıklarını bire yaklaştırır. Bütün grafiklerde, potansiyel aykırı değerler büyük mutlak değer sapma artıklarına karşılık gelir.

Durdurulmuş gözlemler için sapma artıkları sıfır yakınlarında küme olarak yer alır [35].

Durdurma çok küçük olduğunda yani yüzde yirmi beşten az oranda ya da yakın bir değerde ise, bu artıklar sıfır etrafında simetriktir ya da normal dağılıma oldukça yakındır. Yüzde kırktan daha fazla orandaki durdurma için ise sıfır etrafındaki artıklarla geniş noktalar kümesi normallik yaklaşımını bozar. Ama bu dönüşüm artıkların setlerini simetrikleştirmede yardımcıdır [33, 5].

Martingale artıkları model uygun olsa bile çarpıktır ve bu çarpıklık artık grafiklerinin yorumlanmasını zorlaştırır. Sapma artık grafiklerinin yorumlanması ise daha kolaydır. Böylece bu artıklar diğer artıklardan aşırı derecede farklı olan yaşam sürelerine sahip birimlerin belirlenmesinde grafiksel bir araç olarak kullanılabilir. Birçok araştırmacı aykırı değerleri belirlemede sapma artık grafiklerini kullanmışlardır. Aykırı değerleri tanımlamak için herhangi bir hipotez testi olmadan öznel akıl yürütme kullanılmıştır [30].

Sapma artıkları beklenenden daha uzun yaşam süresi olan gözlemler için negatif iken beklenenden daha kısa yaşam süresi olan gözlemler için pozitiftir. Çok büyük ya da çok küçük değerler olması bu değerlerin aykırı değer olduklarının göstergesi olabilir. Bu nedenle bu değerler dikkatle incelenmelidir [38].

Başarısız olan gözlemlerde uyum yeterli ise sapma artıkları ak gürültüye benzer biçimde dağılır. Durdurulmuş gözlemler için ise sapma artıkları sıfır yakınlarında küme olarak yer alır [39].

Açıklayıcı değişkenlere karşı artık grafiği çizilebilir ve olağanüstü örüntüler modelin verilere iyi bir uyumu olmadığını gösterebilir. Ancak dikkat edilmelidir ki durdurulmuş gözlemlerin olması çarpıcı bir örüntü oluşturabilir ve bu modelle ilgili bir problem olduğu anlamı taşımaz [38].

3.4. Schoenfeld Artıkları

Cox orantılı tehlikeler modelinde kullanılan Cox-Snell, Martingale ve sapma artıklarının iki dezavantajı söz konusudur. Bu dezavantajlar, artıkların ağırlıklı olarak gözlenen yaşam süresine bağlı olmaları ve birikimli tehlike fonksiyonunun tahminini gerektirmesidir. Schoenfeld tarafından önerilen, skor artıkları olarak da adlandırılan Schoenfeld artıklarında bu sorunlar giderilmiştir [8]. Bu yönüyle Schoenfeld artıkları, diğer artıklardan önemli bir farklılık göstermektedir. Bu artıklarda, her birimin artığı için tek bir değer yerine, tahmin edilmiş olan Cox orantılı tehlikeler modelinde yer alan her bir açıklayıcı değişken için birer tane olmak üzere değerler kümesi yer almaktadır [1, 18].

Schoenfeld artıkları, değişkenin gerçek değeri ile ağırlıklı risk skorlarının ortalaması arasındaki farktır.

i. birim için Schoenfeld artığı,

$$\widehat{S}_i = X_i - \frac{\sum_{j \in R(t_i)} X_j \exp(\widehat{\beta}' X_j)}{\sum_{j \in R(t_i)} \exp(\widehat{\beta}' X_j)}$$

olarak tanımlanır. Burada \widehat{S}_i , p açıklayıcı değişken sayısı olmak üzere, px1 boyutlu bir vektördür ve $\widehat{S}_i = (\widehat{S}_{i1}, \dots, \widehat{S}_{ip})'$ şeklindedir. i. birim ve k. açıklayıcı değişken için Schoenfeld artıkları,

$$\hat{S}_{ik} = X_{ik} - \frac{\sum_{j \in R(t_i)} X_{jk} \exp(\hat{\beta}' X_j)}{\sum_{j \in R(t_i)} \exp(\hat{\beta}' X_j)}$$

biçimindedir. Burada X_j j. birim için p sabit açıklayıcı değişken vektörüdür. X_{jk} ise j. birimin k. açıklayıcı değişkenin değeridir. Bu nedenle, bu artık, X_{jk} 'nin gözlenen değeri ile t_i zamanında risk altındaki birimler üzerinden açıklayıcı değişken değerlerinin ağırlıklı ortalaması arasındaki farktır. t_i zamanındaki risk kümesinde k. açıklayıcı değişkenli j. birim için kullanılan ağırlık,

$$\frac{\exp(\hat{\beta}' X_j)}{\sum_{j \in R(t_i)} \exp(\hat{\beta}' X_j)}$$

biçimindedir. Bu ifade, kısmi olabilirliği en büyükmeye bu birimin katkısıdır.

Schoenfeld artıkları durdurulmamış gözlemler için tanımlıdır ve ayrıca her bir açıklayıcı değişkenin Schoenfeld artıkları toplamı sıfır olmalıdır [32].

Schoenfeld artıkları sıfırda toplanır. İki düzeyli (0,1) değişkenler için bu artıklar -1 ile 1 arasında değer alır. Bu nedenle, artık grafiğinde iki kuşak olur; bir tanesi $x=1$ için sıfırın üstünde ve diğeri $x=0$ için sıfırın altında yer alır.

Orantılı tehlikeler varsayımını incelemek için Schoenfeld artıklarına dayanan bir test geliştirilmiştir. Belirli bir değişken için Schoenfeld artıkları ile birimlerin yaşam sürelerinin rankı arasındaki korelasyon kullanılarak orantılı tehlikeler varsayımı incelenebilir. Bu teste göre, orantılı tehlikeler varsayımının sağlanması için korelasyonun sıfıra yakın olması beklenmektedir [20].

Orantılı tehlikeler varsayımının geçerliliği zamana bağlı çizilen Schoenfeld artıkları grafiği ile kontrol edilebilir. Çizilen Schoenfeld artıkları grafiği, yatay bir doğru etrafında seyrediyorsa, orantılı tehlikeler varsayımının sağlandığı söylenilir [8].

Hosmer ve Lemeshow regresyon katsayılarının kovaryans matrisine dayalı ölçeklendirilmiş Schoenfeld (scaled Schoenfeld) artıkları grafiğinin orantılı tehlikeler varsayımı için kullanılmasını önermiştir [19].

Ölçeklendirilmiş Schoenfeld artıkları;

$$r_{kt}^* = m \sum_{i=1}^p V_{ik} \hat{S}_{ik}$$

biçimindedir. Burada m toplam başarısız birim sayısını, V ise regresyon katsayılarından tahmin edilmiş kovaryans matrisini göstermektedir.

Ölçeklendirilmiş Schoenfeld artıkları grafiği zamana karşı her bir açıklayıcı değişken için çizilir. Ölçeklendirilmiş Schoenfeld artıkları etkili gözlemleri bulmak için kullanılır.

Winnett ve Sasieni açıklayıcı değişkenler sürekli olduğunda ölçeklenmiş Schoenfeld artıklarının kullanılmasını önermiştir [24, 40].

İlgilenilen olayla erken bir zamanda karşılaşan birimler için artıklar pozitif olma eğilimi gösterirse ve olayla geç bir zamanda karşılaşan birimler için artıklar negatif olma eğilimi gösterirse, tehlike oranı zaman eksenini boyunca sabit olmaz, bu durumda orantılı tehlikeler varsayımı sağlanmaz [20].

Schoenfeld artıklarının toplamı sıfır olmalıdır. Bu artıklar için, büyük örneklerde \hat{S}_i 'nin beklenen değerinin sıfıra eşit olduğu ve artıkların bağımsız olduğu söylenebilir [1].

Bu artıkları kullanırken, t_i zamanında başarısızlık olasılıkları düşük olan birimler, t_i zamanında başarısızlık riski altında olan birimlere göre, daha düşük değerli artıklara sahip olacaklardır. Dolayısıyla, başarısızlık riski altında olan birimlere ilişkin artıklar, başarısızlık riski az olan birimlere ilişkin artıklara oranla daha büyük olacaklardır [18].

Erken başarısızlığa uğrayan büyük ortak değişken değerlerine sahip birimler büyük pozitif Schoenfeld artıkları değerlerine sahiptir. Benzer şekilde erken başarısızlığa uğrayan küçük açıklayıcı değişken değerlerine sahip birimler büyük negatif Schoenfeld artıklarına sahiptir. Bu nedenle, erken başarısızlık sürelerinde mutlak değer olarak büyük artıklar özel bir ilgiye neden olmayabilir. Böylece, bu artıklar erken başarısızlık sürelerinde aykırı ortak değişken değerlerini tespit etmeye yardımcı olur. Ancak orantılı tehlikeler varsayımı sağlanıyorsa değer olarak büyük Schoenfeld artıklarının geç başarısızlık sürelerinde görünmeleri

beklenmez. Bu nedenle, geç başarısızlık sürelerinde artıklar kontrol edilmelidir [33].

3.5. Log-Odds Artıkları

Log-odds artıkları, Nardi ve Schemper tarafından verilmiş ve L_i ile gösterilmiştir [7]. Log-odds artıklarının dağılımı $E(L_i) = 0$ ortalama ve $V(L_i) = (\pi^2 / 3)$ varyans ile lojistik dağılımdır ve

$$L_i = \log \left[S_i(T_i) / \{1 - S_i(T_i)\} \right]$$

biçimindedir. Bilinmeyen yaşam fonksiyonu yerine tahmin edicisi alındığında \hat{L}_i , L_i 'ye yakınsar. Gözlenen yaşam süreleri t_i ($1 \leq i \leq n$) olduğunda, \hat{L}_i 'nin gözlenen değeri,

$$\hat{l}_i = \log \left[\hat{S}_i(t_i) / \{1 - \hat{S}_i(t_i)\} \right]$$

olur.

Durdurulmuş süre için $S_i(t_i^c)$ yaşam olasılığı bilinmemektedir. Dolayısıyla durdurulmuş yaşam süresi artıklarını uyarlamak için, $\hat{S}_i(t_i^c)$ ile koşullu ortanca değeri $\frac{\hat{S}_i(t_i^c)}{2}$ 'nin yer değiştirmesini önerilmiştir. Bu durumda l_i^m log-odds artık değerleri,

$$l_i^c = \log \left[\hat{s}_i(t_i^c) / \{2 - \hat{s}_i(t_i^c)\} \right]$$

biçiminde elde edilir. L_i 'nin koşullu dağılımından beklenen ortalama değerlerin sapması kullanıldığında ise daha karmaşık olmakta ve

$$l_i^m = l_i^c - \frac{1 + \exp(l_i^c)}{\exp(l_i^c)} \log \{ 1 + \exp(l_i^c) \}$$

biçiminde elde edilmektedir.

Durdurulmuş yaşam sürelerinde beklenen ortalama ve ortanca değerlerinin artıkları elde edilirken kısa durdurulmuş yaşam sürelerinden gelen bilgiler, onları

aykırı olarak belirlemek için oldukça kısıtlıdır. Aslında son derece kısa yaşam süreleri için l_i^m (ya da l_i^c) sifıra yakın olur ve beklenen $\widehat{S}_i(t_i) \cong 0.5$ olduğu varsayılır.

Durdurulmuş yaşam süreleri aykırı değerler olmaya adaydırlar. Bu nedenle aykırı değerleri belirlemek için her durdurulmuş yaşam süresi için P_i olasılığını hesaplamak özellikle caziptir. Eğer durdurulmamışsa karşılık gelen artık, en büyük negatif değerlerle ("çok uzun yaşamış" birimlerle ilgili) α -yüzdeli (genellikle $\alpha = 0,05$ veya $\alpha = 0,025$) artıkların üyesi olur. $P_i = \min\left[1.0, \alpha / \widehat{S}_i(t_i^c)\right]$, $S_i(T_i)$ 'nin $\left[0, \widehat{S}_i(t_i^c)\right]$ 'deki uniform dağılımının sonucudur. Böylece erken durdurulan bir birim α' ya eşit ya da α' nın biraz üzerinde P_i olasılığına sahiptir ve çok dikkat çekmez. Ancak uzun durdurma süreli bir birim, α -yüzdeli "çok uzun yaşayan" birimlerin grubunun üyesi olup yüksek olasılığa sahip olabilir. Durdurulmuş birimler birikimli yaşam olasılığı α 'dan küçük iken kesinlikle aykırıdır.

Log-odds artıkları lojistik dağılımlı olduklarından aykırı değerler bu dağılımlara dayanan kesim noktası ile belirlenebilir.

Kesim noktaları "çok erken ölen"=ÇEÖ ve "çok uzun yaşayan"=ÇUY ile gösterilirse log-odds artıkları için kesim noktaları,

$$R_{\text{ÇEÖ,L}} = \{l_i : l_i > w_{1-\alpha}\}$$

$$R_{\text{ÇUY,L}} = \{l_i : l_i < w_\alpha\}$$

biçimindedir. Burada w_α lojistik dağılımda α güven düzeyine karşılık gelen noktadır. Eğer birimlerin artık değerleri kesim noktalarını aşarsa bu gözlemler aykırı değerler olur [7, 41].

3.6. Normal Sapma Artıkları

Normal sapma artıkları Nardi ve Schemper tarafından verilmiş ve N_i ile ifade edilmiştir [7]. Artıkların dağılımı standart normal dağılımdır.

Bilinmeyen yaşam fonksiyonu yerine tahmin edicisi alındığında N_i 'nin olasılıkta yakınsaması \widehat{N}_i ' dir. Φ , normal birikimli dağılım fonksiyonudur.

Böylece,

$$N_i = \Phi^{-1} \{S_i(T_i)\}$$

ve artık değerleri

$$\hat{n}_i = \Phi^{-1} \{\hat{S}_i(t_i)\}$$

biçimindedir. Burada durdurulmuş süreler için $S_i(t_i^c)$ yaşam olasılığı bilinmemektedir.

Durdurulmuş yaşam sürelerinin artıklarının uyarlanması (accommodating) için çeşitli yollar vardır. Gerçek yaşam süresi, gözlenen durdurulmuş yaşam süresinden daha uzundur ve bilinmeyen doğru artıkların dağılımı $S_i(T_i)$ 'nin

$[0, \hat{S}_i(t_i^c)]$ 'de uniform dağılımı ile ilişkilidir. Böylece, $\hat{S}_i(t_i^c)$, koşullu ortanca değeri

$\frac{\hat{S}_i(t_i^c)}{2}$ ile yer değiştirir. Sonuç olarak durdurulmuş süre için N_i ,

$$n_i^c = \Phi^{-1} \left\{ \frac{\hat{S}_i(t_i^c)}{2} \right\}$$

biçimindedir ve burada n_i^c , durdurulmuş süre N_i 'nin ortalamasıdır ya da

$$n_i^m = -\frac{\exp(0.5(n_i^c)^2)}{\sqrt{2\pi} \hat{S}_i(t_i^c)}$$

ile değiştirilmiş olabilir. Burada n_i^m , durdurulmuş süre N_i 'nin ortancasıdır [7, 41].

Durdurulmuş gözlemler için normal sapma artıkları koşullu ortalama ya da ortanca değeriyle yer değiştirir. Ancak bu durum artıkların anormal bir yığılma göstermesine neden olur. Böylece iyi uyum sağlayan model bile normal dağılımdan farklı bir dağılım gösterir [41].

Log-odds artıklarında olduğu gibi normal sapma artıklarında da son derece kısa yaşam süreleri için n_i^m (ya da n_i^c) sifıra yakın olur ve beklenen $\hat{S}_i(t_i) \cong 0.5$ olduğu varsayılır [7].

Normal sapma artıkları normal dağılımlı olduklarından aykırı değerler bu dağılımlara dayanan kesim noktası ile belirlenebilir.

Normal sapma artıklarının kesim noktaları,

$$R_{CEÖ,N} = \{n_i : n_i > z_{1-\alpha}\}$$

$$R_{ÇUY,N} = \{n_i : n_i < z_\alpha\}$$

biçimindedir. Burada z_α standart normal dağılımda α güven düzeyine karşılık gelen noktadır. Benzer kesim noktaları sapma artıkları için de kullanılabilir çünkü aynı referans dağılıma sahiptirler [7, 41].

3.7. Etkili Gözlemlerin Belirlenmesi

Cox orantılı tehlikeler modelinde aykırı değerler üç madde altında incelenebilir:

1. Örneklem ortalamasından büyük oranda farklı bir açıklayıcı değişken değerine sahip olanlar,
2. Parametre tahminlerinde güçlü etkiye sahip olanlar,
3. Kısmi olabilirlik fonksiyon değerinde ve böylece model yeterliğinde güçlü etkiye sahip olanlar.

İlk madde skor artıkları kullanılarak belirlenebilir. Model kestiriminden sonra, her gözlem için skor artıkları ilgili açıklayıcı değişken bakımından hesaplanır ve analiz edilen ortak değişkene karşı artıkların değerlerini gösteren bir grafik çizilir. Aykırı değerler bu tip analizlerle kolaylıkla belirlenebilir. Aykırı değerlerin belirlenmesinden sonraki aşama şüpheli gözlemlerin parametre tahminlerindeki etkilerinin şiddetinin tahmin edilmesidir. Şüpheli gözlem örneklem dışında bırakılarak model yeniden tahmin edilmiştir. Parametre tahminindeki değişim,

$$\Delta\beta_{ki} = \beta_k - \beta_{k(-i)}$$

biçiminde hesaplanır. Burada β_k , k. açıklayıcı değişkenin modelin tüm örneklemdeki parametre kestirimidir ve $\beta_{k(-i)}$ ise i. gözlem çıkarıldıktan sonra hesaplanan örneklemdeki benzer bir değerdir. Δ_i k. elemanının vektörü ile iyi yaklaşılmış olduğu kanıtlanmıştır ve

$$\Delta\beta_i = V(\beta)L_i$$

biçimindedir. Burada L_i i. gözlem için skor artıkları vektörü ve $V(\beta)$ parametre tahminlerinin bir varyans-kovaryans matrisidir. Bu tanım ölçeklendirilmiş skor artıkları yada DFBETA artıkları olarak adlandırılır. Analiz edilen açıklayıcı değişkene karşı bu tip artıkların grafiği etkili gözlemlerin saptanmasında yararlıdır [36].

Eğer bu fark sıfıra yakın ise i. gözlemin tahmindeki etkisi çok küçüktür. Asıl modeldeki tüm değişkenler için bu süreç tekrarlanır. N gözlemlili büyük veri setleri için bu farkları hesaplamak yani bu işlemi N kez tekrarlamak pratik bir yöntem değildir [35].

Bu gözlemlerin kısmi olabilirlik fonksiyonu değerindeki etkilerini tahmin etmek için skor artıkları,

$$ld_i = \Delta\beta' V(\beta_i)^{-1} \Delta\beta_i = L_i' V(\beta) V(\beta)^{-1} V(\beta) L_i = L_i' V(\beta) L_i$$

biçimindedir.

Bu istatistikler, olabilirlik değişim (likelihood displacement) olarak adlandırılır ve örneklemden i. gözlemin örneklem dışında bırakılmasından sonra kısmi olabilirlik fonksiyonunun logaritmasındaki değişiklik ile ilgilenir.

Olabilirlik değişim değerleri i. gözlemin etkisini bu gözlem çıkarıldığında log olabilirlik modelindeki değişim yaklaşımıyla hesaplar.

i. değişken için olabilirlik değişim değeri,

$$2 \left\{ \log L(\hat{\beta}) - \log L(\hat{\beta}_i) \right\}$$

biçimindedir. Burada $\hat{\beta}$ tüm model için hesaplanan tahmini ve $\hat{\beta}_i$ i. gözlem çıkarıldıktan sonra hesaplanan tahmin değerini gösterir. $L(\cdot)$ tüm veriden tahmin edilen kısmi olabilirlik değeridir. $L(\cdot)$ hesaplanırken tüm veri kullanılır ama $\hat{\beta}_i$ parametre tahminleri i. gözlem çıkarılarak elde edilir. $\hat{\beta}$ uygun bir çözüm verir ve olabilirlik değişimi hiçbir zaman negatif değildir.

Ayrıca $L_i' V(\beta) L_i$ matrisi için özdeğerler bulunur. Yüksek özdeğerler ile ilişkili özdeğerler LMAX istatistikleri olarak adlandırılır. Olabilirlik değişim de LMAX da özet istatistiklere (örneğin Martingale artıkları) karşı çizilir. Yüksek derecede etkili gözlemler bu tarz grafiklerle kolayca belirlenebilir [36].

Etkili gözlemlerin belirlenmesinde kullanılan iki alternatif yöntem LMAX ve olabilirlik değişim değerleri, DFBETA değerlerinin aksine gözlemlerin etkisini bütün olarak katsayılar vektöründen ölçer. Böylece çok değişkene karşı sadece tek bir değer elde edilir [35].

Artıkların kullanımı veriye ve araştırmacıya bağlı olsa da geleneksel ve tavsiye edilen kullanım şekilleri vardır. Bunlar;

- Cox-Snell artıkları modele uyumu araştırmak için kullanılır.
- Martingale artıkları modele dahil edilen açıklayıcı değişkenlerin fonksiyonel formunu belirlemede ve bazen de aykırı değerleri belirlemede kullanılır.
- Sapma artıkları modelin doğruluğunu test etmede ve aykırı değerleri belirlemede kullanılır.
- Schoenfeld ve ölçeklendirilmiş Schoenfeld artıkları orantılı tehlikeler varsayımının kontrolünde kullanılır.
- Normal sapma ve log-odds artıkları aykırı değerleri belirlemede kullanılır.
- Olabilirlik değişim ve Lmax değerleri etkili gözlemleri belirlemede kullanılır. DFBETA değerleri de etkili gözlemleri belirlemede kullanılır ancak işlemleri değişkenler bazında teker teker yapar.

4. UYGULAMA

Ankara Onkoloji Hastanesi'nde Ocak 1990 ve Kasım 1995 tarihleri arasında mide kanseri tanısı konulan ve cerrahi tedavi geçiren, yaşları 29 ile 84 arasında değişen 118 hastaya ait veriler Erođlu v.d. tarafından incelenmiş ve Kaplan-Meier yöntemiyle yaşam olasılıkları bulunmuş, orantılı tehlikeler varsayımının sağlanıp sağlanmadığına bakılmaksızın Cox orantılı tehlikeler modeli elde edilmiştir [42]. Yılmaz ve Karasoy çalışmalarında aynı verileri, mide kanserinden ölümü etkileyen faktörleri belirlemek için, orantılılık varsayımı incelendikten sonra Cox orantılı tehlikeler ve orantılı odds modellerinin uygulanmasını göstermek ve elde edilen sonuçları karşılaştırabilmek amacıyla kullanmıştır [43].

Bu çalışmada ise bu veriler, yaşam çözümlemesinde aykırı değerleri belirlemede kullanılan artık türlerinin uygulama alanlarını göstermek ve elde edilen sonuçları karşılaştırabilmek amacıyla kullanılmıştır.

Analizler için STATA 12 programı deneme sürümü kullanılmıştır.

Açıklayıcı değişkenler olarak yaş, cinsiyet, kilo kaybı, anemi, tümörün midedeki lokalizasyonu, lenf nodu diseksiyonunun genişliği, hastalığın evresi ve adjuvan kemoterapi alınmıştır. Bu hastaların yaş ortalaması 56.70'dir ve %58.5'i erkektir. Hastaların %55.1'i hastalığın 3. evresinde ve %19.5'i 4. evresindedir. Hastaların %71.2'si kemoterapi almıştır. Hastaların 52 tanesinde başarısızlık gözlenmiştir. Başarısızlık, ölüm olarak alınmıştır. Açıklayıcı değişkenlerin düzeyleri ve sıklıkları Çizelge 4.1'de verilmiştir.

Çizelge 4.1. Açıklayıcı değişkenler ve düzeyleri

Değişkenler	Düzeyleyler	Sıklıklar (%)
Cinsiyet	1. Erkek 2. Kadın	69 (58.5) 49 (41.5)
Kilo Kaybı	1. Yok 2. Var	90 (76.3) 28 (23.7)
Anemi	1. Yok 2. Var	31 (26.3) 87 (73.7)
Lenf nodu diseksiyonunun genişliği (Diseksiyon)	1. D0-1 2. D2-3	62 (52.5) 56 (47.5)
Tümörün midedeki lokalizasyonu (Lokal)	1. Üst üçlük 2. Orta üçlük 3. Alt üçlük 4. Tüm mide	21 (17.8) 25 (21.2) 62 (52.5) 10 (8.5)
Hastalığın evresi (Evre)	1. Evre1+Evre2 2. Evre3 3. Evre4	30 (25.4) 65 (55.1) 23 (19.5)
Adjuvan kemoterapi (Kemoterapi)	1.Yok 2. Var	34 (28.8) 84 (71.2)
Yaş	56,403 ±1,044	

Yaşam süresine değişkenlerin etkisi araştırılmak istenildiğinde Cox orantılı tehlikeler modelinin kullanılabilmesi için değişkenlerin orantılı tehlikeler varsayımını sağlaması gerekmektedir. Orantılı tehlikeler varsayımın sağlanıp sağlanmadığı araştırılmıştır.

Orantılı tehlikeler varsayımını incelemek için kullanılacak bir yöntem modellere değişkenle birlikte zamana bağlı değişken eklenerek yapılan zamana bağlı açıklayıcı değişken kullanılması yöntemidir. Zamana bağlı değişkenlerin anlamlılığı için bulunan değerler Çizelge 4.2'de verilmiştir.

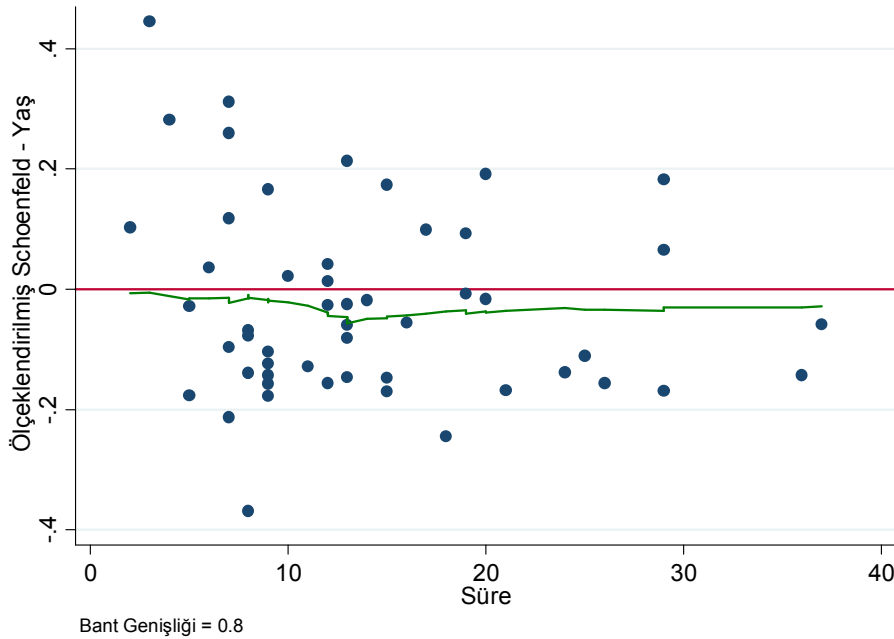
Çizelge 4.2. Zamana bağlı değişkenlerin anlamlılığı için model sonuçları

Değişkenler	Tahmin	Standart Hata	p-değeri
Yaş x T	-0.001	0.002	0.615
Cinsiyet x T	0.003	0.035	0.939
Kilo Kaybı x T	0.032	0.038	0.395
Anemi x T	0.025	0.042	0.549
Diseksiyon x T	0.023	0.039	0.555
Lokal x T	-0.021	0.022	0.979
Evre x T	0.063	0.040	0.118
Kemoterapi x T	0.043	0.037	0.329

Çizelge 4.2 incelendiğinde tüm değişkenler için p-değeri>0.05 olduğundan orantılı tehlikeler varsayımının sağlandığı söylenebilir.

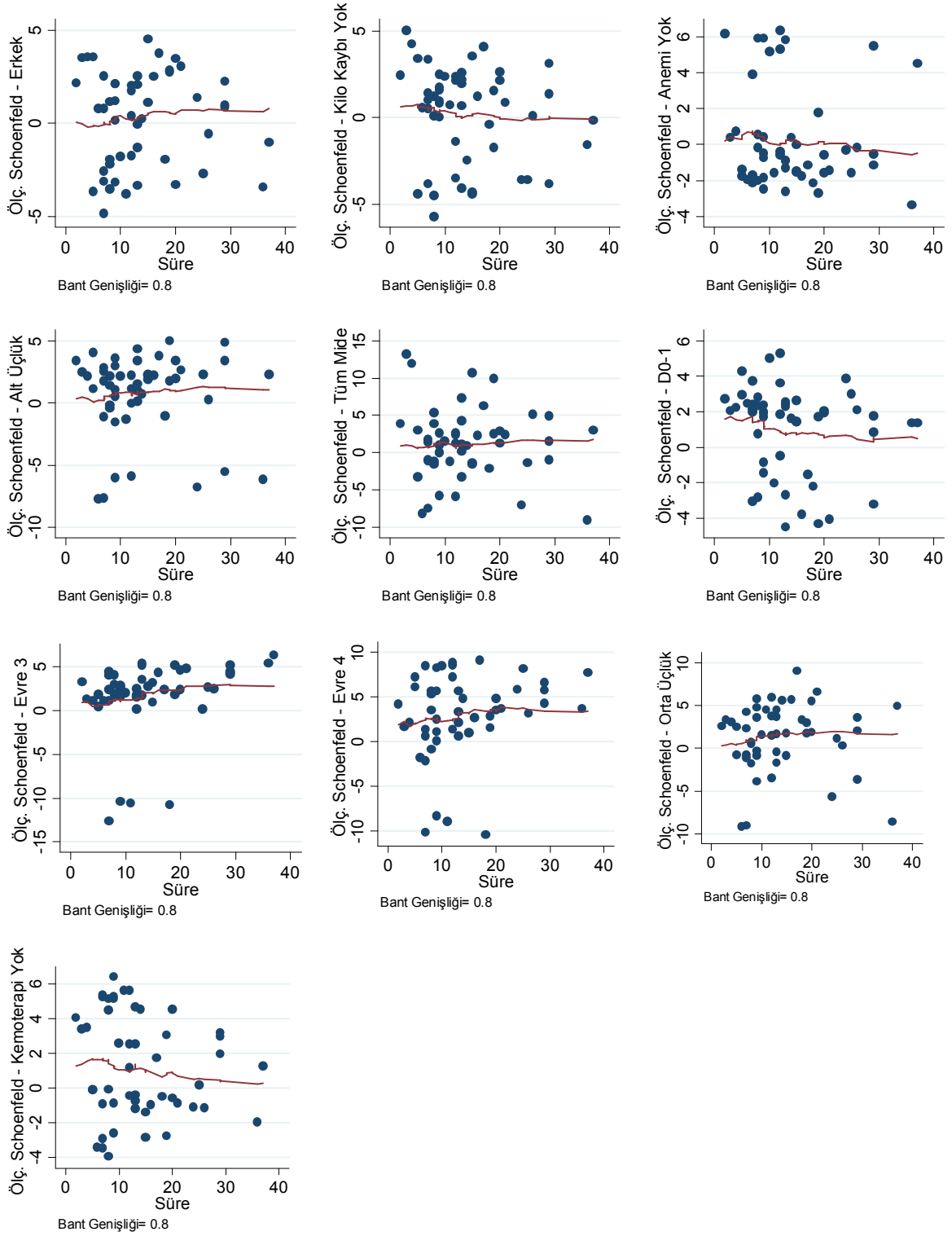
Orantılı tehlikeler varsayımını incelemek için Schoenfeld artıkları da kullanılabilir. Schoenfeld artıkları, açıklayıcı değişkenin gerçek değeri ile ağırlıklı risk skorlarının ortalaması arasındaki farktır.

Süreye karşı çizdirilen grafikte eğri sıfır etrafında yaklaşık olarak doğrusal ise varsayımın sağlandığı sonucuna ulaşılır.



Şekil 4.1. Yaş değişkeni için ölçeklendirilmiş Schoenfeld artığı grafiği

Şekil 4.1 incelendiğinde yaş değişkeni için süreye karşı çizdirilen ölçeklendirilmiş Schoenfeld artıkları eğrisinin sıfır yakınında yaklaşık olarak doğru olduğu dolayısıyla yaş değişkeninin orantılı tehlikeler varsayımını sağladığı açıkça görülmektedir.



Şekil 4.2. Diğer açıklayıcı değişkenler için ölçeklendirilmiş Schoenfeld artığı grafikleri

Şekil 4.2 incelendiğinde neredeyse tüm grafiklerde eğrinin yaklaşık olarak doğrusal olduğu yani değişkenlerin orantılı tehlikeler varsayımını sağladığı görülmektedir.

Çizelge 4.3. Yaşam süresi rankının Schoenfeld artıkları ile ilişkisi

Değişken	p-değeri
Yaş	0,21
Cinsiyet- Erkek	0,6
Kilo kaybı- Yok	0,21
Anemi- Yok	0,55
Diseksiyon- D0-1	0,09
Lokalizasyon- Orta Üçlük	0,37
Lokalizasyon- Alt Üçlük	0,79
Lokalizasyon- Tüm mide	0,81
Evre- Evre3	0,08
Evre- Evre4	0,19
Kemoterapi- Yok	0,28

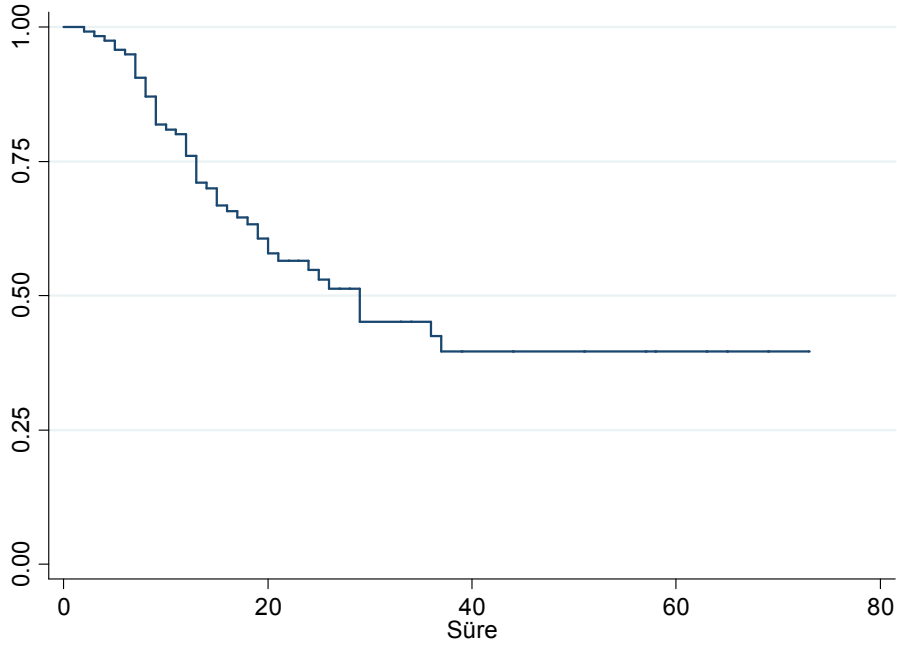
Orantılı tehlikeler varsayımını incelenmek için kullanılan diğer bir yöntem olan, yaşam süresi rankının Schoenfeld artıkları ile ilişkisi de incelenmiştir. Çizelge 4.3 incelendiğinde tüm değişken düzeylerinde p-değeri>0.05 olduğu için orantılı tehlikeler varsayımının sağlandığı görülmüştür.

Çizelge 4.4'te Cox orantılı tehlikeler modelinin sonuçları verilmiştir.

Çizelge 4.4. Cox orantılı tehlikeler modelinin sonuçları

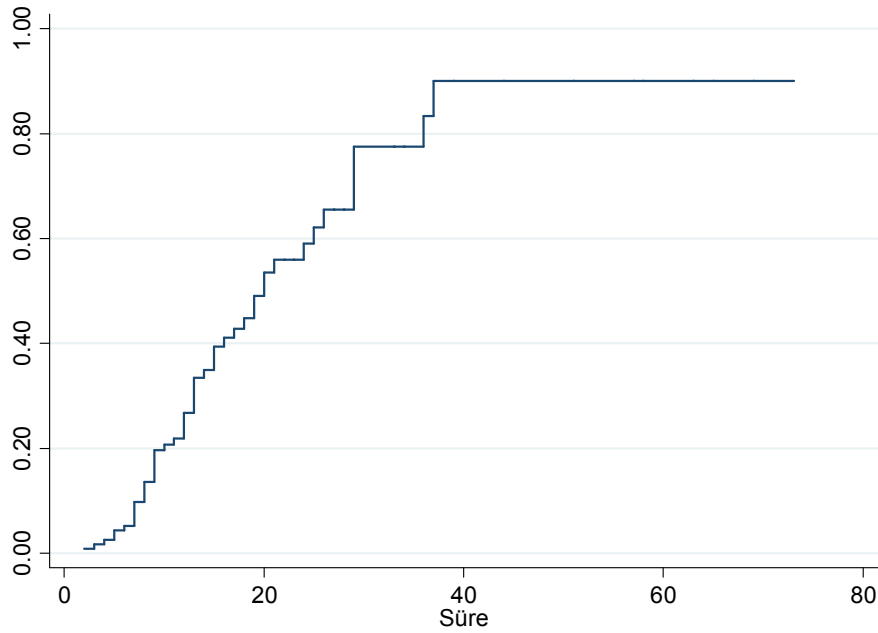
Değişken	Tahmin	Standart Hata	p-değeri	Tehlike Oranı
Yaş	-0,025	0,014	0,080	0,976
Cinsiyet- Erkek	0,301	0,335	0,369	1,351
Kilo kaybı- Yok	0,263	0,352	0,456	1,300
Anemi- Yok	0,177	0,380	0,642	1,193
Diseksiyon- D0-1	1,043	0,351	0,003*	2,837
Lokalizasyon- Orta Üçlük	1,238	0,506	0,140	3,450
Lokalizasyon- Alt Üçlük	0,731	0,442	0,114	2,078
Lokalizasyon- Tüm mide	1,172	0,601	0,051	3,230
Evre- Evre3	1,774	0,559	0,001*	5,897
Evre- Evre4	2,891	0,636	0,000*	18,007
Kemoterapi- Yok	1,070	0,339	0,002*	2,915
-2log(L)	385,364	(p=11)		

*p-değeri< 0.05 olduğundan anlamlıdır.



Şekil 4.3. Kaplan-Meier yaşam eğrisi

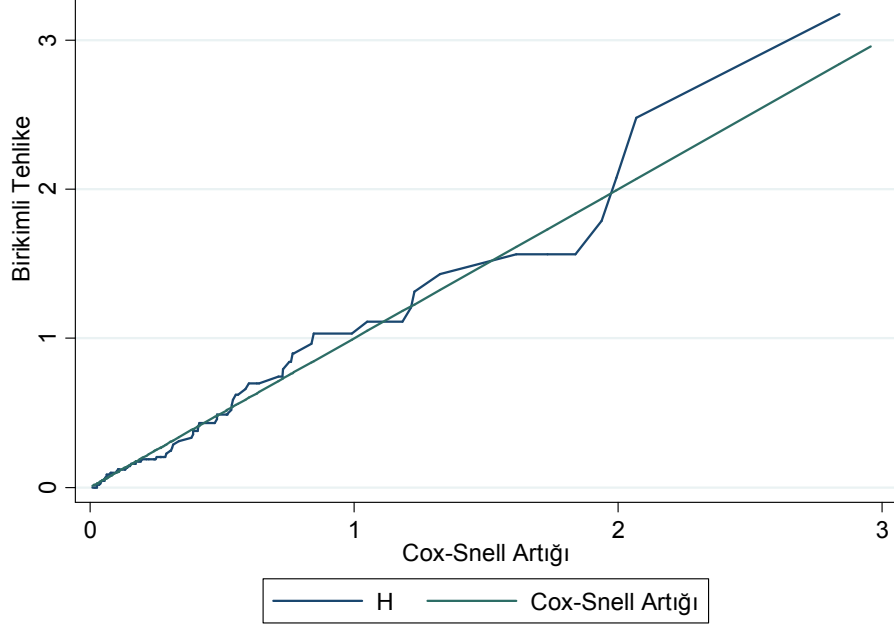
Şekil 4.3 incelendiğinde, süre arttıkça yaşam olasılıklarının düştüğü yaklaşık 40. aydan sonra ise sabitlendiği görülmektedir.



Şekil 4.4. Birikimli tehlike eğrisi

Şekil 4.4'teki birikimli tehlike eğrisi incelendiğinde süre arttıkça başarısızlık riskinin arttığı 40. aydan sonra ise riskin sabitlendiği görülmektedir.

Cox-Snell artıkları modele uyumu araştırmak için kullanışlıdır. Cox orantılı tehlikeler modeli için Cox-Snell grafiği Şekil 4.5'te verilmiştir.

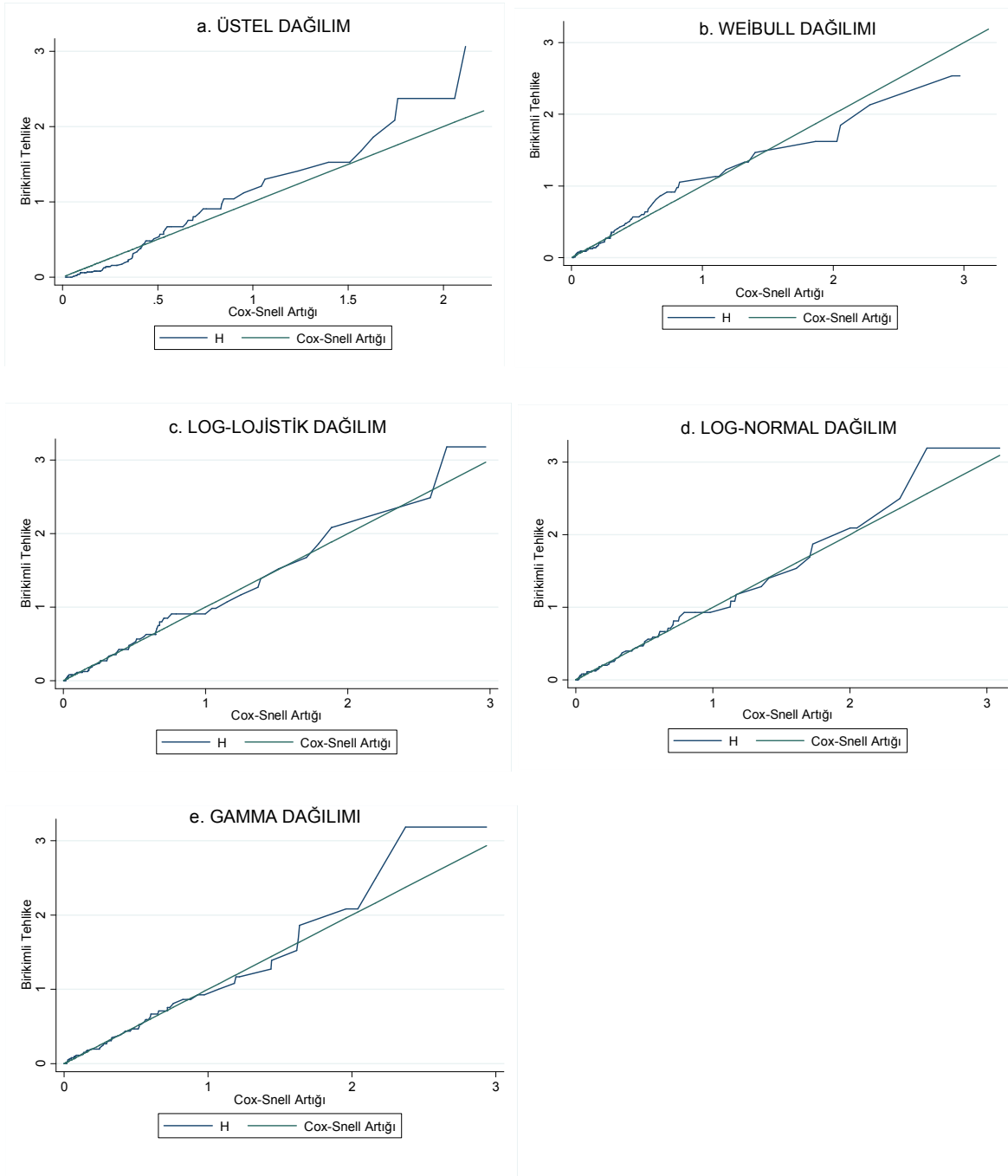


Şekil 4.5. Cox orantılı tehlikeler modeli için Cox-Snell artığı grafiği

Cox orantılı tehlikeler modeli uygun ise birikimli tehlikeye karşı Cox-Snell artığı grafiği yaklaşık olarak bir eğimli bir doğru olur.

Şekil 4.5 incelendiğinde Cox orantılı tehlikeler modelinin veriye uygun olduğunu göstermektedir. Veri orantılı tehlikeler varsayımını sağladığından bu beklenen bir durumdur.

Verilerin dağılımı Easyfit programı ile incelendiğinde bilinen bir dağılıma uygunluk göstermese de parametrik bir dağılım gösterdiği görülmüştür. Bu nedenle, parametrik yaşam çözümlemesi dağılımlarının uygunluğu da Cox-Snell artıkları grafikleri ile incelenmiştir.



Şekil 4.6. a-e. Parametrik modeller için Cox-Snell artığı grafikleri

Şekil 4.6 incelendiğinde log-lojistik modelin ve log-normal modelin daha iyi sonuç verdiği görülmektedir. Bu durumda Akaike bilgi kriterine (AIC) göre karar verilmelidir. Modellere ait AIC değerleri Çizelge 4.5'te verilmiştir.

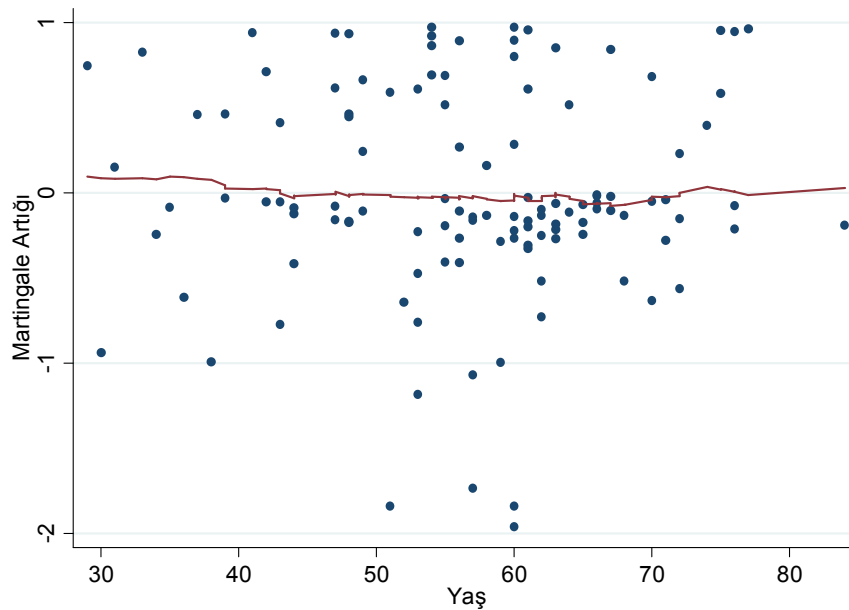
Çizelge 4.5. Modellerin karşılaştırılması

Model	-2log(L)	AIC
Cox Orantılı Tehlikeler	385.364	418.364
Üstel	190.56	214.56
Weibull	174.20	200.20
Log-normal	170.98	196.98
Log-lojistik	172.26	198.26
Genelleştirilmiş Gamma	170.84	198.84

Çizelge 4.5 incelendiğinde AIC kriterlerine göre parametrik yöntemlerden log-normal modelin veriye en uygun olduğu sonucuna ulaşılmıştır.

Martingale artıkları modele dahil edilen açıklayıcı değişkenlerin fonksiyonel formunu belirlemede kullanışlıdır. Bazen orantılı tehlikeler modelinin varsayımlarının sağlanması için değişkenlerin dönüşüme ihtiyacı olabilir. En uygun biçimi belirlemek için model kurulduktan sonra Martingale artıklarına karşı değişkenlerin istenilen dönüşümlerinin grafiği çizdirilir. Eğer dönüşüm uygun ise grafikteki eğri yaklaşık olarak doğrusal olur.

Yaş değişkeni için Martingale artığı grafiği Şekil 4.7'de verilmiştir.

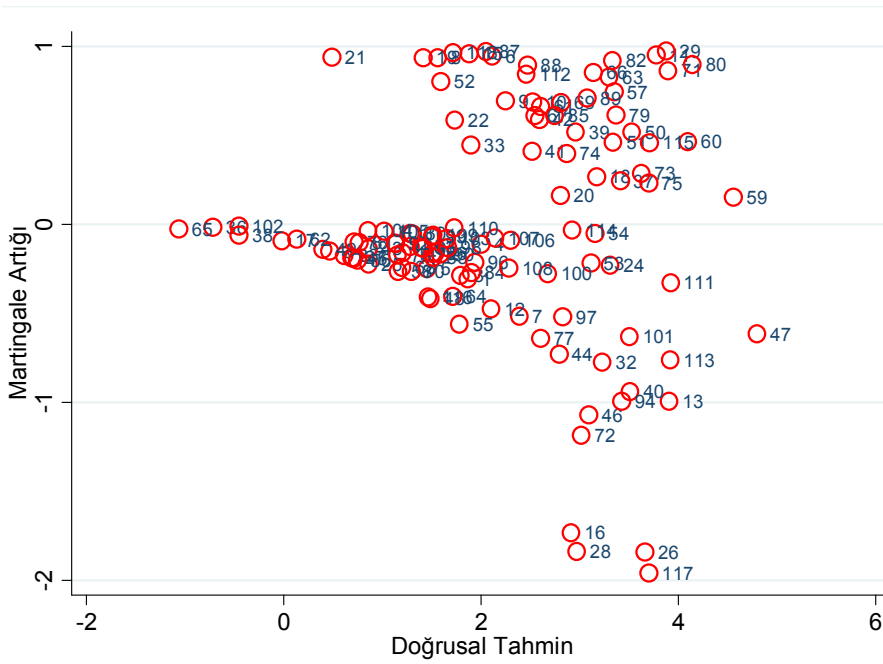


Şekil 4.7. Yaş değişkeni için Martingale artığı grafiği

Şekil 4.7 incelendiğinde eğrinin yaklaşık olarak doğrusal olduğu görülmektedir. Burada yaş değişkenine hiçbir dönüşüm uygulanmamış ve bu hali ile modele uygun olduğu görülmüştür. Eğer eğri doğrusal elde edilmezse değişik dönüşümler uygulanıp bu dönüşümlere karşı Martingale artıkları grafikleri eğri yaklaşık olarak doğrusal olana dek dönüşüm yapılır.

Martingale artıkları zaman boyunca gözlenen değerler ile modelden tahmin edilen değerler arasındaki farklar olarak da tanımlanabilir. Bu durumda Martingale artıklarına karşı doğrusal tahmin grafiği aykırı değerleri belirlemede kullanılabilir.

Doğrusal tahmine karşı çizdirilen Martingale artığı grafiği Şekil 4.8’de verilmiştir.

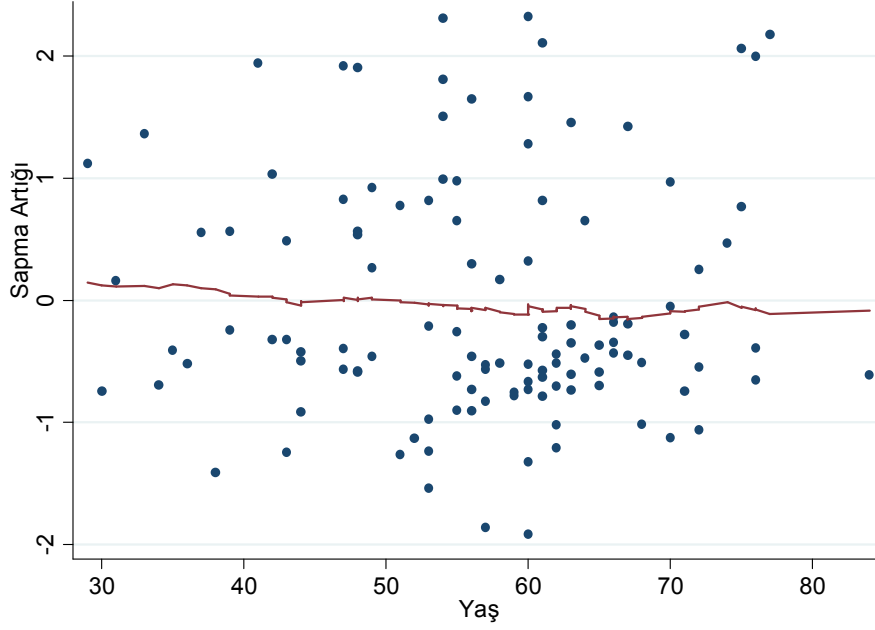


Şekil 4.8. Doğrusal tahmine karşı Martingale artığı grafiği

Şekil 4.8 incelendiğinde 21.,16., 28., 26. ve 117. gözlemlerin aykırı değer olabileceği görülmektedir.

Martingale artıklarının simetrik dağılmaması nedeniyle grafiklerinin yorumlanması zordur. Sapma artıkları Martingale artıklarının dönüşümüdür böylece sıfır etrafında simetrik dağılır. Doğrusal regresyondaki artıklar gibidir. Bu nedenle aykırı değerleri belirlerken ve modelin uyumunu test ederken sapma artıkları tercih edilir.

Yaş değişkeni için sapma artığı grafiği Şekil 4.9’da verilmiştir.



Şekil 4.9. Yaş değişkeni için sapma artığı grafiği

Şekil 4.9 incelendiğinde eğrinin yaklaşık olarak doğrusal olduğu görülmektedir.

Martingale ve sapma artıklarının doğrusal tahmine, yaşam süresine, yaşam sürelerinin sıralamasına ya da gözlem sayısına karşı grafikleri aykırı gözlemlerin belirlenmesinde ve modele uyumun sağlanmasında kullanışlıdır.

Doğrusal tahmine karşı çizdirilen sapma artığı grafiği Şekil 4.10'da verilmiştir.



Şekil 4.10. Doğrusal tahmine karşı sapma artığı grafiği

Şekil 4.10 incelendiğinde 6., 11., 15., 29., 87. ve 118. gözlemlerin aykırı değer olabileceği görülmektedir. Hesaplanan sapma artık değeri Ek 1’de yer almaktadır. Buradan sapma artıklarının referans dağılımı olan normal dağılımın kesim noktalarına göre karşılaştırılma yapıldığında da aynı sonuca ulaşılır.

Martingale ve sapma artıkları grafiklerinin ikisi de aynı bilgiyi vermesine karşın simetrik dağılmasından dolayı artıklarının yorumlanması daha kolaydır. Başarısız olan gözlemlerde uyum yeterli ise sapma artıkları ak gürültüye benzer biçimde dağılır. Durdurulmuş gözlemler için ise sapma artıkları sıfır yakınlarında küme olarak yer alır.

Aykırı değerleri belirlerken önerilen diğer artık türleri log-odds ve normal sapma artıklarıdır. Tüm değişkenler için hesaplanan log-odds ve normal sapma artıkları Ek 2’de verilmiştir. Hesaplanan değerler kritik değerlerle karşılaştırılır. Log-odds artıkları için $\alpha = 0,025$ düzeyinde kesim noktaları +3,66 ve -3,66’dır. Normal sapma artıkları için ise $\alpha = 0,025$ kesim noktaları +1,96 ve -1,96’dır.

Bu değerlerin dışında kalan gözlemler aykırı değer olarak belirlenir. Buna göre bulunan aykırı değerler Çizelge 4.6’da verilmiştir.

Çizelge 4.6. Log-odds ve normal sapma artık değerleri

Gözlem	Süre	Durum	$\widehat{S}_i(t_i)$	$\widehat{S}_i(t_i^c)$	l_i^c	l_i^m	n_i^c	n_i^m
11	3	Ölmüş	0,9831	-	1,7634	1,7634	2,1213	2,1213
29	2	Ölmüş	0,9915	-	2,0682	2,0682	2,3878	2,3878
55	39	Yaşıyor	-	0,1982	-0,6070	0,6541	-0,8482	-2,8845
56	37	Yaşıyor	-	0,1982	-0,6070	-0,0850	-0,8482	-2,8845
58	36	Yaşıyor	-	0,2123	-0,5693	-1,1083	-0,7984	-2,5840
62	34	Yaşıyor	-	0,2256	-0,5356	-1,0776	-0,7534	-2,3487
64	33	Yaşıyor	-	0,2256	-0,5356	-1,0776	-0,7534	-2,3487
65	33	Yaşıyor	-	0,2256	-0,5356	-1,0776	-0,7534	-2,3487
68	29	Yaşıyor	-	0,2256	-0,5356	-1,0776	-0,7534	-2,3487
70	29	Yaşıyor	-	0,2256	-0,5356	-1,0776	-0,7534	-2,3487
80	4	Ölmüş	0,9746	-	1,5836	1,5836	1,9528	1,9528
116	73	Yaşıyor	-	0,1982	-0,6070	-1,1426	-0,8482	-2,8845

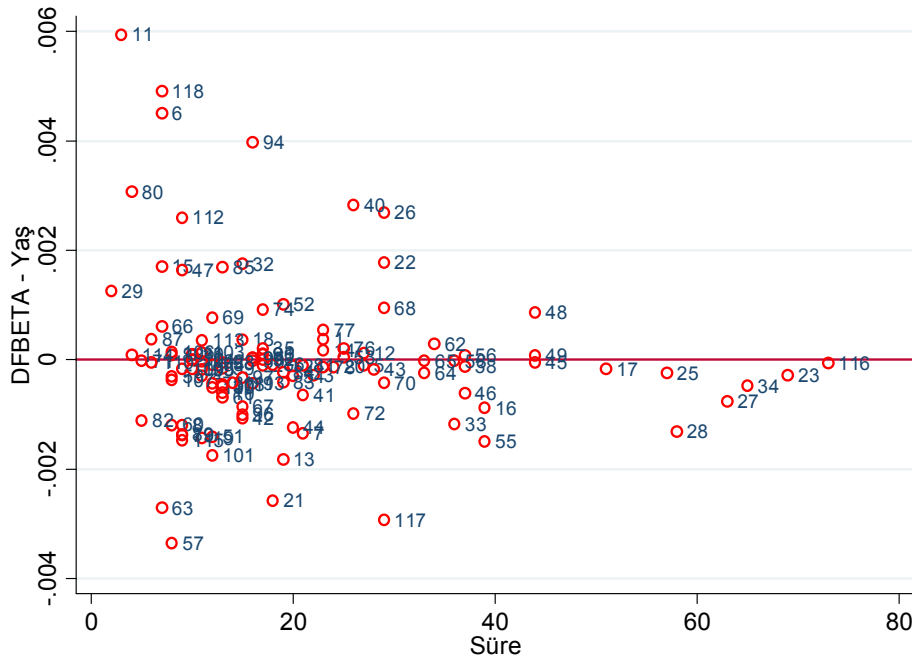
Çizelge 4.6’daki $\widehat{S}_i(t_i)$, yaşam olasılığını, $\widehat{S}_i(t_i^c)$ durdurulmuş gözlemler için hesaplanan yaşam olasılığını göstermek üzere l_i^c , log-odds artık değerleri ve l_i^m , beklenen ortalama değerlerin sapması kullanıldığında hesaplanan log-odds artık değerlerdir. Benzer şekilde n_i^c , normal sapma artık değerleri ve n_i^m , beklenen

ortalama değerlerin sapması kullanıldığında hesaplanan normal sapma artık değerleridir.

Çizelge 4.6 incelendiğinde log-odds artık değerlerine göre aykırı değer bulunmadığı ve normal sapma artık değerlerine göre ise 11., 29., 55., 56., 58., 62., 64., 65., 68., 70., 80. ve 116. gözlemlerin aykırı değer olabileceği görülür.

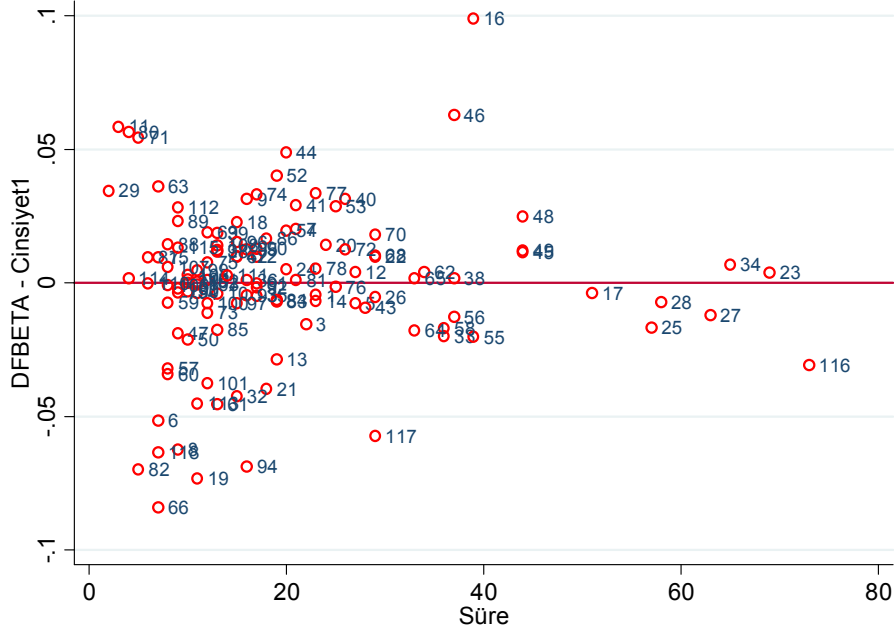
Modelin yeterliliğini değerlendirirken tahmin edilen parametre üzerinde aşırı etkiye sahip gözlemler de belirlenmelidir. Bu analiz, etki (influence or leverage) analizi olarak bilinir.

Etkili gözlemleri test etmek için en çok tercih edilen yöntem DFBETA yöntemidir. Süreye ya da gözlem sayılarına karşı çizilen grafiklerle aşırı etkiye sahip gözlemler belirlenebilir. Daha çok bilgi elde edebilmek için grafikler süreye karşı ve gözlem numaralarıyla her değişken için çizdirilmiştir. Değişkenler ve düzeyleri için çizilen DFBETA grafikleri Şekil 4.11 ve Şekil 4.21 arasında verilmiştir.



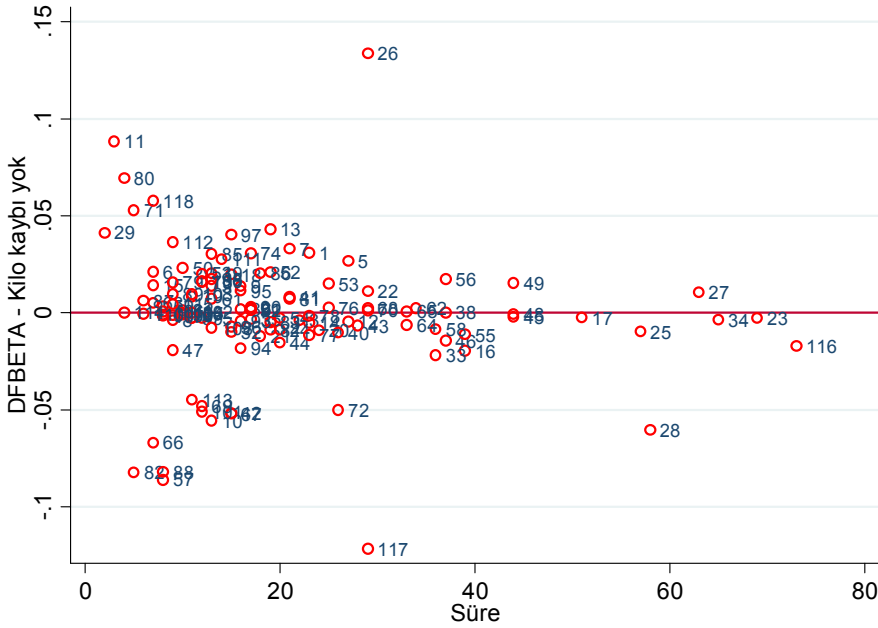
Şekil 4.11. Yaş değişkeni için DFBETA grafiği

Şekil 4.11 incelendiğinde örneğin, 11. gözlem çıkarılırsa yaş değişkeni katsayısını yaklaşık olarak 0,006 azaltacağı ya da yaş değişkeni için tehlike oranını yaklaşık olarak $\exp(-0,006)=0,994$ azaltacağı görülmektedir.



Şekil 4.12. Cinsiyet değişkeni için DFBETA grafiği

Şekil 4.12 incelendiğinde örneğin, 16. gözlem çıkarılırsa cinsiyet katsayısını yaklaşık olarak 0,1 azaltacağı ya da cinsiyet değişkeni için tehlike oranını yaklaşık olarak $\exp(-0,1)=0,9048$ azaltacağı görülmektedir.



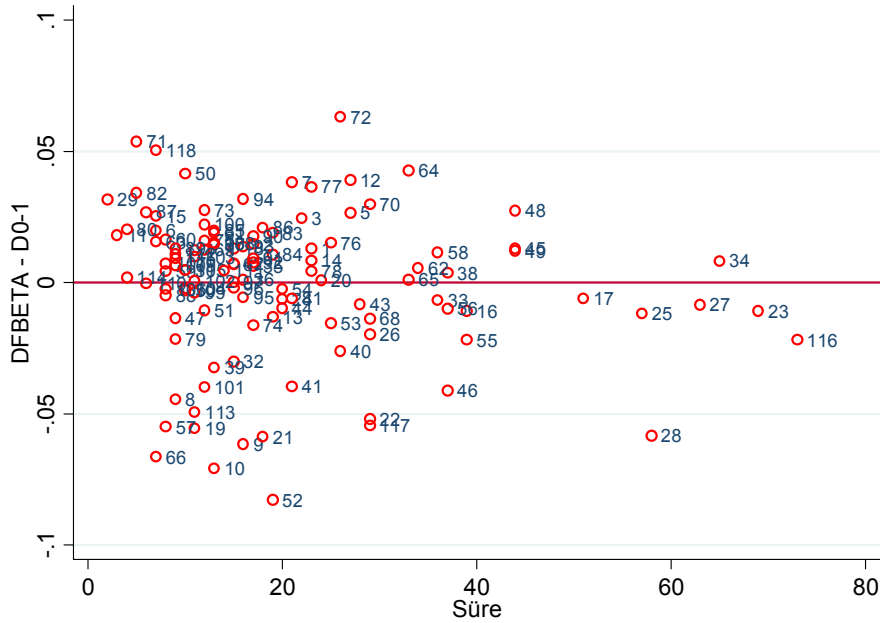
Şekil 4.13. Kilo Kaybı (Semptom) değişkeni için DFBETA grafiği

Şekil 4.13 incelendiğinde örneğin, 26. gözlem çıkarılırsa kilo kaybı değişkeninin katsayısını yaklaşık olarak 0,13 azaltacağı ya da kilo kaybı değişkeni için tehlike oranını yaklaşık olarak $\exp(-0,13)=0,8780$ azaltacağı görülmektedir.



Şekil 4.14. Anemi (Hb) değişkeni için DFBETA grafiği

Şekil 4.14 incelendiğinde örneğin, 29. gözlem çıkarılırsa anemi değişkeninin katsayısını yaklaşık olarak 0,12 azaltacağı ya da anemi değişkeni için tehlike oranını yaklaşık olarak $\exp(-0,12)=0,8869$ azaltacağı görülmektedir. Aynı şekilde 8. gözlem çıkarılırsa anemi değişkeninin katsayısını yaklaşık olarak 0,11 azaltacağı ya da anemi değişkeni için tehlike oranını yaklaşık olarak $\exp(-0,11)=0,8958$ azaltacağı görülmektedir.



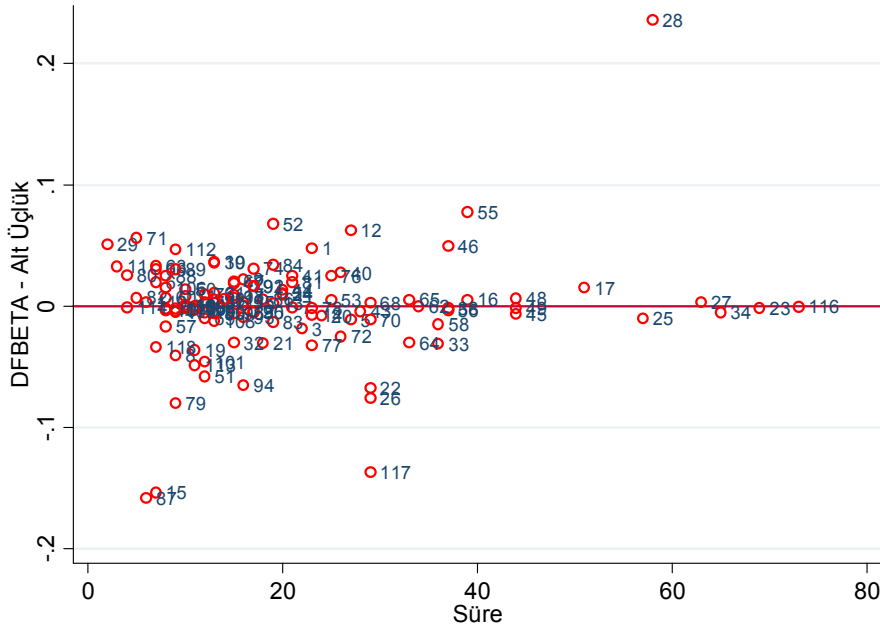
Şekil 4.15. Diseksiyon değişkeni için DFBETA grafiği

Şekil 4.15 incelendiğinde örneğin, 72. gözlem çıkarılırsa diseksiyon değişkeninin katsayısını yaklaşık olarak 0,06 azaltacağı ya da diseksiyon değişkeni için tehlike oranını yaklaşık olarak $\exp(-0,06)=0,9417$ azaltacağı görülmektedir.



Şekil 4.16. Tümörün midedeki lokalizasyonu değişkeninin orta üçlük düzeyi için DFBETA grafiği

Şekil 4.16 incelendiğinde örneğin, 28. gözlem çıkarılırsa lokalizasyon değişkeninin orta üçlük düzeyinin katsayısını yaklaşık olarak 0,24 azaltacağı ya da lokalizasyon değişkeninin orta üçlük düzeyi için tehlike oranını yaklaşık olarak $\exp(-0,24)= 0,7866$ azaltacağı görülmektedir.



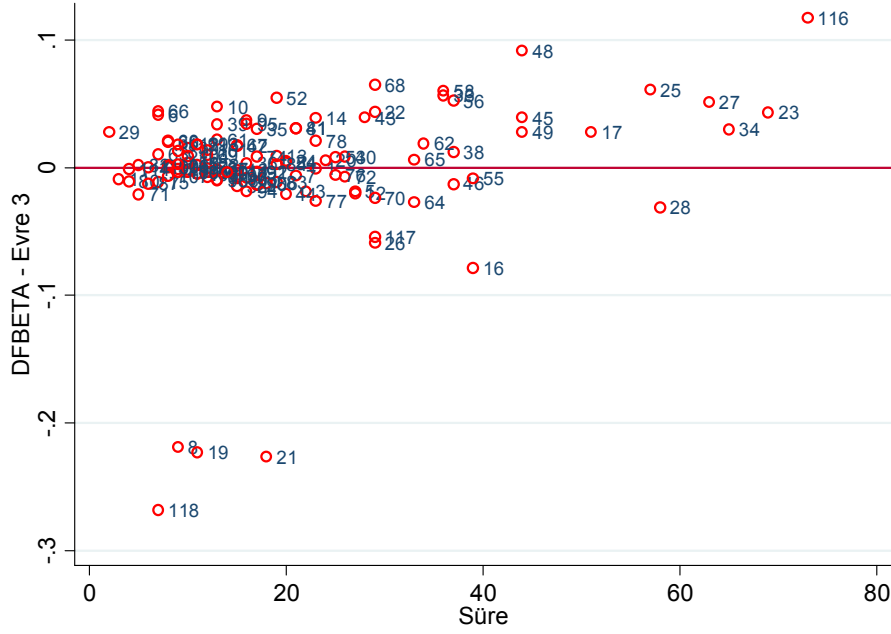
Şekil 4.17. Tümörün midedeki lokalizasyonu değişkeninin alt üçlük düzeyi için DFBETA grafiği

Şekil 4.17 incelendiğinde örneğin, 28. gözlem çıkarılırsa lokal değişkeninin alt üçlük düzeyinin katsayısını yaklaşık olarak 0,24 azaltacağı ya da lokal değişkeninin alt üçlük düzeyi için tehlike oranını yaklaşık olarak $\exp(-0,24)=0,7866$ azaltacağı görülmektedir.



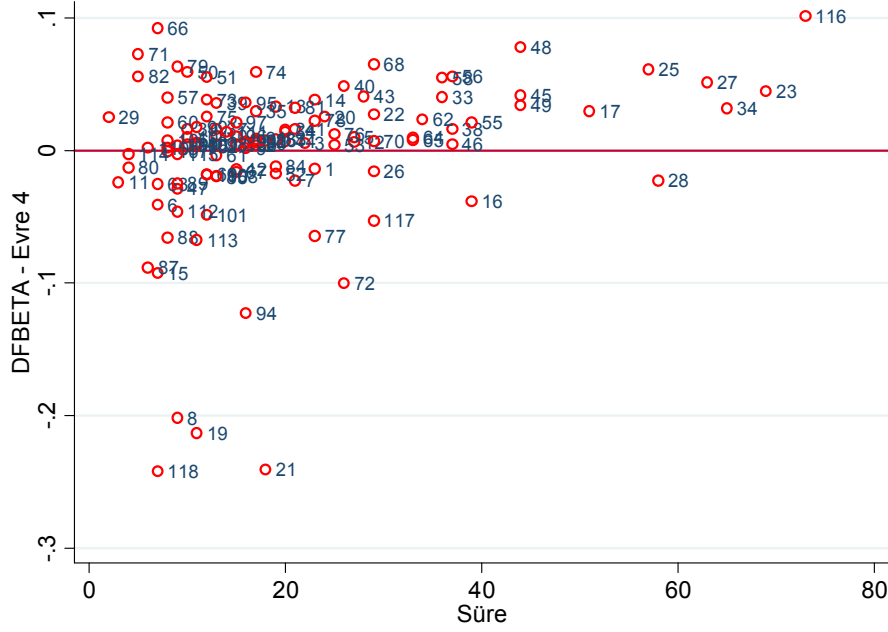
Şekil 4.18. Tümörün midedeki lokalizasyonu değişkeninin tüm mide düzeyi için DFBETA grafiği

Şekil 4.18 incelendiğinde örneğin, 28. gözlem çıkarılırsa lokal değişkeninin tüm mide düzeyinin katsayısını yaklaşık olarak 0,24 azaltacağı ya da lokal değişkeninin tüm mide düzeyi için tehlike oranını yaklaşık olarak $\exp(-0,24)=0,7866$ azaltacağı görülmektedir.



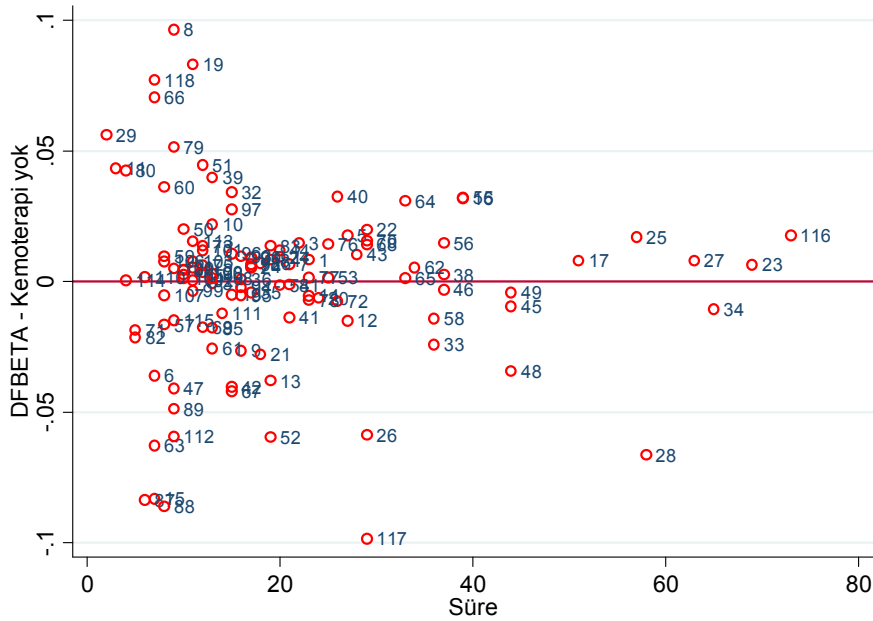
Şekil 4.19. Evre değişkeninin evre 3 düzeyi için DFBETA grafiği

Şekil 4.19 incelendiğinde örneğin, 116. gözlem çıkarılırsa evre değişkeninin evre 3 düzeyinin katsayısını yaklaşık olarak 0,12 azaltacağı ya da evre değişkeninin evre 3 düzeyi için tehlike oranını yaklaşık olarak $\exp(-0,12)=0,8869$ azaltacağı görülmektedir.



Şekil 4.20. Evre değişkeninin evre 4 düzeyi için DFBETA grafiği

Şekil 4.20 incelendiğinde örneğin, 116. gözlem çıkarılırsa evre değişkeninin evre 4 düzeyinin katsayısını yaklaşık olarak 0,1 azaltacağı ya da evre değişkeninin evre 4 düzeyi için tehlike oranını yaklaşık olarak $\exp(-0,1)=0,9048$ azaltacağı görülmektedir.



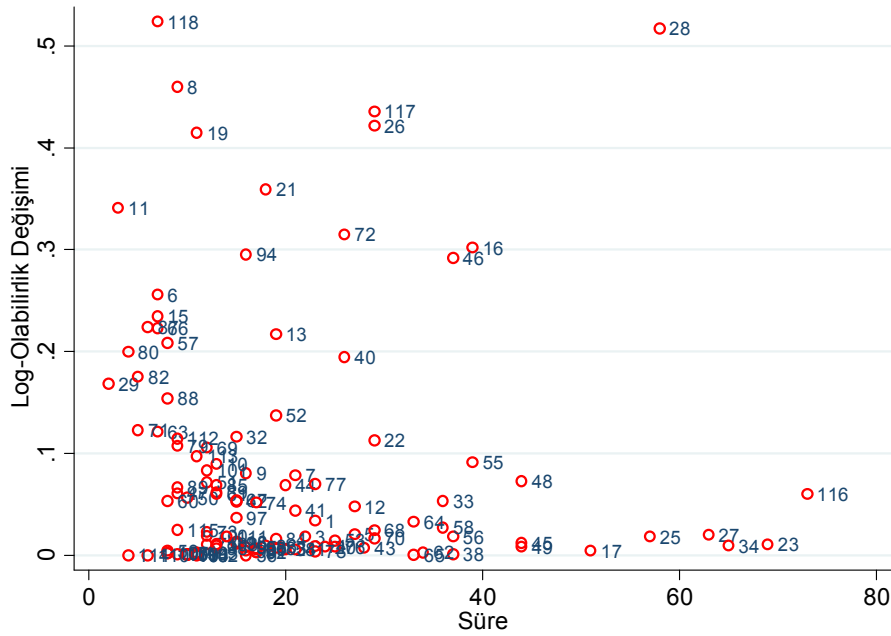
Şekil 4.21. Kemoterapi değişkeni için DFBETA grafiği

Şekil 4.21 incelendiğinde örneğin, 8. gözlem çıkarılırsa kemoterapi değişkeninin katsayısını yaklaşık olarak 0,1 azaltacağı ya da kemoterapi değişkeni için tehlike oranını yaklaşık olarak $\exp(-0,1)=0,9048$ azaltacağı görülmektedir.

DFBETA değerlerinin tek dezavantajı açıklayıcı değişkenlerin sayısı ve örneklem büyüklüğü artıkça test edilecek değerlerin sayısı da artar. Her bir gözlem ve her bir açıklayıcı değişken düzeyi için elde edilen DFBETA değerleri Ek 3'te yer almaktadır.

Etkili gözlemlerin belirlenmesinde iki alternatif yöntem LMAX ve olabilirlik değişim değerleridir. Her iki yöntem de gözlemlerin etkisini bütün olarak katsayılar vektöründen ölçer. Böylece çok değişkene karşı sadece tek bir değer elde edilir. Her gözlem için elde edilen LMAX ve olabilirlik değişim değerleri Ek 4'te yer almaktadır.

Süreye karşı çizilen olabilirlik değişim değerlerinin grafiği Şekil 4.22'de verilmiştir.

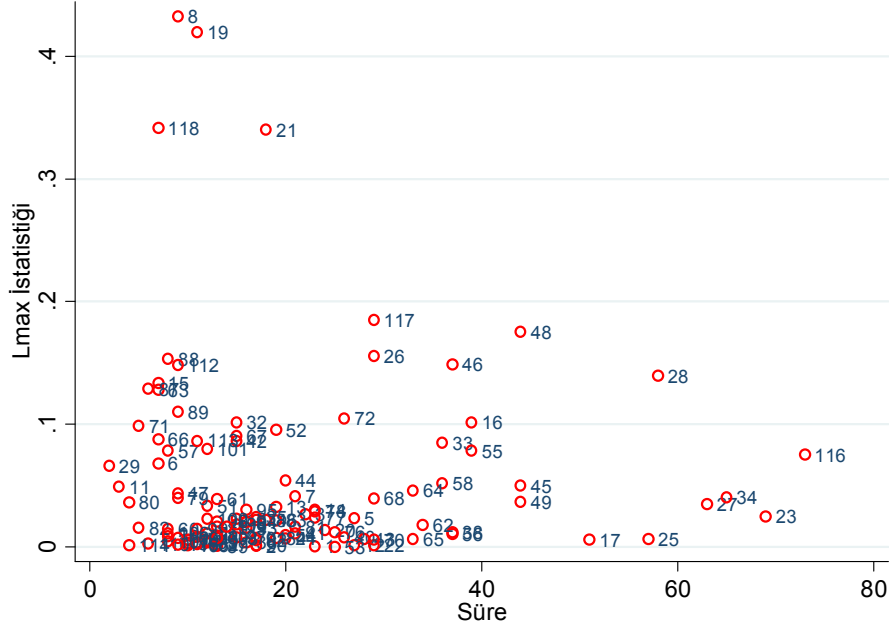


Şekil 4.22. Log-olabilirlik değişim grafiği

Şekil 4.22 incelendiğinde 118., 28., 8., 19.,117., 21. ve 26. gözlemlerin etkili gözlemler olabileceği görülür.

Olabilirlik değişim ve LMAX değerlerinin ikisi de her gözlemin genel etkisini ölçer ama direk olarak birbirleriyle karşılaştırılmazlar. Her ikisi de kendi içinde karşılaştırılmalıdır.

Süreye karşı çizilen LMAX değerlerinin grafiği Şekil 4.23'de verilmiştir.



Şekil 4.23. Lmax istatistiği grafiği

Şekil 4.23 incelendiğinde 8., 19., 118. ve 21. gözlemlerin etkili gözlemler olabileceği görülür.

Log olabilirlik değişimi grafiğinde ayrıca 117., 26. ve 28. gözlemler de etkili gözlemler olabilir sonucuna varılmıştı. Bu halde karar vermek araştırmacıya bırakılmış bir durumdur.

Belirtilen aykırı değer belirleme yöntemleri uygulandığında değişik artık türleriyle farklı sonuçlar elde edilmiştir. Yöntemlerce ortak çıkan değerleri aykırı değerler olarak belirsek 11. ve 29. gözlemler aykırı değerdir. Bu değerler çalışmadan çıkarılıp Cox orantılı tehlikeler modeli yeniden kurulmuş ve sonuçlar Çizelge 4.7’de verilmiştir.

Çizelge 4.7. Aykırı değerler çıkarıldıktan sonra Cox orantılı tehlikeler modeli sonuçları

Değişken	Tahmin	Standart Hata	p-değeri	Tehlike Oranı
Yaş	-0,328	0,014	0,025*	0,967
Cinsiyet- Erkek	0,199	0,341	0,559	1,220
Kilo kaybı- Yok	0,120	0,357	0,736	1,127
Anemi- Yok	0,052	0,394	0,893	1,054
Diseksiyon- D0-1	0,988	0,352	0,005*	2,686
Lokalizasyon- Orta Üçlük	1,177	0,504	0,020*	3,245
Lokalizasyon- Alt Üçlük	0,644	0,464	0,166	1,904
Lokalizasyon- Tüm mide	0,865	0,621	0,163	2,376
Evre- Evre3	1,759	0,561	0,002*	5,809
Evre- Evre4	2,900	0,640	0,000*	18,184
Kemoterapi- Yok	0,961	0,347	0,006*	2,616
-2log(L)	369,706	(p=11)		

*p-değeri < 0.05 olduğundan anlamlıdır.

Çizelge 4.4 ve Çizelge 4.7 incelendiğinde bazı karşılaştırmalar yapılabilir. Aykırı değerler çıkarıldıktan sonra kurulan Cox orantılı tehlikeler modeli ile bu değerler çıkarılmadan önceki Cox orantılı tehlikeler modeli karşılaştırıldığında genel olarak tahmin değerleri artarken buna karşın tehlike oranı azaldığı görülmüştür. Ayrıca, daha çok değişken anlamlı bulunmuştur. İki model eşit parametre sayısına sahip olduğu için -2log(L) değerleri kullanılarak karşılaştırma yapılabilir. Aykırı değer çıkarılmadan önceki modelin -2log(L) değeri 385,364 iken çıkarıldıktan sonraki modelin -2log(L) değeri Çizelge 4.7'den görüldüğü gibi 369,706'dır. Buna göre aykırı değerlerin çıkarılması sonucunda elde edilen modelin daha uygun bir model olduğu söylenebilir.

Cox-Snell artıkları incelemesinde ve AIC kriterlerine göre parametrik yöntemlerden log-normal modelin veriye uygun olduğu sonucuna ulaşılmıştı. Bu nedenle veriye log-normal regresyon modeli uygulanmış ve elde edilen sonuçlar Çizelge 4.8'de verilmiştir.

Çizelge 4.8. Log-normal regresyon modelinin sonuçları

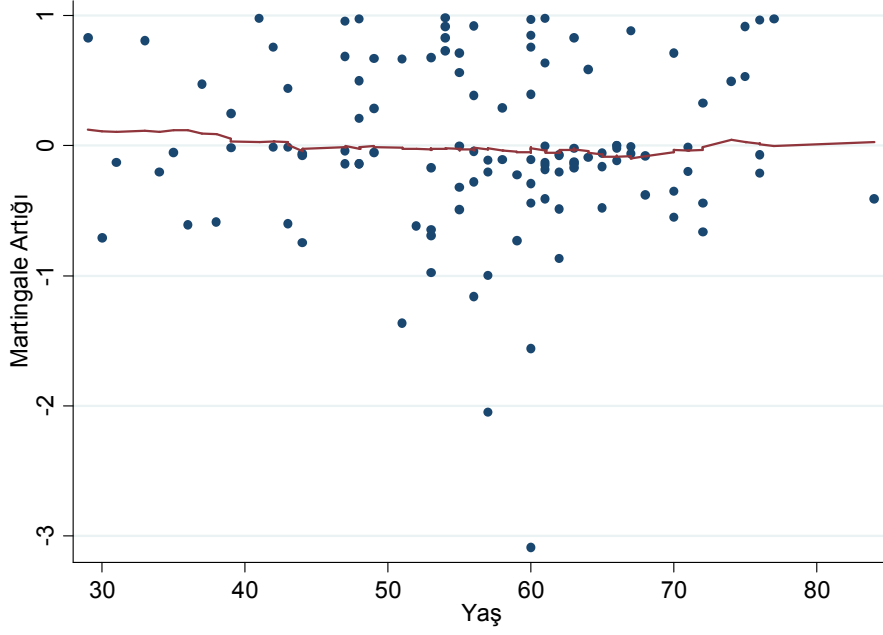
Değişken	Tahmin	Standart Hata	p-değeri
Yaş	0,0139	0,0084	0,096
Cinsiyet- Erkek	-0,106	0,2105	0,616
Kilo kaybı- Yok	-0,212	0,2253	0,347
Anemi- Yok	-0,028	0,2368	0,907
Diseksiyon- D0-1	-0,723	0,2099	0,001*
Lokalizasyon-Orta Üçlük	-0,853	0,3169	0,007*
Lokalizasyon-Alt Üçlük	-0,482	0,2859	0,092
Lokalizasyon-Tüm mide	-0,705	0,3791	0,063
Evre- Evre3	-1,262	0,2951	0,000*
Evre- Evre4	-1,835	0,3457	0,000*
Kemoterapi- Yok	-0,835	0,212	0,000*
Sabit	5,0182	0,6803	0,000*
Lnsigma	-0,23	0,1021	0,024*
Sigma	0,7947	0,0811	
-2log(L)	170,995	(p=11)	

*p-değeri<0.05 olduğundan anlamlıdır.

Log-normal model için Martingale, sapma, log-odds ve normal sapma artık değerleri hesaplanarak aykırı değerleri tespit etmek amaçlanmıştır.

Martingale artıkları modele dahil edilen açıklayıcı değişkenlerin fonksiyonel formunu belirlemede kullanışlıdır. Eğer değişken modele uygun ise grafikteki eğri yaklaşık olarak doğrusal olur.

Log-normal regresyon modelinde yaş değişkeni için Martingale artığı grafiği Şekil 4.24'te verilmiştir.



Şekil 4.24. Log-normal regresyon modelinde yaş değişkeni için Martingale artığı grafiği

Şekil 4.24 incelendiğinde eğrinin yaklaşık olarak doğru olduğu yani yaş değişkenin modele uygun olduğu, bir dönüşüme ihtiyaç olmadığı görülür.

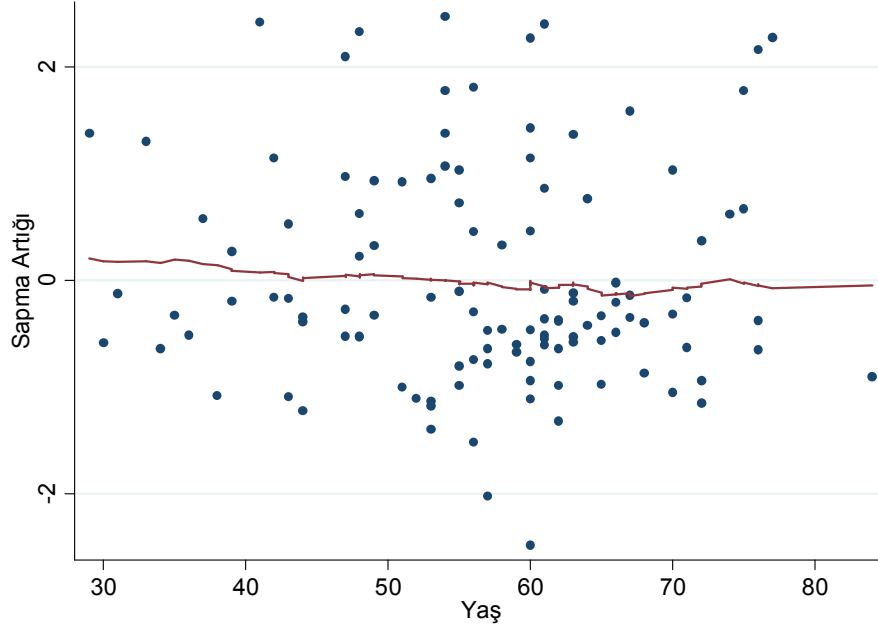
Martingale artıklarına karşı doğrusal tahmin grafiği aykırı değerleri belirlemede kullanılabilir. Log-normal regresyon modelinde doğrusal tahine karşı Martingale artığı grafiği Şekil 4.25'te verilmiştir.



Şekil 4.25. Log-normal regresyon modelinde doğrusal tahine karşı Martingale artığı grafiği

Şekil 4.25 incelendiğinde 16., 21., 29. ve 28. gözlemlerin aykırı değerler olabilecekleri görülür.

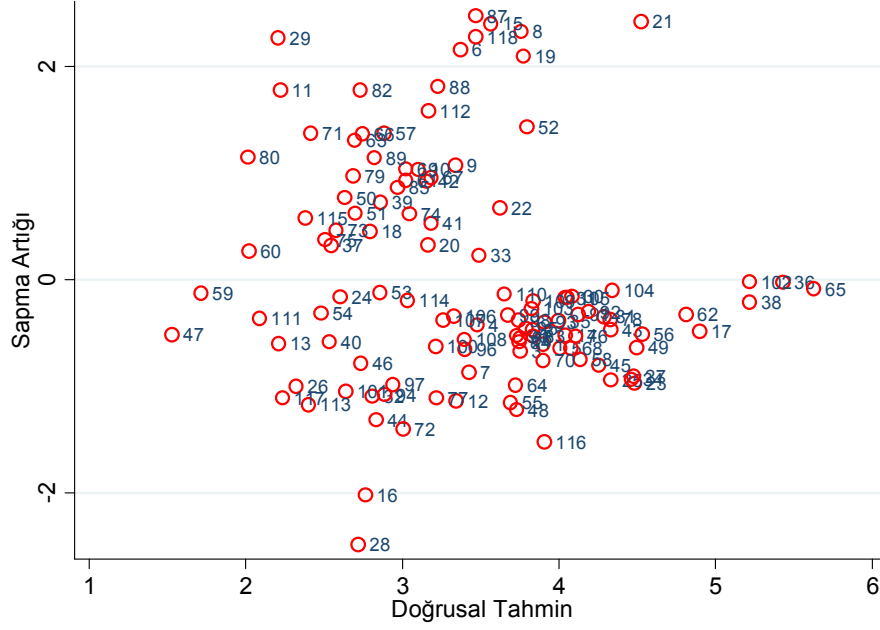
Simetrik bir dağılım göstermesinden dolayı sapma artıklarının aykırı değerleri belirlemede kullanımı daha yaygındır. Log-normal regresyon modelinde yaş değişkeni için sapma artığı grafiği Şekil 4.26'da verilmiştir.



Şekil 4.26. Log-normal regresyon modelinde yaş değişkeni için sapma artığı grafiği

Şekil 4.26 incelendiğinde eğrinin yaklaşık olarak doğrusal olduğu görülmektedir. sapma artıklarının doğrusal tahmine karşı grafikleri aykırı gözlemlerin belirlenmesinde ve modele uyumun sağlanmasında kullanışlıdır.

Log-normal modelde doğrusal tahmine karşı sapma artığı grafiği Şekil 4.27'de verilmiştir.



Şekil 4.27. Log-normal regresyon modelinde doğrusal tahmine karşı sapma artığı grafiği

Şekil 4.27 incelendiğinde 6., 8., 15., 16., 19., 21., 28., 29., 87., 118. ve 128. gözlemlerin aykırı değerler olabileceği görülür. Log-normal regresyon modelinde hesaplanan sapma artık değeri Ek 5'te yer almaktadır. Buradan sapma artıklarının referans dağılımı olan normal dağılımın kesim noktalarına göre karşılaştırılma yapıldığında aynı sonuca ulaşılar.

Tüm değişkenler için hesaplanan log-odds ve normal sapma artıkları Ek 6'da verilmiştir. Hesaplanan değerler kesim noktalarıyla karşılaştırıldığında bu değerlerin dışında kalan gözlemler aykırı değer olarak belirlenir. Buna göre bulunan aykırı değerler Çizelge 4.9'da verilmiştir.

Çizelge 4.9. Log-normal regresyon modelinde log-odds ve normal sapma artık değerleri

Gözlem	Süre	Durum	$\widehat{S}_i(t_i)$	$\widehat{S}_i(t_i^c)$	l_i^c	l_i^m	n_i^c	n_i^m
11	3	Ölmüş	0,983		1,763	1,763	2,121	2,121
16	39	Yaşıyor	-	0,1982	-0,607	-1,143	-0,848	-2,884
17	51	Yaşıyor	-	0,1982	-0,607	-1,143	-0,848	-2,884
23	69	Yaşıyor	-	0,1982	-0,607	-1,143	-0,848	-2,884
25	57	Yaşıyor	-	0,1982	-0,607	-1,143	-0,848	-2,884
27	63	Yaşıyor	-	0,1982	-0,607	-1,143	-0,848	-2,884
28	58	Yaşıyor	-	0,1982	-0,607	-1,143	-0,848	-2,884
29	2	Ölmüş	0,991	-	2,068	2,068	2,387	2,387
34	65	Yaşıyor	-	0,1982	-0,607	-1,143	-0,848	-2,884
38	37	Yaşıyor	-	0,1982	-0,607	-1,143	-0,848	-2,884
45	44	Yaşıyor	-	0,1982	-0,607	-1,143	-0,848	-2,884
48	44	Yaşıyor	-	0,1982	-0,607	-1,143	-0,848	-2,884
49	44	Yaşıyor	-	0,1982	-0,607	-1,143	-0,848	-2,884
55	39	Yaşıyor	-	0,1982	-0,607	-1,143	-0,848	-2,884
56	37	Yaşıyor	-	0,1982	-0,607	-1,143	-0,848	-2,884
58	36	Yaşıyor	-	0,2123	-0,569	-1,108	-0,798	-2,584
62	34	Yaşıyor	-	0,2256	-0,536	-1,078	-0,753	-2,349
64	33	Yaşıyor	-	0,2256	-0,536	-1,078	-0,753	-2,349
65	33	Yaşıyor	-	0,2256	-0,536	-1,078	-0,753	-2,349
68	29	Yaşıyor	-	0,2256	-0,536	-1,078	-0,753	-2,349
70	29	Yaşıyor	-	0,2256	-0,536	-1,078	-0,753	-2,349
116	73	Yaşıyor	-	0,1982	-0,607	-1,143	-0,848	-2,884

Çizelge 4.9 incelendiğinde log-odds artık değerlerine göre aykırı değer bulunmadığını ve normal sapma artık değerlerine göre ise 11., 16., 17., 23., 25., 27., 28., 29., 34., 38., 45., 48., 49., 55., 56., 58., 62., 64., 65., 68., 70. ve 80. gözlemlerin aykırı değer olabileceği görülür.

Yine aynı şekilde yöntemlerce ortak çıkan değerleri aykırı değerler olarak belirsek 16., 28. ve 29. gözlemler aykırı değerdir. Bu değerler çalışmadan çıkarılıp log-normal regresyon modeli yeniden kurulmuş ve sonuçlar Çizelge 4.10'da verilmiştir.

Çizelge 4.10. Aykırı değerler çıkarıldıktan sonra log-normal regresyon modeli sonuçları

Değişken	Tahmin	Standart Hata	p-değeri
Yaş	0,0113	0,0075	0,130
Cinsiyet- Erkek	0,0250	0,1948	0,898
Kilo kaybı- Yok	-0,2218	0,2038	0,285
Anemi- Yok	0,1483	0,2217	0,503
Diseksiyon- D0-1	-0,7043	0,1913	0,000*
Lokalizasyon-Orta Üçlük	-0,6854	0,2871	0,017*
Lokalizasyon-Alt Üçlük	-0,2456	0,2618	0,348
Lokalizasyon-Tüm mide	-0,4317	0,3453	0,211
Evre- Evre3	-1,2598	0,2704	0,000*
Evre- Evre4	-1,7469	0,3136	0,000*
Kemoterapi- Yok	-0,7758	0,1965	0,000*
Sabit	4,7470	0,6141	0,000*
Lnsigma	-0,3389	0,1024	0,001*
Sigma	0,7126	0,0730	
-2log(L)	153,102	(p=11)	

*p-değeri < 0.05 olduğundan anlamlıdır.

Çizelge 4.8 ve Çizelge 4.10 incelendiğinde, aykırı değerler çıkarıldıktan sonra kurulan log-normal regresyon modeli ile bu değerler çıkarılmadan önceki log-normal regresyon modeli karşılaştırıldığında standart hataların genel olarak azaldığı görülmüştür. İki model eşit parametre sayısına sahip olduğu için -2log(L) değerleri kullanılarak karşılaştırma yapılabilir. Aykırı değer çıkarılmadan önceki modelin -2log(L) değeri 170,995 iken çıkarıldıktan sonraki modelin -2log(L) değeri Çizelge 4.10'dan görüldüğü gibi 153,102'dir. Buna göre aykırı değerlerin çıkarılması sonucunda elde edilen modelin daha uygun bir model olduğu söylenebilir.

Log-normal model için adımsal seçim yöntemi uygulanırsa Çizelge 4.11'deki sonuçlar elde edilir.

Çizelge 4.11. Log-normal regresyon modeli için adımsal seçim yöntemi sonuçları

Değişken	Tahmin	Standart Hata	p-değeri
Dissek1 (D0-1)	-0,718	0,211	0,001
Evre3	-1,251	0,276	0,000
Evre4	-1,632	0,327	0,000
Kemoterapi Yok	-0,779	0,214	0,000
sabit	4,994	0,318	0,000
Lnsigma	-0,170	0,102	0,096
Sigma	0,843	0,863	
-2log(L)	183,184		

Çizelge 4.11 incelendiğinde diseksiyon, evre ve kemoterapi değişkenleri yaşam süresini etkileyen faktörler olarak bulunmuştur. Lenf nodu diseksiyonunun değeri D0-1 olanların yaşam süresi D2-3 olanlara göre 2 ($1/\exp(-0,718)$) kat daha kısa, evre 3 olanların evre 1+evre 2 olanlara göre yaşam süresi 3,5 ($1/\exp(-1,251)$) kat, evre 4 olanların evre 1+evre 2 olanlara göre yaşam süresi 5 ($1/\exp(-1,632)$) kat daha kısadır. Kemoterapi almayanların yaşam süresi ise alanlara göre 2 ($1/\exp(-0,779)$) kat daha kısadır.

5. SONUÇLAR

Bu çalışmada, yaşam çözümlemesi ve aykırı değerler hakkında genel bilgiler, kavramlar, fonksiyonlar ve modeller verilmiş yaşam çözümlemesinde aykırı değerleri belirlemede kullanılabilecek yöntemler incelenmiştir.

Aykırı değerlerin belirlenmesi modele uyum için oldukça önemlidir ve yaşam çözümlemesinde aykırı değerleri belirleme yöntemleri artıklara dolayısıyla artıkların analizine dayanmaktadır. Bu amaçla literatürde karşılaşılan artık türleri incelenmiştir. Bu yöntemlerin uygulaması 118 gözlemlili ve 8 açıklayıcı değişkenli mide kanseri verisi üzerinde yapılmıştır.

Uygulamada ilk olarak yaşam çözümlemesinde en çok kullanılan model olan Cox orantılı tehlikeler modeli uygulanmış ve modele uyumlu olduğu başka yöntemlerle birlikte Cox-Snell artıkları ile de görülmüştür. Cox orantılı tehlikeler modelinin temel varsayımı olan orantılı tehlikeler varsayımı da Schoenfeld artıkları ve diğer yöntemlerle incelenmiş ve bu varsayımın sağlandığı görülmüştür. Verilerin dağılımı bilinen bir dağılıma uygunluk göstermese de parametrik bir dağılım gösterdiği görüldüğünden parametrik yaşam çözümlemesi modelleri incelendiğinde log-normal dağılımın veriye uygun bir model olduğu sonucuna ulaşılmıştır.

Aykırı değerleri belirlemek için Martingale, sapma, log-odds ve normal sapma artıkları hesaplanmış ve grafikleri çizdirilmiştir. Aykırı değer olabilecek gözlemler tespit edilmiştir. Martingale artıkları simetrik dağılmadığından grafiklerinin yorumlanması zordur. Ama sapma artıkları simetrik ve referans dağılımı normal dağılım olduğundan anlamlılık düzeyine göre kesim noktası yaklaşık olarak 2 belirlenip sapma artık grafiğinde +2 ve -2 değerlerinin dışındaki gözlemlerin aykırı değerler olabilecekleri sonucuna ulaşılmıştır. Aynı referans dağılımına sahip normal sapma artıkları için de kesim noktası aynı alınmış ve benzer yorumlara ulaşılmıştır. Log-odds artıkları incelendiğinde ise aykırı değerler olabilecek gözlemler bulunamamıştır. Sonuçta diğer üç yöntem tarafından da Cox orantılı tehlikeler modeli için aykırı değer olabileceği düşünülen iki gözlem çalışmadan çıkarılıp model yeniden kurulmuştur. Aykırı değerler çıkarıldıktan sonra sonuçlar karşılaştırıldığında modelin daha iyi sonuç verdiği görülmüştür.

Aykırı deęerlerin belirlenmesi kadar etkili gözlemlerin belirlenmesi de önemli olduğundan etkili gözlemler de belirlenmeye çalışılmıştır. Bu amaçla DFBETA deęerleri hesaplanmış, grafikleri çizdirilmiştir. Ama her deęişken ve düzeyi için hesaplandığından ayrıca örneklem büyüklüğü artıkça test edilecek deęerlerin sayısı da arttığından bir dezavantaj oluşturmaktadır. Etkili gözlemlerin belirlenmesinde kullanılan dięer iki yöntem LMAX ve olabilirlik deęişim deęerleridir. Bu iki yöntem de uygulanmış etkili gözlemler belirlenmiştir.

Veriye uygun olduğu bulunan log-normal regresyon modeli de elde edilmiştir. Log-normal regresyon modeli için de aykırı deęerler belirlenmeye çalışılmıştır. Model kurulduktan sonra Martingale, sapma, log-odds ve normal sapma artıkları hesaplanmıştır. Martingale ve sapma artıkları için grafikler çizdirilmiştir. Aykırı deęer olabilecek gözlemler belirlenmiştir. Cox orantılı tehlikeler modelinde olduğu gibi burada da log-odds artıkları ile aykırı deęer olabilecek gözlemler bulunamamıştır. Aynı şekilde aykırı deęer olabileceęi düşünölen üç gözlem çalışmadan çıkarılıp model yeniden kurulmuş ve aykırı deęerler çıkarıldıktan sonra elde edilen modelin daha iyi olduğu görölmüştür.

Sonuç olarak yaşam çözümlemesinde kullanılan artık türlerinin genel amacı aykırı deęerleri belirlemektir. Ancak, Cox-Snell artıklarının modele uyumda kullanılabileceęi ya da Martingale artıklarının açıklayıcı deęişkenin modele uyum için bir dönüşöme ihtiyacı olup olmadığını belirlemede kullanılabileceęi gibi özel amaçları da vardır.

KAYNAKLAR

- [1] Collett,D., *Modelling Survival Data in Medical Research*, Chapman &Hall/CRC, **1994**.
- [2] Fitrianto, A., Jiin, R. L. T., Several Types of Residuals in Cox Regression Model: An Emprical Study, *International Journal of Mathematical Analysis*, 7,73, 2645-2654, **2013**.
- [3] Cox, D. R., Snell, E. J. , A general definition of residuals, *Journal of the Royal Statistical Society*,30, 2, 248-275, **1968**.
- [4] Barlow ,W. E., Prentice, R. L., Residuals for relative risk regression, *Biometrika*, 75, 65 – 74, 1988.
- [5] Therneau, T. M., Grambcsh, P. M., Fleming, T. R., Martingale-based residuals for survival models, *Biometrika*, 77, 1, 147-160, **1990**.
- [6] Flemming, T. R., Harrington, D. P., *Couting Processes and Survival Analysis*, Wiley, New York, **1991**.
- [7] Nardi, A., Schemper, M., New residuals for cox regression and their application to outlier screening, *Biometrics*, 55, 2, 523-529, **1999**.
- [8] Schoenfeld, D., Partial residuals for the proportional hazards regression model, *Biometrika* , 69, 239-241, **1982**.
- [9] Kleinbaum,D.G., Klein,M.,*Survival analysis: A Self-Learning Text*, Second Edition, Springer, **2005**.
- [10] Elandt-Johnson, R. C., Johnson, N. L., *Survival Models and Data Analysis*, John Wiley& Sons, Inc, New York,**1980**.
- [11] Cox, D.R., Oakes, D., *Analysis of Survival Data*, Chapman and Hall, London, **1984**.
- [12] Allison,P.D., *Survival Analysis Using SAS* , A Practical Guide, SAS Institute Inc., Cary, NC, USA, **1995**.
- [13] Kul, S., The Use of Survival Analysis for Clinical Pathways, *International Journal of Care Pathways* ,14, 23–26, **2010**.
- [14] Gross, A., Clark, V., *Survival Distributions: Reliability Applications in the Biomedical Sciences*, John Wiley And Sons, New York, **1975**.
- [15] Kleinbaum, D.G., *Survival Analysis: A Self-Learning Text*, First Edition Springer, **1996**.
- [16] Lee, E.T., Wang, J.W., *Statistical Methods for Survival Data Analysis*, Third Edition, Wiley, **2003**.
- [17] Sertkaya, D., Ata, N., Sözer, M.T., Yaşam çözümlemesinde zamana bağlı açıklayıcı değişkenli Cox regresyon modeli, *Ankara Üniversitesi Tıp Fakültesi Mecmuası*, 58, 153-158, **2005**.
- [18] Yay, M., Çoker, E., Uysal, Ö., Yaşam analizinde Cox regresyon modeli ve artıkların incelenmesi, *Cerrahpaşa Tıp Dergisi*, 38, 139 – 145, **2007**.
- [19] Hosmer.Jr., D. W., Lemeshow, S., *Applied Survival Analysis: Regression Modeling of Time to Event Data*, Wiley&Sons, New York, **1999**.

- [20] Ata, N., Sertkaya, D., Sözer, M.T., Orantılı tehlike varsayımının incelenmesinde kullanılan yöntemler ve bir uygulama, *Eskişehir Osmangazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, XX, S.1, **2007**.
- [21] Bulut, V., *Türkiye’de İşsizlik Süresini Etkileyen Faktörlerin Yaşam Çözümlemesi İle İncelenmesi*, Yüksek Lisans Tezi, Hacettepe Üniversitesi Fen Bilimleri Enstitüsü, Ankara, **2011**.
- [22] Therneau, T. M., Grambsch, P. M., *Modeling Survival Data: Extending Cox Model*, Springer, New York, **2000**.
- [23] Ata, N., *Yaşam Çözümlemesinde Orantısız Hazard Modeli*, Yüksek Lisans Tezi, Hacettepe Üniversitesi Fen Bilimleri Enstitüsü, Ankara, **2005**.
- [24] Terzi, Y., Cengiz, M.A., Bek, Y., Cox Regresyon Modelinde Oransal Hazard Varsayımının Artıklarla İncelenmesi ve Akciğer Kanseri Hastaları Üzerinde Uygulanması, *Türkiye Klinikleri Tıp Bilimleri Dergisi*, 25, 770-775, **2005**.
- [25] İnceoğlu, F., *Sağkalım Analiz Yöntemleri ve Karaciğer Nakli Verileri İle Bir Uygulama*, Yüksek Lisans Tezi, İnönü Üniversitesi Sağlık Bilimleri Enstitüsü, Malatya, **2013**.
- [26] Tamam, D., *Tam ve Sansürlü Örneklem Durumlarında Weibull Dağılımı için Bazı İstatistikî Sonuç Çıkarımları*, Ankara Üniversitesi Fen Bilimleri Enstitüsü, Ankara, **2008**.
- [27] Ersoy, E., *Sağ Kalım Analizinde Parametre Tahmin Problemlerine Katkılar*, Ondokuz Mayıs Üniversitesi Fen Bilimleri Enstitüsü, Samsun, **2005**.
- [28] Gürsakal, N., *Sosyal Bilimlerde Araştırma Yöntemleri*, Uludağ Üniversitesi Basımevi, Bursa, **2001**.
- [29] Stepanova, M., Thomas, L., Survival analysis methods for personal loan data, *Operations Research*, 50, 2, 277-289, **2002**.
- [30] Noh, N. A., *Detecting Outliers and Influential Observations in Survival Model*, Master Thesis, University of Malaya, Institute of Mathematical Sciences, Kuala Lumpur, **2010**.
- [31] Dağhan, G., *Biyoeşdeğerlik Çalışmalarında Aykırı Değerlerin İncelenmesi*, Yüksek Lisans Tezi, Hacettepe Üniversitesi Fen Bilimleri Enstitüsü, Ankara, **2011**.
- [32] Gharibvand, L., Jeske, D.R., Liao, S., Evaluation of a Hospice Care Referral Program Using Cox Proportional Hazards Model, *Western Users of SAS Software Conference*, Universal City, CA, **2008**.
- [33] Tableman, M., Kim, J.S., *Survival Analysis Using S: Analysis of Time-to-Event Data*, Chapman & Hall/CRC, **2004**.
- [34] Lin, D. Y., Wei, L. J., Ying, Z., Checking the Cox model with cumulative sums of martingale-based residuals, *Biometrika*, 80, 3, 557-572, **1993**.

- [35] Anonim, Postestimation tools for stcox, http://www.stata-press.com/manuals/errata/stata11/i/stcox_postestimation.pdf (Nisan, **2014**).
- [36] Borucka, J., Assesment of Cox Proportional Hazard Model Adequancy Using PROC PHREG and PROC GPLOT, *PhUSE 2010 Berlin Conference*, 17-20 October, Berlin, Germany, p. SP05, **2010**.
- [37] Fox, J., Cox Proportional-Hazards Regression for Survival Data, Appendix to An R and S-PLUS Companion to Applied Regression, **2002**.
- [38] Gharibvand, L. and Fernandez, G., *Advanced Statistical and Graphical features of SAS® PHREG*, SAS GLOBAL Forum 2008 Conference proceedings San Antonio TX , **2008**.
- [39] Klein, J. P., Moeschberger, M. L., *Survival Analysis: Techniques for Censored and Truncated Data.*, Springer, New York, **2003**.
- [40] Winnett, A, Sasieni, P., Miscellanea, A note on scaled Schoenfeld residuals for the proportional hazards model, *Biometrika*, 88, 565-71, **2001**.
- [41] Nardi, A., Schemper, M., Comparing Cox and parametric models in clinical studies, *Statistics in Medicine*, 22, 3597-3610, **2003**.
- [42] Eroğlu, A., Altınok, M., Özgen, K., Sertkaya, D., A Multivariate Analysis of Clinical and Pathological Variables in Survival After Resection of Gastric Cancer, *Türkiye Klinikleri Medical Research*, 15, 1, 15-20, **1997**.
- [43] Yılmaz, A. E., Karasoy D., Proportional Odds Model in Survival Analysis, *GIS 2013: 1. Genc Istatistikciler Sempozyumu*, 10-11 Eylül, Ankara, Türkiye, **2013**.

EKLER

EK1: COX-SNELL, MARTINGALE VE SAPMA ARTIK DEĞERLERİ

Gözlem	Cox-Snell Artığı	Martingale Artığı	Sapma Artığı	Gözlem	Cox-Snell Artığı	Martingale Artığı	Sapma Artığı
1	0,3064	-0,3064	-0,7828	39	0,4807	0,5193	0,6531
2	0,1592	-0,1592	-0,5643	40	1,9382	-0,9382	-0,7436
3	0,2848	-0,2848	-0,7547	41	0,5875	0,4125	0,4887
4	0,1112	-0,1112	-0,4717	42	0,4087	0,5913	0,7791
5	0,2479	-0,2479	-0,7042	43	0,1365	-0,1365	-0,5225
6	0,0528	0,9472	1,9971	44	0,7281	-0,7281	-1,2067
7	0,5165	-0,5165	-1,0164	45	0,1911	-0,1911	-0,6182
8	0,0634	0,9366	1,9087	46	2,0687	-1,0687	-0,8268
9	0,3049	0,6951	0,9927	47	1,6141	-0,6141	-0,5202
10	0,3097	0,6903	0,9816	48	0,4159	-0,4159	-0,9121
11	0,0460	0,9540	2,0615	49	0,1399	-0,1399	-0,5289
12	0,4730	-0,4730	-0,9726	50	0,4813	0,5187	0,6520
13	1,9941	-0,9941	-0,7796	51	0,5366	0,4634	0,5641
14	0,1585	-0,1585	-0,5630	52	0,1968	0,8032	1,2825
15	0,0416	0,9584	2,1074	53	1,2151	-0,2151	-0,2014
16	1,7333	-1,7333	-1,8619	54	1,0496	-0,0496	-0,0488
17	0,0923	-0,0923	-0,4295	55	0,5602	-0,5602	-1,0584
18	0,7305	0,2695	0,2984	56	0,1990	-0,1990	-0,6308
19	0,0618	0,9382	1,9210	57	0,2528	0,7472	1,1205
20	0,8375	0,1625	0,1722	58	0,2650	-0,2650	-0,7281
21	0,0589	0,9411	1,9445	59	0,8476	0,1524	0,1609
22	0,4142	0,5858	0,7689	60	0,5344	0,4656	0,5675
23	0,1735	-0,1735	-0,5891	61	0,3362	0,6638	0,9233
24	1,2279	-0,2279	-0,2126	62	0,0831	-0,0831	-0,4077
25	0,2222	-0,2222	-0,6666	63	0,1723	0,8277	1,3644
26	2,8400	-1,8400	-1,2619	64	0,4047	-0,4047	-0,8997
27	0,1875	-0,1875	-0,6124	65	0,0252	-0,0252	-0,2244
28	1,8384	-1,8384	-1,9175	66	0,1470	0,8530	1,4590
29	0,0252	0,9748	2,3260	67	0,3891	0,6109	0,8162
30	0,0520	-0,0520	-0,3225	68	0,2416	-0,2416	-0,6951
31	0,0508	-0,0508	-0,3187	69	0,3150	0,6850	0,9697
32	0,7723	-0,7723	-1,2428	70	0,2648	-0,2648	-0,7277
33	0,5528	0,4472	0,5396	71	0,1354	0,8646	1,5067
34	0,1501	-0,1501	-0,5478	72	1,1842	-1,1842	-1,5390
35	0,1227	-0,1227	-0,4954	73	0,7130	0,2870	0,3202
36	0,0157	-0,0157	-0,1774	74	0,6014	0,3986	0,4688
37	0,7544	0,2456	0,2692	75	0,7672	0,2328	0,2538
38	0,0597	-0,0597	-0,3455	76	0,1717	-0,1717	-0,5859

EK1: COX-SNELL, MARTINGALE VE SAPMA ARTIK DEĞERLERİ (DEVAM)

Gözlem	Cox-Snell Artığı	Martingale Artığı	Sapma Artığı
77	0,6406	-0,6406	-1,1319
78	0,0967	-0,0967	-0,4399
79	0,3842	0,6158	0,8257
80	0,1011	0,8989	1,6691
81	0,1017	-0,1017	-0,4510
82	0,0770	0,9230	1,8113
83	0,1828	-0,1828	-0,6046
84	0,2686	-0,2686	-0,7330
85	0,3890	0,6110	0,8163
86	0,1633	-0,1633	-0,5715
87	0,0260	0,9740	2,3134
88	0,1050	0,8950	1,6483
89	0,2877	0,7123	1,0330
90	0,1665	-0,1665	-0,5771
91	0,1057	-0,1057	-0,4597
92	0,1061	-0,1061	-0,4608
93	0,1301	-0,1301	-0,5101
94	0,9922	-0,9922	-1,4087
95	0,1328	-0,1328	-0,5153
96	0,2125	-0,2125	-0,6519
97	0,5176	-0,5176	-1,0174
98	0,1312	-0,1312	-0,5122
99	0,0678	-0,0678	-0,3683
100	0,2772	-0,2772	-0,7446
101	0,6312	-0,6312	-1,1235
102	0,0096	-0,0096	-0,1382
103	0,0783	-0,0783	-0,3957
104	0,0334	-0,0334	-0,2583
105	0,0392	-0,0392	-0,2800
106	0,0883	-0,0883	-0,4203
107	0,0757	-0,0757	-0,3892
108	0,2441	-0,2441	-0,6987
109	0,0605	-0,0605	-0,3478
110	0,0189	-0,0189	-0,1943
111	1,3267	-0,3267	-0,2967
112	0,1555	0,8445	1,4261
113	0,7608	-0,7608	-1,2335
114	0,0298	-0,0298	-0,2441
115	0,5403	0,4597	0,5585
116	0,4071	-0,4071	-0,9023
117	2,9589	-1,9589	-1,3222
118	0,0354	0,9646	2,1799

EK 2: LOG-ODDS VE NORMAL-SAPMA ARTIK DEĞERLERİ

Gözlem	Süre	Durum	$\widehat{S}_i(t_i)$	$\widehat{S}_i(t_i^c)$	l_i^c	l_i^m	n_i^c	n_i^m
1	23	Yaşıyor		0,282	-0,405	-0,96	-0,576	-0,576
2	17	Yaşıyor		0,323	-0,322	-0,885	-0,460	-0,460
3	22	Yaşıyor		0,282	-0,405	-0,96	-0,576	-0,576
4	11	Yaşıyor		0,400	-0,176	-0,756	-0,253	-0,253
5	27	Yaşıyor		0,256	-0,462	-1,011	-0,655	-0,655
6	7	Ölmüş	0,906		0,983	0,983	1,315	1,315
7	21	Yaşıyor		0,282	-0,405	0,654	-0,576	-0,576
8	9	Ölmüş	0,818		0,654	0,654	0,910	0,910
9	16	Ölmüş	0,657		0,283	0,283	0,405	0,405
10	13	Ölmüş	0,710		0,389	0,389	0,554	0,554
11	3	Ölmüş	0,983		1,763	1,763	2,121	2,121
12	27	Yaşıyor		0,256	-0,462	0,983	-0,655	-0,655
13	19	Ölmüş	0,606		0,188	0,188	0,270	0,270
14	23	Yaşıyor		0,282	-0,405	0,603	-0,576	-0,576
15	7	Ölmüş	0,906		0,983	0,983	1,315	1,315
16	39	Yaşıyor		0,198	-0,607	0,236	-0,848	-0,848
17	51	Yaşıyor		0,198	-0,607	-0,085	-0,848	-0,848
18	15	Ölmüş	0,668		0,304	0,304	0,435	0,435
19	11	Ölmüş	0,800		0,603	0,603	0,843	0,843
20	24	Ölmüş	0,548		0,083	0,083	0,119	0,119
21	18	Ölmüş	0,633		0,236	0,236	0,339	0,339
22	29	Ölmüş	0,451		-0,085	-0,085	-0,123	-0,123
23	69	Yaşıyor		0,198	-0,607	0,389	-0,848	-0,848
24	20	Ölmüş	0,579		0,138	0,138	0,199	0,199
25	57	Yaşıyor		0,198	-0,607	0,113	-0,848	-0,848
26	29	Ölmüş	0,451		-0,085	-0,085	-0,123	-0,123
27	63	Yaşıyor		0,198	-0,607	-0,183	-0,848	-0,848
28	58	Yaşıyor		0,198	-0,607	0,654	-0,848	-0,848
29	2	Ölmüş	0,992		2,068	2,068	2,388	2,388
30	10	Yaşıyor		0,405	-0,167	0,502	-0,241	-0,241
31	10	Yaşıyor		0,405	-0,167	0,188	-0,241	-0,241
32	15	Yaşıyor		0,334	-0,299	0,053	-0,429	-0,429
33	36	Ölmüş	0,425		-0,132	-0,132	-0,190	-0,190
34	65	Yaşıyor		0,198	-0,607	0,831	-0,848	-0,848
35	17	Yaşıyor		0,323	-0,322	0,831	-0,460	-0,460
36	16	Yaşıyor		0,329	-0,310	0,831	-0,444	-0,444
37	13	Ölmüş	0,710		0,389	0,389	0,554	0,554
38	37	Yaşıyor		0,198	-0,607	0,983	-0,848	-0,848
39	13	Ölmüş	0,710		0,389	0,389	0,554	0,554

EK 2: LOG-ODDS VE NORMAL-SAPMA ARTIK DEĞERLERİ (DEVAM)

Gözlem	Süre	Durum	$\widehat{S}_i(t_i)$	$\widehat{S}_i(t_i^c)$	l_i^c	l_i^m	n_i^c	n_i^m
40	26	Ölmüş	0,513		0,022	0,022	0,032	0,032
41	21	Ölmüş	0,565		0,113	0,113	0,163	0,163
42	15	Ölmüş	0,668		0,304	0,304	0,435	0,435
43	28	Yaşıyor		0,256	-0,462	0,502	-0,655	-0,655
44	20	Yaşıyor		0,289	-0,390	0,260	-0,555	-0,555
45	44	Yaşıyor		0,198	-0,607	0,502	-0,848	-0,848
46	37	Ölmüş	0,396		-0,183	-0,183	-0,263	-0,263
47	9	Ölmüş	0,818		0,654	0,654	0,910	0,910
48	44	Yaşıyor		0,198	-0,607	1,353	-0,848	-0,848
49	44	Yaşıyor		0,198	-0,607	0,389	-0,848	-0,848
50	10	Ölmüş	0,810		0,629	0,629	0,876	0,876
51	12	Ölmüş	0,761		0,502	0,502	0,709	0,709
52	19	Ölmüş	0,606		0,188	0,188	0,270	0,270
53	25	Ölmüş	0,530		0,053	0,053	0,076	-0,754
54	20	Ölmüş	0,579		0,138	0,138	0,199	-0,703
55	39	Yaşıyor		0,198	-0,607	0,654	-0,848	-2,884
56	37	Yaşıyor		0,198	-0,607	-0,085	-0,848	-2,884
57	8	Ölmüş	0,871		0,831	0,831	1,133	-0,869
58	36	Yaşıyor		0,212	-0,569	-1,108	-0,798	-2,584
59	8	Ölmüş	0,871		0,831	0,831	1,133	-0,869
60	8	Ölmüş	0,871		0,831	0,831	1,133	-0,869
61	13	Ölmüş	0,710		0,389	0,389	0,554	-0,655
62	34	Yaşıyor		0,226	-0,536	-1,078	-0,753	-2,349
63	7	Ölmüş	0,906		0,983	0,983	1,315	-1,046
64	33	Yaşıyor		0,226	-0,536	-1,078	-0,753	-2,349
65	33	Yaşıyor		0,226	-0,536	-1,078	-0,753	-2,349
66	7	Ölmüş	0,906		0,983	0,983	1,315	-1,046
67	15	Ölmüş	0,668		0,304	0,304	0,435	-0,656
68	29	Yaşıyor		0,226	-0,536	-1,078	-0,753	-2,349
69	12	Ölmüş	0,761		0,502	0,502	0,709	-0,674
70	29	Yaşıyor		0,226	-0,536	-1,078	-0,753	-2,349
71	5	Ölmüş	0,957		1,353	1,353	1,722	-1,836
72	26	Yaşıyor		0,256	-0,462	-1,011	-0,655	-1,928
73	12	Ölmüş	0,761		0,502	0,502	0,709	-0,674
74	17	Ölmüş	0,646		0,260	0,260	0,373	-0,663
75	12	Ölmüş	0,761		0,502	0,502	0,709	-0,674
76	25	Yaşıyor		0,265	-0,443	-0,993	-0,627	-1,831
77	23	Yaşıyor		0,282	-0,405	-0,960	-0,576	-1,668
78	23	Yaşıyor		0,282	-0,405	-0,960	-0,576	-1,668

EK 2: LOG-ODDS VE NORMAL-SAPMA ARTIK DEĞERLERİ (DEVAM)

Gözlem	Süre	Durum	$\widehat{S}_i(t_i)$	$\widehat{S}_i(t_i^c)$	l_i^c	l_i^m	n_i^c	n_i^m
79	9	Ölmüş	0,818		0,654	0,654	0,910	-0,737
80	4	Ölmüş	0,975		1,584	1,584	1,953	-2,755
81	21	Yaşıyor		0,282	-0,405	-0,960	-0,576	-1,668
82	5	Ölmüş	0,957		1,353	1,353	1,722	-1,836
83	19	Yaşıyor		0,303	-0,361	-0,921	-0,515	-1,503
84	19	Yaşıyor		0,303	-0,361	-0,921	-0,515	-1,503
85	13	Ölmüş	0,710		0,389	0,389	0,554	-0,655
86	18	Yaşıyor		0,316	-0,335	-0,897	-0,478	-1,414
87	6	Ölmüş	0,949		1,269	1,269	1,635	-1,599
88	8	Ölmüş	0,871		0,831	0,831	1,133	-0,869
89	9	Ölmüş	0,818		0,654	0,654	0,910	-0,737
90	17	Yaşıyor		0,323	-0,322	-0,885	-0,460	-1,374
91	17	Yaşıyor		0,323	-0,322	-0,885	-0,460	-1,374
92	17	Yaşıyor		0,323	-0,322	-0,885	-0,460	-1,374
93	16	Yaşıyor		0,329	-0,310	-0,875	-0,444	-1,339
94	16	Yaşıyor		0,329	-0,310	-0,875	-0,444	-1,339
95	16	Yaşıyor		0,329	-0,310	-0,875	-0,444	-1,339
96	15	Yaşıyor		0,334	-0,299	-0,865	-0,429	-1,309
97	15	Yaşıyor		0,334	-0,299	-0,865	-0,429	-1,309
98	13	Yaşıyor		0,355	-0,259	-0,830	-0,372	-1,204
99	11	Yaşıyor		0,400	-0,176	-0,756	-0,253	-1,029
100	12	Yaşıyor		0,380	-0,212	-0,788	-0,304	-1,099
101	12	Yaşıyor		0,380	-0,212	-0,788	-0,304	-1,099
102	11	Yaşıyor		0,400	-0,176	-0,756	-0,253	-1,029
103	11	Yaşıyor		0,400	-0,176	-0,756	-0,253	-1,029
104	10	Yaşıyor		0,405	-0,167	-0,748	-0,241	-1,015
105	10	Yaşıyor		0,405	-0,167	-0,748	-0,241	-1,015
106	8	Yaşıyor		0,436	-0,112	-0,700	-0,162	-0,928
107	8	Yaşıyor		0,436	-0,112	-0,700	-0,162	-0,928
108	13	Yaşıyor		0,355	-0,259	-0,830	-0,372	-1,204
109	9	Yaşıyor		0,409	-0,159	-0,741	-0,229	-1,001
110	6	Yaşıyor		0,474	-0,044	-0,641	-0,064	-0,843
111	14	Ölmüş	0,700		0,367	0,367	0,523	-0,654
112	9	Ölmüş	0,818		0,654	0,654	0,910	-0,737
113	11	Yaşıyor		0,400	-0,176	-0,756	-0,253	-1,029
114	4	Yaşıyor		0,487	-0,022	-0,621	-0,032	-0,819
115	9	Ölmüş	0,818		0,654	0,654	0,910	-0,737
116	73	Yaşıyor		0,198	-0,607	-1,143	-0,848	-2,884
117	29	Ölmüş	0,451		-0,085	-0,085	-0,123	-0,891
118	7	Ölmüş	0,906		0,983	0,983	1,315	-1,046

EK3: DFBETA DEĞERLERİ

n	df1	df2	df3	df4	df5	df6	df7	df8	df9	df10	df11
1	0,000	-0,005	0,031	-0,035	0,013	0,041	0,048	0,042	-0,001	-0,014	0,008
2	0,000	0,010	0,002	-0,011	0,015	0,005	-0,004	0,004	-0,013	0,004	0,005
3	0,000	-0,016	-0,004	-0,024	0,025	-0,003	-0,019	-0,016	-0,019	0,005	0,015
4	0,000	-0,004	-0,003	0,005	0,012	-0,008	-0,003	-0,004	-0,005	0,004	0,005
5	0,000	-0,008	0,027	-0,023	0,027	0,003	-0,011	-0,004	-0,019	0,010	0,018
6	0,005	-0,052	0,021	0,068	0,020	-0,042	0,020	0,003	0,041	-0,041	-0,036
7	-0,001	0,020	0,033	0,008	0,038	-0,007	-0,001	-0,068	-0,006	-0,023	0,007
8	-0,001	-0,062	-0,004	0,103	-0,045	0,064	-0,041	-0,005	-0,219	-0,202	0,096
9	0,000	0,031	0,014	-0,024	-0,062	0,064	0,022	0,017	0,037	0,002	-0,026
10	-0,001	-0,004	-0,055	-0,018	-0,071	0,008	0,036	-0,014	0,047	0,011	0,022
11	0,006	0,059	0,088	0,004	0,018	0,038	0,033	0,220	-0,009	-0,024	0,043
12	0,000	0,004	-0,004	0,013	0,039	0,060	0,063	0,070	-0,020	0,006	-0,015
13	-0,002	-0,029	0,043	-0,017	-0,013	-0,001	0,002	-0,156	0,009	0,033	-0,038
14	0,000	-0,007	-0,005	0,006	0,008	-0,003	-0,007	-0,004	0,039	0,038	-0,005
15	0,002	0,010	0,014	-0,043	0,025	-0,190	-0,154	-0,159	-0,011	-0,093	-0,083
16	-0,001	0,099	-0,019	0,076	-0,011	-0,082	0,005	0,053	-0,078	-0,039	0,032
17	0,000	-0,004	-0,003	0,005	-0,006	0,017	0,015	0,011	0,028	0,030	0,008
18	0,000	0,023	0,020	0,001	0,013	0,012	0,009	0,050	-0,002	0,004	-0,005
19	-0,001	-0,073	0,009	-0,031	-0,056	0,058	-0,036	-0,041	-0,223	-0,213	0,083
20	0,000	0,014	-0,009	0,001	0,001	0,002	-0,008	-0,005	0,005	0,025	-0,006
21	-0,003	-0,040	-0,012	-0,041	-0,059	0,039	-0,031	-0,060	-0,226	-0,241	-0,028
22	0,002	0,010	0,011	0,059	-0,052	-0,051	-0,068	-0,012	0,044	0,027	0,020
23	0,000	0,004	-0,003	-0,013	-0,011	0,007	-0,002	-0,003	0,043	0,045	0,006
24	0,000	0,005	-0,009	0,013	-0,006	0,002	0,010	0,019	0,005	0,016	0,009
25	0,000	-0,017	-0,010	0,012	-0,012	0,004	-0,010	-0,010	0,061	0,061	0,017
26	0,003	-0,005	0,134	0,016	-0,020	0,004	-0,076	0,062	-0,059	-0,016	-0,059
27	-0,001	-0,012	0,011	-0,020	-0,009	-0,011	0,003	-0,016	0,051	0,051	0,008
28	-0,001	-0,007	-0,060	0,039	-0,058	0,242	0,236	0,228	-0,031	-0,023	-0,066
29	0,001	0,035	0,041	0,112	0,032	0,025	0,051	0,051	0,028	0,025	0,056
30	0,000	-0,003	-0,001	0,002	-0,003	0,003	-0,001	0,001	0,015	0,016	0,004
31	0,000	0,000	0,001	0,003	0,005	0,010	0,008	0,010	-0,001	0,005	0,005
32	0,002	-0,042	-0,010	-0,081	-0,030	0,013	-0,030	-0,007	-0,015	0,004	0,034
33	-0,001	-0,020	-0,022	-0,041	-0,007	-0,072	-0,031	-0,061	0,056	0,040	-0,024
34	0,000	0,007	-0,004	-0,012	0,008	-0,003	-0,005	-0,008	0,030	0,032	-0,011
35	0,000	-0,005	-0,004	0,005	0,006	-0,002	-0,005	-0,002	0,030	0,030	-0,004
36	0,000	0,001	0,002	-0,001	0,001	0,001	0,000	0,001	0,003	0,005	0,001
37	0,000	0,012	0,014	-0,002	0,014	0,021	0,002	-0,001	-0,002	0,016	-0,001
38	0,000	0,002	0,000	-0,003	0,004	0,001	-0,001	-0,001	0,012	0,016	0,003
39	0,000	0,019	0,020	0,058	-0,032	0,028	0,036	0,030	0,034	0,036	0,040
40	0,003	0,031	-0,010	0,005	-0,026	0,048	0,027	-0,056	0,009	0,048	0,033

EK3: DFBETA DEĞERLERİ (DEVAM)

n	df1	df2	df3	df4	df5	df6	df7	df8	df9	df10	df11
41	-0,001	0,029	0,008	-0,011	-0,040	0,057	0,025	0,022	0,031	0,016	-0,014
42	-0,001	0,010	-0,052	-0,017	0,007	-0,019	0,019	-0,032	0,017	-0,014	-0,040
43	0,000	-0,009	-0,007	0,006	-0,008	0,004	-0,005	-0,004	0,039	0,041	0,010
44	-0,001	0,049	-0,015	0,030	-0,010	-0,026	0,013	0,031	-0,020	0,007	0,012
45	0,000	0,011	-0,002	0,011	0,013	0,000	-0,007	0,002	0,039	0,042	-0,009
46	-0,001	0,063	-0,014	-0,126	-0,041	-0,053	0,049	0,033	-0,013	0,004	-0,003
47	0,002	-0,019	-0,019	0,001	-0,014	-0,058	-0,005	0,005	-0,003	-0,029	-0,041
48	0,001	0,025	-0,001	0,015	0,027	-0,035	0,006	0,009	0,091	0,078	-0,034
49	0,000	0,012	0,015	0,007	0,012	0,004	-0,002	0,007	0,028	0,034	-0,004
50	0,000	-0,021	0,023	0,050	0,041	0,005	0,014	0,003	0,004	0,059	0,020
51	-0,001	0,001	0,020	-0,006	-0,011	-0,040	-0,058	-0,064	-0,003	0,055	0,045
52	0,001	0,040	0,021	-0,043	-0,083	0,012	0,068	0,023	0,055	-0,018	-0,060
53	0,000	0,029	0,015	0,010	-0,015	0,023	0,005	0,028	0,008	0,004	0,002
54	0,000	0,020	-0,003	0,010	-0,002	0,017	0,010	0,022	0,004	0,014	-0,001
55	-0,001	-0,020	-0,011	-0,054	-0,022	0,088	0,078	0,049	-0,009	0,021	0,032
56	0,000	-0,013	0,017	-0,020	-0,010	0,007	-0,003	-0,006	0,052	0,056	0,015
57	-0,003	-0,032	-0,086	-0,005	-0,055	-0,011	-0,017	0,038	0,002	0,040	-0,017
58	0,000	-0,017	-0,009	-0,027	0,012	-0,012	-0,015	-0,025	0,060	0,055	-0,014
59	0,000	-0,007	0,003	0,000	0,004	-0,002	-0,002	0,012	0,000	0,002	0,010
60	-0,001	-0,034	-0,001	0,051	0,016	-0,008	0,007	-0,024	0,021	0,021	0,036
61	-0,001	-0,045	0,007	-0,034	0,019	-0,035	0,011	-0,056	0,022	-0,004	-0,026
62	0,000	0,004	0,002	0,006	0,005	0,004	0,000	0,007	0,019	0,023	0,005
63	-0,003	0,036	0,005	-0,032	0,016	-0,031	0,033	-0,033	0,011	-0,025	-0,063
64	0,000	-0,018	-0,006	0,024	0,043	-0,001	-0,030	-0,010	-0,027	0,010	0,031
65	0,000	0,002	0,001	-0,002	0,001	0,005	0,005	0,004	0,006	0,008	0,001
66	0,001	-0,084	-0,067	-0,032	-0,067	0,017	0,030	-0,035	0,044	0,092	0,070
67	-0,001	0,010	-0,052	-0,019	0,007	-0,019	0,020	-0,030	0,018	-0,016	-0,042
68	0,001	0,010	0,002	0,012	-0,014	0,016	0,003	0,021	0,065	0,065	0,014
69	0,001	0,019	-0,048	0,082	0,013	0,035	-0,010	0,016	0,010	-0,018	-0,018
70	0,000	0,018	0,001	0,020	0,030	0,008	-0,011	0,010	-0,024	0,007	0,016
71	0,000	0,054	0,053	-0,026	0,054	0,020	0,056	0,032	-0,021	0,073	-0,019
72	-0,001	0,013	-0,050	0,020	0,063	-0,047	-0,025	-0,182	-0,007	-0,100	-0,008
73	0,000	-0,011	0,016	0,029	0,028	0,004	0,010	-0,001	-0,001	0,038	0,014
74	0,001	0,033	0,031	-0,004	-0,016	0,069	0,031	0,043	0,009	0,059	0,006
75	0,000	0,008	-0,003	-0,002	0,016	0,024	0,003	0,005	-0,008	0,025	0,006
76	0,000	-0,001	0,003	0,008	0,015	0,030	0,025	0,030	-0,006	0,012	0,014
77	0,001	0,034	-0,011	0,035	0,036	-0,009	-0,033	0,020	-0,026	-0,065	0,001
78	0,000	0,005	-0,002	-0,008	0,004	0,000	-0,001	-0,001	0,021	0,022	-0,007
79	-0,001	-0,004	0,016	-0,007	-0,022	-0,062	-0,080	-0,081	-0,001	0,063	0,051
80	0,003	0,057	0,069	0,010	0,020	0,032	0,025	0,187	-0,011	-0,013	0,042

EK3: DFBETA DEĞERLERİ (DEVAM)

n	df1	df2	df3	df4	df5	df6	df7	df8	df9	df10	df11
81	0,000	0,001	0,007	0,001	-0,006	0,019	0,020	0,017	0,031	0,032	-0,001
82	-0,001	-0,070	-0,083	-0,034	0,034	-0,035	0,007	-0,078	0,002	0,056	-0,021
83	0,000	-0,007	-0,005	0,012	0,019	0,001	-0,013	-0,005	-0,011	0,008	0,014
84	0,000	-0,006	-0,009	0,011	0,011	0,033	0,034	0,033	0,002	-0,012	0,008
85	0,002	-0,018	0,030	-0,011	0,020	-0,016	-0,006	0,073	-0,001	-0,020	-0,018
86	0,000	0,017	0,020	0,008	0,021	-0,003	0,005	0,014	-0,012	0,007	0,007
87	0,000	0,010	0,006	-0,040	0,027	-0,194	-0,158	-0,174	-0,012	-0,088	-0,084
88	0,000	0,014	-0,082	-0,038	-0,005	-0,052	0,025	-0,040	0,020	-0,066	-0,086
89	-0,001	0,023	0,010	-0,027	0,013	-0,029	0,030	-0,015	0,013	-0,025	-0,049
90	0,000	0,013	0,003	0,012	0,018	0,008	-0,005	0,013	-0,013	0,006	0,009
91	0,000	0,000	0,002	0,005	0,009	0,019	0,016	0,019	-0,003	0,009	0,009
92	0,000	-0,002	0,002	-0,010	0,008	0,018	0,016	0,015	-0,004	0,007	0,007
93	0,000	-0,005	-0,004	0,009	0,014	0,001	-0,009	-0,005	-0,008	0,007	0,010
94	0,004	-0,069	-0,018	-0,107	0,032	-0,047	-0,065	-0,040	-0,019	-0,123	-0,002
95	0,000	0,013	0,011	0,005	-0,006	0,008	0,003	0,011	0,034	0,036	-0,005
96	-0,001	0,015	-0,007	0,015	-0,002	0,011	-0,006	0,006	-0,009	0,011	0,011
97	0,000	-0,008	0,040	0,011	0,000	-0,018	0,007	0,011	-0,006	0,021	0,028
98	0,000	0,012	0,017	-0,012	0,015	-0,003	0,005	0,007	-0,010	0,004	0,003
99	0,000	0,005	-0,003	0,003	-0,004	0,002	0,000	0,002	0,018	0,018	-0,003
100	-0,001	-0,008	0,016	0,008	0,022	-0,028	-0,005	-0,010	-0,004	-0,019	0,003
101	-0,002	-0,037	-0,051	0,024	-0,040	-0,017	-0,046	-0,040	0,014	-0,049	0,012
102	0,000	0,000	0,000	-0,001	0,001	0,000	0,000	0,000	0,002	0,003	0,000
103	0,000	-0,001	0,010	0,004	0,010	0,004	-0,003	0,004	-0,005	0,006	0,008
104	0,000	0,001	0,000	-0,003	-0,002	0,002	0,001	0,001	0,009	0,010	0,001
105	0,000	0,003	0,000	0,003	0,005	0,002	-0,001	0,001	-0,003	0,003	0,003
106	0,000	-0,001	0,001	0,003	-0,003	0,018	0,015	0,018	0,001	0,007	0,008
107	0,000	0,006	-0,002	-0,006	0,007	-0,001	-0,003	-0,002	-0,007	-0,001	-0,005
108	0,000	0,014	-0,008	0,016	0,015	-0,002	-0,012	0,003	-0,010	-0,019	0,001
109	0,000	-0,002	-0,001	0,004	0,007	0,001	-0,004	-0,002	-0,004	0,004	0,005
110	0,000	0,000	0,000	0,001	0,000	0,004	0,003	0,003	0,000	0,002	0,002
111	0,000	0,003	0,028	0,003	0,005	-0,015	0,006	0,002	-0,004	0,014	-0,012
112	0,003	0,028	0,036	-0,042	0,011	-0,026	0,046	0,024	0,018	-0,046	-0,060
113	0,000	-0,045	-0,045	0,023	-0,049	-0,015	-0,049	-0,023	0,018	-0,067	0,015
114	0,000	0,002	0,000	0,002	0,002	0,000	-0,001	0,002	-0,001	-0,003	0,000
115	-0,001	0,013	0,004	-0,007	0,009	0,019	-0,002	-0,009	0,003	-0,003	-0,015
116	0,000	-0,031	-0,017	0,013	-0,022	-0,026	0,000	-0,020	0,117	0,101	0,018
117	-0,003	-0,057	-0,122	0,041	-0,055	-0,055	-0,137	-0,063	-0,054	-0,053	-0,099
118	0,005	-0,063	0,058	-0,035	0,050	0,055	-0,034	0,011	-0,268	-0,242	0,077

EK4: OLABİLİRLİK DEĞİŞİM VE LMAX DEĞERLERİ

Gözlem	Süre	Durum	Id	lmax
1	23	Yaşiyor	0,0340	0,0006
2	17	Yaşiyor	0,0059	0,0010
3	22	Yaşiyor	0,0185	0,0265
4	11	Yaşiyor	0,0027	0,0036
5	27	Yaşiyor	0,0203	0,0233
6	7	Ölmüş	0,2559	0,0681
7	21	Yaşiyor	0,0783	0,0412
8	9	Ölmüş	0,4601	0,4325
9	16	Ölmüş	0,0807	0,0207
10	13	Ölmüş	0,0896	0,0166
11	3	Ölmüş	0,3409	0,0491
12	27	Yaşiyor	0,0480	0,0015
13	19	Ölmüş	0,2170	0,0326
14	23	Yaşiyor	0,0086	0,0304
15	7	Ölmüş	0,2346	0,1336
16	39	Yaşiyor	0,3020	0,1015
17	51	Yaşiyor	0,0043	0,0057
18	15	Ölmüş	0,0135	0,0230
19	11	Ölmüş	0,4145	0,4199
20	24	Ölmüş	0,0084	0,0139
21	18	Ölmüş	0,3594	0,3404
22	29	Ölmüş	0,1126	0,0015
23	69	Yaşiyor	0,0105	0,0248
24	20	Ölmüş	0,0063	0,0071
25	57	Yaşiyor	0,0182	0,0066
26	29	Ölmüş	0,4215	0,1558
27	63	Yaşiyor	0,0199	0,0351
28	58	Yaşiyor	0,5168	0,1395
29	2	Ölmüş	0,1683	0,0660
30	10	Yaşiyor	0,0011	0,0032
31	10	Yaşiyor	0,0008	0,0033
32	15	Yaşiyor	0,1160	0,1015
33	36	Ölmüş	0,0534	0,0850
34	65	Yaşiyor	0,0092	0,0404
35	17	Yaşiyor	0,0053	0,0241
36	16	Yaşiyor	0,0001	0,0035
37	13	Ölmüş	0,0096	0,0024
38	37	Yaşiyor	0,0013	0,0118
39	13	Ölmüş	0,0617	0,0009

Gözlem	Süre	Durum	Id	lmax
40	26	Ölmüş	0,1947	0,0079
41	21	Ölmüş	0,0440	0,0110
42	15	Ölmüş	0,0527	0,0863
43	28	Yaşiyor	0,0072	0,0065
44	20	Yaşiyor	0,0688	0,0542
45	44	Yaşiyor	0,0122	0,0498
46	37	Ölmüş	0,2916	0,1487
47	9	Ölmüş	0,0605	0,0438
48	44	Yaşiyor	0,0726	0,1754
49	44	Yaşiyor	0,0084	0,0366
50	10	Ölmüş	0,0569	0,0058
51	12	Ölmüş	0,0715	0,0336
52	19	Ölmüş	0,1372	0,0956
53	25	Ölmüş	0,0146	0,0001
54	20	Ölmüş	0,0057	0,0097
55	39	Yaşiyor	0,0909	0,0786
56	37	Yaşiyor	0,0181	0,0107
57	8	Ölmüş	0,2083	0,0785
58	36	Yaşiyor	0,0270	0,0520
59	8	Ölmüş	0,0045	0,0087
60	8	Ölmüş	0,0532	0,0148
61	13	Ölmüş	0,0602	0,0391
62	34	Yaşiyor	0,0029	0,0181
63	7	Ölmüş	0,1212	0,1281
64	33	Yaşiyor	0,0330	0,0461
65	33	Yaşiyor	0,0004	0,0065
66	7	Ölmüş	0,2230	0,0879
67	15	Ölmüş	0,0541	0,0904
68	29	Yaşiyor	0,0246	0,0394
69	12	Ölmüş	0,1056	0,0038
70	29	Yaşiyor	0,0166	0,0061
71	5	Ölmüş	0,1227	0,0988
72	26	Yaşiyor	0,3148	0,1048
73	12	Ölmüş	0,0229	0,0048
74	17	Ölmüş	0,0514	0,0222
75	12	Ölmüş	0,0107	0,0112
76	25	Yaşiyor	0,0081	0,0119
77	23	Yaşiyor	0,0700	0,0240
78	23	Yaşiyor	0,0038	0,0287

EK4: OLABİLİRLİK DEĞİŞİM VE LMAX DEĞERLERİ (DEVAM)

Gözlem	Süre	Durum	Id	lmax
79	9	Ölmüş	0,1072	0,0397
80	4	Ölmüş	0,1994	0,0365
81	21	Yaşıyor	0,0055	0,0163
82	5	Ölmüş	0,1750	0,0156
83	19	Yaşıyor	0,0073	0,0214
84	19	Yaşıyor	0,0159	0,0072
85	13	Ölmüş	0,0689	0,0093
86	18	Yaşıyor	0,0093	0,0219
87	6	Ölmüş	0,2241	0,1292
88	8	Ölmüş	0,1541	0,1536
89	9	Ölmüş	0,0665	0,1101
90	17	Yaşıyor	0,0065	0,0015
91	17	Yaşıyor	0,0031	0,0071
92	17	Yaşıyor	0,0035	0,0053
93	16	Yaşıyor	0,0042	0,0156
94	16	Yaşıyor	0,2953	0,0202
95	16	Yaşıyor	0,0076	0,0303
96	15	Yaşıyor	0,0115	0,0183
97	15	Yaşıyor	0,0364	0,0078
98	13	Yaşıyor	0,0069	0,0205
99	11	Yaşıyor	0,0019	0,0147
100	12	Yaşıyor	0,0186	0,0229
101	12	Yaşıyor	0,0832	0,0797
102	11	Yaşıyor	0,0000	0,0021
103	11	Yaşıyor	0,0020	0,0081
104	10	Yaşıyor	0,0005	0,0056
105	10	Yaşıyor	0,0005	0,0012
106	8	Yaşıyor	0,0020	0,0109
107	8	Yaşıyor	0,0023	0,0047
108	13	Yaşıyor	0,0111	0,0071
109	9	Yaşıyor	0,0009	0,0073
110	6	Yaşıyor	0,0001	0,0026
111	14	Ölmüş	0,0184	0,0167
112	9	Ölmüş	0,1148	0,1482
113	11	Yaşıyor	0,0975	0,0865
114	4	Yaşıyor	0,0002	0,0015
115	9	Ölmüş	0,0251	0,0017
116	73	Yaşıyor	0,0598	0,0752
117	29	Ölmüş	0,4352	0,1852
118	7	Ölmüş	0,5236	0,3418

EK5: LOG-NORMAL REGRESYON MODELİNDE MARTINGALE-LIKE VE SAPMA ARTIK DEĞERLERİ

Gözlem	Martingale-Like Artığı	Sapma Artığı	Gözlem	Martingale-Like Artığı	Sapma Artığı
1	-0,184	-0,606	41	0,438	0,527
2	-0,111	-0,471	42	0,663	0,922
3	-0,227	-0,673	43	-0,110	-0,469
4	-0,091	-0,427	44	-0,870	-1,319
5	-0,206	-0,641	45	-0,323	-0,804
6	0,963	2,159	46	-1,000	-0,784
7	-0,379	-0,870	47	-0,608	-0,516
8	0,975	2,330	48	-0,748	-1,223
9	0,728	1,071	49	-0,205	-0,640
10	0,712	1,032	50	0,585	0,768
11	0,918	1,778	51	0,500	0,622
12	-0,648	-1,138	52	0,846	1,432
13	-0,728	-0,601	53	-0,126	-0,121
14	-0,137	-0,523	54	-0,354	-0,319
15	0,979	2,403	55	-0,665	-1,154
16	-2,051	-2,025	56	-0,132	-0,514
17	-0,119	-0,488	57	0,831	1,374
18	0,387	0,453	58	-0,279	-0,747
19	0,957	2,096	59	-0,131	-0,126
20	0,292	0,326	60	0,246	0,270
21	0,980	2,421	61	0,668	0,933
22	0,530	0,672	62	-0,054	-0,329
23	-0,476	-0,975	63	0,810	1,304
24	-0,171	-0,162	64	-0,490	-0,990
25	-0,443	-0,941	65	-0,004	-0,085
26	-1,363	-1,003	66	0,829	1,368
27	-0,410	-0,905	67	0,679	0,957
28	-3,091	-2,486	68	-0,206	-0,642
29	0,971	2,270	69	0,713	1,035
30	-0,013	-0,159	70	-0,290	-0,761
31	-0,015	-0,171	71	0,831	1,376
32	-0,600	-1,095	72	-0,979	-1,400
33	0,209	0,226	73	0,394	0,462
34	-0,442	-0,940	74	0,497	0,617
35	-0,074	-0,385	75	0,327	0,372
36	0,000	-0,029	76	-0,141	-0,531
37	0,288	0,321	77	-0,616	-1,110
38	-0,022	-0,209	78	-0,069	-0,373
39	0,562	0,726	79	0,686	0,973
40	-0,707	-0,587	80	0,757	1,148

EK5: LOG-NORMAL REGRESYON MODELİNDE MARTINGALE-LIKE VE SAPMA ARTIK DEĞERLERİ (DEVAM)

Gözlem	Martingale-Like Artığı	Sapma Artığı
81	-0,061	-0,348
82	0,918	1,778
83	-0,141	-0,530
84	-0,171	-0,585
85	0,634	0,862
86	-0,149	-0,546
87	0,982	2,473
88	0,923	1,809
89	0,756	1,143
90	-0,138	-0,526
91	-0,054	-0,328
92	-0,045	-0,299
93	-0,080	-0,400
94	-0,585	-1,082
95	-0,106	-0,461
96	-0,213	-0,653
97	-0,487	-0,987
98	-0,072	-0,380
99	-0,056	-0,334
100	-0,198	-0,629
101	-0,552	-1,050
102	0,000	-0,020
103	-0,037	-0,272
104	-0,005	-0,102
105	-0,014	-0,167
106	-0,060	-0,345
107	-0,071	-0,378
108	-0,160	-0,565
109	-0,020	-0,199
110	-0,010	-0,140
111	-0,408	-0,363
112	0,882	1,586
113	-0,691	-1,175
114	-0,019	-0,196
115	0,473	0,579
116	-1,159	-1,522
117	-1,562	-1,115
118	0,972	2,280

EK 6: LOG-NORMAL REGRESYON MODELİNDE LOG-ODDS VE NORMAL-SAPMA ARTIK DEĞERLERİ

Gözlem	Süre	Durum	$\widehat{S}_i(t_i)$	$\widehat{S}_i(t_i^c)$	l_i^c	l_i^m	n_i^c	n_i^m
1	23	Yaşıyor		0,282	-0,405	-0,960	-0,576	-1,668
2	17	Yaşıyor		0,323	-0,322	-0,885	-0,460	-1,374
3	22	Yaşıyor		0,282	-0,405	-0,960	-0,576	-1,668
4	11	Yaşıyor		0,400	-0,176	-0,756	-0,253	-1,029
5	27	Yaşıyor		0,256	-0,462	-1,011	-0,655	-1,928
6	7	Ölmüş	0,906		0,983	0,983	1,315	1,315
7	21	Yaşıyor		0,282	-0,405	-0,960	-0,576	-1,668
8	9	Ölmüş	0,818		0,654	0,654	0,910	0,910
9	16	Ölmüş	0,657		0,283	0,283	0,405	0,405
10	13	Ölmüş	0,710		0,389	0,389	0,554	0,554
11	3	Ölmüş	0,983		1,763	1,763	2,121	2,121
12	27	Yaşıyor		0,256	-0,462	-1,011	-0,655	-1,928
13	19	Ölmüş	0,606		0,188	0,188	0,270	0,270
14	23	Yaşıyor		0,282	-0,405	-0,960	-0,576	-1,668
15	7	Ölmüş	0,906		0,983	0,983	1,315	1,315
16	39	Yaşıyor		0,198	-0,607	-1,143	-0,848	-2,884
17	51	Yaşıyor		0,198	-0,607	-1,143	-0,848	-2,884
18	15	Ölmüş	0,668		0,304	0,304	0,435	0,435
19	11	Ölmüş	0,800		0,603	0,603	0,843	0,843
20	24	Ölmüş	0,548		0,083	0,083	0,119	0,119
21	18	Ölmüş	0,633		0,236	0,236	0,339	0,339
22	29	Ölmüş	0,451		-0,085	-0,085	-0,123	-0,123
23	69	Yaşıyor		0,198	-0,607	-1,143	-0,848	-2,884
24	20	Ölmüş	0,579		0,138	0,138	0,199	0,199
25	57	Yaşıyor		0,198	-0,607	-1,143	-0,848	-2,884
26	29	Ölmüş	0,451		-0,085	-0,085	-0,123	-0,123
27	63	Yaşıyor		0,198	-0,607	-1,143	-0,848	-2,884
28	58	Yaşıyor		0,198	-0,607	-1,143	-0,848	-2,884
29	2	Ölmüş	0,992		2,068	2,068	2,388	2,388
30	10	Yaşıyor		0,405	-0,167	-0,748	-0,241	-1,015
31	10	Yaşıyor		0,405	-0,167	-0,748	-0,241	-1,015
32	15	Yaşıyor		0,334	-0,299	-0,865	-0,429	-1,309
33	36	Ölmüş	0,425		-0,132	-0,132	-0,190	-0,190
34	65	Yaşıyor		0,198	-0,607	-1,143	-0,848	-2,884
35	17	Yaşıyor		0,323	-0,322	-0,885	-0,460	-1,374
36	16	Yaşıyor		0,329	-0,310	-0,875	-0,444	-1,339
37	13	Ölmüş	0,710		0,389	0,389	0,554	0,554
38	37	Yaşıyor		0,198	-0,607	-1,143	-0,848	-2,884
39	13	Ölmüş	0,710		0,389	0,389	0,554	0,554

EK6: LOG-NORMAL REGRESYON MODELİNDE LOG-ODDS VE NORMAL-SAPMA ARTIK DEĞERLERİ (DEVAM)

Gözlem	Süre	Durum	$\widehat{S}_i(t_i)$	$\widehat{S}_i(t_i^c)$	l_i^c	l_i^m	n_i^c	n_i^m
40	26	Ölmüş	0,513		0,022	0,022	0,032	0,032
41	21	Ölmüş	0,565		0,113	0,113	0,163	0,163
42	15	Ölmüş	0,668		0,304	0,304	0,435	0,435
43	28	Yaşıyor		0,256	-0,462	-1,011	-0,655	-1,928
44	20	Yaşıyor		0,289	-0,390	-0,946	-0,555	-1,608
45	44	Yaşıyor		0,198	-0,607	-1,143	-0,848	-2,884
46	37	Ölmüş	0,396		-0,183	-0,183	-0,263	-0,263
47	9	Ölmüş	0,818		0,654	0,654	0,910	0,910
48	44	Yaşıyor		0,198	-0,607	-1,143	-0,848	-2,884
49	44	Yaşıyor		0,198	-0,607	-1,143	-0,848	-2,884
50	10	Ölmüş	0,810		0,629	0,629	0,876	0,876
51	12	Ölmüş	0,761		0,502	0,502	0,709	0,709
52	19	Ölmüş	0,606		0,188	0,188	0,270	0,270
53	25	Ölmüş	0,530		0,053	0,053	0,076	0,076
54	20	Ölmüş	0,579		0,138	0,138	0,199	0,199
55	39	Yaşıyor		0,198	-0,607	-1,143	-0,848	-2,884
56	37	Yaşıyor		0,198	-0,607	-1,143	-0,848	-2,884
57	8	Ölmüş	0,871		0,831	0,831	1,133	1,133
58	36	Yaşıyor		0,212	-0,569	-1,108	-0,798	-2,584
59	8	Ölmüş	0,871		0,831	0,831	1,133	1,133
60	8	Ölmüş	0,871		0,831	0,831	1,133	1,133
61	13	Ölmüş	0,710		0,389	0,389	0,554	0,554
62	34	Yaşıyor		0,226	-0,536	-1,078	-0,753	-2,349
63	7	Ölmüş	0,906		0,983	0,983	1,315	1,315
64	33	Yaşıyor		0,226	-0,536	-1,078	-0,753	-2,349
65	33	Yaşıyor		0,226	-0,536	-1,078	-0,753	-2,349
66	7	Ölmüş	0,906		0,983	0,983	1,315	1,315
67	15	Ölmüş	0,668		0,304	0,304	0,435	0,435
68	29	Yaşıyor		0,226	-0,536	-1,078	-0,753	-2,349
69	12	Ölmüş	0,761		0,502	0,502	0,709	0,709
70	29	Yaşıyor		0,226	-0,536	-1,078	-0,753	-2,349
71	5	Ölmüş	0,957		1,353	1,353	1,722	1,722
72	26	Yaşıyor		0,256	-0,462	-1,011	-0,655	-1,928
73	12	Ölmüş	0,761		0,502	0,502	0,709	0,709
74	17	Ölmüş	0,646		0,260	0,260	0,373	0,373
75	12	Ölmüş	0,761		0,502	0,502	0,709	0,709
76	25	Yaşıyor		0,265	-0,443	-0,993	-0,627	-1,831
77	23	Yaşıyor		0,282	-0,405	-0,960	-0,576	-1,668
78	23	Yaşıyor		0,282	-0,405	-0,960	-0,576	-1,668

EK 6: LOG-NORMAL REGRESYON MODELİNDE LOG-ODDS VE NORMAL-SAPMA ARTIK DEĞERLERİ (DEVAM)

Gözlem	Süre	Durum	$\widehat{S}_i(t_i)$	$\widehat{S}_i(t_i^c)$	l_i^c	l_i^m	n_i^c	n_i^m
79	9	Ölmüş	0,818		0,654	0,654	0,910	0,910
80	4	Ölmüş	0,975		1,584	1,584	1,953	1,953
81	21	Yaşıyor		0,282	-0,405	-0,960	-0,576	-1,668
82	5	Ölmüş	0,957		1,353	1,353	1,722	1,722
83	19	Yaşıyor		0,303	-0,361	-0,921	-0,515	-1,503
84	19	Yaşıyor		0,303	-0,361	-0,921	-0,515	-1,503
85	13	Ölmüş	0,710		0,389	0,389	0,554	0,554
86	18	Yaşıyor		0,316	-0,335	-0,897	-0,478	-1,414
87	6	Ölmüş	0,949		1,269	1,269	1,635	1,635
88	8	Ölmüş	0,871		0,831	0,831	1,133	1,133
89	9	Ölmüş	0,818		0,654	0,654	0,910	0,910
90	17	Yaşıyor		0,323	-0,322	-0,885	-0,460	-1,374
91	17	Yaşıyor		0,323	-0,322	-0,885	-0,460	-1,374
92	17	Yaşıyor		0,323	-0,322	-0,885	-0,460	-1,374
93	16	Yaşıyor		0,329	-0,310	-0,875	-0,444	-1,339
94	16	Yaşıyor		0,329	-0,310	-0,875	-0,444	-1,339
95	16	Yaşıyor		0,329	-0,310	-0,875	-0,444	-1,339
96	15	Yaşıyor		0,334	-0,299	-0,865	-0,429	-1,309
97	15	Yaşıyor		0,334	-0,299	-0,865	-0,429	-1,309
98	13	Yaşıyor		0,355	-0,259	-0,830	-0,372	-1,204
99	11	Yaşıyor		0,400	-0,176	-0,756	-0,253	-1,029
100	12	Yaşıyor		0,380	-0,212	-0,788	-0,304	-1,099
101	12	Yaşıyor		0,380	-0,212	-0,788	-0,304	-1,099
102	11	Yaşıyor		0,400	-0,176	-0,756	-0,253	-1,029
103	11	Yaşıyor		0,400	-0,176	-0,756	-0,253	-1,029
104	10	Yaşıyor		0,405	-0,167	-0,748	-0,241	-1,015
105	10	Yaşıyor		0,405	-0,167	-0,748	-0,241	-1,015
106	8	Yaşıyor		0,436	-0,112	-0,700	-0,162	-0,928
107	8	Yaşıyor		0,436	-0,112	-0,700	-0,162	-0,928
108	13	Yaşıyor		0,355	-0,259	-0,830	-0,372	-1,204
109	9	Yaşıyor		0,409	-0,159	-0,741	-0,229	-1,001
110	6	Yaşıyor		0,474	-0,044	-0,641	-0,064	-0,843
111	14	Ölmüş	0,700		0,367	0,367	0,523	0,523
112	9	Ölmüş	0,818		0,654	0,654	0,910	0,910
113	11	Yaşıyor		0,400	-0,176	-0,756	-0,253	-1,029
114	4	Yaşıyor		0,487	-0,022	-0,621	-0,032	-0,819
115	9	Ölmüş	0,818		0,654	0,654	0,910	0,910
116	73	Yaşıyor		0,198	-0,607	-1,143	-0,848	-2,884
117	29	Ölmüş	0,451		-0,085	-0,085	-0,123	-0,123
118	7	Ölmüş	0,906		0,983	0,983	1,315	1,315

ÖZGEÇMİŞ

Kimlik Bilgileri

Adı Soyadı : Nuray TUNCER
Doğum Yeri : Ankara
Medeni Hali : Bekar
E-posta : nuray.tuncer@vdk.gov.tr
Adresi : Vergi Denetim Kurulu Risk Analiz Müdürlüğü/ Ankara

Eğitim

Lise : 2000-2004 Halide Edip Lisesi (Yabancı Dil Ağırlıklı)
Lisans : 2005-2010 Hacettepe Üniversitesi İstatistik Bölümü
Lisans : 2007-2011 Anadolu Üniversitesi İşletme Bölümü (Açıköğretim)
Yüksek Lisans : 2011-2014 Hacettepe Üniversitesi İstatistik Bölümü

Yabancı Dil ve Düzeyi

İngilizce, İyi

İş Deneyimi

2011- 2012 Gelir İdaresi Başkanlığı Gelir Uzman Yardımcısı
2012- 2014 Vergi Denetim Kurulu Başkanlığı Gelir Uzman Yardımcısı
2014- ... Vergi Denetim Kurulu Başkanlığı Gelir Uzmanı

Deneyim Alanları

Tezden Üretilmiş Projeler ve Bütçesi

Tezden Üretilmiş Yayınlar

Tezden Üretilmiş Tebliğ ve/veya Poster Sunumu ile Katıldığı Toplantılar