

T.C.
VAN YÜZÜNCÜ YIL ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ

**SAĞKALIM ANALİZİNDE AÇIKLANAN VARYANS ORANININ
İNCELENMESİ**

Sebahattin ÇELİK
BİYOİSTATİSTİK ANABİLİM DALI
(TIP PROGRAMI)
YÜKSEK LİSANS TEZİ

DANIŞMAN
Prof. Dr. Sıddık KESKİN

VAN-2018

T.C.
VAN YÜZÜNCÜ YIL ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ

**SAĞKALIM ANALİZİNDE AÇIKLANAN VARYANS ORANININ
İNCELENMESİ**

Sebahattin ÇELİK
BİYOİSTATİSTİK ANABİLİM DALI
(TIP PROGRAMI)
YÜKSEK LİSANS TEZİ

DANIŞMAN
Prof. Dr. Sıddık KESKİN

VAN-2018

T.C.
VAN YÜZÜNCÜ YIL ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ

**SAĞKALIM ANALİZİNDE AÇIKLANAN VARYANS ORANININ
İNCELENMESİ**

Sebahattin ÇELİK
BİYOİSTATİSTİK ANABİLİM DALI
(TIP PROGRAMI)
YÜKSEK LİSANS TEZİ

Prof. Dr. Sıddık KESKİN
(Jüri Başkanı)

Prof. Dr. Sinan ÇALIK
(Üye)

Prof. Dr. H. Eray ÇELİK
(Üye)

TEZ KABUL TARİHİ
03/08/2018

TEŐEKKÜR

Çalıőmanın planlanması ve yürütülmesinde yardımcı olan Yüzüncü Yıl Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı öğretim üyesi, tez danışmanım Sayın Prof. Dr. Sıddık KESKİN'e çok teşekkür ederim. Biyoistatistik' e ilgimin ve öğrenmemin en önemli motivasyon kaynağı olan Prof.Dr. Hayrettin Okut hocama teşekkür ederim. Biyoistatistik derslerini beraber aldığım ve dostluklarından zevk aldığım tüm yüksek lisans öğrencisi arkadaşlarıma, ayrıca, her zaman, her kararımda bana destek olan eşime ve kızım Lorin' e teşekkür ederim.



İÇİNDEKİLER

KABUL VE ONAY	III
TEŞEKKÜRLER.....	IV
İÇİNDEKİLER.....	V
SİMGELER VE KISALTMALAR.....	VI
ŞEKİLLER LİSTESİ	VII
TABLolar LİSTESİ.....	VIII
1. GİRİŞ.....	1
2. GENEL BİLGİLER.....	3
2.1. Tanım.....	3
2.1.1. Sansürlü veri.....	3
2.2. Tarihçe.....	5
2.2.1 Demografi ve yaşam istatistiklerinin temelleri.....	6
2.2.2 Sağkalım analizinin yirminci yüzyıldaki gelişimi	7
2.3. Sağkalım analiz yöntemleri.....	8
2.3.1 Yaşam tabloları (Life tables).....	9
2.3.2. Kaplan – Meier yöntemi.....	12
2.3.3. Farklı sağ-kalım dağılımlarının karşılaştırılması	15
3. GEREÇ VE YÖNTEM.....	19
3.1. Gereç	19
3.2. Yöntem.....	20
3.2.1 Cox'un oransal riskler modeli	20
3.2.2 Risk (hazard) fonksiyonu	20
3.2.3. Cox oransal risk modeli.....	23
3.2.4.Cox oransal risk modelinde katsayıların tahmin.....	24
3.2.5. Risk oranının (RO) hesaplanması.....	25
3.2.6. Cox modelinde açıklanan varyans	26
3.2.7. Cox OR modelinde açıklanan varyans ölçüsünün özellikleri.....	27
3.2.8. Korn-Simon ölçüm tekniği	28
3.2.9. Kent ve O'Quigley'in bağımlılık ölçüm tekniği.....	30
3.2.10. Schemper'in V1 veV2 ölçüm yöntemleri.....	32
3.2.11. Schemper-Henderson ölçüm yöntemi.....	33
4. BULGULAR.....	35
5. TARTIŞMA VE SONUÇ.....	39
ÖZET.....	42
SUMMARY	43
KAYNAKLAR	44
ÖZGEÇMİŞ	46
EK 1. SCHEMPER-HENDERSON VE KENT-O,QUINLY İÇİN SAS KODLARI.	47
EK 2. ETİK KURUL BELGESİ.....	49
EK 3. TEZ ORJİNALLİK RAPORU.....	50

SİMGELER VE KISALTMALAR

Σ	: Toplam sembolü
n	: Örneklem içindeki birey sayısı
Y	: Bağımlı değişken
X	: Bağımsız değişken
β	: Regresyon katsayısı
X_j	: Sürekli bir değişken
\bar{y}_j	: j değişkeni için genel ortalama
e	: Hata terimi
ML	: En çok olabilirlik (Maximum likelihood)
R^2	: Korelasyon katsayısı (belirleme katsayısı)

ŞEKİLLER LİSTESİ

Şekil 1. Sansürlü sağ-kalım süresinin gösterimi	4
Şekil 2. Eklemeli sağkalım oranları	12
Şekil 3. Kaplan-Meier sağkalım grafiđi	14
Şekil 4: Sabit ölüm hızı grafiđi.	21
Şekil 5. Zamanla artan ölüm hızı grafiđi	22
Şekil 6. Zamanla azalan ölüm hızı grafiđi	22
Şekil 7. Zamanla deđişkenlik gösteren ölüm hızı modeli (Log-normal)	22
Şekil 8. Kaplan-Meier yöntemine göre sağkalım grafiđi.	37



TABLULAR LİSTESİ

Tablo 1. Sansürlü veri olmadığında yaşam tablosu	10
Tablo 2. Sansürlü veri ile yaşam tablosu	11
Tablo 3. Kaplan-Meier yöntemi ile sağ-kalım hesaplanması	13
Tablo 4. Yemek borusu kanserli olguların verilerinin genel bir özeti	19
Tablo 5. Cox regresyonda cerrahi yönleme ilişkin sonuçlar	35
Tablo 6. Cox regresyonda cinsiyet ve cerrahi yönleme ilişkin sonuçlar	35
Tablo 7. Sağkalım süresine etkili olduğu düşünülen faktörlere ilişkin analiz sonuçları (En çok olabilirlik analizi)	36
Tablo 8. Schemper-Henderson yöntemine göre hata tahminleri ve açıklanan varyans oranları	37
Tablo 9. Kent ve O'Quigley yöntemi ile açıklanan varyans oranı	38

1. GİRİŞ

Sağkalım analizi, istatistiğin bir dalı olup, bir veya daha fazla olayın belirli bir zaman süreci içerisinde beklenen ömrünü veya süresini analiz etmede kullanılan yöntemlerden birisidir. Biyolojide organizmaların ölümünü veya mekanik sistemlerde herhangi bir cihazın bozulma veya arıza verme olasılığını belirlemede yaygın olarak kullanılmaktadır. Sağkalım analizi, mühendislik bilimlerinde güvenilirlik teorisi ya da güvenilirlik analizi olarak adlandırılırken, ekonomide süreç modelleme veya Süreç analizi, sosyolojide ise olay tarihi analizi olarak adlandırılmaktadır. Sağkalım analizi belirli bir geçen zaman sürecinde, popülasyondaki sağkalım oranını bulmayı ve buna etki eden değişkenleri belirlemeyi amaçlar. Sağkalım, ölme ya da bozulma ile ilişkili olarak, buna etkili olan çoklu nedenleri belirlemede kullanılabilir. Bunun için de yaşam süresinin ya da ömür tanımlanması gerekmektedir. Biyolojik olaylarda ömür, yaşam süresi olarak tanımlanabilirken, mekanik olaylarda güvenilirlik süresi ya da bozulma süresi olarak tanımlanabilir. Biyolojik olaylarda, ayrıca kalp krizi geçirme riski, herhangi bir organ yetmezliği riski, evli çiftlerin boşanma riski, tutukluların hapisaneden kaçma riski veya hapisaneden çıkmış birisinin yeniden hapisaneye girme riski bu analiz yöntemi ile değerlendirilebilir.

Sağkalım analizlerinde bazı terimlerin açıklanmasında yarar görülmektedir. Olay denilince, hastalığın ortaya çıkması, hastalığın yeniden nüksetmesi veya hastanın ölmesi akla gelmektedir. Zaman denilince, ameliyattan sonra nükse kadar geçen zaman, tedaviye başlanan andan ölümün görülmesine kadar geçen zaman gibi herhangi bir gözlem sürecinde olayın ortaya çıkmasına kadar geçen “süre” akla gelmektedir.

Açıklanan varyasyonu belirlemede kullanılan ölçüler, bir modele ilişkin bilginin miktarını ve özellikle modelin gelecekte yeni gözlemlere ilişkin yapılacak tahminlemedeki doğru tahmin yapma oranını belirlemek amacıyla kullanılan ölçülerdir. Bu ölçüler, standart regresyon modellerinin yanı sıra, genelleştirilmiş doğrusal modeller ve sağkalım analizinde de kullanılmaktadır.

Sağkalım analizinde, yaşam süresi üzerine etkili olan faktörleri belirlemek için Cox regresyon analizi kullanılmaktadır. Sansürlü verileri analiz eden Cox regresyonda bağımlı değişkendeki varyasyonun (Sağkalım analizlerinde bağımlı değişken zamandır) ne kadarının açıkladığı, süregelen bir tartışma olmuştur (Akawaza, 1997; Schemper ve Stare, 1996). Doğrusal ve Lojistik regresyon modellerinde çoklu korelasyon katsayısının karesi (R^2)

model tarafından açıklanan varyansın bir ifadesi olarak kullanılırken, bu ölçünün (R^2) analogu olan bazı yöntemler, Cox regresyonu için de önerilmektedir (Schemper ve Stare,1996). Bağımlı deęiřkendeki varyasyonun ne kadarının, kurulan Cox regresyon modeli ile açıklandığını inceleyen farklı yaklaşımlar bulunmaktadır (Graf ve Schumacher,1995)

Bu çalışmada, sağkalım analizlerine genel olarak değinildikten sonra, bu analizlerde açıklanan varyans oranını inceleyen testler irdelenecektir.



2. GENEL BİLGİLER

2.1. Tanım

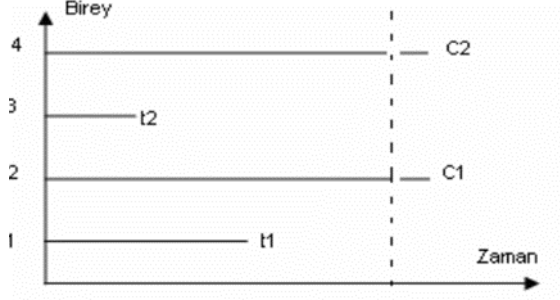
Genel ve en basit ifadeyle, Sağkalım analizi (SA), herhangi bir olayın görülmesine kadar geçen zamanı analiz etmektedir. Bu zaman, sağlık bilimlerinde, bir hastanın ölümüne ya da hastalığının tekrar etmesine kadar geçen süre iken, ekonomide bir iş yerinin faaliyet gösterebildiği süre olmaktadır. Sağkalım analizinde olayın ortaya çıkma zamanı ile birlikte, bu zamana etki eden faktörler de incelenebilmektir (Cox, 1972). Bazen ilgilenilen olayın (örneğin; hastanın ölümü, hastalığın nüks etmesi, alınan bir cihazın bozulması, evlilerin boşanması vb.) görülmesine kadar geçen süreyi etkileyen özellikleri incelemek için de, farklı özellikteki grupların aynı olayı deneyimlemesine kadar geçen süreyi karşılaştırmak için de sağkalım analizlerine ihtiyaç duyulmaktadır.

Standart regresyon analizi modellerinden farklı olarak sağkalım analizleri, sansürlü ve sansürlü olmayan gözlemlere ilişkin bilgilerden yararlanarak, önemli model parametrelerini tahmin edebilmektedir. Sağkalım analizinde cevap değişkeni iki kısımdan oluşmaktadır. Birinci kısım, olayın meydana geldiği ya da ortaya çıktığı zamanı ifade ederken, İkinci kısım, olayın ortaya çıkıp çıkmama durumunu göstermektedir.

Sağkalım verilerine, neden standart regresyon modeli veya lojistik regresyon modelinin uygulanamadığı aşağıda açıklanacaktır.

2.1.1. Sansürlü veri

Sansürlü veri (Censored data); çalışmaya dahil edilen bireylerin bazılarında herhangi bir nedenle veri elde edilememesi durumunda ortaya çıkmaktadır. Kısaca, sağkalım süresi tam olarak bilinmeyen bireylere ilişkin verilerdir. Sansürlü veriler genel olarak 3 şekilde ortaya çıkmaktadır. Birincisi; çalışmanın bitiş noktasına kadara gözlem yapılamaması, ikincisi çalışma bitmeden deneklerden bilgi alınamaması, üçüncüsü ise başka bir nedenle deneğin çalışmadan çekilmesidir.



Şekil 1. Sansürlü sağkalım süresinin gösterimi

Şekil 1’de görüldüğü üzere, bireylerin ölüm olayları incelenmiş ve belirlenen süre boyunca t_1 ve t_2 kişisi ölmüş (kesin veri, tarih belli) ancak C1 ve C2 kişileri belirlenen zamanda halen yaşamakta olduğu ve bu bireylerin ne zaman öleceği bilinmediği için bu kişilerin verileri sansürlü veri olarak kabul edilmektedir.

Sansürlü veri, yalnızca sağkalım ile ilgili değil, başka alanlarda da gözlenebilmektedir. Örneğin, herhangi bir konsere olan ilginin satılan bilet âdetince belirlendiği düşünüldüğünde, biletlerin bitmesi durumunda, satılan bilet adedi sansürlenmiş olmaktadır. Zira, bilet olması durumunda halen satış devam edebilirdi. Benzer şekilde, herhangi bir cihazın ölçü yaptığı (mesela tartı aleti) en alt ve en üst seviye dışındaki gözlemler de sansürlenmiş olmaktadır.

Sansürlü verinin; sağdan sansürlü veri, soldan sansürlü veri ve aralık sansürlü veri olmak üzere 3 tipi vardır. Başarısızlık olarak adlandırılan (ölüm, bozulma, hastalığın yeniden nüks etmesi vb.) olay, çalışma için belirlenen zaman aralığında gerçekleşmezse, bireyin yaşam süresinin uzunluğu, çalışmanın bitiş zamanının sağ tarafında kalır. Böyle bir durumda, bu bireyin yaşam süresi kesin olarak bilinmeyecektir. Bu tip sansürlü veri “sağdan sansürlü veri” olarak adlandırılır. Bireyin yaşam süresi, çalışmanın sonlanma zamanından küçük olması durumunda, soldan sansürlü veri olarak ifade edilir. Aralık sansürlü veri, genelleştirilmiş bir sansürlü veri çeşididir. Genellikle takip gerektiren olaylarda kullanılır. Çalışmaya konu olan olayın meydana gelme süresi bir aralıkta ifade edilir.

Araştırmacılar, sansürlü verileri kayıp veri olarak almakta ya da sansürlü veriye sıfır kodunu vermektedir. Bunun yanı sıra, minimum değer, maksimum değer ya da ortalama bir değer verilmektedir (Soley Bori, 2013). Sansürlü gözlem sayısının az olması durumunda, bu adı geçen yöntemlerden birinin denenmesi uygun olabilir. Ancak, sansürlü gözlem sayısının çok olması durumunda, bu yöntemler belirgin yanlılığa (bias) neden olabilmektedir (Soley Bori, 2013).

Lojistik regresyon analizi ile risk faktörlerinin, ilgili cevap değişkeni (örneğin hastalığın olup olmaması) üzerine etkileri incelenebilmektedir. Ancak, araştırılan konu *zaman* ise regresyon analizleri uygun yöntem olmamaktadır. Zira, zaman değişkeni ile ilgili bilgi tam değildir. Belli bir hastalıktan dolayı ölüm süreleri araştırılmak isteniyorsa ve izlem süresi mesela beş yıl ise ve bu süre boyunca çalışmaya alınan bütün gözlemler (hastalar) öldüyse, bu durumda sağkalım sürelerini etkileyen faktörler lojistik regresyon analizleri ile incelenebilmektedir. Ancak, hastaların bir kısmı belirlenen takip süresinin sonunda halen yaşayabilmektedir. Yani, yukarıda da açıklandığı gibi verilerin sansürlü olduğu durumlarda, lojistik regresyon modelleri yerine sağkalım analizlerinin kullanılması daha uygun olacaktır.

2.2. Tarihçe

Sağkalım analizinin temelini anlamak ve tarihi süreçteki değişimini belirlemek üzere, önce demografi ve yaşam istatistiklerinin tarihi sürecine kısaca değinilecektir.

Yirminci yüzyıl sonlarına ve yirmi birinci yüzyıl başlarına kadar geçen sürede, istatistik bilimi genel olarak aşağıda sıralanan 4 alan etrafında gelişmiştir:

- Olasılık teorisi ve matematiksel istatistik
- Hataların ve belirsizliğin incelenmesi (epidemiyolojik çalışmalarda hatanın irdelenmesi)
- Deneysel çalışmaların düzenlenmesi ve örnekleme çalışmaları
- Gözlemsel çalışmaların verilerinin toplanması, özetlenmesi ve yorumlanması

Birçok istatistik tarihçisi, modern *olasılık teorisinin* asıl kaynağını bazı şans oyunlarına dayandırmaktadır (Stigler, 1968; Westergaard, 1932). Şans oyunlarında karşılaşılan problemlerle, Pacioli (1494), Cardano (1539), ve Forestani (1603) uğraşmış ancak tam bir çözüm Pierre De Fermat (1601-1665) ve Blaise Pascal (1623-1662) tarafından sağlanmıştır (Feinleib, 2005).

Verilerin toplanması ve gözlem değerlerinin tablolar halinde sunulmasına sistematik yaklaşıma, *Francis Bacon* (1561-1626) önemli katkı sunmuştur. *Novum Organum* (1620) adlı çalışmasında, Aristo metoduna dayalı gelişen orta çağ skolastik düşüncüyü eleştirmiştir. Bacon'un yaklaşımından en erken etkilenen alanlar *demografi* ve *yaşam istatistikleri* (vital statistics) alanları olmuştur (Feinleib, 2005).

2.2.1. Demografi ve yaşam istatistiklerinin temelleri

Antik çağlardan beri, askeri amaçlaGr için vergi ve haraç toplama amacıyla, belirgin aralıklarla nüfus ile birlikte diğer varlıkların sayımı ve değerlendirilmesi yapılmaktaydı. Ancak, Roma İmparatorluğu'nun yıkılmasından sonra, tüm devleti kapsayan geniş sayımlar 18. yüzyıla kadar yapılmamış ve bu süre içerisinde, daha çok kiliseler tarafından, belirgin aralıklarla ölümler, doğumlar, evlilikler, vaftiz ve defin işlemlerine ilişkin kayıtlar tutulmuştur. Londra, bu tür verilerin 1538 'li yıllarda ilk kez düzenli olarak tutulduğu şehirdir. Yine yaklaşık aynı dönemlerde Londra'da, veba salgınlarına karşı erken uyarı sistemi amacıyla yapılan “*tarama*” programları yapılmıştır (Hald, 1990). O dönemlerde, olası veba salgınını erken fark etmek için, kilise kâtipleri haftalık ölüm sayısını ve sebeplerini bildirmiştir. Bununla ilişkili olarak Londra'nın politik ve entelektüel camiasında etkili olan John Graunt (1662),“*Ölüm faturaları üzerine politik ve doğal gözlemler*” (Natural and Political Observations Made Upon the Bills of Mortality) adında önemli bir kitap yayınlamıştır. Bu kitap birçok istatistik tarihçisi tarafından, “demografik verilerin istatistik analizi için mükemmel bir kitap” olarak nitelendirilmektedir (Hald, 1990; Greenwood, 1941). Hald (1990) kitabında, Graunt'un katkılarını şöyle özetlemektedir; “Graunt, ölüm olayını, sebepleri açısından değerlendirerek, aynı büyüklüğü farklı yöntemler ile tahmin etmiş ve istatistik oranların sabitliğine ve sağkalım tablolarına (istatistik mantığı çerçevesinde) yeni standartlar getirmiştir. Graunt'un çalışması 3 farklı araştırma tipine yol açmıştır. Bunlar; Politika aritmetiği, İstatistik oranların sabitliği testi ve Sağkalım beklentisinin hesaplanması ile yaşam olasılığının hesaplanmasıdır (Greenwood, 1941).

Popülasyon genelinde yaşa göre ne kadar insanın yaşadığının bilinmediği yıllarda, Graunt oldukça gelişmiş yöntemleri kullanmış ve sağ-kalım tablolarını geliştirerek mantıklı tahminlerde bulunmuştur. Graunt'un çalışmalarının geniş etkisi ile önce Paris'te (1667) daha sonra Avrupa'nın diğer şehirlerinde, *ölüm faturalarına* benzer uygulamalar başlamıştır. Graunt'un yaşam tabloları, daha sonra Christiaan Huygens (1629–1695) ve kardeşi Ludwig (1631–1699) tarafından gündeme getirilmiştir. Ludwig, (Hollandalı) yaşam tablolarından olasılık hesabını yapmıştır ki bu olasılık hesabını İsveçli Nicholas Bernoulli (1687–1759) de bağımsız olarak yapmıştır. Bu araştırmacılar ile birlikte, Edmond Halley (1656–1742), Breslau şehrindeki doğum ve merasim verilerini kullanarak, Deparcieux (1703–1768) ise Fransa'da ki verileri kullanarak sigortacılık mesleği için Aktüerya (Aktüer; finansal riskleri değerlendirebilen, çözümler öneren ve her çözümün uzun dönemdeki sonuçlarını

irdeleyebilen bir profesyoneldir) bilimini kurmuşlardır. Bu katkılar daha sonra Abraham DeMoivre (1667–1754), Thomas Simpson (1710–1761), Benjamin Gompertz (1779–1865) ve William Makeham (1826–1891) tarafından geliştirilmiştir (Greenwood, 1941).

Ülke genelinde ölüm tabloları 1766 yılına kadar yapılamamıştır. Per Wargentın (1717–1783), ilk kez İsveç'teki doğumları ve ölümleri numaralandırarak, mortalite tablolarını yayınlamış ve daha sonra bu tablolardan yararlanılarak, toplum genelinde kadınların ölüm hızının erkeklerden az olduğu hesaplanmıştır (Feinleib, 2005).

Almanya'da Luterian Ortodoks inancına göre kanıt niteliğindeki tüm ulaşılabilen verilerden, yaşam istatistiklerini toplayan Johann Peter Suessmilch (1707–1767), şunu belirtmiştir; “sağ kalım beklentisi sabittir ve ölüm oranlarını iyileştirmek için yapılabilecek pek bir şey yoktur” (Feinleib, 2005). Suessmilch'in çalışmalarından etkilenen Thomas Robert Malthus'un (1766–1834) katkıları ile birlikte yapılan geniş çabalar sonucunda 19. yüzyılda devlet istatistik kurumları oluşmuştur (Feinleib, 2005).

18. yüzyılda matematiksel istatistik alanında görülen gelişmelerden ikisi, sağlık istatistiği alanı için oldukça önemlidir. Bunlardan birincisi, Bernoulli ailesinden Daniel'in (1700–1782) binomiyal dağılıma normal yaklaşımı ortaya koymuş olmasıdır. Bu yaklaşımı ile doğumda cinsiyet oranlarının sabitliğini ve mortalite hızını sabit sürekli fonksiyonlar kabul ederek, ölüm oranlarını incelemiştir. Daniel Bernoulli'nin bu çabaları ona, 1760 yılların önemli bir halk sağlığı sorunu olan çiçek hastalığının ortadan kaldırılmasına imkânı vermiştir. İkincisi ise Laplace'nin 1786 yılında önerdiği; rastgele olmayan örnekleme ile toplumun (popülasyonun) toplam nüfusunun tahmin edilebileceği öngörüsüdür. Bu öngörü günümüzde de kullanılan *oran tahminine* benzemektedir. Şöyle ki; “herhangi bir bölgedeki toplam nüfus, o bölgedeki yıllık doğum oranları ile orantılıdır ve bu sabit oran, bu bölgeden alınmış bir alt bölge örneğinden hesaplanabilir”. Aslında bu varsayıma benzer bir mantığı, daha önce Graunt kullanmıştır (Bollet, 1973; Feinleib, 2005).

2.2.2. Sağkalım analizinin yirminci yüzyıldaki gelişimi

Sağkalım analizi, ikinci dünya savaşından sonra güncel hale gelmeye başlamıştır. Bunun en önemli nedenlerinden birisi, bu analiz yönteminin, yalnızca sağlık ya da politika alanına özel bir yöntem bilimi olmadığına farkına varılmış olması olarak yorumlanmaktadır (Feinleib, 2005). Özellikle savunma sanayinde, üretilen silahların güvenilirliği ile ilgili analizler artmaya başlamıştır. Üretilen bir ürünün ömrünün ne kadar olduğu ve garanti

belgesinde belirtilen sürenin yeterli ve güvenilir olup olmadığı gibi sorular, sağkalım veya güvenilirlik analizi (Reliability analysis) konularını ortaya çıkarmıştır (Anonymous 2013). Sağkalım analizi için 1958 yılı önemli bir dönem noktası olarak kabul edilmektedir. Amerika’da yaşayan iki araştırmacı, Edward L. Kaplan ve Paul Meier, günümüzde sıklıkla kullanılan ve parametrik olmayan yöntemler arasında yer alan Kaplan-Meier yöntemini geliştirmişlerdir (Kaplan ve Meier, 1958). Kaplan-Meier yöntemi, (Kaplan-Meier survival analysis) sansürlü veri setinden, belli bir zaman için sağ kalım olasılığını hesaplamaktadır.

Kaplan-Meier yönteminde, eksik görülen, sağ-kalım süresini ya da beklenen olayın gerçekleşmesine kadar geçen zamanı, etkileyen faktörlerin belirlenememesi sorunu daha sonra David Cox tarafından ele alınmış (Cox, 1972) ve kendi adıyla bilinen Cox regresyon analizi yöntemi geliştirilmiştir. Yöntemde aynı zamanda, ölüme kadar geçen süreyi etkileyen faktörler bakımından gruplar arasında fark olup olmadığı da incelenebilmektedir (Cox, 1972).

Cox regresyon modelinin belirlediği faktörlerin, bağımlı değişkendeki toplam değişimin ne kadarını açıkladığı sorusu sonraki yıllarda gündeme gelmiş ve birçok araştırmacı (Harrell, 1982; Kent ve O’Quigley 1988; Kornve Simon 1990; Schemper,1990) farklı yöntemler önermiştir. Ancak günümüzde halen yaygın kabul görmüş bir yöntemin olmadığı söylenebilir

2.3. Sağkalım Analizi Yöntemleri

Sağkalım analizinin amacı kısaca aşağıdaki gibi özetlenebilir:

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

Sağkalım analizi için ilgilenilen amaca göre önerilen yaklaşımlar aşağıdaki gibi özetlenebilir:

- Başarısızlık sürelerinin bağımlı değişken olarak kabul edildiği ve bu değişkeni açıkladığı düşünülen diğer değişkenlerin de çözümlemede yer aldığı modellerin kullanıldığı yaklaşım.

Bu yaklaşım için iki model düşünülebilir. Bunlar; Bağımsız değişkenlerin risk fonksiyonu üzerinde çarpımsal etkiye sahip olduğu orantılı risk modelleri ve açıklayıcı değişkenlerin yaşam süresinin logaritması üzerinden toplanmasıyla yaşam süresi üzerinde çarpımsal etkiye sahip olduğu log doğrusal modellerdir (Kleinbaum ve Klein, 2005).

- Çeşitli parametrik yaşam dağılımlarını kullanarak tahminlerde bulunmak ve hipORez testlerini hazırlamak,
- Herhangi bir dağılım varsayımına dayanmayan parametrik olmayan süreçleri kullanarak tahminlerde bulunmak (Collett, 2003)

Sağkalım analizi, ilgilenilen herhangi bir olayın ortaya çıkmasına kadar geçen sürenin incelenmesinde kullanılmakta olup, analizde geçen bazı önemli kavramlar aşağıdaki gibidir:

Geçen zaman (süre): Araştırmanın başından ilgilenilen olay gerçekleşinceye kadar ya da olayın gerçekleştiği andaki zamanı (yıl, ay, hafta, gün, saat, vb.) ifade etmektedir.

Olay (event): Ölüm, hastalığa yakalanma, hastalığın kötüye gitmesi, iyileşme, işe dönme, hapishaneden kaçma, boşanma, çürüme ve bozulma gibi ilgilenilen herhangi bir olayı ifade etmektedir (Kleinbaum ve Klein, 2005) .

Yaşayan bir organizma ya da cansız bir nesne için belirli bir başlangıç zamanı ile ilgilenilen olay arasında geçen zamana “sağkalım (yaşam) süresi” ya da “başarısızlık süresi” denilmektedir. Her bir birime ait yaşam süresi, tanımı gereği pozitif yönde sürekli bir değere sahiptir ve herhangi bir birim için başarısızlık ancak bir kez oluşabilir. İlgilenilen olaya göre farklı şekillerde ortaya çıkabilen bu değişkenin (sağkalım süresi) duyarlı olarak ölçülmesi için;

- İzleme alınan her bireyin ya da nesnenin, çalışmaya alındığı başlangıç zamanı şüpheye yer vermeyecek şekilde açık ve net olmalıdır.
- İzlem süresini tanımlayacak bir zaman birimi (gün, ay, yıl vb.) kabul edilmelidir.
- Çalışmaya alınan her birimin başarısızlık (olay anı) zamanı doğru ve net olmalıdır (Kleinbaum, ve Klein, 2005).

2.3.1 Yaşam tabloları (Life tables)

Yaşam tablolarının (Life tables) ilk örnekleri 1660’lı yıllarda Graunt tarafından kullanılmış ve yaşa göre herhangi bir popülasyonda ne kadar insan yaşadığı tahmin edilmeye

çalışılmıştır (Greenwood, 1941). Daha sonraki yıllarda da yaygın olarak sadece nüfus biliminde kullanılmıştır. Ancak 1940'lı yıllardan sonra Botanik ve Zooloji gibi birçok farklı alanda kullanılmaya başlanmıştır (Anderson, 1981).

Yaşam (sağkalım) tabloları toplumun doğum, ölüm, nüfus, göç, ölüm hızı, yaşam hızı gibi özelliklerini incelemek için kullanılmaktadır. Özellikle sigortacılık alanında aktüerya, belli bir toplumda herhangi bir bireyin yaşam süresini ve o süreye kadar yaşama ihtimalini hesaplamak için başvurulan bir yöntemdir.

Tıp alanında genelde merak edilen olay “ölüm” ya da “hastalığın nüksü” olmaktadır. Hastanın tanı aldıktan sonra ne kadar süre daha yaşama olasılığı olduğu genellikle hem toplum hem de bireyin kendisi tarafından merak edilir. Bu bağlamda herhangi bir hastalığın tanısından sonraki yaşam süresi, yaşam tabloları yöntemi ile incelenebilir. Çalışmaya alınan hasta popülasyonunun başlama zamanı (izleme hangi aralıkta girdiği), çalışmadan çıkan ya da kayıp olan ya da ölen hasta sayıları ve zamanları, yaşam tabloları ile belirlenebilmektedir. Yine yaşam tabloları ile eklemeli yada birikimli (cumulative) yaşam olasılıkları hesaplanarak, sağkalım grafikleri oluşturulabilmektedir.

Bu yöntemde, öncelikle araştırmacı tarafından takip için *zaman aralığı* belirlenir. Daha sonra sonuçlar bu belirlenen zaman aralıkları çerçevesinde değerlendirilir.

I_t = t zamanının başlangıcında yaşayan birey (hasta) sayısı,

d_t = Belirlenen zaman aralığı içerisinde ölen birey sayısı,

q_t = Belirtilen zaman aralığı süresince ölme ihtimali = d_t / I_t ,

p_t = Belirtilen zaman aralığı süresince yaşama ihtimali = $1 - q_t$,

olmak üzere, sansürlü veri olmadığı duruma ilişkin yaşam tablosu örneği Tablo 1' de verilmiştir.

Tablo 1. Sansürlü veri olmadığında yaşam tablosu

İzlemin başlamasından itibaren geçen süre (yıl)	Başlangıçta mevcut kişi sayısı (I_t)	Ölüm sayısı (d_t)	Aralık süresince ölme ihtimali (q_t)	Yaşama ihtimali (p_t)	P_t
1 (0-1 Aralığı)	200	20	0.1 ($d_t / I_t = 20/200$)	0.9 ($p_t = 1 - q_t$) ($1 - 0.1$)	0.9 (0.9×1.0)
2 (1-2 Aralığı)	180*	30	0.17	0.83	0.75 (0.83×0.9)
3 (2-3 Aralığı)	150**	40	0.27	0.73	# 0.548 (0.73×0.75)

* $200 - 20 = 180$ (2. aralık başındaki yaşayan birey sayısı) ** $180 - 30 = 150$ (3. aralık başındaki yaşayan birey sayısı)

#: $P_2 = p_2 \times P_1$

Belirtilen aralık sonundaki (ya da bir sonraki aralık başındaki) eklemeli sağkalım olasılığı $P_t = p_t \times P_{(t-1)}$ olmaktadır. Çalışmanın başlangıcında yani birinci aralıkta eklemeli sağkalım olasılığı (P_1), tüm hastalar yaşadığından %100'dür ($P_1 = p_1 \times P_0$ olur ki bu da en başta $1.0 \times 1.0 = 1$ dir). İkinci aralıkta; $P_2 = p_2 \times P_1 = 0.9 \times 1.0 = 0.9$ olmaktadır. Üçüncü yılın sonunda (yani dördüncü yılın başında) ise sağkalım olasılığı; $0.73 \times 0.75 = 0.548$ olmaktadır.

En yalın haliyle yaşam tabloları, yukarıda özetlendiği gibi olsa da gerçek yaşamda böyle olmamaktadır. İzleme alınan bireylerin bir kısmı kendi istekleri ile çalışmadan çıkmakta, bir kısmı gelişen komplikasyonlar ya da yan etkiler nedeniyle çalışma dışı bırakılmakta, bir kısmı ise izlem sonunda halen yaşamakta olduğu için hesaplamaya alınamamaktadır. Bu tür sansürlü verilerin olduğu yaşam tablosu Tablo 2' de özetlenmiştir.

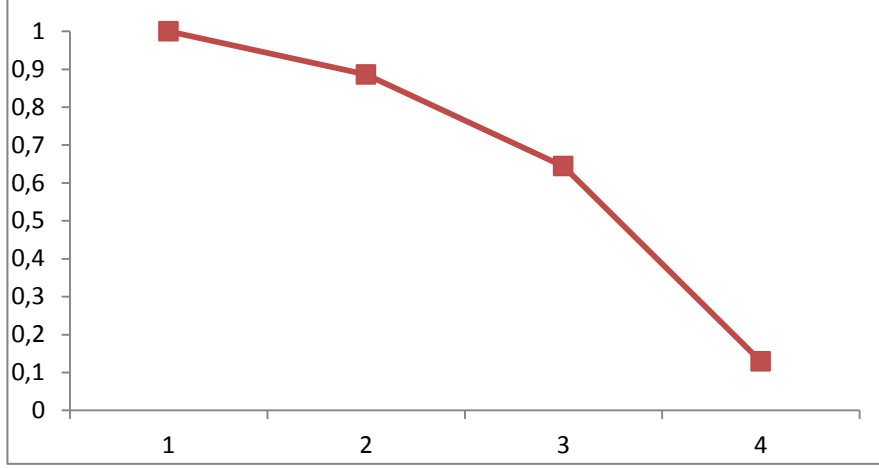
Tablo 2. Sansürlü veri ile yaşam tablosu

Aralık	I_t	d_t	w_t	I'	$q_t (= d_t/I')$	p_t	P_t
1	200	20	50	175(200-50/2)	0.114 (20/175)	0.886	0.886
2	130	30	40	110	0.273	0.727	0.644
3	60	40	20	50	0.80	0.20	0.129

Yaşam tablosu yönteminde, izlem aralığında sansürlenene bireylerin ortalama yarısı kadarının aralık süresince takipte olduğu kabul edilir. Dolayısıyla ilgilenilen aralık süresince ölüm riski altında olan toplam birey sayısı, sansürlü birey sayısının yarısı kadar azalmış olarak kabul edilir.

w_t , aralık süresince çalışmadan çıkan/ayrılan (sansürlü vaka sayısı) sayısı ve I , aralık süresince risk altında olan birey sayısı olmak üzere; $I' = I_t - w_t/2$ olur. Bu sayıya düzeltilmiş birey sayısı (risk altındaki) denilmektedir. Yeni ölme risk ise; $q_t = d_t/I$ olmaktadır.

Bu yaşam tablosunda verilen 3 yıllık takip sonunda, eklemeli sağkalım oranları grafiği Şekil 2' de gösterilmiştir.



Şekil 2. Eklemeli sağkalım oranları

Yaşam tablosu yönteminin varsayımları aşağıdaki gibi özetlenebilir.

Herhangi bir aralıkta ölen hastalar dışındaki hastalar, bir sonraki sınıfa eşit ölüm riski veya yaşam olasılığı ile geçer.

İzlemdeki tüm birimlere (çalışmadan çıkanlara, kayıplara ya da takipte kalanlara) aynı ortamın sağlanmış olması gerekir.

Beklenen olay (örneğin ölüm), düzgün olarak önceden belirlenen zaman aralıkları içerisinde görülmelidir.

Çalışmadan çıkan/kayıp birimlerin de çalışmadan çıkma zamanları belirlenen aralıklar içerisinde olmalıdır.

Yaşam tablosu yönteminin örneklem genişliği büyük olan çalışmalarda kullanılması önerilmektedir (Cox, 1972).

2.3.2. Kaplan – Meier yöntemi

Kaplan ve Meier (1958) tarafından geliştirilen bu yöntem, sağkalım süresinin tahmin edilmesinde, yaşam tablolarına benzerlik göstermektedir (Kaplan ve Meier, 1958). Yaşam tablolarından en belirgin farkı; aralık tahminleri yerine, olayın gerçekleşmesine kadar geçen gerçek sürenin kullanılmasıdır. Yani, yaşam tablosunda hesaplamalarda aralıklara karşılık gelen olay sayılarının frekans dağılımı kullanılırken, Kaplan-Meier (K-M) tahmin edicisinde, tam olarak geçen süre verileri kullanılmaktadır.

İlgilenilen olayın gerçekleşme olasılığı, belirlenen zamanda görülen olay sayısının, risk altında olan toplam sayıya bölünmesi ile hesaplanır. Olaydan önce çalışmadan çıkan birey varsa bu sansürlü birey sayısı, toplam risk altında olan sayıdan çıkarılır. Bu yöntemde hayatta

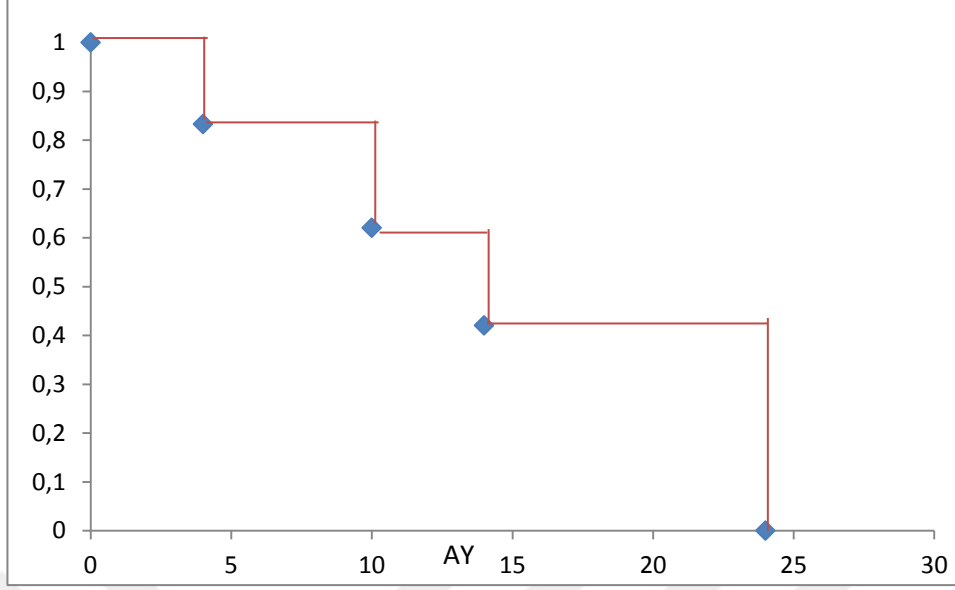
kalma süresi her bir ölüm gerçekleştiğinde hesaplandığı için tedaviden geri çekilen ya da takibi yapılamayan hastalar dikkate alınmamaktadır. K-M tahmin edicisinin az bireyin olduğu çalışmalarda da rahatlıkla kullanılabilmesi, yaşam tablolarına göre ayrı bir üstünlüğü olarak kabul edilmektedir.

K-M yöntemi aşağıdaki küçük örnekte kısaca açıklanmıştır. Mide kanseri tanısı almış 6 hastanın 60 ay boyunca takip edildiği varsayalım. Sonuç değişkeni olan hastaların ölümü, 4 hastada gözlenmiş ve olayın gözlenme zamanları; 4, 10, 14 ve 24. aylar olsun. İki hasta ise sansürlü ve birincisi için sansür 10. aydan önce, ikincisi için ise 24. aydan önce olsun. Oluşturulacak tablo, yaşam tablolarına benzeyecektir. Ancak bu tabloda olayın gerçekleşme zamanları (4, 10, 14, 24. aylar) tam olarak alınacağı için toplam 4 sıra olacaktır (Tablo 3). Yaşam tablosunda belirttiği gibi diğer parametreler (q_t , p_t ve P_t) benzer şekilde hesaplanır. Eklemeli sağkalım olasılığı (P_t), olayın gerçekleştiği zamana kadar (diyelim ki bu örnekte ilk ölüm 4. ayda) hesaplanır. Olayın gözlendiği zamana kadar olan süre içindeki eklemeli sağkalım oranı, bir önceki olayın görülmesinde hesaplanan sağkalım oranına eşit olmaktadır. Bu şekilde hesaplama yapıldığı için, K-M yönteminde sağkalım grafiği basamaklı bir şekil almaktadır (Şekil 3). İzlemde ölüm gözlenmediği sürece, K-M grafiğinde, eklemeli sağkalım olasılığı, x eksenine paralel düz bir çizgi olmaktadır. Ölüm görüldüğünde ise bu paralel çizgi, yeni hesaplanan olasılığın olduğu basamağa dikey olarak düşmektedir.

Tablo 3. Kaplan-Meier yöntemi ile sağ-kalım hesaplanması

Ölümün görüldüğü zaman	Yaşayan	d_t	q_t	p_t	P_t
4	6	1	0.167	0.833	0.833
10	4*	1	0.25	0.75	0.62
14	3	1	0.33	0.67	0.42
24	1**	1	1.00	0.00	0.00

*: Dördüncü ayda bir ölüm gözükteği ve 10. aydan önce de bir kayıp olduğundan, 10. Ayda hayatta ve takip edilen hasta sayısı 4 tür.**: Bir diğer kayıp ise 24. Aydan önce gözlendiği için 24. ayda sadece 1 bir hasta takip edilmiş ve o da bu ay içerisinde ölmüştür.



Şekil 3. Kaplan-Meier sağkalım grafiği

Şekil 3'te görüldüğü gibi, 0. ayda sağkalım olasılığı 1 (%100) iken 24. ayın sonunda bu olasılık sıfır (%0) olmuştur.

Klinik çalışmalarda daha çok K-M tahmin edicisi kullanılır. Zira, K-M yöntemi ile eğer gruplar benzer ise, sağkalım olasılıkları karşılaştırılabilir, az denek sayısı olsa dahi çıkarım yapılabilir.

Genel K-M eşitliği (product limit formülü):

$$\hat{S}(t(j)) = \hat{S}(t(j-1)) \times \prod_{r(T > t(j) | T \geq t(j))} \quad (1)$$

Eşitlik 1' de, herhangi bir zamandaki (rastgele alınan, T) sağkalım olasılığı, belirlenen zamandan bir önceki eklemeli sağkalım olasılığı ($t_{(j-1)}$) ile belirlenen (t_j) zaman aralığındaki sağ-kalım olasılığının çarpımına eşittir.

(1) no'lu sağkalım eşitliği, aynı zamanda;

$$\hat{S} = \prod_{i=1}^j \Pr[T > t_i | T \geq t_i] \quad (2)$$

Olarak da yazılabilir. Eşitlik (1)'e göre örnekten 14. Aydan fazla sağkalım olasılığı hesaplandığında;

$$S(14) = S(10) \times p(14) = 0.62 \times 0.67 = 0.42 \text{ olur.}$$

2.3.3. Farklı sağkalım grafiklerinin karşılaştırılması

Farklı sağkalım grafiklerinin karşılaştırılması ile daha çok sağlık bilimlerinde karşılaşılmaktadır. Aynı olayın incelendiği farklı gruplar arasında (en az 2 grubun) T zaman değişkeni bakımından fark olup olmadığı sağkalım analizi incelenebilmektedir. Farklı gruplara ait sağkalım süreleri, gruplara ait saçılım grafiği (scatter-plOR) elde edilerek karşılaştırılabilir. Ancak bu sübjektif ve yüzeysel bir karşılaştırma olmaktadır (Elisa ve Wenyuwang,2003; İnceoğlu, 2013). Sansürlü verilere sahip sağkalım analizlerinde, gruplar arası farkın önemliliği bazı istatistik yöntemlerle test edilebilmektedir. Bu yöntemlerin yaygın kullanılanlar aşağıda verilmiştir.

- Log-Rank testi
- Mantel-Cox testi
- Breslow-Wilcoxon testi
- Tarone-Ware testi.
- Cox's F-test

Bu çalışmada K-M yöntemi ile elde edilen iki sağkalım grafiğinin, istatistik olarak farklı olup olmadığı, en yaygın test olan Log-Rank testi ile incelenmiştir.

Log-Rank Testi: İki gruba ait sağkalım grafiğinin aynı popülasyona ait olup olmadığı, log- Rank testi ile test edilebilmektedir. Bu test için hipotezler aşağıdaki gibidir.

H_0 : İki grup arasında sağ kalım süreleri açısından fark yoktur. Her ne kadar sağ kalım grafikleri farklı görülse de bu fark tesadüften ileri gelmektedir.

Alternatif hipotez ise;

H_1 : İki grup arasında sağkalım süreleri açısından fark vardır. Sağkalım grafikleri farklı popülasyonları temsil etmektedir.

Log-rank testi Ki-Kare testinin geniş örnek (large-sample) için tasarlanmış bir versiyonu olarak bilinmektedir. K-M eğrilerini karşılaştırmak için Ki-Kare test istatistiği (χ^2) kullanmaktadır (Kleinbaum ve Klein, 2005). Log-rank testi, Ki-Kare testinin farklı versiyonları olan diğer testler gibi, gözlenen ve beklenen değerleri kullanmaktadır. Log-rank testinde başarısızlık zamanları, her iki grubun verileri kullanılarak küçükten büyüğe doğru

yazılır. Daha sonra (3) no'lu eşitlik kullanılarak her başarısızlık zamanı kategorisi (hücre) için beklenen başarısızlık (ölüm) oranları yazılır.

$$e_{1j} = \left(\frac{n_{1j}}{n_{1i}+n_{2j}} \right) * (m_{1j} + m_{2j}) \quad (3)$$

Eşitlik 3' te:

e: beklenen “başarısızlık sayısı”

n_{1j}: Birinci grup “j” hücrede risk altındaki birey sayısı

n_{1j} + n_{2j} : “j” hücrede her iki grupta toplam risk altında olan birey sayısı

m_{1j}: Birinci grupta “j” hücrede başarısız olan ya da ölen birey sayısı

m_{2j}: İkinci grupta “j” hücrede başarısız olan ya da ölen birey sayısıdır.

(3) no' eşitlik ile ikinci gruba ait beklenen sayılar

$$e_{2j} = \left(\frac{n_{2j}}{n_{1i}+n_{2j}} \right) * (m_{1j} + m_{2j}) \quad (4)$$

olarak hesaplanır. Daha sonra başarısızlık zamanı, her iki grup için gözlenen sayıdan (4) no'lu eşitlik ile hesaplanan beklenen sayı çıkarılarak (G-B) hesaplanır. Hesaplanan “G-B” değerlerinin toplamı ayrı ayrı her iki grup için de hesaplanır;

$$G - B = \sum_{j=1}^{başarısızlıksayısı} (m_{ij} - e_{ij}) \quad (5)$$

$I=$ 1. ve 2. Grup

(5) no'lu eşitlik ile işlem sonucu hesaplanan beklenen ve gözlenen değerlerin toplamı, herhangi bir grup için alındıktan sonra long-rank istatistiği aşağıdaki şekilde hesaplanır:

$$\log - rankistatistiği = \frac{(G-B)^2}{VAR(G-B)} \quad (6)$$

(6) no'lu eşitliğin paydasında verilen varyans ise her grup için

$$Var(G-B) = \sum_j^{başarısızlıksayısı} \frac{n_{1j} * n_{2j} (m_{1j} + m_{2j}) (n_{1j} + n_{2j} - m_{1j} - m_{2j})}{(n_{1j} + n_{2j})^2 (n_{1j} + n_{2j} - 1)} \quad (7)$$

Eşitliği ile hesaplanır. Bulunan log-rank istatistiği, k grup sayısı olmak üzere, k-1 serbestlik dereceli Ki-Kare dağılımı gösterir ve hesaplanan log-rank değeri, k-1 serbestlik dereceli χ^2 tablo değeri ile karşılaştırılır. Hesaplanan değer tablo değerden büyük ise H₀ hipotezi ret edilir.

Varyans hesabı yapılmadan *Ki-Kare yaklaştırma yöntemi* ile yaklaşık bir log-rank değeri;

$$\chi^2 = \sum_1^{grupsayısı} \frac{(G-B)^2}{B_j} \quad (8)$$

Eşitliği ile hesaplanabilmektedir (Kleinbaum ve Klein, 2005); (8) no'lu eşitlikte hesaplanan değer, varyans ile hesaplanan değere oldukça yakındır.

İkiden fazla grup karşılaştırmalarında, log rank testine ait yapılan hesaplamalar biraz daha karmaşık ve uzun olmaktadır. Ancak, bu hesaplamalar, günümüzde bilgisayar programları ile rahatlıkla yapılabilmektedir.

Log-Rank testine alternatif testler: K-M yönteminde grupların sağkalım sürelerini test etmek için log-rank testine alternatif bazı testler önerilmiştir ve bu testler bilgisayar programlarında bulunmaktadır. Alternatif test yöntemlerinin hepsi de log-rank testinin değişik versiyonlarıdır. Bunlardan yaygın olanları “Wilcoxon” ve “Tarone-Ware” testleridir.

Log-rank testinde test istatistiğini oluşturma için her grupta gözlenen değerlerin toplamından beklenen değer toplamı (G-B) çıkarılmakta idi. Bu şekilde basitçe yapılan toplama işleminde, her başarısızlık zamanına eşit ve bir (1) *ağırlık* (weight) değeri verilmektedir. Alternatif testlerde ise başarısızlık zamanlarına verilen *ağırlık* (w ile gösterilmektedir) farklı olmaktadır.

Wilcoxon (Breslow-Wilcoxon) testi: Bu yöntemde, gözlenen değer eksi beklenen değer (G-B) risk altındaki (n_j) birey sayısı ile ağırlıklandırılmaktadır.

$$Testistatistiği = \frac{(\sum_j w(t_j) * (m_{ij} - e_{ij}))^2}{var(\sum_j w(t_j) * (m_{ij} - e_{ij}))} \quad (9)$$

Eşitliği ile hesaplanır (9) no'lu eşitlikte, w_{ij} ile gösterilen *ağırlıklandırma* Wilcoxon testinde, “j” zamanda risk altındaki birey sayısı kadar olmaktadır. Böylece risk altındaki bireyin fazla olduğu çalışmanın başlangıç zamanındaki bilgiye daha fazla önem verilmektedir. Uygulanan tedavinin ya da yöntemin etkinliği, ilk başlarda daha fazla ve zamanla azalıyor ise, bu yöntem daha uygun olmaktadır.

Tarone-Ware testi: Tarone-Ware testi de Wilcoxon testi gibi uygulamanın ilk dönemlerine daha fazla önem vermektedir. Bunun için gözlenen ve beklenen değerler arası

fark deęeri, o başarısızlık anındaki (t_j) risk altında olan birey sayısının karekoku ile arpmaktadır;

$$W_{tij} = \sqrt{n_j} \quad (10)$$

Hangi testin kullanılacağına alıřmanın bařında karar verilmelidir. Karar verilirken de karřılařtırılan uygulamanın ya da tedavinin alıřmanın hangi dnemlerinde daha nemli olduęuna dikkat edilmelidir.



3. GEREÇ VE YÖNTEM

3.1. Gereç

Çalışmada, uygulama materyali olarak; 2014 yılında Van Yüzüncü Yıl Üniversitesi Dursun Odabaş Tıp Merkezi'nde tanı, tedavi ve takibi yapılmış yemek borusu (Özofagus ca) kanseri olan 31 hastanın verileri (Çelik ve ark. 2017) kullanılmıştır. Çalışmaya açıklayıcı değişkenler:

Cinsiyet: (Erkek, Kadın)

Cerrahi yöntem: (Açık cerrahi yöntem ve kapalı cerrahi (minimal invazif yöntem) yöntem)

Neoadjuvan tedavi durumu: Hastanın ameliyat öncesi kemORerapi alıp almadığı

Tümör lokalizasyonu: Tümörün özofagustaki yeri

Histoloji: Patolojik inceleme sonrası tümörün histolojik tipi

Derece (Grade): Tümörün histopatolojik olarak diferasyon derecesi

Patolojik evreleme (p TNM): Patoloji incelenmesi sonrası tümörün evresi

Total çıkarılan lenf nodu sayısı (Total_LN): Tümör dokusu ile beraber çıkarılan lenf nodu sayısı olarak dahil edilmiştir.

Yemek borusu kanserli hastaların yaşam süresi üzerine etkisi olan faktörlere ilişkin özet bilgi Tablo 4'te verilmiştir.

Tablo 4. Yemek borusu kanserli olguların verilerinin genel bir özeti.

Bağımlı değişken: sağkalım süresi		N	%
Analize dahil edilen hastalar	Olay ^a	10	32,3
	Sansürlü	21	67,7
	Genel	31	100,0
Çalışmada çıkan hastalar	Eksik verili vaka	0	0,0
	Negatif zamanlı vaka	0	0,0
	En erken olaydan önce gözlenen sansürlü vaka	0	0,0
	Genel	0	0,0
Toplam		31	100,0

3.2. Yöntem

3.2.1. Cox'un oransal riskler modeli

Bağımlı değişken veya cevap değişkeninin “zaman” gibi sansürlü bir değişken olması durumunda, açıklayıcı değişkenler (bağımsız değişken) ile bağımlı değişken arasındaki ilişki doğrusal regresyon modelleri ile açıklanamamaktadır. Daha önce yaşam tabloları ve K-M yöntemlerinde anlatılan sağkalım veya ölüm olasılığını etkileyen değişkenlerin incelenmesi, Cox tarafından geliştirilen ve yaygın olarak kullanılan Cox regresyon yöntemi ile yapılmaktadır (Cox, 1972). Cox'un (1972) Oransal Risk (OR) Modeli, (Proportional Hazard Regression Model) ölüm hızlarının, sağkalım süresini etkileyen değişken değerlerinin log-doğrusal fonksiyonu olduğu varsayımından yola çıkarak geliştirilmiştir. Yaşamsal verilerde prognostik değişkenler arasında oransal bir etkinin olduğu varsayılmaktadır. Cox OR modeli, ölüm hızı üzerinden yapıldığı için, önce ölüm hızı (Hazard ratio) fonksiyonu irdelenmiştir.

3.2.2. Risk (hazard) fonksiyonu

Olumsuz olay olarak nitelendirilen (sağlık bilimlerinde bu çoğunlukla “ölüm” olarak verilir) olayın gerçekleşmesine kadar geçen süre “ T ” değişkeni ile temsil edilir. Sağkalım analizinde rastgele alınan bir “ T ” değerinin önceden belirlenen bir zamandan (“ t ”) büyük olma olasılığının hesaplanması eşitlik (11)'e göre yapılır. Sağkalım fonksiyonu ile herhangi bir bireyin belirlenen bir zamandan daha fazla yaşama olasılığı hesaplanır. Risk fonksiyonu ise; belirli bir bireyin önceden belirlenen herhangi bir zamandaki ölüm oranı ya da daha doğru bir ifade ile ani ölüm hızı olarak tanımlanır (Kleinbaum ve Klein, 2005). Bu fonksiyon;

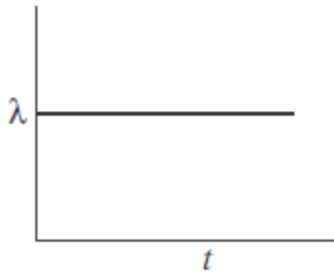
$$F(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t < T < t + \Delta t | T \geq t)}{\Delta t} \quad (11)$$

olarak ifade edilir. Hazard oranı, “ t ” zamana kadar yaşadığı bilinen bireyin, birim zamandaki (“ t ” zamanındaki) ani ölüm hızıdır. Bu nedenle, hazard oranı yerine “ani ölüm hızı” ifadesi de kullanılmaktadır (Kleinbaum ve Klein, 2005). Eşitlik (11)'deki Δt çok küçük bir zaman aralığını belirtir. Bu küçük zaman aralığının sıfıra yaklaştığı ama sıfır olmadığı kabul edilir. Rastgele alınan “ T ” değişkeninin adı geçen “ t ” ve “ $t + \Delta t$ ” aralığına düşme oranı, hazard oranını verir. Dikkat edileceği üzere; sağkalım fonksiyonu belirlenen “ t ” zamanından sonra yaşama olasılığını verirken, risk fonksiyonu, “ t ” anındaki ani ölüm (başarısızlık) hızını vermektedir. Yine, eşitlik (11)'de görüldüğü üzere, hazard oranı matematiksel olarak şartlı bir olasılıktır. “ T ” değişkeninin daha önceden belirlenen “ t ” zamanına eşit yada büyük olduğu

şartından sonra, “T” nin “t” “t+Δt” aralığına düşme ihtimali araştırılır. Eşitlik (11)’de pay kısmındaki değer şartlı olasılığı verirken, paydadaki Δt çok küçük zaman aralığını belirtir. Dolayısıyla, payın, paydaya bölünmesi ile zaman birimdeki olasılık değeri elde edilir ki bu olasılıktan daha çok “anlık hız” anlamına gelmektedir. Bulunan “oranın” 0 ile 1 arasında bir değer olmaması da, bu değerın orandan daha çok aralığı “0” ile “∞” arasında değişen “hıza” ait olduğunu belirtir. Bulunan ani ölüm hızı, zamanın ölçü birimine bağlı olarak değişebilir. Örneğin; pay kısmının ¼ (P), zaman aralığının da 12 saat (yarım gün) alınması durumunda, günlük ölüm hızı hesaplanmak istenirse bu değer; $h(t) = (1/4)/(1/2) = 0,5/\text{gün}$ olarak bulunur. Günlük ölüm hızı yerine haftalık ölüm hızı hesaplanırsa; $h(t) = (1/4)/(1/14) = 3,5/\text{hafta}$ olarak bulunur.

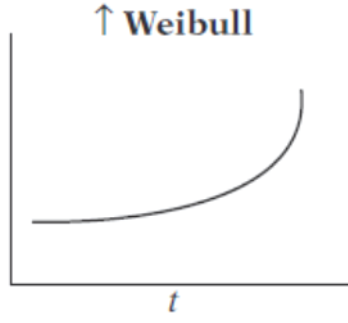
Risk fonksiyonu, negatif bir değer almamakta olup, sıfır ile artı sonsuz arasında herhangi bir değeri alabilir.

Klinik çalışmalarda, sağlıklı bireylerin bazı değişkenlerin etkisi altındaki yaşam süreleri çalışma periyodu süresince incelenir. Bu çalışma periyodu boyunca sağlıklı birey için “ölüm hızı” (Hazard oranı) sabit olmaktadır. Ölüm hızı “λ” gibi sabit bir değer ise, “t” değerinin kaç olduğunun bir önemi olmayacaktır (Şekil 4).



Şekil 4. Sabit ölüm hızı grafiği

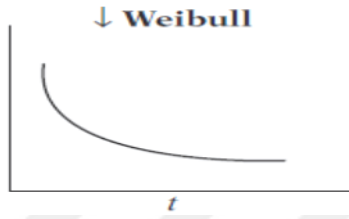
Sağkalım analizinde $h(t)$ fonksiyonu, Şekil 4’ te ki gibi sabit bir değer olduğunda, sağkalım modeline “*üstel*” (exponential model) model denilmektedir. Risk fonksiyonu zamanla artan bir fonksiyon ise buna “*artan Weibull*” modeli denilmektedir.



Şekil 5. Zamanla artan ölüm hızı grafiği

Şekil 5’te gösterildiği gibi zamanla artan ölüm riski genelde tedaviye cevap vermeyen hastalarda gözlenir. Böyle bir modelde zaman geçtikçe hastanın beklenen olumsuz olayın potansiyel riski artmaktadır.

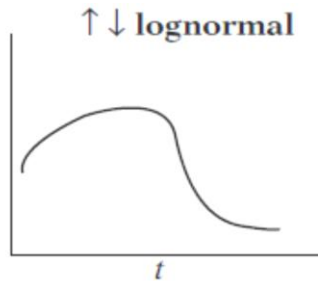
Risk fonksiyonunun azalan bir fonksiyon olması durumu Şekil 6’da verilmiştir.



Şekil 6. Zamanla azalan ölüm hızı grafiği

Şekil 6’da verilen “azalan Weibull” modeli, daha çok cerrahi uygulanmış hastaların iyileşme sürecinde izlenir. Ameliyat sonrası ilk günlerde ölüm riski yüksek iken zamanla bu risk azalmaktadır.

Diğer bir sağkalım modeli; “lognormal sağkalım” modelidir. Bu modelde de ölüm riski, ilk başta zamanla artar ancak, belli bir süre sonra bu risk azalmaya başlar (Şekil 7).



Şekil 7. Zamanla değişkenlik gösteren ölüm hızı modeli (Log-normal)

Bu duruma, kanserli hastaların nüks ihtimali örnek verilebilir. Tedaviden sonra ilk yıllarda nüks olasılığı zamanla artarken, belli bir süreden sonra -ki bu genel de 2 yıldır- bu olasılık azalmaktadır.

Sağkalım ve hazard fonksiyonları arasındaki ilişki aşağıdaki gibi özetlenebilir;

- Sağkalım fonksiyonu, herhangi bir kesitsel çalışmanın sağkalım süresini inceler ve eklemeli yaşam olasılığını vermektedir. Risk fonksiyonu ise daha çok ani ölüm hızını incelemektedir.
- Risk fonksiyonu belirli veri setine hangi modelin (Weibull, exponential, lognormal) uygun olacağına karar vermek için de kullanılabilir.
- Risk fonksiyonu, veri setinin sağkalım analizinde kullanılacak modelleme için bir araç olarak kullanılır. Sağkalım modelleri matematiksel olarak risk fonksiyonları şeklinde yazılır.

Sağkalım ($S(t)$) ya da risk fonksiyonlarından ($h(t)$) herhangi birisi bilindiğinde diğeri tahmin edilebilmektedir. Örneğin $h(t)$ sabit bir λ değeri ise $S(t) = e^{-\lambda t}$ üstel bir fonksiyon olarak ifade edilir. Eşitlikteki “e” sabit değeri *exp* olarak gösterilir. Daha genel bir ifade olarak, sağkalım fonksiyonu ile risk fonksiyonu arasındaki matematiksel ilişki aşağıdaki gibi ifade edilir.

$$S(t) = \exp\left[-\int_0^t h(u)du\right] \quad (12)$$

$$h(t) = -\left[\frac{dS(t)/dt}{S(t)}\right] \quad (13)$$

3.2.3. Cox oransal risk modeli

Risk fonksiyonu modeli kullanılarak Cox OR modeli,

$$h(t, \mathbf{X}) = \mathbf{h}_0(t) * e^{\sum_{i=1}^p \beta_i X_i} \quad (14)$$

$\mathbf{X} = X_1, X_2, X_3, \dots, X_p$ (Açıklayıcı değişkenler).

Olarak yazılır (Kleinbaum DG, Klein M 2005). Eşitlik (14)' te verilen, Cox OR modelindeki \mathbf{X} değişkenleri, herhangi bir bireyin sağkalım süresini etkileyen açıklayıcı ya da

tahminde kullanılan değişkenlerdir. \mathbf{X} tüm açıklayıcı değişkenlerin toplamı olarak, bazen \mathbf{x} vektörü olarak da tanımlanmaktadır (Kleinbaum ve Klein, 2005). Eşitlik, bazı özelliklere sahip bireyin “ t ” zamanındaki ani ölme olasılığını vermektedir. Eşitlikteki h_0 temel riski (baseline hazard) vermektedir. Eğer; $X_1=X_2=X_3\dots=X_p= 0$ ise Eşitlik (6) daki e üstü toplam değer sıfır olur. $e^0=1$ olacağından $h(t, X) = h_0$ olur ki bu da temel (baseline hazard function) risk fonksiyonuna eşit kabul edilir.

Cox OR modeli “ t ” anındaki risk pORansiyelinin, 2 ölçümün çarpımına eşit olduğunu belirtmektedir. Birincisi temel (baseline) risk olan h_0 diğeri de doğal logaritma değeri olan e ’nin üstel değeridir. Burada $\sum_{i=1}^p \beta_i x_i$ açıklayıcı (prognostik) değişkenlerin katsayılarının çarpımının toplamını ifade etmektedir. Eşitlikteki önemli bir özellik, temel h_0 fonksiyonunda herhangi bir X değişkeninin olmaması ve üstel dağılım ($e^{\sum_{i=1}^p \beta_i x_i}$) da da “ t ” nin olmamasıdır. Bu nedenle, standart Cox’un oransal risk modelinde zamandan bağımsız değişkenler (time-independent X) olduğu belirtilmektedir. Cox OR modellerinde zamandan bağımsız değişkenlerin kullanıldığı varsayılır.

Cox OR modelinin diğer bir özelliği de h_0 fonksiyonunun tanımlı olmamasıdır. Dolayısıyla $h_0(t)$ özellikli bir fonksiyon olmaması bu modeli semi-parametrik yapmaktadır. Semi-parametrik bir model olmasına rağmen Cox OR modelinin bu kadar yaygın kullanılmasının bazı nedenleri vardır: Öncelikle, Cox OR modeli ile farklı veri setleri için iyi katsayılar, risk oranları ve düzeltilmiş sağkalım eğrileri (adjusted survival curves) elde edilebilmektedir (Kleinbaum ve Klein, 2005). Cox OR modeli parametrik modeller ile elde edilen parametrelere yakın tahminlerde bulunduğu için sağlam (robust) bir model olarak kabul edilir (Elisa ve Wenyuwang, 2003). Ayrıca Cox OR modelinde, e ’nin üssü alındığı için katsayıların toplamı kısmı negatif olsa da, sonuç 0 ile sonsuz arasında bir değer çıkmaktadır. Bu da, tanım gereği risk oranının negatif olmaması nedeniyle istenen durumdur (Elisa ve Wenyuwang, 2003).

3.2.4. Cox oransal risk modelinde katsayıların tahmini

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

$$L(\beta) = \frac{\prod_1^k \exp(\beta' X_i)}{\sum_{j \in R_j} \exp(\beta' X_j)} \quad (15)$$

Eşitlik (15)' te 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.

Kısmi olabilirlik fonksiyonu ($L(\beta)$) elde edildikten sonra, bir sonraki aşamada bu fonksiyon maksimize edilir. Fonksiyonu maksimum yapmak için de genelde L fonksiyonunun doğal logaritması alınır ve bu $\ln L$ fonksiyon, modeldeki her parametre için kısmi türevi alınarak sıfıra eşitlenir:

$$\frac{\partial \ln L}{\partial \beta_i} = 0, i=1, 2, \dots, p \text{ (değişken sayısı kadar).}$$

Elde edilen eşitlikler tekrarlama (iteration) yapılarak çözülür ve tahmini regresyon katsayıları ($\hat{\beta}$) elde edilir.

3.2.5. Risk oranının (RO) hesaplanması

Maksimum olabilirlik (Maksimum Likelihood, ML) yöntemi ile Cox OR modelindeki katsayılar hesaplandıktan sonra, iki risk oranının karşılaştırılması, yani Cox modelindeki risk oranının (Hazard Rate, RO) hesaplanması yapılır.

Genel anlamda, risk oranı, bir bireyin hesaplanan orantılı ölüm tehlikesi (pORansiyeli) hızının, başka bir bireyin orantılı ölüm tehlikesi hızına bölünmesi ile elde edilen değerdir (Kleinbaum ve Klein, 2005). Karşılaştırılan iki bireyin birbirinden farkı, bu bireylerin ölüm pORansiyeli üzerinde etkili olduğu düşünülen değişken (\mathbf{X} 'ler) değerlerinin farklı olmasından kaynaklanmaktadır. Risk oranı bir bireyin $h(t, \mathbf{X}^*)$ fonksiyonu sonucunun, diğer bireyin $h(t, \mathbf{X})$ fonksiyonu sonucuna bölünmesi ile aşağıdaki gibi ifade edilir.

$$\widehat{HR} = \frac{h(t, \mathbf{x}^*)}{h(t, \mathbf{x})}, \quad \mathbf{x}^* = (x^{*1}, x^{*2}, \dots, x^{*p}) \text{ ve } \mathbf{X} = (x_1, x_2, \dots, x_p) \quad (16)$$

Eşitlik (16)'deki pay ve payda kısımlarına, risk fonksiyonları eşitlik (14)'teki gibi yazılırsa:

$$HR = \frac{h_0(t) * e^{\sum_{i=1}^p \beta_i X_i^*}}{h_0(t) * e^{\sum_{i=1}^p \beta_i X_i}} \quad (17)$$

elde edilir. Eşitlik (17)'de temel (h_0) fonksiyon sadeleştirilirse; un

$$\widehat{HR} = \exp[\sum_{i=1}^P \beta_i * (X_i^* - X_i)] \quad (18)$$

olarak yazılır. Buna göre risk oranı, çalışmaya alınan bireylerin, herhangi bir değişkene ait değerleri farkının ($x^* - x$) katsayı vektörü ile çarpılıp toplanması ve bunun e üssü alınarak ifade edilmesidir.

3.2.6. Cox modelinde açıklanan varyans

Çoklu regresyon analizinde kullanılan R^2 , sansürlü verilerin olduğu sağkalım analizine uygulanamamaktadır. Standart regresyon analizinde kullanılan R^2 benzer şekilde, sağkalım analizlerinde, özellikle Cox OR modeli için ölçüler geliştirilmiştir.

Sağkalım analizlerinde, Cox'un orantısal risk regresyon modelleri ile prognostik faktörlere bağlı rölatif risklerin, noktasal ya da güven aralığı tahminleri yapılmakta ve bu tahminlere özgü " p değeri bulunmaktadır. Bir veri setini, sadece noktasal tahminler ya da aralık tahminleri ile değerlendirip, bir değişkenin, hastanın ya da hastalığın seyri ile ilgili ne kadar bilgi verdiğini analiz etmek, önemli bilgelerin gözden kaçmasına neden olabilmektedir (Schemper, 1996). Oysaki herhangi bir hastanın tedavisinde ya da hastalığının seyrinde etkili olan faktörleri bilmek önemlidir. Hastalıkta etkili olan faktörü bilmek, hastaları sınıflandırmak (adı geçen faktöre sahip olmak bakımından riskli, az riskli ya da çok riskli gibi) ve tedaviyi planlamak açısından kullanılmaktadır. Üzerinde durulan özelliğin, bağımlı değişkendeki gözlenen değişimde ne kadar etkili olduğu belirlenebilir. Standart çoklu regresyon analizinde, çoklu korelasyon katsayısının karesi (coefficient of determination, R^2) bağımlı değişkendeki varyasyonun bağımsız değişkenler ile açıklanabilen kısmını göstermektedir ve buna Belirleme Katsayı denilmektedir (Nagelkerke, 1991). Standart regresyon modelinde açıklanan varyans, artıklar (residuals) üzerinden yapılan kareler toplamı ile bulunmaktadır.

$$R^2 = 1 - \frac{\text{hata kareler toplamı}(HKT)}{\text{total kareler toplamı}(TKT)} = \frac{TKT - HKT}{TKT} \quad (19)$$

Regresyon modelinin açıkladığı varyans oranının artması için, hata kareler toplamının azalması gerekmektedir.

Kullanılan regresyon modeline göre artıklardaki azalma mantığı, 2 özelliğin birlikteliğini kapsamaktadır (Korn ve Simon, 1991). Bunlar:

- Modelin veri setine uygulanabilirliği ve
- Modeldeki kovaryantların farklı sonuçları ayırabilmesidir.

Birinci madde, modelin veriler ile tutarlılığını “goodness of fit” ifade ederken, ikinci madde “açıklanan riski” ifade etmektedir. Goodness of fit, modelin kullanışlı olup olmadığını, başka kovaryantlar kullanılarak daha iyi modeller oluşturulup oluşturulamayacağını belirtmemektedir. Bazı kovaryantları modele dâhil ederek ya da bazılarını çıkararak nasıl daha iyi tahminlerin yapılabileceği “açıklanan risk” kavramı ile değerlendirilmektedir (Korn ve Simon, 1991).

Açıklanan varyans oranı (AVO) eşitlik (19)’dan yola çıkarak, aşağıdaki gibi yazılabilmektedir (Schemper ve Stare, 1996);

$$AVO = \frac{\sum D(y_i) - \sum D(y_i|x_i)}{\sum D(y_i)} \quad (20)$$

Eşitlik (20)’de D , merkezi bir ölçütten (doğrusal regresyonda genelde ortalama iken sağkalım analizinde medyandır), gözlenen y_i değerinin uzaklığını ifade etmektedir. Bu uzaklık i . birey ($1 \leq i \leq n$) için verilen herhangi bir modelde ve x_i kovaryant vektöründe, şartlı ya da şartsız olabilmektedir. $D(y_i)$ değerleri daha açık yazılırsa; $D(y_i) = (y_i - \bar{y})^2$ ve $D(y_i|x_i) = (y_i - \tilde{y}_i)^2$ olur. Buradan, doğrusal regresyondaki (Eşitlik 19) ile benzer sonuç elde edilir.

3.2.7. Cox OR modelinde açıklanan varyans ölçüsünün özellikleri

Cox OR modelinde açıklanan varyansı hesaplamak için önerilen ölçme yöntemleri (Kent ve O’Quigley, 1988; Korn ve Simon, 1990; Schemper, 1990; Schemper ve Stare, 1996; Schemper-Henderson, 2000) için Schemper ve Stare, (1996) aşağıda verilen özellikleri taşımaktadır:

- Hesaplama ile bulunan tahmini ölçme değeri, sansürlü verilerden fazla etkilenmemelidir. Zira, aynı ölçme yönteminin sansürlü veride de kullanılabilme durumu vardır. Yine de veri setinde sansürlü verinin olması yapılacak olan tahminin doğruluğunu veya keskinliğini azaltması beklenir.

- Cox modelinde, tahmin edilen katsayılar zamandan bağımsız olduğu için, varyans analizinde kullanılacak olan ölçme yönteminde, sonucun zaman skalasının monORonik transformasyonundan etkilenmemesi beklenir.

- Ölçme yöntemi açık bir şekilde yorumlanabilmeli ve sezgisel olarak anlaşılır olmalıdır.

- Eğer bir veri seti alternatif olarak standart regresyon analizi ile incelenebilirse, önerilen ölçme sonucu, klasik ölçmede kullanılan R^2 sonucu ile benzer olmalıdır. Bu özelliğe Schemper ve Stare (1996) “standart çoklu regresyon R^2 ile sayısal tutarlılık” denilmesini önermektedir.

- Önerilen ölçme yöntemi, hatalı yapılmış modellerde de tatmin edici olmalıdır. Schemper ve Stare (1996), bu özellik için “yanlış veya hatalı düzenlenmiş modellere karşı sağlamlık-dayanıklılık” denilmesini önermektedir.

- Cebirsel hesaplamaların zorluğu önemli kabul edilmese de, diğer kriterler açısından aynı performansı gösteren iki ölçme yöntemi için, basit hesaplamalar ile yapılanı değerlendirmeye alınması önerilmektedir.

Stare (1996), yukarıdaki belirtilen özelliklerden yalnızca ilk ikisinin, Cox modelinin vazgeçilmezi olduğunu, diğerlerinin ise tercihen olabileceğini belirtmiştir. Buna göre önerilen ölçüm, sansürlü verilerin varlığından ve zaman skalasında yapılan değişimlerden etkilenmemelidir.

3.2.8. Korn-Simon ölçüm tekniği

Korn ve Simon (1990), standart çoklu regresyon analizinde açıklanan varyansı incelemek için kullanılan belirleme katsayısının (R^2) sağkalım verilerinde de kullanılmasını önermişlerdir. Korn-Simon ölçme yöntemi olarak adlandırılan yöntemde, gözlenen değerlerdeki sapmalar (D) çoklu regresyondaki ortalamadan uzaklık yerine, logaritmik değerlerin ortalamalarından uzaklıkları yada medyandan olan uzaklık şeklinde hesaplanmaktadır. Buna göre sapmalar (ya da hatalar);

$$D(y_i) = |y_i - \text{medyan } y| \text{ ve ya } D(y_i) = (\log y_i - \overline{\log y})^2 \quad (21)$$

eşitliği ile hesaplanmaktadır. Sağkalım analizinde, belli bir zaman (T_0) dan sonrası ile ilgili yaşam süreleri ile ilgilenilmiyorsa, yukarda verilen sapma eşitlikleri (eşitlik 21) ile Korn-Simon yönteminin kullanılabilmesi belirtilmektedir (Schemper ve Stare J, 1996). Korn-

Simon yönteminde gerçek y değerleri yerine, tahmini $\hat{y}_i = \min(y_i, T_0)$ değer kullanılarak uzaklıklar (sapmalar) hesaplanmaktadır.

Sağkalım analizinin, sansürlü veri içermesi nedeniyle bazı y değerleri bilinmemektedir. Buna çözüm için, Korn-Simon, gerçek sağkalım zamanlarının yerine, tahmini sağkalım fonksiyonundan elde edilen tahmini değerlerin kullanılmasını önermektedir (Korn ve Simon, 1990). Model uyumunda gerçek değerler kullanılırken, herhangi bir bireyde gözlenen hata (sapma) o birey için sağkalım fonksiyonundan hesaplanan tahmini bir \hat{y} değeri alınarak hesaplanmaktadır. Bir birey için, en uygun model bulunduktan sonra, merkezi ölçütten sapmayı hesaplamak için aşağıdaki eşitlik önerilmektedir:

$$D(y_i|x_i) = \int (y - E(y|X_i))^2 dF(y|x_i) \quad (22)$$

Eşitlik (22)' de, $F(y|x_i) = 1 - S(y|x_i)$, iken $E(y|x_i)$ x bilindiğinde hesaplanan beklenen y değeridir. $S(y|x_i)$ ise verilere uygun model bulunduktan sonraki sağkalım fonksiyonudur. Buna göre yokluk (null) modeli basitçe yazılabilmektedir:

$$D(y_i) = \int (y - E(y))^2 dF_0(y) \quad (23)$$

Yine, $F_0(y) = 1 - S_0(y)$ dir. $S_0(y)$, yokluk hipotezinde sağkalım eğrisini ifade etmektedir. Eşitlik (22)' de tüm bireyler için hesaplanan hatalar, herhangi bir kofaktör olmadığı için aynı olacaktır. Korn ve Simon (1990) aşağıda verilen eşitlikler ile hataların hesaplanacağını belirtmişlerdir:

$$D(y_i|x_i) = \int_0^{T_0} t^2 dF(t|x_i) + T_0^2 S(T_0|x_i) - \left[\int_0^{T_0} t dF(t|x_i) + T_0 S(T_0|x_i) \right]^2 \quad (24)$$

Eşitlik (24)' te

$$\int_0^{T_0} t^2 dF(t|x_i) = \sum_{t_j \leq T_0} t_j^2 [S(t_{j-}|x_i) - S(t_j|x_i)]$$

$$\int_0^{T_0} t dF(t|x_i) = \sum_{t_j \leq T_0} t_j [S(t_{j-}|x_i) - S(t_j|x_i)]$$

olarak hesaplanmaktadır.

$S(t_j|x_i)$ tahmini sağkalım eğrisi iken, $S(t_{j-}|x)$ gerçek t_j zamanının hemen öncesindeki sağkalım oranıdır. Uzaklıklar hesaplandıktan sonra, eşitlik (20) kullanılarak açıklanan varyans hesaplanmaktadır (Korn ve Simon, 1990).

3.2.9. Kent ve O'Quigley'in bağımlılık ölçüm tekniği

Kent ve O'Quigley (1988) bağımsız değişken ile bağımlı değişken arasındaki “bağımlılık” ölçütü (measure of dependence) olarak kullanılan bilgi-kazanma (information gain-based) temelli bir yaklaşımı sağkalım analizlerine uyarlamışlardır.

Öncelikle, $X (x_1, x_2, \dots, x_p)$ bağımsız değişkenler vektörü ile tanımlanan T değişkeni, monORon transformasyon ile belirtilmektedir. Bu dönüşüm, katsayıları bulmak için kullanılan “kısmi olasılık” fonksiyonunun bir sonucu olarak görülmektedir. Dolayısıyla monORon dönüşümü artıran herhangi bir, $\phi(\cdot)$ fonksiyonu, $T^* = \phi(T)$ aynı şekilde Cox modelindeki aynı katsayılar ile uyumlu olmaktadır.

Φ fonksiyonu, Cox modelindeki h_0 modelinin bir formu olarak

$$h_0^\phi(t) = \alpha e^{\mu} t^{\alpha-1} \quad (25)$$

şeklinde yazılır. Eşitlik (25)'te μ herhangi bir değer olabilir ve $\alpha > 0$ olmalıdır.

Kent ve O'Quigley yöntemine X değişkenlerine ait katsayılar eklenince;

$h(t) = \alpha e^{\mu} t^{\alpha-1} e^{\beta x}$ olmaktadır. Buradan, sağkalım fonksiyonu yeniden

$S(t) = e^{-t^\alpha e^{\mu+\beta^T x}} = \exp(-t^\alpha e^{\mu+\beta^T x})$ olarak yazılmaktadır.

Bağımlı (T^*) değişkenine $Y^* = \text{Log}T^*$ dönüşümü yapıldıktan sonra

$$Y^* = -\sigma\mu - \sigma\beta^T x + \sigma\varepsilon, \quad (26)$$

olur.

$\sigma = \alpha^{-1}$ ve ε hata dağılımı, tanımlanmış olasılık yoğunluk fonksiyonu $f(\cdot)$ olarak modellenmektedir. Eğer $f(\cdot)$ normal standart yoğunluk fonksiyonu ise bu durumda $\mu = 0$, $\sigma = 1$ olacaktır. Böylece Eşitlik (26), standart doğrusal regresyon fonksiyonu olmaktadır. Ancak,

sağkalım analizine daha uygun olan, Weibull dağılımını veren *Gumbel* yoğunluk fonksiyonu için

$T=e^Y$ alınmaktadır.

$\emptyset = (\beta, \mu, \sigma^2)$ model parametreleri ve $\emptyset_1 = (\beta_1, \mu_1, \sigma_1^2)$ verilerin gerçek parametreleri olduğunu varsayılarak, $H_0 : \beta = 0$ iken $H_1 : \beta$ için sınırlama yoktur. Burada amaç, X_1 ve Y değişkeni arasındaki bağımlılık derecesini, regresyonda X değişkeni varken ölçmektir (Kent ve O'Quigley, 1988). Olabilirlik fonksiyonunun maksimize edilmesi ile elde edilen θ değerinden yola çıkarak ve Gumbel yoğunluk fonksiyonu kullanılarak oluşturulan fonksiyon;

$$\emptyset(\theta, \theta_1) = \iint \log\{f(y|x', \theta)\} f(y|x', \theta_1) dy G(dx) \quad (27)$$

olarak yazılır. Kullback ve Leibler (1951), bilgi kazanımı (information gain)

ise

$r = 2((\emptyset(\theta_1; \theta_1) - \emptyset(\theta_0; \theta_1))$ olarak ifade edilir.

X ve Y değişkenleri arasında bir bağımlılık ölçüsü olarak Kent (1983);

$$I_{IG}^2 = 1 - e^{-r} \quad (28)$$

eşitliğini önermiştir.

Daha sonra Kent ve O'Quigley (1988), yukarıda anlatılan yaklaşımlardan yola çıkarak, Weibull modeline yaklaşma yaparak;

$$\rho_{w,a}^2 = \frac{A}{A+1} \quad (29)$$

önermişlerdir. Eşitlik (29)'da, $A = \beta^T \Omega \beta$ dır. Ω ise X değişkeninin kovaryans matrisi olup, örneklemin kovaryans matrisinden hesaplanmaktadır.

Kent ve O'Quigley (1988), bu ölçütleri için aşağıdaki özellikleri belirtmişlerdir.

- $0 < \rho_{IG}^2 < 1$, $\|\beta\| \rightarrow \infty$ olur ise $\rho_{IG}^2 \rightarrow 1$ yaklaşır. Eğer H_0 doğru ise o zaman ρ_{IG}^2 sıfır olur.
- ρ_{IG}^2 , X ve Y değişkenlerinin doğrusal dönüşümlerinden etkilenmeyip, yalnızca katsayılarından ve X değerlerinin marjinal dağılımından etkilenir. μ ve σ ' den etkilenmez.

- Sansürlü verinin olmadığı bir $\{(x, y); i = 1, \dots, n\}$ örnekte, r (kazanılan bilgi) ölçümü, H_0 ve H_1 hipotezlerini değerlendirmek için kullanılan olabilirlik logaritmasının n^{-1} ile çarpılması ile bulunmaktadır. Bu durumda ρ_{IG}^2 :

$$\rho_{IG}^2 = 1 - e^{(-LR/n)} \quad (30)$$

olmaktadır.

3.2.10. Schemper'in V_1 ve V_2 ölçüm yöntemleri

Schemper (1990), Cox OR modelinde açıklanan varyansı incelemek için 2 yöntem tanımlamıştır. Kullanılan ölçüm, daha basit ve yorumlanmasının kolay olması nedeniyle diğer istatistikçiler tarafından da önerilmektedir (Stare, 1996).

Korn-Simon yönteminde olduğu gibi; Schemper yönteminde de, klasik regresyonda açıklanan varyansı inceleyen yöntem temel alınmaktadır. Bilindiği üzere Eşitlik (20)' de verilmiş olan; $AVO = \{\sum D(y_i) - \sum D(y_i|x_i)\} / \sum D(y_i)$ ifadesi, Açıklanan varyans oranı olup, bağımlı değişkende (y_i) gözlenen toplam sapmanın ne kadarının model tarafından açıklandığı belirtmektedir.

Doğrusal modellerde, $D(y_i)$ ve $D((y_i|x_i))$ sırasıyla; $(y_i - \bar{y})^2$ ve $(y_i - \bar{y}_i)^2$ ile hesaplanmaktadır. Sağkalım analizlerinde, sansürlü veri olduğundan ve hataların dağılımı da genellikle eğik (skewness) olduğundan, \bar{y} ve \bar{y}_i Cox modelinde elde edilememektedir. Bunun için, tahmin edilen medyan sağkalım süresi kullanılmaktadır. Ancak, sansürlü verinin çok fazla olduğu şartlı ya da şartsız tahminlerde medyan değer de çok kullanışlı olmamaktadır. Sonuçta, sağkalım analizinde ister şartlı ister şartsız durum olsun, kullanılacak tahmin yöntemi, sağkalım olasılıkları olmaktadır (Schemper, 1990). Buradan yola çıkarak, Schemper sağkalım olasılıklarına dayanan yöntemini önermektedir.

V_1 yöntemi: Herhangi bir i bireyinin sağkalım sürecini göstermek için S_{ij} kullanılmaktadır. Bireyi sembolize eden i ; $1 \leq i \leq n$ ve bu bireyin zamanını ifade eden t_j ; $1 \leq j \leq k_i$ şeklinde gösterilmektedir. k_i , sansürlü veri için sansürlenme tarihine kadar olan ölümlerin sayısı iken, sansürlü veri için örnekteki ölüm sayısıdır. S_{ij} yi, Schemper (1990) aşağıdaki şekilde önermiştir.

$S_{ij} = 1$, i bireyi t_j zamanında yaşıyor ise (yani sansürlü veri ise)

$S_{ij} = 0.5$, i bireyi tam olarak t_j zamanında ölmüş ise

$S_{ij}= 0$, i bireyi t_j zamanından önce ölmüş ise

Şartlı ve şartsız hata ölçümlerinin hesaplanması için

$$D(y_i) = k_i^{-1} \sum |S_{ij} - \hat{S}_j|, D(y|x_i) = k_i^{-1} \sum |S_{ij} - \hat{S}_{ij}|, \quad (31)$$

Eşitliği önerilmektedir. Eşitlik (31)' de \bar{S}_j , Kaplan-Meier sağkalım fonksiyonu ile t_j zamanında tahmin edilen değeri ifade etmektedir. Benzer şekilde, \bar{S}_{ij} ise Cox modelinde X_i gibi kovaryantların olduğu bir modelde oluşturulan sağkalım fonksiyonunda i bireyi için elde edilen değeri vermektedir. Böylece hesaplanan sapmaların mutlak değeri kullanılarak, ortalama tahmini hatayı yorumlamada $D(y_i)$ ve $D(y|x_i)$ değerlerinin kullanılması mümkün olmaktadır.

V₂ yöntemi: Birinci (V₁) öneri için geçerli olan varsayımların aynısını V₂ için de kabul ettikten sonra, Schemper, ortalama hataların (sapmaların) karesinin alternatif bir yöntem olarak alınabileceğini belirtmiştir (Schemper 1990).

$$D(y_i) = (k_i^{-1} \sum |S_{ij} - \hat{S}_j|)^2 \text{ ve } D(y|x_i) = (k_i^{-1} \sum |S_{ij} - \hat{S}_{ij}|)^2 \quad (32)$$

S_{ij} sağkalım fonksiyonu için Kaplan-Meier tahmin edicisi kullanılmaktadır. Kaplan-Meier sağkalım analizinde ise sansürlü veriler hesaplamada göz önüne alınmayarak, herhangi bir birey için ortalama değerler üretilmektedir. Bu ortalama değerler, hata kareler ortalamasını da minimize etmektedir. Hem bu özellik nedeniyle hem de mevcut ilgili terminoloji ile uyum içinde olmak adına Schemper, V₂ yönteminin tercih edilmesinin daha uygun olacağını belirtmiştir (Schemper ve Stare, 1996).

3.2.11. Schemper-Henderson ölçüm yöntemi

Schemper ve Henderson 2000 yılında, sağkalım analizinde açıklanan varyans için farklı bir yaklaşım önermişlerdir. Bu yaklaşım SAS programında yer almaktadır. Önerilen yöntemin üstünlükleri aşağıdaki gibi özetlenebilir:

Cox OR modelinin bir şartı olan “zaman ölçüm skalasındaki monoton değişikliklere karşı sabit olma, değişmeme”, kısmen daha kolay hesaplanabilme, gözlenen sonuçlar ile adapte edilmiş sonuçların açık ve tam karşılaştırma ve yanlış bir şekilde tanımlanmış

modellere karşı sağlam olma (Schemper ve Henderson, 2000). Sağkalım analizinde açıklanan varyans, zamanla değişen bir fonksiyon olarak görülmekte ve yapılan varyans ölçümlerinin veri setindeki takip süresi ya da seçilen zaman aralığından etkilenmesinin kaçınılmaz olduğunu düşünmektedir (Graf ve Schumacher, 1995). V_1 ve V_2 ölçüm yöntemlerine yöneltilen eleştirilere karşı Schemper ve Henderson (2000), “*her ne kadar takip süresi ve ya seçilen zaman aralığı açıklanan varyans ölçümlerini etkilese de, tatmin edici bir ölçüm yönteminin, maksimum takip süresi içerisinde ve seçilen zaman aralığı boyunca, asimptomatik olarak sansürlü veriden etkilenmemesi gerektiğini*” ifade etmişlerdir. Pratikte, toplam takip süresince (0, T) sağkalım ile ilgilenildiği belirtilmekte ve marjinal ve şartlı fonksiyon;

$$D(\tau) = 2 \int_0^\tau S(t)\{1 - S(t)\}f(t)dt / \int_0^\tau f(t)dt \quad (33)$$

$$D_x(\tau) = 2 \int_0^\tau E_x [S(t|x)\{1 - S(t|X)\}] f(t)dt / \int_0^\tau f(t)dt \quad (34)$$

Olarak yazılmaktadır (Schemper ve Henderson, 2000). Eşitlik (33) ve (34)'ten aşağıdaki eşitliğe ulaşmıştır;

$$\hat{V} = \{\hat{D} - \hat{D}_x\} / \hat{D} \quad (35)$$

$$\hat{V}_w = w^{-1} \sum_{j=1}^m \hat{G}(t_j)^{-1} d_j X \left[\frac{\{\hat{M}(t_{(j)}) - \hat{M}(t_{(j|x)})\}}{\hat{M}(t_{(j)})} \right] \quad (36)$$

Eşitlikleri yazılmış ve \hat{V}_w yerine \hat{V} nin kullanılmasının daha uygun olacağı belirtilmiştir (Schemper ve Henderson, 2000).

4. BULGULAR

Özofagus kanserli hastalarda açık ve minimal invazif cerrahinin hastaların sağkalım süresi üzerine etkileri Cox regresyon analizi ile incelenmiş ve sonuçlar Tablo 5’ te özetlenmiştir.

Tablo 5. Cox regresyonda cerrahi yönetime ilişkin sonuçlar

	Katsayı	Standart hata	p	Exp (B)	Exp(B) % 95,0 Güven Aralığı	
					Alt	Üst
Cerrahi yöntem	-,107	,844	,899	0,899*	,172	4,69

*. HR = Açık cerrahi (1) hazard/minimal invazif (0) hazard

*: Cerrahi yöntemin (açık cerrahi = 1, minimal invazif = 0) katsayısı (β_i).

Sağkalım üzerine cerrahi yöntem seçeneğinin etkisi Cox OR modelinde yerine yazıldığında

$$HR = \exp \left[\sum_1^1 \beta(x^* - x) \right] = \exp[-0,107 * (1 - 0)] = e^{-0,107}$$

katsayı 0.899 olarak bulunur. Buna göre açık cerrahi yöntemlerinin daha az riskli olduğu görülse de, bu fark istatistik olarak anlamlı bulunmamıştır ($p = 0,899$). Ayrıca, katsayı için güven aralığı da (0.172 – 4.69) “1” değerini içermektedir.

Sağkalım üzerinde etkisi olabileceği düşünülen bir diğer değişken olan cinsiyette modele eklenmiş ve sonuçlar Tablo 6’ te özetlenmiştir.

Tablo 6. Cox regresyonda cinsiyet ve cerrahi yönetime ilişkin sonuçlar

	Katsayı	Standart hata	p.	Exp(B)	Exp(B), %95 Güven aralığı	
					Alt	Üst
Cerrahi yöntem	-,139	,852	,870	0,870	,164	4,618
Cinsiyet	,207	,800	,796	1,230	,256	5,907

Tablo 5’ te görüldüğü üzere, yalnızca cerrahi yöntemin modele dahil edilmesi ile katsayı -0,107 (Tablo 5) olarak bulunmuşken, cinsiyet ile birlikte modele dahil edildiğinde cerrahi yönetime ait katsayı -0.139 olmuştur (Tablo 6).

Cinsiyet modele dâhil edildiğinde, cerrahi yöntem değişkeninin katsayısının negatif yönde bir miktar arttığı gözlenmiştir. Diğer yandan; cinsiyet modele dâhil edildiğinde HR

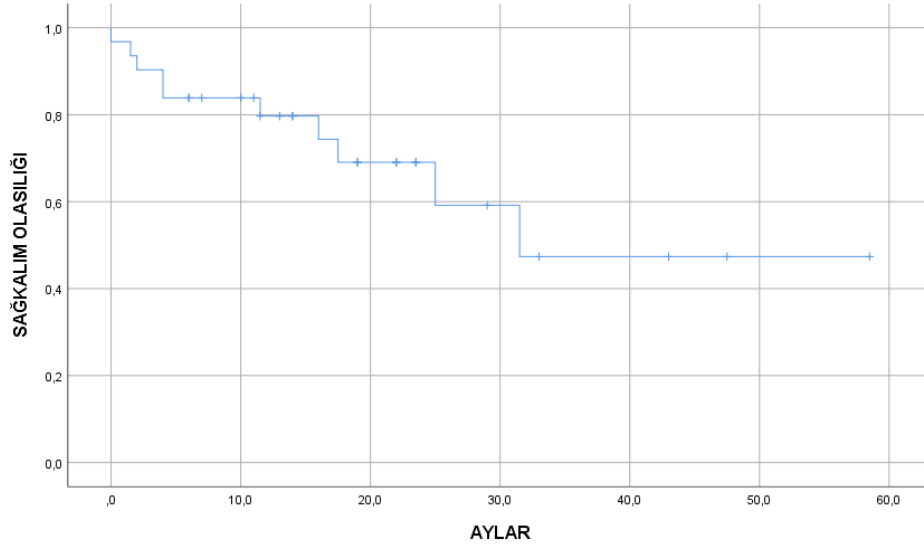
azalmıştır (0,899'dan 0,870'e düşmüştür). Ancak cinsiyet model dahil edildiğinde de cerrahi tiplerinin yaşam süresi üzerine istatistik olarak anlamlı etkisinin olmadığı görülmektedir. Bu durumda hangi modelin daha iyi olduğu, ya da hangi HR'nın kullanılması gerektiği sorunu ortaya çıkmaktadır. Herhangi bir kovaryant alınmadan oluşturulan ilk modele “kaba (rude) model” denilmektedir. Kaba model ile yapılan güven aralığı tahmini (%95 CI) 0,172 ile 4,69 (fark = 4,518) arasında iken ikinci model ile yapılan noktasal tahmin aralığı 0,164 ile 4,618, (fark = 4,454) arasında görülmektedir. İkinci model ile yapılan tahmin aralığı daha dar bir aralıktır.

Schemper-Henderson yöntemi, sıklıkla kullanılan yöntem olduğundan ve SAS programında tanımlı olduğundan, sağkalım analizi SAS' ta yapılmış ve analiz sonuçları Tablo 7' de verilmiştir. Modele dahil edilen değişkenlere ilişkin oluşturulan sağ-kalım (Kaplan-Meier) grafiği ise Şekil 8' de verilmiştir.

Tablo 7. Sağkalım süresine etkili olduğu düşünülen faktörlere ilişkin analiz sonuçları (En çok olabilirlik analizi)

Parametre		SD	Katsayı	SE	Ki-Kare	p	Risk oranı	%95 güven aralığı		Etiket
<i>Neoadjuvan</i>	0	1	-3.7472	1.3443	7.7688	0.0053	0.024	0.002	0.329	NEOADJ 0
<i>Cerrahi yöntem</i>	4	1	-1.3033	0.7653	2.8998	0.0886	0.272	0.061	1.217	SURG_ METHOD 4
<i>Tümör lokalizasyonu</i>	0	1	-12.5314	1703	0.0001	0.9941	0.000	0.000	.	TM_ Location 0
<i>Tümör lokalizasyonu</i>	1	1	1.1310	0.8392	1.8161	0.1778	3.099	0.598	16.053	TM_ Location 1
<i>Tümör lokalizasyonu</i>	2	1	0.0936	0.7470	0.0157	0.9002	1.098	0.254	4.748	TM_ Location 2
<i>Histoloji</i>	0	1	3.2034	1.6696	3.6809	0.0550	24.617	0.933	649.335	HISTOLOGY 0
<i>GRADE</i>	1	1	0.4304	1.0221	0.1774	0.6736	1.538	0.207	11.403	GRADE 1
<i>GRADE</i>	2	1	0.6542	0.9054	0.5222	0.4699	1.924	0.326	11.347	GRADE 2
<i>GRADE</i>	3	0	0	GRADE 3
<i>p_TNM</i>		1	-0.5189	0.2018	6.6059	0.0102	0.595	0.401	0.884	
<i>TORAL_LN</i>		1	0.0267	0.0212	1.5781	0.2090	1.027	0.985	1.071	

SD: Serbestlik derecesi, SH: Standart hata



Şekil 8. Kaplan-Meier yöntemine göre sağkalım grafiği

Tablo 7'de görüldüğü üzere 9 değişkenden *Neoadjuvan tedavi durumu* ve *Total çıkarılan Lenf nodu sayısı* değişkenlerine ait katsayılar dışında, diğer değişkenlere ait katsayıların istatistik olarak önemli olmadığı görülmüştür. Histoloji değişkenine ait katasının p değeri ise 0.05' e yakındır. *Neoadjuvan tedavi durumu* değişkenine ait risk oranı 0.024 olarak bulunmuş olup, buna göre Neoadjuvan tedavi almayan hastalarda, alan hastalara oranla ölüm riskinin 0.024 kat daha fazla olacağı söylenebilir. Benzer şekilde total çıkarılan Lenf Nodu sayısında bir birimlik artışa karşılık risk oranınının 1.027 kat artacağı söylenebilir.

Schemper-Henderson yöntemi, açıklanan varyans oranı için sıklıkla kullanılmakta ve SAS programında yer almaktadır. Tüm değişkenlerin modele dahil edilmesiyle, Schemper-Henderson yöntemine göre hata tahminleri ve açıklanan varyans oranları Tablo 8' de verilmiştir.

Tablo 8. Schemper-Henderson yöntemine göre hata tahminleri ve açıklanan varyans oranları

	Hata tahminleri		Açıklanan varyans oranı (%)
	Kovaryant değişkeni içermeyen model	Kovaryant değişkeni içeren model	
Tüm değişkenleri içeren model	0.3301	0.1709	48.24
Neoadjuvan değişkeni çıkarıldığında oluşan model	0.3301	0.2134	35.36

Tablo 8 de görüldüğü üzere, kovaryant değişken modele dahil edilmediğinde, hata tahmini 0.3301 olarak bulunmuş iken, kovaryant değişkeni modele dahil edildiğinde, bu değer 0.1709 olarak olmuştur. Kovaryant değişkeni modele dahil edilmediğinde, temel modele göre

hata oranı hesaplanmaktadır. Hata oranının düşük olması, modelin uyumluluğu açısından daha çok tercih edilen durumdur. Bu nedenle kovaryant değişkenini içeren modelin tercih edilebileceği söylenebilir. Modele ilişkin açıklanan varyans oranı ise %48.24 olarak bulunmuştur. Buna göre cevap değişkeni olan sağkalım süresindeki gözlenen varyasyonun veya farklılığın, %48.24 lük kısmının 9 değişken içeren model ile açıklanabileceğini söylenebilir.

Modelden *Neoadjuvan* değişkeni çıkarıldığında, modeldeki hata değeri 0.1709 dan 0.2134' a çıkarken, açıklanan varyans oranı ise %48.24 ten % 35.36' ya düşmüştür (Tablo 8). Hata değerindeki artışın yaklaşık 4 katı kadar açıklanan varyans oranında azalma gözlenmiştir.

Kent ve O'Quigley tarafından önerilen yonteme ilişkin açıklanan varyans oranının hesaplanmasına, herhangi bir istatistik paket programında rastlanmamıştır. Bu nedenle kodlama (Makro yazımı, Ek-1) kullanılarak, yonteme ilişkin hesaplama yapılmıştır.

Bir önceki yonteme benzer şekilde; tüm değişkenler modele dahil edilerek ve Neoadjuvan değişkeni modelden çıkarılarak her iki durumda açıklanan varyans oranı Kent ve O'Quigley yontemine göre incelenmiş ve sonuçlar Tablo 9' de verilmiştir.

Tablo 9. Kent ve O'Quigley yontemi ile açıklanan varyans oranı

	Tüm değişkenleri içeren model	Neoadjuvan değişkeni çıkarıldığında oluşan model
Kriter	-2 LOG L	-2 LOG L
Kovaryant değişkeni içermeyen model	98.830	98.830
Kovaryant değişkeni içeren model	75.132	83.543
W_1	3.46E-16	3.46E-16
W_2	4.85E-12	7.22E-13
W_3	7.15E+01	0.0004792957
W_4	0.065	0.065
R^2 (unadjusted)	0.534	0.389
R^2 (maximum)	0.958	0.959
Genel R^2	0.557	0.406

Tablo 9 de görüldüğü üzere, tüm değişkenleri içeren modelde R^2 değeri % 55.7 olarak bulunmuşken, Neoadjuvan değişkeni çıkarıldığında oluşan model de % 40.6 olarak bulunmuştur. Benzer şekilde düzeltilmemiş R^2 değerleri de her iki model için sırasıyla; % 53.4 ve % 38.9 olarak bulunmuştur.

5. TARTIŞMA VE SONUÇ

İnsan sađlığını ilgilendiren alıřmaların hemen hemen hepsinde genel ama, kiřinin sađlığını artırmak ve bunu yaparken kiřinin sađlıklı, daha uzun yařamasına katkıda bulunmaktır. Sađkalım sũresini arařtıran alıřmaların da zel amacı, bireyin sađkalım sũresini etkileyen deđiřkenleri tanımlamak, bu yolla da tedavi geliřtirmek ya da mevcut tedaviyi daha iyi hale getirmektir. Bireyin sađkalım sũresinde etkili olan faktrlerin tanımlanması, bilinmeyen faktrler nedeniyle olduka zor olmaktadır. Sađkalım sũresi ũzerinde etkisi olduđu dũřũnũlen deđiřkenlerin, kiřinin yařam sũresini ne kadar etkilediđini Cox OR modeli incelemektedir (Cox, 1972). Cox OR modeli ile prognostik faktrlerin oransal etkileri incelenirken, sađkalım sũresindeki toplam varyasyonun ne kadarının oluřturulan model ile aıklandıđı sũregelen bir tartiřma olmuřtur (Graf ve Schumacher, 1995).

Bu alıřmada da yemek borusu kanserli hasta verileri kullanılarak, sađkalım analizleri ve Cox OR modelinde aıklanan varyansın incelemesi yapılmıřtır. Kanserli hastalarda, hastalısız sađkalım sũresi ve toplam sađkalım sũresi, birok faktre bađlı olmakla birlikte, bu alıřmada zellikle literatũrde tanımlanmıř bazı faktrler (Neoadjuvan kemORerapi, cerrahi yntem, tũmrũn yeri, tũmrũn tipi, tũmrũn evresi vb.) incelenmiřtir. Hastalar 5 yıl (60 ay) takip edilmiř ve toplam sađkalım sũreleri irdelenmiřtir. Takip sũresince 10 hasta lmũř (olay) 21 hasta ise yařamıřtır (sansũrlũ). Bu vakalardan elde edilen sansũrlũ verilerden, sađkalım sũresindeki varyasyonun, oluřturulan Cox OR modelleri ile ne kadarının aıklandıđı incelenmiřtir.

Bilindiđi gibi sansũrlũ verilerin olduđu analizlerde regresyon analiz yntemlerinde kullanılan R^2 yntemleri ile bađımlı deđiřkende gzlenen varyasyon incelenememektedir. Bađımlı deđiřkenin zaman olduđu durumlarda, bađımsız deđiřkenlerin farklı bireylerde gzlenen farklı zamanların nedenini aıklaması beklenmektedir. Ancak, neden farklı sađkalım sũrelerinin olduđunu aıklayabilen tek deđiřken bulunamamaktadır. Bilinen ya da tahmin edilen ve bilinmeyen birok faktrũn, sađ-kalım sũrelerini etkilediđi dũřũnũlmektedir. Dolayısıyla, bilinen ya da tahmin edilen modelle belirlenen faktrlerin, ilgili bađımlı deđiřkendeki toplam varyasyonun (deđiřkenliđin) ne kadarlık bir kısmını aıkladıđı nemli bir sorun olmuřtur.

alıřmada, bir hastanın ameliyat ncesi kemoterapi (Neoadjuvan) almamıř olması, almıř olan bireye gre lũm riskini 0,024 kat artırmıřtır. Her ne kadar medyan yařam sũreleri

açısından anlamlı bir fark bulunmamış olsa da (Çelik ve ark, 2017), sağkalım süreleri üzerinde etkisi olduğu düşünülen Neoadjuvan kemoterapinin katsayısı anlamlı bulunmuştur. Neoadjuvan kemoterapi ile beraber, diğer bağımsız değişkenlerin, bağımlı değişken olan sağkalım süresindeki gözlenen varyasyonun ne kadarını açıkladığı irdelenmiştir.

Cox OR modeli sonuçları ile prognostik faktörlere ilişkin noktasal tahminler, güven aralığı ve bunlara ilişkin p değerleri verilmektedir. Sağlık araştırmaları yapan bir kişi, eğer veri setini ve sonuçlarını sadece bu ölçümler ile (nokta tahmini, güven aralığı) sınırlı tutar ise prognostik faktörlerin bireyin ya da tüm deneklerin üzerindeki etkisinin genişliği ile ilgili ek bilgileri kaçırabilir. Aynı şekilde, kronik bir hastalığın seyri ile ilgili ne kadar bilgi sahibi olduğu ile ilgili ciddi bilgiler kaçırılabilir (Schemper ve Stare, 1996). Bu tür sorunlar, bağımsız prognostik değişkenler ile açıklanmaya çalışılan bağımlı değişkende gözlenen varyasyonun ölçülmesi ile kısmen çözülebilir (Korn ve Simon, 1990). Genel doğrusal regresyon modellerinde, açıklanan varyasyon çoklu korelasyon katsayısının (R^2) karesi ile incelenmektedir (Nagelkerke, 1991). Cox modeli için açıklanan varyasyona benzer yöntemler son yıllarda önerilmiştir (Kom ve Simon, 1990; Schemper, 1990; Kent and O'Quigley, 1988; Harrell, 1986; Schemper ve Henderson, 2000). Bu alan halen gelişmelere açık olsa da, bu çalışmada şimdiye dek bilinen ve genel kabul görmüş çalışmaların değerlendirilmesi, sınıflandırılması yapılmış ve iki farklı yöntem (Schemper ve Henderson, 2000; Kent ve O'Quinley, 1988) ile bir uygulama yapılmıştır. Her önerinin kendine has avantaj ve dezavantajları olsa da, günümüzdeki yazılım programları ile hepsinin kodlamalarının yazılıp sonuçlarının elde edilmesi mümkün olabilmektedir. Bu çalışmada, SAS programında hazır olarak bulunan Schemper-Henderson yöntemi ve kodlaması (EK-1) yazılıp kullanılan Kent ve O'Quinley yöntemleri ile inceleme yapılmıştır. Schemper – Henderson (2000) ölçüm yönteminde V olarak tanımlanan açıklanan varyans oranı, yaşayan ile ölen bireylerin (1/0) arasındaki uzaklıklar ve kovaryan değişkeni olmadan ve kovaryant değişkeni ile sağkalım (fitted survival curve) eğrileri arasındaki uzaklıklar hesaplanarak yapılmaktadır. Kent ve O'Quigley (1988) yönteminde ise; bağımsız değişken ile bağımlı değişken arasındaki “bağımlılık” ölçütü olarak kullanılan bilgi-kazanma (information gain-based) temelli bir yaklaşım sağkalım analizlerine genelleştirilerek kullanılmıştır.

Öncelikle, Schemper – Henderson ölçüm yöntemi ile (Tablo 8). Tüm prognostik değişkenler modele dahil edildiğinde, açıklanan varyans oranı %48.24 olarak bulunmuş iken, bu oran Neoadjuvan değişkeni modelden çıkarılıp tekrar analiz yapıldığında %26.7 oranında azalma göstererek %35.36 olmuştur. Kent- O'Quinley yöntemi kullanılarak analiz yapıldığında

ise; tüm deęişkenler modelde iken açıklanan varyans oranı % 55.74 olarak bulunmuşken, Neoadjuvan deęişkeni modelden çıkarıldığında ise % 27.2 oranında azalıp %40.60 olmuştur. Bu oranlar arasındaki farklılık dięer ölçüm yöntemlerinde de rapor edilmektedir (Akazawa, 1997). Genel doğrusal regresyon analizine göre Cox OR regresyonda açıklanan varyans oranı genel de daha düşük bulunmaktadır. Mevcut çalışmada da, prognostik olduęu düşünölen deęişkenlerin, özofagus kanserli hastalarda sağkalım sürelerindeki varyasyonun ancak yaklaşık yarası kadarını açıkladıęı söylenebilir. Önerilen her iki ölçüm yönteminde de, ameliyat öncesi verilen kemoterapinin (Neoadjuvan) modelde olup olmasının, varyasyonun önemli bir kısmını (sırasıyla %12,88 ve % 15,14) azalttıęı saptanmıştır.

Sonuç olarak, Cox OR modelinde sonuçların yorumlanmasına geçmeden önce açıklanan varyasyonun incelenmesinin daha yararlı olacağı ve modelin uygunluęu hakkında bilgi vereceęi, böylelikle de sonuçlar hakkında eksik/fazla yorumu engelleyeceęi düşünölmektedir. Ölçüm yöntemlerinin sonuçlarının heterojen olması nedeniyle, açıklanan varyans oranının en az iki yöntem ile incelenmesinin daha uygun olacağı düşünölmektedir.

ÖZET

Çelik S, Sağkalım analizinde açıklanan varyans oranının incelenmesi. Van Yüzüncü Yıl Üniversitesi, Sağlık Bilimleri Enstitüsü, Biyoistatistik Anabilim Dalı Yüksek Lisans Tezi, Van 2018. Sağkalım analizi, Sağlık Bilimlerinde önemli ve yaygın kullanılan yöntemlerden birisidir. Yaşam tabloları ile başlayıp, Kaplan-Meier Yöntemi ile devam etmekte olan sağkalım analizlerinde, günümüzde Cox'un oransal risk modelinde açıklanan varyans oranının incelenmesi tartışma konusu olmuştur. Sansürlü verilerin olduğu sağkalım analizlerinde, bağımlı değişken olan sağkalım süresini etkileyen değişkenlerin, tahmin edilmesinde ve katsayılarının belirlenmesinde Cox'un önerdiği oransal risk modeli kullanılmaktadır. Oluşturulan Cox modelinin, bağımlı değişkende gözlenen toplam varyasyonun ne kadarını açıkladığı, prognostik faktörleri belirleme açısından önemlidir. Standart regresyon analizi yöntemlerinde çoklu korelasyon katsayısının (R^2) karesi model tarafından açıklanan varyans oranını belirtmektedir. Sağkalım analizinde ise bu ölçütün (R^2) analogu olan bazı yöntemler önerilmektedir. Bu çalışmada, Cox regresyon analizinde açıklanan varyansı inceleyen, Korn-Simon, Kent-O,Quinley, Schamper V_1 - V_2 ve Schemper-Henderson yöntemleri irdelenmiştir. Adı geçen yöntemler içerisinde sık kullanılan iki yöntem olan Schemper-Henderson ve Kent-O,Quinley yöntemleri kullanılarak, özofagus kanserli hastaların sağkalım verileri ile bir uygulama yapılmıştır. Sonuç olarak, sansürlü verilerin olduğu Cox modelinde, açıklanan varyansı inceleyen mevcut yöntemlerin sonuçlarının heterojen olduğu belirtilmiş ve Cox modelinin sonuçlarının yorumlanmasına geçmeden önce en az 2 farklı ölçüm yöntemi ile modelin açıkladığı varyasyonun belirtilmesi gerektiği önerilmiştir.

Anahtar kelimeler: Cox regresyon modeli, açıklanan varyans oranı, oransal risk, sansürlü veri

SUMMARY

Celik S, Analysis of explained variation in survival analysis. Van Yuzuncu Yil University, Institute of Health Sciences, Master Thesis in Department of Biostatistics, Van 2018. Survival analysis methods are very important and frequently applied methods in health sciences. Date back to life tables and continuing with Kaplan- Meier method survival analysis currently has a discussion about explained variation in Cox proportional hazard model. Analyses of prognostic factors and coefficients in studies of survival and other censored outcomes are most often based on the Cox proportional hazards regression model. It is paramount importance to determine the variation of a dependent variable explained by prognostic factors. In the context of the general linear model, the square of the multiple correlation coefficient (R^2) is the measure of explained variation. For the Cox model, analogous measures of explained variation have been proposed in recent years. In this study, some measures (Korn-Simon, Kent- O,Quinley, Schamper V_1 - V_2 ve Schemper-Henderson) that have been proposed for explained variation in Cox model was evaluated. An application with esophageal cancer data was performed by using two most common methods, Schemper-Henderson and Kent-O,Quinley. Consequently, heterogeneous results were found by different measure methods. Therefore, at least two different measures should be used to determine explained variation by model and then results should be evaluated.

Key words: Cox regression model, explained variation, propotional hazard, censored data

KAYNAKLAR

- Akazawa K (1997). Measures of Explained Variation for a Regression Model Used in Survival Analysis. *J Med Syst.* 21, 4, 229-238.
- Anderson DR, Wywiałowski AP, Burnham KP (1981). Tests of assumptions underlying life table methods for estimating parameters from cohort data. *Ecology*, 62, 4, 1121-1124.
- Anonymous (2013). https://en.wikipedia.org/wiki/Reliability_engineering Erişim tarihi:14.5.2016.
- Celik S, Dogan E, Arslan H, Batur A, Kiziltan R, Toktas O, Yilmaz O, Kotan MC (2017). Changing Trends and Experience with Esophageal Cancer Surgery in a Single University Hospital: Are The Results Similar or NOR? *Bagcilar Medical Bulletin*, 2, 1, 1-8.
- Collett D (2003). *Modelling Survival Data in Medical Research*, Chapman & Hall/CRC 2. Edition.
- Cox DR (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society*, 34, 187-220.
- Elisa TL ve Wenyuwang J (2003). *Statistical Methods for Survival Data Analysis*. 3rd Edition. by John Wiley & Sons, New Jersey.
- Graf E ve Schumacher M (1995). An investigation on measures of explained variation in survival analysis. *The Statistician* 44, 4, 497-507.
- Greenwood M (1941). Medical statistics from Graunt to Farr. *Biometrika*, 32 101-127.
- Hald A (1990). *A History of Probability and Statistics and Their Applications Before 1750*. Wiley, New York.
- Harrell FE, Califf RM, Pryor DB, Lee KL, and Rosati RA (1982). Evaluating the yield of medical tests. *JAMA*, 247, 2543-2546.
- Harrell FE (1986). The PHGLM procedure. In *SUGI Supplemental Library User's Guide, Version 5th Ed*. Hastings RP, Cary, North Carolina : SAS Institute Inc. 437-466.
- Feinleib M (2005). <http://www.medicine.mcgill.ca/epidemiology/hanley/c609/Material/EoB-heathStatisticsHistory.pdf> (Erişim tarihi: 05.06.2017).
- Kaplan EL ve Meier P (1958). Nonparametric Estimation from Incomplete Observations. *J Am Stat Assoc.* 53, 282, 457-481.
- Kent JT (1983). Information gain and a general measure of correlation *Biometrika*, 70, 163-173.
- Kent JT ve O'Quigley J (1988). Measures of dependence for censored survival data. *Biometrika*, 75, 525-34.
- Kleinbaum DG, Klein M (2005). *Survival analysis: A Self-Learning Text*, 2nd Edition. Springer, New York.

<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.472.2636&rep=rep1&type=pdf>
(erişim tarihi: 09.08.2018).

Korn EL, Simon R (1990). Measures of explained variation for survival data. *Statistics in Medicine* 9, 487-503.

Korn EL, Simon R (1991). Explained residual variation, explained risk and goodness of fit. *Am. Statistn*, 45, 201-206

Kullback S, Leibler RA. (1951). On information and sufficiency. *Ann Math Stat.* 22, 72-86.

Nagelkerke NJD. (1991) A Note on the general definition of the coefficient of determination. *Biometrika* 78, 691-692.

İnceoğlu F (2013). Sağlık analiz yöntemleri ve karaciğer nakli verileri ile bir uygulama. T.C.İnönü Üniversitesi Sağlık Bilimleri Enstitüsü yayınlanmamış Yüksek lisans tezi Malatya.

Schemper M (1990). The explained variation in proportional hazards regression. *Biometrika*, 77, 216-8.

Schemper M ve Stare J (1996). Explained variation in survival analysis. *Statistics in Medicine*, 15, 1999-2012.

Schemper M ve Henderson R (2000). Predictive Accuracy and Explained Variation in Cox Regression. *Biometrics*, 56, 249-255.

Stare J (1996). On Choosing a Measure of Explained Variation for Cox Models. *Developments in Data Analysis. Metodološki zvezki* 12, Ljubljana: FDV.

Stigler SM (1968). *The History of Statistics. The Measurement of Uncertainty Before 1900.* Harvard University Press, Cambridge, Mass.

Soley-Bori M (2013). Dealing with missing data: Key assumptions and methods for applied analysis. Technical Report No. 4 <https://www.bu.edu/sph/files/2014/05/Marina-tech-report.pdf>. (Erişim tarihi: 04.05.2018).

Bollet AJ (1973). Pierre Louis: the numerical method and the foundation of quantitative medicine, *American Journal of Medical Science* 266, 92–101.

Westergaard H (1932). *Contributions to the History of Statistics.* King, London

ÖZGEÇMİŞ

Van Erciş'te 1980 yılında doğdu. İlkokul eğitimini Erciş'de, ortaokul ve lise eğitimini Van'da tamamladı. Hacettepe Üniversitesi Tıp Fakültesi'nden (İngilizce bölümü) 2005 yılında tıp doktoru ünvanı aldı. Aynı yıl Hacettepe Üniversitesi Hastanelerinde Genel Cerrahi AD'ında Araştırma Görevlisi olarak çalışmaya başladı. 2011 yılında, Genel Cerrahi ihtisasını Hacettepe Üniversitesi'nde tamamladı. Devlet hizmet yükümlülüğü görevini 2011-2014 yılları arasında Van Bölge Eğitim-Araştırma Hastanesi'nde tamamladıktan sonra, 2014 Temmuz ayında Van Yüzüncü Yıl Üniversitesi Tıp Fakültesi Genel Cerrahi AD'da Dr.Öğr. Üyesi olarak göreve başladı. Van Yüzüncü Yıl Üniversitesi, Sağlık Bilimleri Enstitüsü, Biyoistatistik Anabilim Dalı'na 2015 yılında Yüksek Lisans eğitimine başladı.

EK 1. SCHEMPER-HENDERSON VE KENT-O,QUINLY İÇİN SAS KODLARI

```
run;

ODS RTF;
ODS GRAPHICS ON;
TITLE "modell1 MODELDE NEOADJSURG_METHODM_LOCATIONHISTOLOGY
GRADE p_TNMTORAL_LN VARIKEN";
ods out=FitStatistics;

proc phreg data = ozefag plORs=survival ev OUTEST=modell1;
class NEOADJ      SURG_METHODM_LOCATIONHISTOLOGY  GRADE
;
model SURVIVAL*STATUS(1) =
NEOADJ      SURG_METHODM_LOCATIONHISTOLOGY  GRADE p_TNMTORAL_LN
/ risklimits ties=efron;
ODS output FitStatistics=fit;* (keep=Table where=(effect="COL1"));
ods output GlobalTests=glob;
output out=aa;
*assess var=(AGE) / resample;
run;

TITLE "      MODELDE SURG_METHOD      TM_LOCATIONHISTOLOGY  GRADE p_TNM
TORAL_LN VARIKEN";
proc phreg data = ozefag plORs=survival ev;
class SURG_METHODM_LOCATIONHISTOLOGY  GRADE
;
model SURVIVAL*STATUS(1) =
SURG_METHODM_LOCATIONHISTOLOGY  GRADE p_TNMTORAL_LN
/ risklimits ties=efron;
ODS output FitStatistics=fit1;* (keep=Table where=(effect="COL1"));
ods output GlobalTests=glob1;
run;
ODS GRAPHICS OFF;
ODS RTF CLOSE;
title "Kent-Oquigley";
data modelld;
set fit1;
length str $64.;
if lowercase(criterion) =: '-2';
w01 = exp(-1*WithoutCovariates/2);
w02 = exp(-1*WithCovariates/2);
w03 = exp(-1*(WithoutCovariates-
WithCovariates)/2);
w04 = 2/31;
r2unadj = 1 - w03**w04;
r2max = 1 - w01**w04;
genrsq02 = r2unadj/r2max;
put 'The unadjusted Generalized
R-Square value is ' r2unadj ;
put 'The Generalized (Kent-Oquigley)
R-Square value is ' genrsq02 ;
output modelld;
run;
/*
%MACRO KENTOQNR(
```

```
DATA=ozefag, TIME=SURVIVAL, STATUS=STATUS,  
COV=NEOAJD SURG_METHODTM_LOCATIONHISTOLOGY GRADE, CENSLIST  
=1);  
%put _user_;  
run;  
%MEND KENTOQNR;  
%KENTOQNR;  
call execute ("%KENTOQNR");  
RUN;  
  
proc sql;  
select * from sashelp.vmacro where scope="KENTOQNR"  
;quit;
```



EK 2. ETİK KURUL BELGESİ



T.C.
YÜZÜNCÜ YIL ÜNİVERSİTESİ
Girişimsel Olmayan Klinik Araştırmalar
Etik Kurul Başkanlığı

Sayı : B.30.2.YYU.0.01.00.00/ 21

Tarih: 21/02/2018

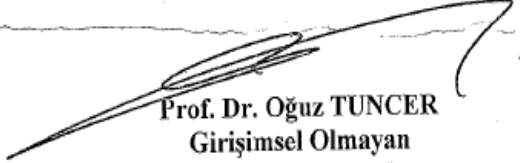
Konu :Etik Kurul Başvurunuz

Sn. Prof.Dr. Sıddık KESKİN



İlgi: 08.02.2018 tarih ve bila sayılı yazınız.

İlgi yazı ile Etik Kurulumuza sunulan “Sağkalım Analizinde Açıklanan Varyansın İncelenmesi” isimli proje özetiniz 16.02.2018 tarihinde yapılan Etik Kurul toplantımızda görüşüldü. Yapılan görüşmede çalışmanın proje özetinin içeriği tamamiyle istatistiksel analiz değerlendirmesi olduğundan, adı geçen çalışmanın etik kurulla doğrudan ilişkisi saptanmamıştır.

Bilgilerinize rica ederim.


Prof. Dr. Oğuz TUNCER
Girişimsel Olmayan
Klinik Araştırmalar
Etik Kurul Başkanı

EK 3. TEZ ORJİNALLİK RAPORU

YÜZÜNCÜ YIL ÜNİVERSİTESİ SAĞLIK BİLİMLERİ ENSTİTÜSÜ LİSANSÜSTÜ TEZ ORJİNALLİK RAPORU	
Tarih: 24/07/2018	
<p>Tez Başlığı / Konusu: Sağkalm analizinde açıklanan varyans oranının incelenmesi</p> <p>Yukarıda başlığı/konusu belirlenen tez çalışmamın Kapak sayfası, Giriş, Ana bölümler ve Sonuç bölümlerinden oluşan toplam 39 sayfalık kısmına ilişkin, 24/07/2018 tarihinde şahsım/tez danışmanım tarafından TURNITIN intihal tespit programından aşağıda belirtilen filtreleme uygulanarak alınmış olan orijinallik raporuna göre, tezimin benzerlik oranı %5 (beş) tir.</p> <p>Uygulanan filtreler aşağıda verilmiştir:</p> <ul style="list-style-type: none">- Kabul ve onay sayfası hariç,- Teşekkür hariç,- İçindekiler hariç,- Simge ve kısaltmalar hariç,- Gereç ve yöntemler hariç,- Kaynakça hariç,- Alıntılar hariç,- Tezden çıkan yayınlar hariç,- 7 kelimeden daha az örtüşme içeren metin kısımları hariç (Limit match size to 7 words) <p>Yüzüncü Yıl Üniversitesi Lisansüstü Tez Orijinallik Raporu Alınması ve Kullanılmasına İlişkin Yönergeyi inceledim ve bu yönergede belirtilen azami benzerlik oranlarına göre tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.</p> <p>Gereğini bilgilerinize arz ederim.</p> <p style="text-align: center;"> 24.07.2018</p>	
Adı Soyadı: Sebahattin Çelik Öğrenci No: 159302025 Anabilim Dalı: Biyoistatistik Programı: TIP Statüsü: Y.Lisans x Doktora <input type="checkbox"/>	
DANIŞMAN ONAYI UYGUNDUR	ENSTİTÜ ONAYI UYGUNDUR
 Prof. Dr. Sıddık KESKİN	