

Aralıklı Sansürlü Veriler için Sağkalım Modelleri

İlknur ERÖZ¹, Nihal ATA TUTKUN^{*2}

^{1,2} Hacettepe Üniversitesi, Fen Fakültesi, İstatistik Bölümü, 06580, Ankara, Türkiye

(Alınış / Received: 29.11.2019, Kabul / Accepted: 07.07.2020, Online Yayınlanma / Published Online: 20.08.2020)

Anahtar Kelimeler

Aralıklı sansürlü,
Cox regresyon modeli,
Hızlandırılmış başarısızlık
süresi modeli,
Orantılı tehlikeler,
Sağdan sansürlü

Özet: Sansürleme, sağkalım analizi diğer istatistiksel yöntemlerden ayıran en önemli özelliktir. Sansürleme türleri ise sağdan, soldan ve aralıklı sansürleme olarak sınıflandırılmaktadır. Sağkalım analizinde kullanılan modellerin çoğu gözlemlerin sağdan sansürlü olduğu veri kümeleri için geliştirilmiştir. Aralıklı sansürlü veriler ile ilgili çalışmalar da son yıllarda hız kazanmıştır. Birimin başarısızlık süresi belli bir aralıkta gerçekleşiyorsa aralıklı sansürlü gözlemler söz konusudur. Bu veriler genellikle, ilgilenilen gözlemlerin sürekli izlenemediği durumda oluşur. Sadece ilgilenilen olayın iki gözlem periyodu arasında meydana geldiği bilinmektedir. sağkalım çözümlemesinde yaygın olarak kullanılan Cox regresyon modeli ve parametrik sağkalım modelleri aralıklı sansürlü veriler için de geliştirilmiştir. Bu çalışmada aralıklı sansürlü veriler için sağkalım modelleri incelenmiş ve literatürde yer alan Primer Biliyer Siroz verisi hem aralıklı sansürlü hem de sağdan sansürlü olarak ele alınarak analiz edilmiştir.

Survival Models with Interval Censored Data

Keywords

Interval censoring,
Cox regression model,
Accelerated failure time
model,
Proportional hazards,
Right censoring

Abstract: Censoring is the most important feature that distinguishes survival analysis from other statistical methods. The types of censoring are classified as right, left and interval censoring. Most of the models used in survival analysis were developed for data sets in which observations were right censored. Studies on interval censoring gains speed in recent years. If the failure time of the unit occurs within a certain range, there are interval censored observations. These data usually occur when observations of interest are not constantly monitored. It is only known that the event of interest takes place between two observation periods. Cox regression model and parametric survival models used commonly in survival analysis have been developed for interval censored data. In this study, survival models for interval censored data are examined and primary biliary cirrhosis data is analyzed both for right censored and interval censored data structure.

1. Giriş

Sağkalım analizi, sağkalım süresinin istatistiksel modellemesi ve çözümlemesi ile ilgilenmektedir. Sağkalım süresi, belirli bir başlangıç noktasından önceden belirlenmiş bir olay meydana gelene kadar geçen süreyi belirtmek için kullanılır [1]. Literatürde sağkalım süresi yerine "başarısızlık süresi" terimi de kullanılmaktadır. Sağkalım analizi, çoğunlukla biyoistatistik alanında kullanılmaktadır. Araştırmacılar, belirli bir tedavinin başarısızlık süresi üzerindeki etkisini incelemek için genellikle parametrik olmayan yöntemlerden Kaplan-Meier sağkalım olasılığı tahminlerini ve log-rank testini kullanırlar. En yaygın kullanılan sağkalım modeli ise Cox regresyon modelidir [2]. Sansürlü gözlemlerin olması ve başarısızlık riskinin ya da tehlike fonksiyonunun modellemede esnek olması Cox regresyon modeli (CRM)'nin tercih edilme nedenidir. Hızlandırılmış başarısızlık süresi (HBS) modelleri ise

sağkalım süresinin parametrik bir dağılıma uygun olması durumunda kullanılmaktadır.

Klasik istatistiksel yöntemlerden farklı olarak sağkalım analizinin en önemli özelliklerinden biri sansürlü gözlemlerdir. Sansürlü gözlem, başarısızlığın meydana gelme süresi hakkında kısmen bilgi vermektedir ve tamamlanamamış gözlem olarak da adlandırılabilir. Yani, gözlem süresi içerisinde başarısızlığın meydana gelmemesi durumunda sansürlü gözlemden söz edilebilir. Bu durumda başarısızlık, gözlemlenen sansürleme süresinden daha sonraki bir sürede gerçekleşmektedir. Ayrıca başarısızlık süresi, gözlemlenen bir birimin bazı nedenlerden dolayı gözlemden çıkması ya da birimlerin araştırmaya farklı zamanlarda dâhil olması gibi nedenler ile tamamlanamamış ise sansürlü gözlemden söz edilebilir.

Sansürleme türleri sağdan, soldan ve aralıklı sansürleme olmak üzere üç grupta incelenebilir. Sağkalım analizinde genellikle sağdan sansürlü veri yapısı için çalışmalar yapılmıştır. Bununla birlikte sağdan ve soldan sansürlü, aralıklı sansürlünün özel biçimleridir. Aralıklı sansürlü, ilgilenilen olayın gözlemlenemediği ve sadece iki zaman aralığının içinde ortaya çıktığının bilindiği durumlarda söz konusudur. Aralıklı sansürlü veri yapısı, bir aralıktaki birimlerin başarısız oldukları kesin zaman noktaları ile ilgili belirsizliği yansıtmaktadır. Aralıklı sansürlü veri türleri, Durum I aralıklı sansürlü veriler (mevcut durum verileri) ve Durum II (genel) aralıklı sansürlü veriler olmak üzere ikiye ayrılmaktadır. Durum I aralıklı sansürlü veri yapısı, sadece sağkalım süresi bilgisi ile gözlem süresinin öncesinde ya da sonrasında meydana gelen soldan veya sağdan sansürlü veri yapılarıdır. Bu nedenle, Durum I aralıklı sansürlü veri yapısı, sansürlü ve $(a, b]$ aralığında aralıklı sansürlü gözlemlere bakılmaksızın, Durum II aralıklı sansürlü veri yapısının özel bir durumu olarak gözlemlenmektedir. Durum I aralıklı sansürlü veri yapısında, birimler sadece bir kez gözlemlenmekte ve sadece olayın gözlemlenmeden önce gerçekleşip gerçekleşmediği bilinmektedir. Bu tür veriler özellikle demografik çalışmalarda yaygındır.

Kesin başarısızlık süresinin bilinmediği durumlarda meydana gelen olayın, sadece başarısızlığın rasgele bir zaman aralığında, sol uç noktasından önce ya da sağ uç noktasından sonra olduğu biliniyorsa Durum II aralıklı sansürlü veri yapısı söz konusudur. Bu durumda her gözlem keyfi olarak sansürlenmektedir. Durum II aralıklı sansürlü veri yapısına, Durum I aralıklı sansürlü veri yapısına göre daha sık rastlanmaktadır.

Aralıklı sansürlü gözlemler için kullanılacak sağkalım modelleri sağdan sansürlü veri yapısı için kullanılan modellerden farklı bir yapıya sahiptir. Aralıklı sansürlü veri yapısı için CRM'ye ilişkin çalışmalar Finkelstein ve Wolfe [3], Finkelstein [4] ile başlamış; Huang ve Wellner [5], Huang [6], Satten [7], Younes ve Lachin [8], Huang ve Rossini [9], Goggins vd. [10], Geskus ve Groeneboom [11], Pan [12], Xue vd. [13], Zeng vd. [14], Sen ve Banerjee [15], Heller [16] ile devam etmiştir. Aralıklı sansürlü veri yapısı için parametrik sağkalım modelleri ise Rabinowitz vd. [17], Betensky vd. [18], Oller vd. [19], Zhang vd. [20], Lawless ve Babineau [21], Tian ve Cai [22], Zhang vd. [23], Zhu vd. [24] tarafından ele alınmıştır. Radhey ve Totawattage [25] çalışmasında aralıklı sansürlü veri yapısı için parametrik olmayan, yarı parametrik ve parametrik sağkalım modellerinin uygulamasını aynı veri üzerinde göstermiştir. Aralıklı sansürlü veri yapısının istatistiksel paket programlarındaki çözümlenmeleri üzerine çalışmalar literatürde mevcuttur. R paketindeki çalışmalar Gomez vd. [26], Fay [27], Anderson-Bergman [28], Delord [29], Bogaerts vd. [30]; SAS paketindeki çalışmalar So vd.

[31], Guo vd. [32], Bogaerts vd. [30]; STATA paketindeki çalışma ise Yang [33] tarafından yapılmıştır.

Bu çalışmanın amacı, aralıklı sansürlü veri yapısı için CRM'yi ve parametrik sağkalım modellerini incelemek ve sağdan sansürlü veri yapısı için kullanılan klasik yöntemler ile karşılaştırma yapmaktır.

Çalışmanın ikinci bölümünde sansürleme türleri ve aralıklı sansürlü veriler için CRM ve parametrik sağkalım modelleri hakkında genel bilgiler verilmiştir. Çalışmanın üçüncü bölümünde ise literatürde bulunan Primer Biliyer Siroz (PBS) [34] verisi kullanılarak sağdan ve aralıklı sansürlü veri yapısı için CRM ve parametrik sağkalım modelleri incelenmiş, sonuçlar yorumlanmıştır.

2. Materyal ve Metot

Sağkalım süresi verilerinin çözümlenmesi için özel yöntemlere ihtiyaç duyulmasının birkaç nedeni vardır. Sağkalım süresi verilerinin en önemli özelliği sansürlü birimlerin varlığıdır ve özellikle sağdan sansürlü birimleri içeren birçok çalışma literatürde yer almaktadır. Sağdan sansürlü veri yapısı için kullanılan yöntemler genel olarak aralıklı sansürlü veriler için geçerli değildir.

Aralıklı sansürleme ile ilgilenilen birimlerin veya sağkalım süresi süreçlerinin sürekli gözlem altında olmadığı kastedilmektedir. Sonuç olarak, başarısızlık ya da sağkalım süresi her zaman tam olarak gözlemlenemez ya da sağdan sansürlenemez. Aralıklı sansürlü bir birim için, sadece olayın gerçekleştiği bir aralık bilinmektedir. Sağdan sansürlü başarısızlık süreleri, aralıklı sansürlü başarısızlık süresinin özel bir durumu olarak kabul edilebilir, dolayısıyla aralık tek bir noktaya indirgenir veya sağdan sınırsız olur. Daha genel olarak, aralıklı sansürlü bir birim, birbiriyle çakışmayan birkaç aralığın birleşimi olarak tanımlanabilir [35].

Aralıklı sansürlü başarısızlık süresi verileri demografik, epidemiyolojik, finansal, tıbbi ve sosyolojik çalışmalar dâhil olmak üzere birçok alanda ortaya çıkmaktadır. Sağlık çalışmalarında, aralıklı sansürlü veri yapısı çeşitli biçimlerde ortaya çıkabilmektedir. Örneğin, bir birey hastalık durumundaki olası değişiklikleri klinik olarak gözlemlemek üzere programlanmış bir veya daha fazla gözlem zamanını kaçırabilir. Alternatif olarak, bireyler önceden belirlenmiş gözlem zamanlarından ziyade kendilerine uygun zamanlarda tedavi gördükleri merkezleri ziyaret edebilir. Her iki durumda da, sağlık durumlarındaki değişiklik ile ilgili veriler aralıklı olarak elde edilmektedir. Tüm birimler önceden belirlenmiş gözlem planını tam olarak takip etse bile sürekli bir değişken olduğu varsayılarak, ilgilenilen durum değişikliğinin tam olarak

gerçekleştiği süre net olarak bilinemez. Son durumda, başarısızlık süresi verileri bir gruptur, yani gözlemlenen her birim için aralıklı sansürlü veri yapısı, çakışmayan aralıkların toplamıdır.

Literatürde aralıklı sansürleme özelliği gösteren veri kümelerine fazla rastlanılmamaktadır. Bu durum veri kayıt sistemi ile ilgili eksikliklerden kaynaklanmaktadır. Literatürde yer alan bazı örnekler ise Hoel ve Walburg [36], Finkelstein ve Wolfe [3], Carvalho vd. [37], Kim vd.[38], Betensky ve Finkelstein [39], Goggins ve Finkelstein [40], Meyns vd. [41], Goethals vd. [42], Hough [43], Karvanen vd. [44] çalışmalarında yer almaktadır.

2.1. Aralıklı sansürlü veri yapısı

Bir sağkalım analizi çalışmasında bir birimin başarısızlık süresini gösteren T negatif olmayan bir rasgele değişken olsun. Eğer T tam olarak gözlemlenemez, L sağkalım süresinin sol uç noktası ve R sağkalım süresinin sağ uç noktası olmak üzere sadece [L,R] aralığında gözlemlenirse, T sağkalım süresine sahip bir birim aralıklı sansürlüdür denir ve $L \leq R$ olacak biçimde

$$T \in (L, R] \quad (1)$$

ile ifade edilir. Burada, $L=R$ kesin bir birim ve $R=\infty$ sağdan sansürlü bir birim olduğunu belirtir. Aralıklı sansürlü veri yapısı Durum I ver Durum II olmak üzere iki biçimde ele alınmaktadır.

Durum I Aralıklı Sansürlü Veri Yapısı: "Durum I aralıklı sansürlü sağkalım süresi verileri" terimi, tüm gözlemlenen aralıklarda ya sıfır zamanını ya da sonsuz zamanı içeren aralıklı sansürlü sağkalım süresi verilerinin belirtilmesi için kullanılmaktadır [6, 45, 46]. Diğer bir deyişle, her bir birim ya soldan ya da sağdan sansürlüdür, yani ya $L=0$ ya da $R=\infty$ 'dur. Durum I aralıklı sansürlü veri yapısı, her birim sadece bir kez gözlemlendiğinde oluşur ve ilgilenilen olay için gözlemlenen tek bilgi, olayın gözlem süresinden daha önce gerçekleşip gerçekleşmediğidir. Eşitlik 1'deki aralıklar yerine, Durum I aralıklı sansürlü veri yapısı için daha uygun bir gösterim, C gözlem süresi ve I gösterge fonksiyonu olmak üzere, $\{C, \delta = I(T = C)\}$ biçimindedir.

Durum I aralıklı sansürlü veri yapısı, genellikle demografik çalışmalardan ortaya çıkmış bir terim olan "mevcut durum verileri (current status data)" olarak da adlandırılır. Örneğin, öldürücü nitelikte olmayan tümörler üzerinde yapılan kesitsel çalışmalar sıklıkla Durum I aralıklı sansürlü veri yapısı üreten çalışmalardır.

Durum II Aralıklı Sansürlü Veri Yapısı: Hem L hem de $(0, \infty)$ olan en az bir aralık içeren aralıklı sansürlü veri yapısı, genel veya Durum II aralıklı sansürlü veri yapısı olarak adlandırılmaktadır [45-48]. Başka bir

deyişle, Durum II aralıklı sansürlü veri yapısı, sıfırdan farklı bazı sonlu aralıklar içeren aralıklı sansürlü veri yapısıdır. Durum II aralıklı sansürlü bir birimi göstermenin başka bir yolu, her bir birimin iki kez gözlemlendiğini varsayan, U ve V rastlantı değişkenleri olmak üzere, $U \leq V$ 'yi 1 olasılığı ile sağlayan aşağıdaki gösterimi kullanmaktır:

$$\{U, V, \delta_1 = I(T \leq U), \delta_2 = I(T \leq V), \delta_3 = 1 - \delta_1 - \delta_2\} \quad (2)$$

Bu gösterim teorik çıkarsamalarda sıklıkla kullanılır. $U=V=C$ alarak, Durum I aralıklı sansürlü veri yapısı Eşitlik 2.2 ile de tanımlanabilmektedir. Yu vd. [49] bu formülü kesin (sağkalım süreleri belli olan) birimleri içerecek biçimde genelleştirmiştir.

Eşitlik 2.2'nin diğer bir genelleştirilmiş biçimi, her bir birim için K rasgele bir tamsayı olmak üzere $U_1 \leq U_2 \leq \dots \leq U_K$ olan gözlem süreleri noktalarına sahip bir kümenin olduğunu varsaymaktır. Gözlemlenen bilgiler $U_0 = 0$ olmak üzere aşağıdaki biçimdedir:

$$\{(K, U_j, \delta_j = I(U_{j-1} < T \leq U_j)), j = 1, \dots, K\} \quad (3)$$

Bu formül ile tanımlanan sağkalım süresi verisi genellikle K durumu ya da karma durum aralıklı sansürlü veri (mixed case interval censored data) olarak adlandırılmaktadır [50, 51]. Eşitlik 3, Eşitlik 2'nin özel bir durumudur ve periyodik gözlemlenmeye sahip boylamsal çalışmalardan ortaya çıkan aralıklı sansürlü başarısızlık süresi verilerinin doğal bir gösterimidir.

Olabilirlik fonksiyonu, Eşitlik 1, Eşitlik 2, Eşitlik 3 ile verilen her üç gösterim için de aynıdır. Eşitlik 2 ve Eşitlik 3'teki gösterimlere sahip, bu biçimde toplanan veya verilen aralıklı sansürlü veri yapılarının uygulamada bulunması çok yaygın değildir. Bununla birlikte, teorik çıkarsamaların asimptotik özelliklerinin türetilmesi için sıklıkla ihtiyaç duyulan T'nin bağımsızlığı gibi bazı varsayımları koymak Eşitlik 1'deki gösterimine göre çok daha kolaydır. Veriler Eşitlik 2 veya Eşitlik 3'teki gösterimlerindeki gibi verilir ise, Eşitlik 1'deki gibi bir gösterimin elde edilmesi kolaydır. Bununla birlikte, gözleme süreci hakkında ek bilgi olmaksızın Eşitlik 1'i Eşitlik 3'e dönüştürmek imkânsızdır ve Eşitlik 1'de verilen gözlemleri Eşitlik 2'ye dönüştürmek de kolay değildir.

2.2. Aralıklı sansürlü veri yapısı için cox regresyon modeli

2.2.1. Durum I aralıklı sansürlü veri yapısı için cox regresyon modeli

Durum I aralıklı sansürlü veriler hayvan kanserojenliği deneyleri, demografik çalışmalar, ekonometri, epidemiyolojik çalışmalar ve güvenilirlik çalışmaları gibi birçok alanda ortaya çıkmaktadır. Gizli tümörlerde kanserojenlik denemeleri gibi bazı

durumlarda Durum I aralıklı sansürlü veriler, tümörün başlangıç süresi gibi ilgilenilen değişkenler hakkında mevcut olan tek bilgidir [52]. Yani, değişkenler doğrudan ölçülemez. Bir dönüm noktası olayı üzerindeki kesitsel çalışmalardan kaynaklananlar gibi bazı durumlarda Durum I aralıklı sansürlü veriler, olaya ilişkin zamanla ilgili kesin bilgileri veren verilere göre daha kolay ve daha güvenilir bilgiler sağlar. Bu gibi durumlara bir örnek, ilgilenilen olayın kronik hastalığın başlangıcı olduğu epidemiyolojik çalışmalardır [53-55]. Bir başka örnek ise söz konusu olayın örneğin ilk hamilelik veya evlilik olabileceği demografik çalışmalar olabilmektedir [56, 57].

Yarı parametrik bir sağkalım modeli olan CRM incelenirken en sık kullanılan yaklaşım yarı parametrik EÇO tahminidir. Bu olabilirlik yaklaşımı oldukça basittir, ancak kullanımı kolay değildir, çünkü olabilirlik fonksiyonu, sonlu boyutlu regresyon parametrelerinin ve sonsuz boyutlu bir nüans parametresinin birikimli temel tehlike fonksiyonu ya da temel sağkalım fonksiyonunun bir fonksiyonudur. Sonuç olarak, regresyon parametrelerini ve nüans parametresini aynı anda tahmin etmek gerekir. Bu da kısmi olabilirlik fonksiyonunun uygulandığı CRM'yi kullanan sağdan sansürlü veri yapısının regresyon çözümlemesinden farklıdır. İkinci yaklaşım olarak regresyon parametrelerine ilişkin çıkarsama için, nüans parametresini içermeyen ve özellikleri martingale teorisi ile kolayca elde edilebilen bir kısmi olabilirlik fonksiyonu elde edilebilir. Fakat, Durum I aralıklı sansürlü veriler için kısmi olabilirlik yaklaşımı mevcut değildir, bu sebeple tam olabilirlik fonksiyonunun kullanılması gerekmektedir.

Sonsuz boyutlu bir nüans parametresini içeren olabilirlik fonksiyonu ile uğraşmamak için, uygulamada genellikle sieve EÇO yöntemi kullanılır. Bu yöntemde, sonsuz boyutlu nüans parametresi sonlu boyutlu parametreler dizisine yaklaştırılmaktadır; yani, orijinal parametre uzayı, artan sonlu boyutlu alt uzayların (sieve'lerin) bir dizisi ile yaklaştırılır. Burada ele alınan bir problem için, yarı parametrik sağkalım modelinin regresyon parametresinin β ve birikimli temel tehlike fonksiyonunun $H_0(t)$ ile tanımlandığını varsayalım. Herhangi sonlu boyutlu bir örneklem için β ve $H_0(t)$ tahminleri, β ve sieve'ler için parametre uzaylarının çarpma üzerindeki olabilirlik fonksiyonunu maksimum yapmak suretiyle bulunabilir. Başka bir deyişle, sieve yöntemiyle sadece sonlu boyutlu bir parametre uzayı ile çalışmak gerekir. Sieve yönteminin başka bir avantajı, birikimli temel tehlike fonksiyonu $H_0(t)$ 'nin tahmin edicisinin, orijinal parametre uzayı üzerindeki tüm olabilirliği maksimize ederek verilen tahmin ediciden daha hızlı bir yakınsama oranına sahip olmasıdır [9].

Sınırsız boyutlu bir nüans parametresiyle uğraşmaktan kaçınabilecek başka bir yaklaşım, β 'nin

temel tahminini bazı tahmin denklemlerine dayandırmaktır. Burada β ile ilgili bazı tahmin denklemlerinin yansızlık gibi iyi özelliklere sahip olduğu ve nüans parametresini içermediği varsayılmaktadır. Bu yaklaşımın en büyük avantajı, tam ya da sieve EÇO yaklaşımına göre çok daha kolay bir şekilde uygulanabilmesidir. Ayrıca, bu yaklaşımdan elde edilen regresyon parametreleri tahminlerinin özellikleri nispeten daha kolay bir biçimde tespit edilebilmektedir.

n bağımsız birimden oluşan bir sağkalım analizi çalışması ve bu çalışmada, i . birim için iki rasgele değişkenin olduğunu varsayalım. $i = 1, \dots, n$ olmak üzere, T_i sağkalım süresini, C_i birimler ile ilgili gözlem süresini ve X_i açıklayıcı değişkenler vektörünü ifade etmektedir. T_i 'lerin dağılımı, regresyon parametresi β ve temel birikimli tehlike fonksiyonu $H_0(t)$ veya temel sağkalım fonksiyonu $S_0(t) = \exp(-H_0(t))$ ile belirlenir. Ayrıca, β ve $H_0(t)$ veya $S_0(t)$ hakkında çıkarım yapmak için, sadece Durum I aralıklı sansürlü verilerin olduğu ve $\{(C_i, \delta_i = I(T_i \leq C_i), X_i); i = 1, \dots, n\}$. Biçiminde ifade edildiği varsayımıyla, her birimin sadece bir kez C_i 'de gözlemlendiği ve C_i 'de sadece ilgilenilen sağkalım olayının C_i 'den önce mi yoksa C_i 'de mi gerçekleştiği bilinir.

Durum I aralıklı sansürlü veriler, sağkalım süresi verilerinin çözümlenmesinde en sık kullanılan regresyon modeli olan CRM ile analiz edilebilmektedir. X_i , T_i ve C_i 'nin bağımsız olduğu varsayımına dayanarak birikimli tehlike fonksiyonu

$$H(t; X_i) = H_0(t) \exp(X_i' \beta)$$

ve olabilirlik fonksiyonu

$$L(\beta, H_0) = \prod_{i=1}^n \exp[-(1 - \delta_i) e^{X_i' \beta} H_0(C_i)] [1 - \exp(-e^{X_i' \beta} H_0(C_i))]^{\delta_i} \quad (4)$$

olarak ifade edilir. Temel sağkalım fonksiyonu S_0 ve regresyon parametresi β ile olabilirlik fonksiyonu Eşitlik 2.5'teki gibidir:

$$L(\beta, S_0) = \prod_{i=1}^n [(S_0(C_i))^{(1-\delta_i) \exp(X_i' \beta)}] [1 - (S_0(C_i))^{\exp(X_i' \beta)}]^{\delta_i} \quad (5)$$

Huang [6] ve Huang ve Wellner [5] çalışmalarında β ve $S_0(t)$ için EÇO tahminlerine yer verilmiştir.

2.2.2. Durum II aralıklı sansürlü veri yapısı için cox regresyon modeli

Durum I aralıklı sansürlü veri yapısı ile karşılaştırıldığında, Durum II aralıklı sansürlü veri yapısının ilgilenilen temel sağkalım süresi hakkında daha fazla bilgi sağladığı açıktır. Ayrıca, Durum II aralıklı sansürlü veri yapısı için, Durum I aralıklı sansürlü veri yapısında olduğu gibi, sadece bir değişken yerine gözlem sürelerini temsil eden iki

veya daha fazla değişken ile ilgilenmek gerekmektedir. Durum II aralıklı sansürlü veri yapısının regresyon analizi, hem hesaplama hem de teoride Durum I aralıklı sansürlü veri yapısından daha karmaşıktır.

n tane bağımsız birimden oluşan ve aralıklı sansürlü veri yapılarının ortaya çıkmasına neden olan bir çalışma ele alınsın. İlgili sağkalım süreleri için

$$\{(L_i, R_i], X_i; i = 1, \dots, n\} \quad (6)$$

tanımlansın. Bu durumda $(L_i, R_i]$, gözlemlenen i. birim için sağkalım olayının gerçekleştiği aralığı belirtir ve $X_i; i = 1, \dots, n$ biriminden gelen değişkenlerin p-boyutlu vektörünü temsil eder. $S(t; X)$, X değişkenleri için sağkalım fonksiyonunu gösterir. Buna göre olabilirlik fonksiyonu

$L = \prod_{i=1}^n [S(L_i, X_i) - S(R_i, X_i)]$ ile orantılıdır. Burada $L_i < R_i, i=1, \dots, n$ varsayımı yapılmaktadır.

$S(t; X)$ 'in CRM ile ifade edildiği varsayılmaktadır. Yukarıda verilen olabilirlik fonksiyonunun logaritması, regresyon parametresi β ve temel sağkalım fonksiyonu S_0 olmak üzere $I(\beta, S_0) = \sum_{i=1}^n \log\{[S_0(L_i)]^{\exp(X_i' \beta)} - [S_0(R_i)]^{\exp(X_i' \beta)}\}$ biçimindedir. β ve S_0 için EÇO tahmin edicileri ve bunların bazı asimptotik özellikleri Finkelstein [4] ve Huang ve Wellner'in [5] tarafından incelenmiştir.

2.3. Aralıklı sansürlü veri yapısı için parametrik sağkalım modelleri

n tane bağımsız birimden oluşan bir çalışma düşünölsün. $i=1, \dots, n$ olmak üzere, i birim için sağkalım süresi T_i ile gösterilsin ve T_i 'nin sağkalım fonksiyonu $S(t, \theta)$ olan bir parametrik modele sahip olduđu varsayölsün. Burada $\theta = (\theta_1, \dots, \theta_p)$ bilinmeyen parametreleri belirtir. Ayrıca, sadece aralıklı sansürlü veri yapılarının olduđu ve $\{(L_i, R_i], X_i; i = 1, \dots, n\}$ biçimine sahip olduđu varsayölsün. Buradaki $(L_i, R_i]$, T_i 'nin bulunduđu aralığı gösterir ve $X_i; i = 1, \dots, n$ olmak üzere i birimle ilişkili ortak değişken vektörüdür. Tüm $i=1, \dots, n$ için $L_i < R_i$ olduđu varsayölsün, olabilirlik fonksiyonu $L(\theta) = \sum_{i=1}^n L_i(\theta) = \prod_{i=1}^n [S(L_i, \theta) - S(R_i, \theta)]$ ile orantılıdır.

2.3.1. Genel parametrik modeller ile çıkarım

θ ile ilgili çıkarımlara standart bir yaklaşım, $L(\theta)$ 'yı maksimum yapan θ değerini EÇO yöntemi ile tahmin etmek ve $L(\theta)$ 'dan elde edilen skor ve olabilirlik oranı (LR) test istatistiklerini kullanmaktır. Burada bağımsız sansürleme varsayımıyla genel olarak, standart büyük örneklem olabilirlik teorisi uygulanır. Özellikle, EÇO tahmin edicisinin dağılımları için asimptotik yaklaşımlar, skor test istatistiğı ve LR test istatistiğı kullanılır ve bunlar basit çıkarım yaklaşımları sağlar. Skor test istatistiğı ve Fisher bilgi

matrisi, $U(\theta) = \sum_{i=1}^n U_i(\theta) = \sum_{i=1}^n \frac{\partial}{\partial \theta} L_i(\theta)$ ve $I(\theta) = \sum_{i=1}^n I_i(\theta) = \sum_{i=1}^n \frac{\partial^2}{\partial \theta \partial \theta'} L_i(\theta)$ olarak verilsin. Belirli koşullar altında, θ 'nın EÇO tahmin edicisi olan $\hat{\theta}$ tutarlıdır. Ayrıca, n büyük olduğunda, $U(\hat{\theta})=0$ için tek bir çözüm vardır ve dağılımına ortalama θ ve kovaryans matrisi $I^{-1}(\theta)$ olan çok değişkenli normal dağılım ile yaklaşılabılır. Başka bir deyişle, $\hat{\theta} \sim N(\theta, I^{-1}(\theta))$, θ için hipotezleri test etmek ve aralık tahminlerini türetmek için kullanılabilir. $\hat{\theta}$ 'yı belirlemek için, herhangi bir kök bulma prosedürü veya Newton-Raphson algoritması kullanılabilir.

$I(\theta_0)$ 'ın tekil olmadığı varsayölsün. Newton-Raphson algoritmasında, θ 'nın başlangıç değeri θ_0 $\theta_1 = \theta_0 - I^{-1}(\theta_0) U(\theta_0)$ ile yakınsayana kadar tekrar ederek güncellenir. θ_0 biliniyorsa, $H_1: \theta = \theta_0$ hipotezini test etmek için, p serbestlik derecesi ile asimptotik χ^2 dağılımı olan $U'(\theta_0) - I^{-1}(\theta_0) U(\theta_0)$ skor test istatistiğinin kullanılması uygundur.

$\theta = (\theta_1', \theta_2')$ eşitliğinde yer alan θ 'nın θ_1 ve θ_2 bileşenlerinin, sırasıyla k ve p-k boyutlarına sahip olduğunu varsayalım. Buna göre uygulamada, θ_{10} bilindiğinde $H_2: \theta_1 = \theta_{10}$ daha bilinen bir hipotez haline gelir. Bu durumda, θ aynı biçimde $U(\theta), U'(\theta) = [U_1'(\theta_1, \theta_2), U_2'(\theta_1, \theta_2)]$ olarak parçalanır. Burada U_1 ve U_2 sırasıyla θ_1 ve θ_2 'ye karşılık gelen k ve p-k boyutlarına sahiptir. $\hat{\theta}_{20}, U_2(\theta_{10}, \theta_2) = 0$ çözümü ile verilen θ_2 tahminini; $I^{11}(\theta), \theta_1$ ile $I^{-1}(\theta)$ 'ya karşılık gelen alt matrisi gösterir. Bu durumda, H_2 testi $\hat{\theta} = (\hat{\theta}_1', \hat{\theta}_2')$ olduğunda aşağıda verilen istatistiklere dayanabilir:

$$(\hat{\theta}_1 - \theta_{10})' [I^{11}(\theta_{10}, \hat{\theta}_{20})]^{-1} (\hat{\theta}_1 - \theta_{10}) \quad (7)$$

veya

$$U_1'(\theta_{10}, \hat{\theta}_{20}) I^{11}(\theta_{10}, \hat{\theta}_{20}) U_1(\theta_{10}, \hat{\theta}_{20}) \quad (8)$$

Her iki istatistik de k serbestlik derecesi ile bir asimptotik χ^2 dağılımına sahiptir. Yukarıda verilen yöntemlerde, bazen gözlemlenmesi zor olabilen Fisher bilgi matrisi $I(\theta)$ belirlenmelidir. Dezavantajı olmayan bu yaklaşım, LR test istatistiğine dayanmaktadır. H_1 ve H_2 hipotezlerini test etmek için $-2 \log[LR(\theta_0)]$ ya da $-2 \log[LR(\theta_{10} - \hat{\theta}_{20})]$ istatistiğı kullanılabilir. Bu istatistikler, sırasıyla p ve k serbestlik dereceleri ile asimptotik χ^2 dağılımına sahiptir [58].

2.3.2. Aralıklı sansürlü veriler için hızlandırılmış başarısızlık süresi modeli

Gözlemlenen aralıklı başarısızlık süresi verisi, $\{K_i, (U_{ij}, \delta_{ij} = I(U_{ij-1} < T_i \leq U_{ij}))_{j=1}^{K_i}, X_i; i = 1, \dots, n\}$ ile ifade edilsin. Burada, i. birim için açıklayıcı değişkenler vektörü X_i , birim sayısı K_i , sağkalım süresi T_i ve $U_{i0} = 0$ olan gözlemlenen zaman noktaları $U_{i1} < \dots < U_{iK_i}$ ile gösterilmiştir. n bağımsız birimden

oluşan çalışmadaki her birim, sağkalım süresinden bağımsız olduğu varsayılan bir dizi zaman noktasında gözlemlenmektedir. T_i 'lerin HBS modeline sahip olduğu, X_i bilindiğinde U_{ij} 'lerin ve T_i 'nin bağımsız olduğu ve sadece birinin regresyon parametresi ile ilgilendiği varsayılmaktadır.

$\log T = X'\beta + W$ eşitliğindeki W dağılımı F ile gösterilsin ve p -boyutlu b vektörü için $U_{ij}(b) = \log(U_{ij}) - X_i'b, j = 1, \dots, K; i = 1, \dots, n$ tanımlansın. Ayrıca, $Y_{ij} = I(T_i \leq U_{ij})$ biçiminde ifade edilsin ve her i için U_{iL} ve U_{iR} sırasıyla, $Y_{ij}=0$ için son gözlem süresi ve $Y_{ij}=1$ için ilk gözlem süresi olan iki U_{ij} olsun. Yani, (U_{ij}, U_{iR}) , T_i 'nin gözlemlendiği aralıktır. Buna göre olabilirlik fonksiyonu,

$$L(\beta, F) = \prod_{i=1}^n [F(U_{iR}(\beta)) - F(U_{iL}(\beta))] \quad (9)$$

biçiminde yazılır. Burada $U_{iL}(b) = \log(U_{iL}) - X_i'b$ ve $U_{iR}(b) = \log(U_{iR}) - X_i'b$ 'dir. β ile ilgili çıkarım doğrusal sıralı istatistiklere dayanan bir tahmin denklemi ile elde edilmektedir [58].

3. Sayısal Örnek

Çalışmamızda aralıklı sansürlü ve sağdan sansürlü veri yapısı için hem CRM'yi hem de parametrik sağkalım modellerini incelemek amacıyla, literatürde yer alan Primer Biliyer Siroz, PBS, (Primary Biliary Cirrhosis, PBC) verisi kullanılmıştır. 312 hastaya ait PBS veri kümesine ilişkin veriler Ocak 1924-Mayıs 1984 yılları arasında Mayo Clinic'de toplanmıştır [34]. 312 hastanın 158'ine D-penisilin ilacı verilmiş ve 154 hasta ise placebo grubu olarak seçilmiştir. İlk 6 ay ve 6. aydan sonra her yıl olmak üzere, serum bilirubin düzeyleri tekrarlı biçimde ölçülmüştür [59].

Literatürde PBS verisi kullanılarak farklı sağkalım çözümlemesi yöntemleri incelenmiş olup, bu çalışmalar Dil (2019)'da detaylı olarak verilmiştir. PBS verisi aralıklı olarak ölçülen değerleri içerdiği için veri yapısı, aralıklı sansürlü olarak ele alınabilmektedir. Bu çalışma, sağkalım analizi için aralıklı sansürleme literatüründe PBS verisi kullanılarak yapılan ilk çalışma olma özelliğine sahiptir.

Çalışmadaki sonuçlar, Stata 15'in stcox, streg ve stintreg komutlarından ve R programının icenReg ve icensBKL paketlerinden yararlanılarak elde edilmiştir.

Bu çalışmada başarısızlık ölüm olarak tanımlanmıştır. sağkalım süresi, hastaların gözlemlenmeye başlamasından itibaren ölüm gerçekleşene kadar geçen süre (yıl) olarak belirlenmiştir. 140 (%44.9) hastanın öldüğü ve 172 (%55.1) hastanın ise durdurulduğu görülmüştür. Veri kümesinde hastaların yaşı, cinsiyeti, kullandıkları ilaç, asit, hepatomegali, deride bozukluk, ödem, histolojik evre, log serum bilirubin (logb) (mg/dl), log albumin

(logalbumin), log protrombin süresi (logprotime) değişkenleri yer almaktadır. Kullanılan kategorik değişkenler ve düzeyleri Tablo 1'de verilmiştir.

Tablo 1. Kategorik değişkenler ve düzeyleri

Değişken	Düzyeler	n	%
İlaç	0-Placebo	158	50.6
	1-D-penisilin	154	49.4
Cinsiyet	0-Kadın	276	88.5
	1-Erkek	36	11.5
Asit	0-Yok	257	82.4
	1-Var	55	17.6
Hepatomegali	0-Yok	165	52.9
	1-Var	147	47.1
Deride Bozukluk	0-Yok	208	66.7
	1-Var	104	33.3
	0-Var	165	52.9
Ödem	1-İlaça rağmen ödem(Odem1)	80	25.6
	2-Yok (Odem2)	67	21.5
	1-Evre 1	11	3.5
Histolojik Evre	2-Evre 2	45	14.4
	3-Evre 3	102	32.7
	4-Evre 4	154	49.4

Çalışmada yaş, logb, logalbumin ve logprotime değişkenleri sürekli değişken olarak ele alınmış olup, ortalamalar, yaş için 49.999 ± 0.600 , logb için 1.196 ± 0.075 , logalbumin için 1.105 ± 0.011 ve logprotime için 2.465 ± 0.009 olarak elde edilmiştir.

PBS, ilk 6 ay ve 6. aydan sonra her yıl tekrarlı olarak ölçülmüş bir veri kümesidir. Bu çalışmada, sağdan sansürlü ve aralıklı sansürlü veri yapılarını aynı veri kümesi üzerinde inceleyebilmek için veride bazı düzenlemeler yapılmıştır. Sağdan sansürlü olarak oluşturulan veri yapısı için sağkalım süresi, gözlem süresinin başlangıcından bitiş noktasına kadar olan toplam süre olarak ele alınmıştır. Aralıklı sansürlü olarak oluşturulan veri yapısı için ise sadece son aralık ele alınmış ve çözümleme bu aralık dikkate alınarak yapılmıştır.

3.1. Sağdan sansürlü veri yapısı için model sonuçları

3.1.1. Sağdan sansürlü veri yapısı için cox regresyon modeli sonuçları

PBS veri kümesi için CRM, p değeri <0.001 olduğundan %95 güven düzeyinde anlamlıdır. CRM için AIC değeri 1290.177 ve BIC değeri 1342.579 olarak elde edilmiştir. Tablo 2'de CRM sonuçları verilmiştir. CRM'de değişken düzeylerini yorumlayabilmek için düzeylerden birinin referans kategorisi olarak alınması gerekmektedir. Bu sebeple çalışmamızda, referans kategorisi tüm değişkenler için birinci düzey olarak alınmıştır.

Tablo 2'de verilen p değerleri incelendiğinde yaş ($p<0.001$), asit ($p=0.012$) ve log(b) ($p<0.001$) değişkenlerinin %95 güven düzeyinde anlamlı ($p<0.05$) olduğu görülmüştür. CRM sonuçlarına göre,

vücudunda asit olan hastaların vücudunda asit olmayan hastalara göre yaklaşık 1.8483 kat ($e^{0.6143}=1.8483$) daha fazla riskli olduğu söylenebilmektedir. Yaş değişkenindeki 1 birimlik artış, başarısızlık riskini 1.0572 birim artırır. Log(b) değişkenindeki 1 birimlik artış ise başarısızlık riskini 2.3268 birim artırmaktadır.

CRM'nin veri kümesine uygulanabilmesi için orantılı tehlikeler (OT) varsayımının sağlanması gerekmektedir. Bu çalışmada, ilk olarak Schoenfeld artıkları yöntemi kullanılarak OT varsayımı incelenmiş ve sonuçlar Tablo 3'de verilmiştir.

Yokluk hipotezi "OT varsayımı sağlanmaktadır" biçiminde kurulmaktadır ve buna göre elde edilen sonuçlar incelendiğinde, tüm değişkenlerin ($p>0.05$) OT varsayımını sağladığı %95 güven düzeyinde görülmüştür. Grafikselsel yöntemlerden $-\ln(-\ln S(t))$ grafiği ve zamana bağlı açıklayıcı değişkenler yöntemi kullanılarak da OT varsayımı incelenmiş ve bu varsayımın PBS verisi için sağlandığı sonucuna ulaşılmıştır.

Buna göre PBS verisi için OT varsayımı sağlanmaktadır ve CRM sonuçlarının kullanılması ve yorumlanması uygundur.

3.1.2. Sağdan sansürlü veri yapısı için parametrik sağkalım modelleri sonuçları

PBS veri kümesini çözümlemek için OT biçimindeki parametrik sağkalım modelleri olarak üstel, Weibull ve Gompertz; HBS biçimindeki parametrik sağkalım modelleri olarak üstel, Weibull, Gamma, log-normal ve log-lojistik sağkalım modelleri kullanılmıştır. LR ile modellerin anlamlılığı test edilmiş ve istatistiksel olarak tüm modeller anlamlı ($p<0.05$) bulunmuştur. Tablo 4'de model seçiminde kullanılan AIC ve BIC değerleri verilmiştir. Tablo 4'deki sonuçlara göre, OT biçimindeki parametrik sağkalım modellerinden Gompertz OT modeli, HBS biçimindeki parametrik sağkalım modellerinden ise Weibull HBS modelinin en uygun modeldir. Şekil 1'de verilen Cox-Snell artıklarına ait grafikler ile uygun olarak belirlenen bu modellerin doğruluğu ispatlanmıştır.

Tablo 2. Sağdan sansürlü veri yapısı için CRM sonuçları

Değişken	$\hat{\beta}$	Std. hata	p değeri	$\exp(\hat{\beta})$	$\exp(\hat{\beta})$ için %95 güven aralığı
Yaş	0.0556	0.0093	0.000	1.0572	1.0381, 1.0767
İlaç	0.0125	0.1772	0.944	1.0126	0.7155, 1.4331
Cinsiyet	-0.2792	0.2493	0.263	0.7564	0.4640, 1.2331
Asit	0.6143	0.2453	0.012	1.8483	1.1428, 2.9893
Hepatomegali	-0.0997	0.2302	0.665	0.9051	0.5764, 1.4212
Deride Bozukluk	0.2830	0.2314	0.221	1.3271	0.8433, 2.0886
Ödem1	-0.0520	0.2470	0.833	0.9493	0.5850, 1.5405
Ödem2	-0.0717	0.2556	0.779	0.9308	0.5640, 1.5362
Logb	0.8445	0.1035	0.000	2.3268	1.8997, 2.8500
Logalbumin	-0.2479	0.4498	0.581	0.7804	0.3232, 1.8845
Logprotime	-0.9506	0.6810	0.163	0.3865	0.1017, 1.4683
Histolojik Evre2	-1.1032	0.9222	0.232	0.3318	0.0544, 2.0223
Histolojik Evre3	-1.0829	0.7692	0.159	0.3386	0.0750, 1.5290
Histolojik Evre4	-0.7767	0.7608	0.307	0.4599	0.1035, 2.0429

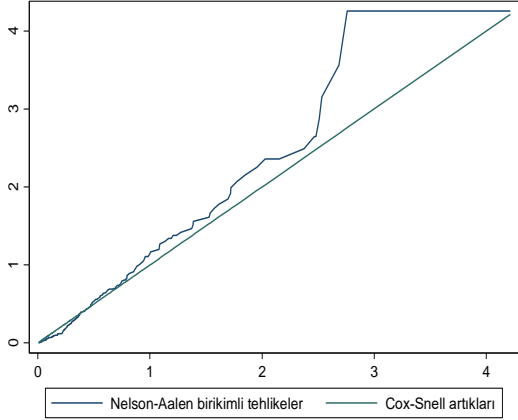
Tablo 3. OT varsayımı için Schoenfeld artıkları yöntemine ait sonuçlar

Değişken	r^*	Ki-kare	Std. Hata	p değeri
Yaş	0.0052	0.00	1	0.950
İlaç	0.0082	0.01	1	0.922
Cinsiyet	0.0383	0.25	1	0.618
Asit	-0.0547	0.46	1	0.500
Hepatomegali	-0.0292	0.14	1	0.711
Deride Bozukluk	-0.0615	0.65	1	0.422
Ödem1	0.0132	0.03	1	0.859
Ödem2	-0.0836	1.13	1	0.289
Logb	-0.0016	0.00	1	0.983
Logalbumin	-0.0674	0.64	1	0.425
Logprotime	0.0571	0.88	1	0.348
Histolojik Evre2	-0.1014	1.38	1	0.239
Histolojik Evre3	-0.0622	0.56	1	0.454
Histolojik Evre4	-0.0498	0.36	1	0.549

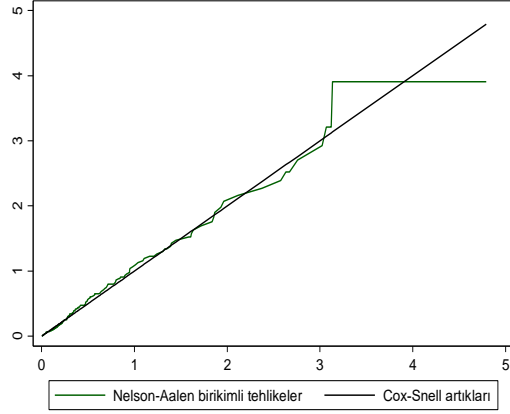
Tablo 5'te PBS veri kümesi için kullanılan Gompertz OT modelinden elde edilen sonuçlar verilmiştir. Tablo 5'te verilen p değerleri incelendiğinde yaş ($p=0.000$), asit ($p=0.013$) ve log b ($p<0.001$) değişkenleri %95 güven düzeyinde anlamlı ($p<0.05$) bulunmuştur. Gompertz OT modeli sonuçlarına göre, vücudunda asit olan hastaların vücudunda asit olmayan hastalara göre yaklaşık 1.8356 kat

($e^{0.6074}=1.8356$) daha fazla riskli olduğu söylenebilmektedir. Yaş değişkenindeki 1 birimlik artış, başarısızlık riskini 1.0577 birim artırır. Log(b) değişkenindeki 1 birimlik artış ise, başarısızlık riskini 2.3385 birim artırmaktadır.

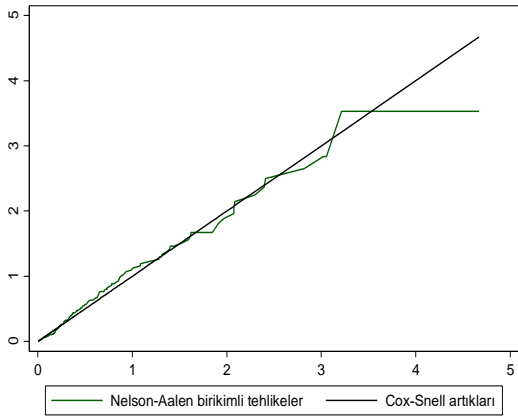
Tablo 6'da Weibull HBS modeli için elde edilen sonuçlar verilmiştir:



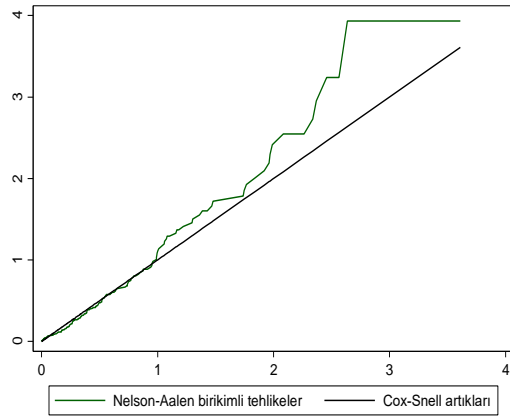
(a) Üstel



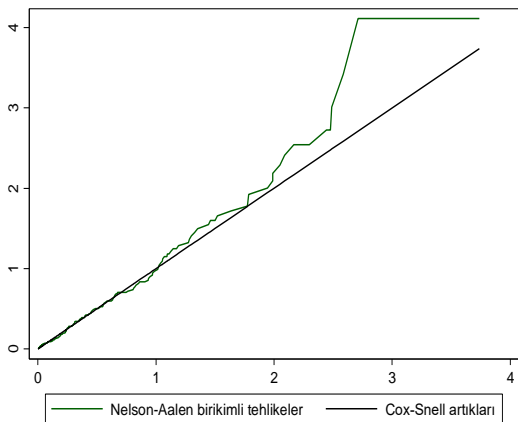
(b) Weibull



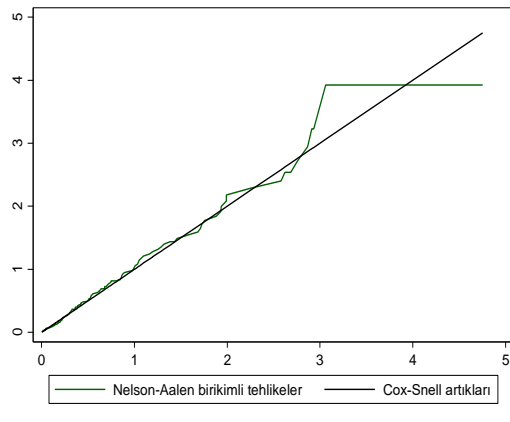
(c) Gompertz



(d) Log-normal



(e) Log-lojistik



(f) Gamma

Şekil 1. Cox-Snell artıkları grafikleri

Tablo 4. Sağdan sansürlü veri yapısı için parametrik sağkalım modellerine ait AIC ve BIC değerleri

Kriter	OT Biçimindeki Parametrik Sağkalım Modelleri		
	Üstel	Weibull	Gompertz
AIC	571.089	560.978	559.398
BIC	627.234	620.866	619.286

Kriter	HBS Biçimindeki Parametrik Sağkalım Modelleri				
	Üstel	Weibull	Log-lojistik	Log-normal	Gamma
AIC	571.089	560.978	562.507	572.313	562.233
BIC	627.234	620.866	622.395	632.201	625.864

Tablo 5. Sağdan sansürlü veri yapısı için Gompertz OT modeli sonuçları

Değişken	$\hat{\beta}$	Std. hata	p değeri	$\exp(\hat{\beta})$	$\exp(\hat{\beta})$ için %95 güven aralığı
Yaş	0.0561	0.0093	0.000	1.0577	1.0386, 1.0771
İlaç	0.0045	0.1767	0.980	1.0045	0.7105, 1.4202
Cinsiyet	-0.2810	0.2493	0.260	0.7550	0.4631, 1.2308
Asit	0.6074	0.2451	0.013	1.8356	1.1353, 2.9679
Hepatomegali	-0.1058	0.2296	0.645	0.8996	0.5736, 1.4109
Deride Bozukluk	0.2730	0.2312	0.238	1.3139	0.8352, 2.0671
Ödem1	-0.0700	0.2464	0.776	0.9324	0.5753, 1.5113
Ödem2	-0.0897	0.2552	0.725	0.9142	0.5544, 1.5077
Logb	0.8495	0.1034	0.000	2.3385	1.9096, 2.8636
Logalbumin	-0.2900	0.4502	0.520	0.7483	0.3097, 1.8084
Logprottime	-0.9034	0.6740	0.180	0.4052	0.1081, 1.5185
Histolojik Evre2	-1.0975	0.9224	0.234	0.3337	0.0547, 2.0345
Histolojik Evre3	-1.0535	0.7686	0.170	0.3487	0.0773, 1.5727
Histolojik Evre4	-0.7595	0.7619	0.319	0.4679	0.1051, 2.0827

Tablo 6. Sağdan sansürlü veri yapısı için Weibull HBS modeli sonuçları

Değişken	$\hat{\beta}$	Std. hata	p değeri	$\exp(\hat{\beta})$	$\exp(\hat{\beta})$ için %95 güven aralığı
Yaş	-0.0420	0.0072	0.000	0.9589	-0.0562, -0.0279
İlaç	-0.0089	0.1361	0.948	0.9911	-0.2756, 0.2579
Cinsiyet	0.2157	0.1914	0.260	1.2407	-0.1595, 0.5909
Asit	-0.4717	0.1889	0.012	0.6239	-0.8418, -0.1016
Hepatomegali	0.0741	0.1766	0.675	1.0769	-0.2720, 0.4203
Deride Bozukluk	-0.2173	0.1782	0.223	0.8047	-0.5667, 0.1321
Ödem1	0.0572	0.1895	0.763	1.0589	-0.3143, 0.4287
Ödem2	0.0378	0.1948	0.846	1.0385	-0.3440, 0.4196
Logb	-0.6408	0.0814	0.000	0.5269	-0.8004, -0.4812
Logalbumin	0.2397	0.3476	0.490	1.2709	-0.4416, 0.9210
Logprottime	0.7064	0.5194	0.174	2.0267	-0.3115, 1.7243
Histolojik Evre2	0.8430	0.7141	0.238	2.3233	-0.5567, 2.2427
Histolojik Evre3	0.8042	0.5936	0.175	2.2349	-0.3592, 1.9677
Histolojik Evre4	0.5814	0.5883	0.323	1.7885	-0.5718, 1.7345

Başarısızlık süresini etkileyen faktörleri belirlemek için kullanılan Weibull HBS modeli için elde edilen sonuçlar incelendiğinde yaş ($p < 0.001$), asit ($p = 0.012$) ve $\log(b)$ ($p < 0.001$) değişkenlerinin %95 güven düzeyinde anlamlı ($p < 0.05$) olduğu görülmüştür.

Vücudunda asit olmayan hastaların ortalama sağkalım süresi, vücudunda asit olan hastaların ortalama sağkalım süresine göre yaklaşık 1.6 ($1/\exp(-0.4717)$) kat daha uzundur. Yaş değişkenindeki 1 birimlik artışın ortalama sağkalım süresini 0.9589 birim ve $\log(b)$ değişkenindeki 1 birimlik artışın ise ortalama sağkalım süresini 0.5269 birim azalttığı söylenebilmektedir.

3.2. Aralıklı sansürlü veri yapısı için model sonuçları

Bu bölümde veriler aralıklı sansürlü birimler olarak düzenlenmiş ve buna göre analiz edilmiştir.

3.2.1. Aralıklı sansürlü veri yapısı için cox regresyon modeli sonuçları

PBS veri kümesindeki veriler aralıklı sansürlü olarak ele alındığında CRM, p değeri < 0.001 olduğu için %95 güven düzeyinde anlamlıdır. CRM için AIC değeri 2828.114, BIC değeri 2880.516 olarak elde edilmiştir. Tablo 7'de PBS veri kümesindeki aralıklı sansürlü

veri yapısı için CRM sonuçları verilmiştir. Tablo 7’de verilen p değerlerine göre yaş ($p<0.001$) ve $\log(b)$ ($p<0.001$) değişkenleri %95 güven düzeyinde anlamlı ($p<0.05$) bulunmuştur. Bu sonuçlara göre, yaş değişkenindeki 1 birimlik artış, başarısızlık riskini 1.0293 birim artırır. $\log(b)$ değişkenindeki 1 birimlik artış ise, başarısızlık riskini 1.5686 birim artırmaktadır.

Sağdan sansürlü veri yapısı için OT varsayımı incelenebilirken, aralıklı sansürlü veri yapısı için inceleme yapılamamaktadır. Aralıklı sansürlü veriler için kullanılan CRM’nin bir dezavantajı, OT varsayımını inceleyen bir yazılımın henüz olmamasıdır.

3.2.2. Aralıklı sansürlü veri yapısı için parametrik sağkalım modelleri sonuçları

PBS veri kümesini çözümlemek için OT biçimindeki parametrik sağkalım modelleri olarak üstel, Weibull ve Gompertz; HBS biçimindeki parametrik sağkalım modelleri olarak üstel, Weibull, Gamma, log-lojistik ve log-normal sağkalım modelleri kullanılmıştır. LR ile modellerin anlamlılığı test edilmiş ve istatistiksel olarak tüm modeller anlamlı ($p<0.05$) bulunmuştur. Tablo 8’de model seçiminde kullanılan AIC ve BIC değerleri verilmiştir. Tablo 8’deki sonuçlara göre OT biçimindeki parametrik sağkalım modelleri içerisinde Gompertz OT modelinin, HBS biçimindeki parametrik sağkalım modelleri içerisinde Gamma HBS modelinin en uygun model olduğu görülmüştür

($p=0.040$), ödem2 ($p=0.014$) ve $\log b$ ($p<0.001$) değişkenleri %95 güven düzeyinde anlamlı ($p<0.05$) bulunmuştur. Gompertz OT modeli sonuçlarına göre, vücudunda asit olan hastaların vücudunda asit olmayan hastalara göre yaklaşık 1.4421 kat ($e^{0.3661}=1.4421$) daha fazla riskli olduğu söylenebilmektedir. Vücudunda ödem olan hastaların vücudunda ödem olmayan hastalara göre yaklaşık 1.62 ($1/\exp(-0.4832)$) kat daha fazla riskli olduğu görülmüştür. Yaş değişkenindeki 1 birimlik artış, başarısızlık riskini 1.0302 birim artırır. $\log b$ değişkenindeki 1 birimlik artış ise başarısızlık riskini 1.5641 birim artırmaktadır.

Başarısızlık süresini etkileyen faktörleri belirlemek için kullanılan Gamma HBS modeli için elde edilen sonuçlar incelendiğinde yaş ($p<0.001$), ödem2 ($p=0.003$) ve $\log b$ ($p<0.001$) değişkenlerinin %95 güven düzeyinde önemli ($p<0.05$) olduğu görülmüştür.

Vücudunda ödem olmayan hastaların ortanca sağkalım süresi, vücudunda ödem olan hastaların ortanca sağkalım süresine göre yaklaşık 1.3221 ($\exp(0.2792)$) kat daha uzundur. Yaş değişkenindeki

Tablo 7. Aralıklı sansürlü veri yapısı için CRM sonuçları

Değişken	$\hat{\beta}$	Std. hata	p değeri	$\exp(\hat{\beta})$	$\exp(\hat{\beta})$ için %95 güven aralığı
Yaş	0.0289	0.0070	0.000	1.0290	1.0153, 1.0435
İlaç	-0.0319	0.1476	0.829	0.9686	0.7253, 1.2936
Cinsiyet	-0.3225	0.2518	0.200	0.7244	0.4422, 1.1865
Asit	0.2858	0.2923	0.328	1.3310	0.7504, 2.3601
Hepatomegali	0.1184	0.1410	0.401	1.1260	0.8538, 1.4841
Deride Bozukluk	0.1947	0.1751	0.266	1.2150	0.8620, 1.7124
Ödem1	-0.0747	0.1769	0.673	0.9281	0.6561, 1.3126
Ödem2	-0.4606	0.2691	0.869	0.6309	0.3723, 1.0691
Logb	0.4502	0.0806	0.000	1.5690	1.3394, 1.8371
Logalbumin	0.3782	0.4218	0.370	1.4600	0.6386, 3.3364
Logprotime	-0.8705	0.5685	0.126	0.4187	0.1374, 1.2761
Histolojik Evre2	0.1600	0.4462	0.720	1.1730	0.4894, 2.8140
Histolojik Evre3	-0.1294	0.4314	0.764	0.8786	0.3772, 2.0464
Histolojik Evre4	-0.1036	0.4253	0.808	0.9016	0.3917, 2.0751
				Üstel	Weibull
					Gompertz
AIC	1492.407			1393.990	1337.837
BIC	1548.552			1453.878	1397.725

Tablo 8. Aralıklı sansürlü veri yapısı için parametrik sağkalım modellerine ait AIC ve BIC değerleri

Kriter	HBS Biçimindeki Parametrik Sağkalım Modelleri				
	Üstel	Weibull	Log-lojistik	Log-normal	Gamma
AIC	1492.407	1393.990	2135.830	1492.066	1372.076
BIC	1548.552	1453.878	2195.718	1551.954	1435.707

Tablo 9’da Gompertz OT ve Tablo 10’da Gamma HBS modellerine ait sonuçlar verilmiştir. Tablo 9’da verilen p değerleri incelendiğinde yaş ($p<0.001$), asit

Tablo 9. Aralıklı sansürlü veri yapısı için Gompertz OT modeli sonuçları

Değişken	$\hat{\beta}$	Std. hata	p değeri	$\exp(\hat{\beta})$	$\exp(\hat{\beta})$ için %95 güven aralığı
Yaş	0.0298	0.0066	0.000	1.0302	1.0170, 1.0436
İlaç	-0.0184	0.1204	0.879	0.9818	0.7753, 1.2431
Cinsiyet	-0.3672	0.1940	0.058	0.6927	0.4736, 1.0132
Asit	0.3661	0.1787	0.040	1.4421	1.0160, 2.0468
Hepatomegali	0.1205	0.1345	0.370	1.1281	0.8667, 1.4685
Deride Bozukluk	0.1768	0.1410	0.210	1.1934	0.9053, 1.5732
Ödem1	-0.0920	0.1602	0.566	0.9121	0.6664, 1.2484
Ödem2	-0.4832	0.1970	0.014	0.6168	0.4192, 0.9075
Logb	0.4473	0.0633	0.000	1.5641	1.3815, 1.7708
Logalbumin	0.3525	0.3565	0.323	1.4226	0.7074, 2.8609
Logprottime	-0.8039	0.5249	0.126	0.4476	0.1600, 1.2522
Histolojik Evre2	0.2247	0.4354	0.606	1.2519	0.5333, 2.9389
Histolojik Evre3	-0.0964	0.4098	0.814	0.9081	0.4068, 2.0273
Histolojik Evre4	-0.0445	0.4081	0.913	0.9565	0.4299, 2.1284

Tablo 10. Aralıklı sansürlü veri yapısı için Gamma HBS modeli sonuçları

Değişkenler	$\hat{\beta}$	Std. hata	p değeri	$\exp(\hat{\beta})$	$\exp(\hat{\beta})$ için %95 güven aralığı
Yaş	-0.0147	0.0033	0.000	0.9854	-0.0212, -0.0083
İlaç	0.0425	0.0568	0.454	1.0434	-0.0687, 0.1538
Cinsiyet	0.1677	0.0880	0.057	1.1826	-0.0047, 0.3402
Asit	-0.0947	0.0825	0.251	0.9096	-0.2564, 0.0669
Hepatomegali	-0.0484	0.0643	0.452	0.9528	-0.1744, 0.0777
Deride Bozukluk	-0.0403	0.0669	0.547	0.9605	-0.1713, 0.0907
Ödem1	0.0610	0.0772	0.429	1.0629	-0.0903, 0.2123
Ödem2	0.2792	0.0930	0.003	1.3221	0.0970, 0.4614
Logb	-0.1769	0.0316	0.000	0.8379	-0.2388, -0.1149
Logalbumin	-0.1120	0.1680	0.505	0.8940	-0.4412, 0.2172
Logprottime	0.3120	0.2808	0.267	1.3662	-0.2384, 0.8624
Histolojik Evre2	-0.0639	0.1944	0.742	0.9381	-0.4450, 0.3171
Histolojik Evre3	0.0520	0.1820	0.775	1.0534	-0.3048, 0.4087
Histolojik Evre4	0.0309	0.1813	0.865	1.0314	-0.3245, 0.3862

1 birimlik artışın ortanca sağkalım süresini 0.9854 birim ve logb değişkenindeki 1 birimlik artışın ise ortanca sağkalım süresini 0.8379 birim azalttığı söylenebilmektedir.

4. Tartışma ve Sonuç

Sağkalım analizinde sansürleme türlerine göre kullanılan sağkalım modellerinin teorik yapısı değişiklik göstermektedir. Sağdan sansürlü veri yapısı ile sıkça karşılaşılmakla birlikte, literatürde soldan ve aralıklı sansürlü veri yapıları ile de karşılaşmak mümkündür. Aralıklı sansürlü veri yapısı, bir aralıktaki birimlerin başarısız oldukları kesin zaman noktaları ile ilgili belirsizliği yansıtmaktadır. Aralıklı sansürlü veri yapısı çoğunlukla demografik, epidemiyolojik, finansal, tıbbi ve sosyolojik çalışmalarda yer almaktadır. Literatürde aralıklı sansürlü veri yapısı meme kanseri, kalıcı dişin ortaya çıkış süresi, AIDS klinik deneyler, yoğurt için raf ömrü, cep telefonu satın

alma, vb. veri kümeleri ile çalışılmıştır (Bogaerts ve diğ., 2018). Aralıklı sansürleme, sağdan sansürlemeden farklı bir model altyapısı gerektirir. Bu çalışmada, aralıklı sansürlü veri yapısı için CRM ve parametrik sağkalım modelleri incelenmiştir

Uygulamada literatürde yer alan Primer Biliyer Siroz verisi için ilk kez aralıklı sansürlü veri yapısı dikkate alınarak modelleme yapılmıştır. PBS veri kümesi sağdan sansürlü veri yapısı ve aralıklı sansürlü veri yapısı ayrı ayrı dikkate alınarak incelenmiştir. Sağdan sansürlü veri yapısı için OT varsayımının sağlandığı, dolayısıyla CRM'nin veri kümesine uygulanabilir olduğu sonucuna varılmıştır. PBS veri kümesi için, CRM'de sağdan sansürlü veri yapısına göre yaş, asit, logb değişkenleri anlamlı çıkmıştır. Sağdan sansürlü veri yapısı için OT modellerinden Gompertz ve HBS modellerinden ise Weibull en iyi sonucu vermiştir. Hem Gompertz OT modeli hem de Weibull HBS modeli sonuçlarına göre yaş, asit, logb değişkenleri anlamlı bulunmuştur. PBS veri kümesi için, aralıklı

sansürlü veri yapısına göre CRM için yaş ve logb değişkenleri anlamlı çıkmıştır. Aralıklı sansürlü veriler için OT modellerinden Gompertz ve HBS modellerinden ise Gamma en iyi sonucu vermiştir. Aralıklı sansürlü veri yapısı için parametrik sağkalım modellerinden Gompertz OT modeli için yaş, asit, ödem2 ve log(b) değişkenleri; Gamma HBS modeli için yaş, ödem2 ve log(b) değişkenleri anlamlı bulunmuştur. Ayrıca sağdan sansürlü veri yapısı ile aralıklı sansürlü veri yapısına bağlı olarak tehlike oranları da farklılık göstermektedir. Bu sonuçlara göre incelenecek veri kümesi için sansürleme türüne doğru bir biçimde karar verilmesi ve modellemelerin buna göre yapılıp, yorumlanmasının önemli olduğu ortaya çıkmıştır. Aralıklı sansürlü veri yapısı mevcut ise, daha fazla bilgi içermesi açısından buna ait modellerin ve sonuçların kullanılması daha uygun olacaktır.

Kaynakça

- [1] Barlow, R. E., Proschan, F. 1975. Importance of System Components and Fault Tree Events. *Stochastic Processes and their Applications*, 3(2), 153-173.
- [2] Cox, D. R. 1972. Regression Models and Life-Tables. *Journal of the Royal Statistical Society, Series B*, 34(2), 187-220.
- [3] Finkelstein, D. M., Wolfe, R. A. 1985. A Semiparametric Model for Regression Analysis of Interval-Censored Failure Time Data. *Biometrics*, 41(4), 993-945.
- [4] Finkelstein, D. M. 1986. A Proportional Hazards Model for Interval-Censored Failure Time Data. *Biometrics*, 42(4), 845-854.
- [5] Huang, J., Wellner, J. A. 1995. Asymptotic Normality of The NPMLE of Linear Functionals for Interval Censored Data, Case 1. *Statistica Neerlandica*, 49(2), 153-163.
- [6] Huang, J. 1996. Efficient Estimation for the Proportional Hazards Model with Interval Censoring. *The Annals of Statistics*, 24(2), 540-568.
- [7] Satten, G. A. 1996. Rank-based Inference in the Proportional Hazards Model for Interval Censored Data. *Biometrika*, 83(2), 355-370.
- [8] Younes, N., Lachin, J. 1997. Linked-based Models for Survival Data with Interval and Continuous Time Censoring. *Biometrics*, 53(4), 1199-1211.
- [9] Huang, J., Rossini, A. J. 1997. Sieve Estimation for The Proportional Odds Failure-Time Regression Model With Interval Censoring. *Journal of the American Statistical Association*, 92:439, 960-967.
- [10] Goggins, W. B., Finkelstein, D. M., Schoenfeld, D. A., Zaslavsky, A. M. 1998. A Markov Chain Monte Carlo EM Algorithm for Analyzing Interval Censored Data Under The Cox Proportional Hazards Model. *Biometrics*, 54(4), 1498-507.
- [11] Geskus, R., Groeneboom, P. 1999. Asymptotically Optimal Estimation of Smooth Functionals for Interval Censoring, Case 2. *The Annals of Statistics*, 27(2), 627-674.
- [12] Pan, W. 2000. A Multiple Imputation Approach to Cox Regression with Interval-Censored Data. *Biometrics*, 56(1), 199-203.
- [13] Xue, H., Lam, K. F., Cowling, B. J., de Wolf, F. 2006. Semi-parametric Accelerated Failure Time Regression Analysis with Application to Interval-Censored HIV/AIDS data. *Statistics in Medicine*, 25(22), 3850-3863.
- [14] Zeng, D., Cai, J., Shen, Y. 2006. Semiparametric Additive Risks Model for Interval-Censored Data. *Statistica Sinica*, 16, 287-302.
- [15] Banerjee, M., Sen, B. 2007. A Pseudolikelihood Method for Analyzing Interval Censored data. *Biometrika*, 94(1), 71-86.
- [16] Heller, G. 2011. Proportional Hazards Regression with Interval Censored Data Using An Inverse Probability Weight. *Lifetime Data Analysis*, 17(3), 373-85.
- [17] Rabinowitz, D., Tsiatis, A. A., Aragon, J. 1995. Regression with Interval-Censored Data. *Biometrika*, 82(3), 501-513.
- [18] Betensky, R. A., Rabinowitz, D., Tsiatis, A. A. 2001. Computationally Simple Accelerated Failure Time Regression for Interval Censored Data. *Biometrika*, 88(3), 703-711.
- [19] Oller, R., Gómez, G., Calle, M. L. 2004. Interval Censoring: Model Characterizations for The Validity of The Simplified Likelihood. *The Canadian Journal of Statistics*, 32(3), 315-326.
- [20] Zhang, Z., Sun, L., Zhao, X., Sun, J. 2005. Regression Analysis of Interval Censored Failure Time Data with Linear Transformation Models. *The Canadian Journal of Statistics*, 33(1), 61-70.
- [21] Lawless, J. F., Babineau, D. 2006. Models for Interval Censoring and Simulation-Based Inference for Lifetime Distributions. *Biometrika*, 93(3), 671-686.
- [22] Tian, L., Cai, T. 2006. On The Accelerated Failure Time Model for Current Status and Interval Censored Data. *Biometrika*, 93(2), 329-342.
- [23] Zhang, Z., Sun, L., Sun, J., Finkelstein, D. M. 2007. Regression Analysis of Failure Time Data with Informative Interval Censoring. *Statistics in Medicine*, 26(12), 2533-46.
- [24] Zhu, L., Tong, X., Sun, J. 2008. A Transformation Approach for The Analysis of Interval-Censored Failure Time Data. *Lifetime Data Analysis*, 14(2), 167-78.

- [25] Radhey, S. S., Totawattage, D.P. 2013. The Statistical Analysis of Interval-Censored Failure Time Data with Applications. *Journal of Statistics*, 3, 155-166.
- [26] Gomez, G., Calle, M. L., Oller, R., Langohr, K. 2009. Tutorial on Methods for Interval-Censored Data and Their Implementation in R. *Statistical Modelling*, 9(4), 259-297.
- [27] Fay, M. P. 2014. interval: Weighted logrank tests and NPMLE for interval-censored data [Computer software manual]. Retrieved from <https://CRAN.R-project.org/package=interval> (R package version 1.1-0.1), 2014.
- [28] Anderson-Bergman, C. 2017. icenReg: Regression Models for Interval Censored Data in R. *Journal of Statistical Software*, 81(12), 1-23.
- [29] Delord, M. 2017. MIICD: Data Augmentation and Multiple Imputation for Interval Censored Data [Computer software manual]. Retrieved from <http://CRAN.R-project.org/package=MIICD> (R package version 2.4).
- [30] Bogaerts, K., Komàrek, A., Lesaffre, E. 2018. Survival analysis with interval-censored data. A practical approach with examples in R, SAS, and BUGS. 1st edition. CRC Press by Taylor and Francis Group, 584s.
- [31] So, Y., Johnston, G., Kim, S. 2010. H. Analyzing Interval Censored Survival Data. *SAS Global Forum Paper* 257.
- [32] Guo, C., So, Y., Johnston, G. 2014. Analyzing Interval-Censored Data with The ICLIFETEST Procedure. *Paper SAS279-2014*.
- [33] Yang, X. 2017. Analyzing interval-censored survival-time data in Stata. https://www.stata.com/meeting/baltimore17/slides/Baltimore17_Yang.pdf. (Erişim Tarihi: 10.03.2018)
- [34] Murtaugh, P., Dickson, E. R., Dam, G. V., Malinchoc, M., Grambsch, P. M., Langworthayn A. and Gips, C. H. 1994. Primary Biliary Cirrhosis: Prediction of Short-term Survival Based on Repeated Patient Visits. *Hepatology*, 20, 126-134.
- [35] Turnbull, B. W. 1976. The Empirical Distribution with Arbitrarily Grouped Censored and Truncated Data. *Journal of the Royal Statistical Society: Series B*, 38(3), 290-295.
- [36] Hoel, D. G., Walburg, H. E. 1972. Statistical Analysis of Survival Experiments. *Journal of National Cancer Institute*, 49, 361-372.
- [37] Carvalho, J. C., Ekstrand, K. R., Thylstrup, A. 1989. Dental Plaque and Caries on Occlusal Surfaces of First Permanent Molars in Relation to Stage of Eruption. *Journal of Dental Research*, 68(5), 773-779.
- [38] Kim, M. Y., De Gruttola, V. G., Lagakos, S. W. 1993. Analyzing Doubly Censored Data with Covariates, with Application to AIDS. *Biometrics*, 49(1), 13-22.
- [39] Betensky, R. A., Finkelstein, D. M. 1999. A Non-Parametric Maximum Likelihood Estimator for Bivariate Interval Censored Data. *Statistics in Medicine*, 18(22), 3089-3100.
- [40] Goggins, W. B., Finkelstein, D. M. 2000. A Proportional Hazards Model for Multivariate Interval-Censored Failure Time Data. *Biometrics*, 56, 940-943.
- [41] Meyns, B., Jashari, R., Gewillig, M., Mertens, L., Komàrek, A., Lesaffre, E., Budts, W., Daenen, W. 2005. Factors Influencing The Survival of Cryopreserved Homografts. The Second Homograft Performs as Well as the First. *European Journal of Cardio-thoracic Surgery*, 28(2), 211-216.
- [42] Goethals, K., Ampe, B., Berkvens, D., Laevens, H., Janssen, P., Duchateau, L. 2009. Modeling Interval-Censored, Clustered Cow Udder Quarter Infection Times through the Shared Gamma Frailty Model. *Journal of Agricultural, Biological, and Environmental Statistics*, 14(1), 1-14.
- [43] Hough, G. 2010. Sensory shelf life estimation of food products. 1st edition. Boca Raton: CRC Press, 264s.
- [44] Karvanen, J., Rantanen, A., Luoma, L. 2014. Survey Data and Bayesian Analysis: A Cost-Efficient Way to Estimate Customer Equity. *Quantitative Marketing and Economics*, 12(3), 305-329.
- [45] Groeneboom, P., Wellner, J. A. 1992. Information bounds and nonparametric maximum likelihood estimation. Basel: Birkhäuser-Verlag, 130s.
- [46] Huang, J., Wellner, J. A. 1997. Interval Censored Survival Data: A Review of Recent Progress. *Proceedings of the First Seattle Symposium in Biostatistics: Survival Analysis*, eds. Lin, D. and Fleming, T. Springer-Verlag, New York, 123-169.
- [47] Sun, J. 1998. Interval Censoring. *Encyclopedia of Biostatistics*, John Wiley, 1st Edition, 2090-2095.
- [48] Sun, J., Zhao, Q., Zhao, X. 2005. Generalized Log Rank Tests for Interval-Censored Failure Time Data. *Scandinavian Journal of Statistics*, 32(1), 49-57.
- [49] Yu, Q., Li, L., Wong, G.. 2000. On Consistency of Self-Consistent Estimator of Survival Functions with Interval-Censored Data. *Scandinavian Journal of Statistics*, 27(1), 35-44.
- [50] Schick, A., Yu, Q. 2000. Consistency of the GMLE with Mixed Case Interval-Censored Data. *Scandinavian Journal of Statistics*, 27(1), 45-55.

- [51] Wellner, J. A. 1995. Interval censoring case 2: alternative hypotheses. *Analysis of Censored Data* (Pune, 1994/1995), eds. H. L. Koul and J. V. Deshoande, IMS Lecture Notes, Monograph Series 27, 271-219.
- [52] Dinse, G. E., Lagakos, S. W. 1983. Regression Analysis of Tumor Prevalence Data. *Applied Statistics*, 32, 236-248.
- [53] Keiding, N. 1991. Age-specific Incidence and Prevalence: A Statistical Perspective (with discussion). *Journal of the Royal Statistical Society, Series A*, 154, 371-412.
- [54] Keiding, N., Begtrup, K., Scheike, T. H., Hasibeder, G. 1996. Estimation from Current Status Data in Continuous Time. *Lifetime Data Analysis*, 2, 119-129.
- [55] Shiboski, S. C., Jewell, N. P. 1992. Statistical Analysis of the Time Dependence of HIV Infectivity Based on Partner Study Data. *Journal of the American Statistical Association*, 87, 360-372.
- [56] Diamond, I. D., McDonald, J. W., Shah, I. H. 1986. Proportional Hazards Models for Current Status Data: Application to the Study of Differentials in Age at Weaning in Pakistan. *Demography*, 23, 607-620.
- [57] Diamond, I. D., McDonald, J. W. 1991. The analysis of current status data. *Demographic Applications of Event History Analysis*, eds. Trussel, J., Hankinson, R. and Tilton, J. Oxford University Press: Oxford, U.K.
- [58] Sun, J. 2006. The statistical analysis of interval-censored failure time data. *Statistics for Biology and Health*, Springer, New York, 304s.
- [59] Dil, E. 2019. Boylamsal ve yaşam verilerinin parametrik bileşik modellemesi. Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, 64s.