

Sağdan Sansürlü Gözlemlerin Yerleşiminin Hazard Oranı Tahminine Etkisi

Effect of Right Censored Observations Locations on Hazard Rate Estimates

Yrd.Doç.Dr. Seval KUL,^a
Prof.Dr. Arzu KANIK^b

^aBiyostatistik AD,
Gaziantep Üniversitesi Tıp Fakültesi,
Gaziantep

^bBiyostatistik AD,
Mersin Üniversitesi Tıp Fakültesi,
Mersin

Geliş Tarihi/Received: 31.01.2011
Kabul Tarihi/Accepted: 05.07.2011

Yazışma Adresi/Correspondence:
Yrd.Doç.Dr. Seval KUL
Gaziantep Üniversitesi Tıp Fakültesi,
Biyostatistik AD, Gaziantep,
TÜRKİYE/TURKEY
sevalkul@gantep.edu.tr

ÖZET Amaç: Hayatta kalma analizlerinde sansürlü gözlemlerin rasgele dağıldığı ve veri seti içinde her hangi bir zamanda rasgele olarak oluşabileceği varsayılır. Fakat bazı klinik çalışmalarda sansürlü gözlemlerin büyük bir kısmı belirli bir zaman aralığı içinde oluşabilir. Bunun birçok nedeni olabilir (hastaların tedaviden memnun olmaması, tedavinin yeterli gelmemesi, basında bir ilaç ya da tedavi hakkında çıkan olumsuz haberler, hastane enfeksiyonu varlığı... vb). Bu makalede sansürlü gözlemlerin veri setinde rasgele olarak oluşmadığı, belirli bir aralıkta olduğu durumda hayatta kalma fonksiyonu ve hazard oranı tahminlerine etkisinin araştırılması amaçlanmıştır. **Gereç ve Yöntemler:** Çalışmanın simülasyon kısmında, sabit hazard, doğrusal artan ve doğrusal azalan hazard olmak üzere üç farklı dağılımdan örnek genişliği 100 ve 250 için hayatta kalma fonksiyonu ve hazard değerleri hesaplanarak dağılımlar oluşturulmuştur. Veri seti içinde 4 farklı oranda (0.20, 0.30, 0.40, 0.50) sansürlü gözlem oluşturulmuş ve sansürlü gözlemler veri seti içinde dört farklı bölgeye yerleştirilmiştir. Hazard fonksiyonlarının tahmininde Kernel düzgünleştirme yöntemi kullanılmıştır. Simülasyon programları R programlama dilinde yazılmıştır ve her bir simülasyon 1000 defa tekrarlanmıştır. **Bulgular:** Sansürlü gözlemlerin veri seti içindeki yerleşiminin bazı tahminlerde tamamen rasgele oluşturulmuş sansür yerleşimine göre istatistiksel olarak anlamlı farklılıklar yarattığı gözlenmiştir. Bunun yanı sıra sansürlü gözlemlerin oranının artırılmasının tahminleri kötüleştirmediği bir kez daha ortaya konulmuştur. Ayrıca küçük örneklerde ve tamamlanmış gözlem sayısının az olduğu bölgelerde tahminlerin kötüleştiği gözlenmiştir. **Sonuç:** Sansürlü gözlemlerin veri seti içinde değiştirilmesinden oluşan farklılıkların klinik olarak anlamlı farklılıklar oluşturmayacağına karar verilmiştir.

Anahtar Kelimeler: Hayatta kalma analizleri; hazard tahmini; Kernel düzgünleştirme yöntemi; sansürlü gözlemler

ABSTRACT Objective: Purpose: In survival analysis censored observations are assumed randomly disturbed and censoring can occur any time in data set randomly. However in some medical studies, most of the right censoring can occur within very short time interval. This can be result of many reasons (e.g., dissatisfaction of the patient from the treatment, inefficient therapy, negative news about remedy or interventions and high risk of hospital infection). In this study, we aimed to investigate the effect of nonrandom occurrence of censored data on the survival and hazard rate estimates. **Material and Methods:** In the simulation part, we simulated data from 3 different distributions; survival functions derived from constant, linear increasing and linear decreasing hazard functions including censored observation with two sample sizes ($n=100$, $n=250$). Four different proportions (0.20, 0.30, 0.40, 0.50) of censoring was constructed and censoring was located at four different parts of data. Kernel smoothing estimates were used. Simulation programs were implemented in R language and each simulation was repeated 1000 times. **Results:** We found location of the censored observations in data set has significant effect on some estimates. Additionally we showed that, increasing the proportions of the censored observation in life time data makes survival and hazard function estimates worse. Moreover, estimates were worse at the time when there are few observations left under risk. **Conclusion:** As a result of our simulations we were concluded that although there was a statistically significant difference at some points, differences were too small to make clinically significant effect.

Key Words: Survival analysis; hazard estimation; Kernel smoothing; censored observations

Hazard fonksiyonu yaşam sürelerinin karakterize edilmesinde kullanılan en önemli fonksiyonlardan biridir.^{1,2} Hazard fonksiyonun tahmin edilmesinde ise en yaygın olarak kullanılan istatistiksel yöntem Kernel tahmin yöntemidir.³⁻⁵ Kernel yönteminin uygulanabilmesi için yaşam sürelerinin dağılımına ilişkin herhangi bir dağılım ön şartı bulunmamaktadır fakat hayatta kalma analizlerinde genellikle sansürlü gözlemlerin veri seti içinde herhangi bir zamanda rasgele olarak oluşacağı varsayılır. Pratikte bazı özel nedenlerle bu varsayım ihlal edilebilir. Örneğin ağrılı bir operasyondan sonra kullanılan iki ağrı kesicinin etkinliğinin karşılaştırılmak istendiği bir klinik denemede, ağrı kesicilerden birinin kullanımından kısa bir süre sonra hastaların memnuniyetsizliği nedeni ile ikinci bir ağrı kesici verilmesi veya ağrı kesicinin dozunun artırılması durumunda aynı zaman aralığında birçok hastanın yeni ilaç kullanımı nedeni ile sansürlü gözlem olarak kabul edilmek zorunda kalması mümkündür. Bu tür durumlarla ilaç denemelerinde de sıklıkla karşılaşılabileceği gibi özellikle placebo kullanıldığı durumlarda araştırmacının karşılaşılabileceği bir durumdur.^{6,7} Rasgele olmayan sansürlü gözlemlere verilebilecek bir başka örnek ise menopoz döneminde östrojen hormonu eksikliğine bağlı olarak ortaya çıkabilecek çeşitli şikayetleri gidermek amacıyla kullanılan hormon replasman tedavisinin meme kanseri riskini arttırdığına dair basında çıkan haberler dolayısıyla birçok kadın hekime danışmadan tedaviyi aynı dönemde bırakma eğilimine girecek ve dolayısıyla izlemde olan kadınların büyük bir kısmı aynı dönemde çalışmadan çıkacaktır.

Sansürlü gözlemlerin rasgele olmayan bir dağılıma sahip olarak karşımıza çıkması mümkün olduğu halde literatürde sansürlü gözlemlerin veri seti içindeki yerleşiminin Hazard oranı tahminleri üzerine etkisinin araştırıldığı bir çalışma henüz mevcut değildir. Bu çalışmada sansürlü gözlemlerin veri setinde farklı bölgelerde bulunduğu durumda Hazard oranının tahminine etkisinin araştırılması amaçlanmıştır. Bu amaçla farklı dağılım tipleri ve sansürlü gözlem miktarları için sansürlü gözlemlerin dağılım içinde yerleşim yerleri değiştirilmiş ve tahminler karşılaştırılmıştır.

GEREÇ VE YÖNTEMLER

HAZARD FONKSİYONU

T hayatta kalma zamanı iken, $h(t)$ hazard fonksiyonu koşullu başarısızlık oranını verir. Bu fonksiyondan elde edilen değer aralığın başında hayatta olan bireylerde kısa bir zaman aralığı içinde ilgilenilen olayın gözlenmesi olasılığıdır.^{1,2}

$$h(t) = \frac{\lim_{\Delta t \rightarrow 0} P\{t \text{ zamanına kadar hayatta kalmış bir bireyde } (t, t + \Delta t) \text{ aralığında olayın gözlenmesi}\}}{\Delta t}$$

Hazard fonksiyonu, artan, azalan ve sabit gibi birçok farklı şekilde gözlenebilir.

KERNEL DÜZGÜNLEŞTİRME YÖNTEMİ

Kernel tahmin yöntemi parametrik olmayan fonksiyonların tahmininde yaygın olarak kullanılan bir yöntemdir. F gibi bir fonksiyonun Kernel tahmin edicisi x gibi bir noktada $[x-b, x+b]$ gibi bir aralıkta verilerin bölgesel ağırlıklandırılmış ortalamasıdır. Buradaki b, Kernel tahmininde bant genişliği (bandwidth) olarak adlandırılır. Sansürlü gözlemler için hazard tahmininde Kernel tahmin yönteminin kullanılması ilk defa Ramlau-Hansen⁴ tarafından gündeme getirilmiş ve diğer araştırmacılar tarafından sabit bir bant genişliği içim sağdan sansürlü gözlemlerin hazard fonksiyonunun tahmininde kullanılmak üzere aşağıdaki tahmin edici tanımlanmıştır.⁵ $K(\cdot)$ Kernel ya da ağırlıklandırma fonksiyonu, b bant genişliği olmak üzere hazard fonksiyonunun tahmininde kullanılan eşitlik aşağıdaki gibidir.

$$\hat{h}(t) = \frac{1}{b} \sum_{i=1}^n K\left(\frac{t-t_{(i)}}{b}\right) \frac{\delta_i}{n-i+1} \quad 0 < t_{(1)} < t_{(2)} < \dots < t_{(n)}$$

δ_i sansür durumunu göstermek üzere, $t_{(n)}$ n. sıralı ölüm zamanlarıdır.

BANT GENİŞLİĞİNİN BELİRLENMESİ

Kernel tahminin başarısındaki en önemli faktör bant genişliğinin belirlenmesidir.⁴ Literatürde sansürlü gözlemlerin varlığı durumunda kullanılan Kernel düzgülendirme yöntemi için birçok bant genişliği belirleme yönteminden bahsedilmiştir.⁸⁻¹²

SİMÜLASYON ÇALIŞMASI

Bu çalışmada farklı dağılımlar ve sansürlü gözlem miktarları için sansürlü gözlemlerin dağılım içinde yerleşim yerleri değiştirilerek hazard oranı tahminlerine sansürlü gözlemlerin veri içindeki yerleşiminin etkisi araştırılmıştır. Sonuçların karşılaştırılabilmesi amacıyla Hess ve ark. tarafından yayınlanmış olan makalede tanımlanmış olan dağılımlardan faydalanılmıştır.¹³ Sabit, lineer artan ve lineer azalan polynomial hazard şekilleri kullanılmıştır. $h(t) = \lambda_0 h_0(t)$ hazard fonksiyonunda $n=100$ için $S(t=90)$ 0.1 ve $n=250$ için $S(t=90)$ 0.04 değerlerinden faydalanarak λ_0 değerleri hesaplanmıştır. Bu değerler $t=90$ zamanında 10 hastanın risk altında olması demektir. Sabit hazard için, $h_0(t)=1$, lineer artış gösteren hazard $h_0(t) = at + c$ için hazard fonksiyonunun a, c değerleri $h_0(0)=0$ ve $h_0(100)=1$ olacak şekilde belirlenmiştir. Lineer azalış gösteren hazard için ise $h_0(t)=at-c$ hazard fonksiyonunun a, c değerleri $h_0(0)=1$ ve $h_0(100)=0$ olacak şekilde belirlenmiştir (Tablo 1, Tablo 2).

Simülasyonun bir sonraki aşamasında standart uniform dağılıma sahip rasgele u değerleri üretilmiştir. Ve bu u değerleri hazard fonksiyonunun in-

tegrali yardımıyla elde edilmiş, hayatta kalma fonksiyonun tersinde $t_i = S^{-1}(u_i)$ yerine konulmuştur. Sansürlü gözlemlerin yerinin ve oranının değiştirilebilmesi için uniform dağılımdan faydalanılmıştır.

Tahminlerin performansının değerlendirilmesinde gerçek fonksiyonla tahminleri arasındaki farkların ortalama uzaklığı olan L_2 hatası kullanılmıştır.

$$L_2 = \frac{\sum (Y - f(x))^2}{n}$$

Hazard fonksiyonu tahminlerinde Kernel düzleştirme yöntemi ve Kernel fonksiyonu olarak ise Epanechnikov Kernel kullanılmıştır. Bant genişliği olarak L_2 hatasını minimum yapan bant genişliği kullanılmıştır. Sonuçların literatürde var olan en detaylı simülasyon çalışması¹³ ile karşılaştırılabilmesi için bulunan hata değerleri 10^8 le çarpılarak logaritması alınmıştır.

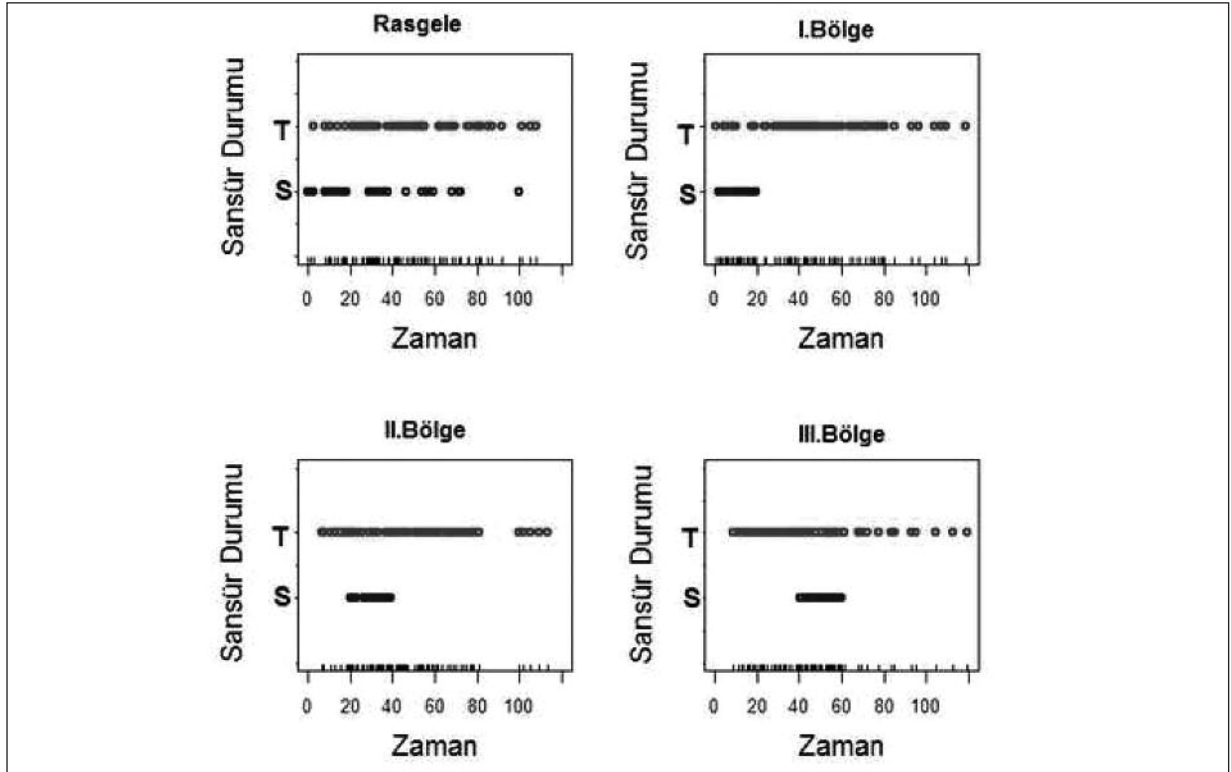
Simülasyonlar R programlama dilinde yazılmıştır. Kernel tahminlerinin başarılarını göstermek amacıyla 1000 tekrarlı olarak programlar çalıştırılmış ve her tahmin noktası için tahminlerin ortalama

TABLO 1: Simülasyonda kullanılan dağılımlara ilişkin hayatta kalma fonksiyonları ve fonksiyonların tersleri.

	$S(t)$	$S^{-1}(t)$
Sabit Hazard	$S(t) = \exp(-\lambda_0 t)$	$S^{-1}(t) = -\frac{\ln(t)}{\lambda_0}$
Lineer Artan	$S(t) = \exp\left(-\frac{\lambda_0 t^2}{200}\right)$	$S^{-1}(t) = \sqrt{-\frac{200 \ln(t)}{\lambda_0}}$
Lineer Azalan	$S(t) = \exp\left(-\lambda_0 \left(t - \frac{t^2}{200}\right)\right)$	$S^{-1}(t) = 100 - \sqrt{\frac{200 \ln(t) + 10000}{\lambda_0}}$

TABLO 2: Simülasyonda kullanılan dağılımlara ilişkin hazard fonksiyonu ve katsayıları.

Hazard Oranı	$h(t) = h_0 \lambda_0(t)$	$n = 100$	$n = 250$
Sabit Hazard	$h(t) = \lambda_0$	$\lambda_0 = 0.0256$	$\lambda_0 = 0.0357$
Lineer Artan	$h(t) = \frac{\lambda_0 t}{100}$	$\lambda_0 = 0.0568$	$\lambda_0 = 0.0794$
Lineer Azalan	$h(t) = \begin{cases} \lambda_0 \left(1 - \frac{t}{100}\right) \\ \lambda_0 \end{cases}$	$\lambda_0 = 0.04651$	$\lambda_0 = 0.0650$



ŞEKİL 1: Doğrusal artan hazard oranı, $n=100$, sansürlü gözlem oranı 0.30 iken sansürlü gözlemlerin yerleşimi.

ması alınmıştır. Tekrar sayılarını ve örnek genişliği belirtmek için sırası ile K ve N harfleri kullanılmıştır.

Makalenin simülasyon çalışmasında 20, 30, 40 ve 50 sansür yüzdeleri için, aynı dağılıma sahip veri setinin içinde sansürlü gözlemler 4 farklı bölgeye yerleştirilmiştir. Doğrusal artan hazard oranına ait fonksiyon tahminleri için oluşturulmuş veriler Şekil 1'de verilmiştir. Şekil 1'de sansür yerleşim grafiğinde tamamlanmış veriler T harfi ile sansürlü gözlemler ise S harfi ile gösterilmiştir.

Simülasyon sonucunda hesaplanan tahmin hataları tek yönlü varyans analizi ile karşılaştırılmış ve alt grup karşılaştırmaları için LSD çoklu karşılaştırma yöntemi kullanılmıştır.

BULGULAR

Hazard oranı tahminlerine ait tüm sonuçlar örnek genişliğinin 100 ve 250 olduğu durumlarda Şekil 2 ve 3'de verilmiştir.

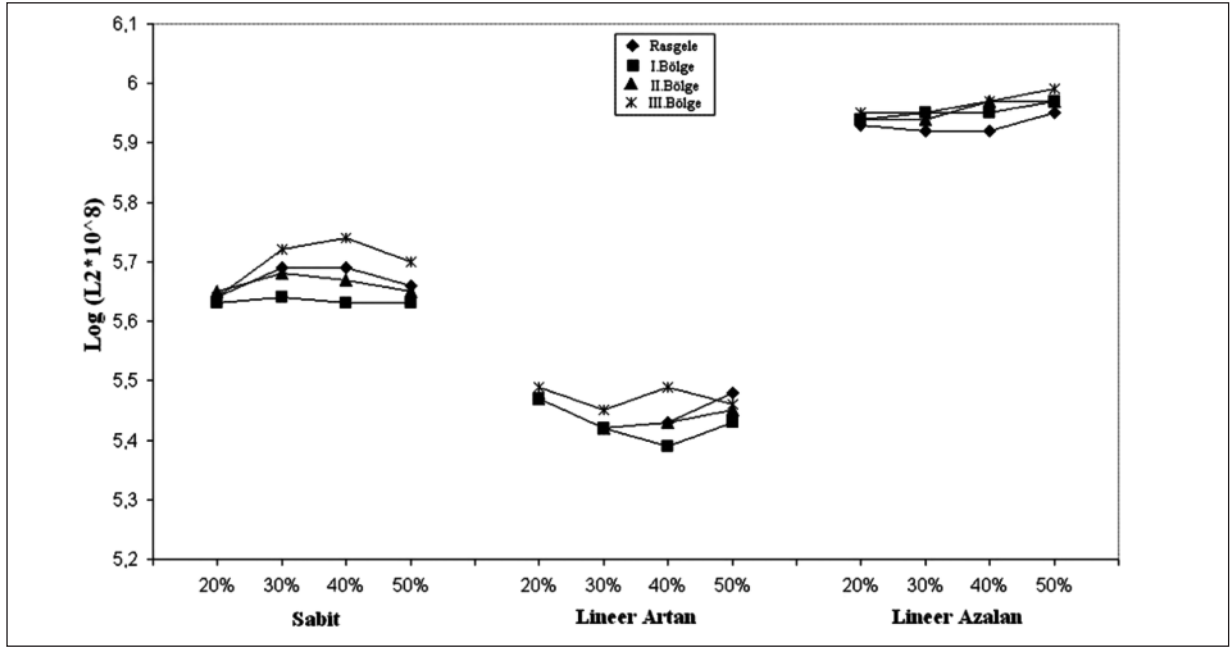
Hayatta kalma analizlerinde sansürlü gözlemlerin çalışmanın başında ortaya çıkması araştırma-

cılar tarafından büyük bir sorun olarak görülmekte ve tahminler için bilgi kaybının daha fazla olduğunu düşündürmektedir. Fakat çalışmanın başında izlemde olan birey sayısı fazla olduğundan sansürlü gözlemlerin hepsinin dahi çalışmanın başında oluşması tahminleri sansürlerin sonda oluşması kadar etkilememiştir.

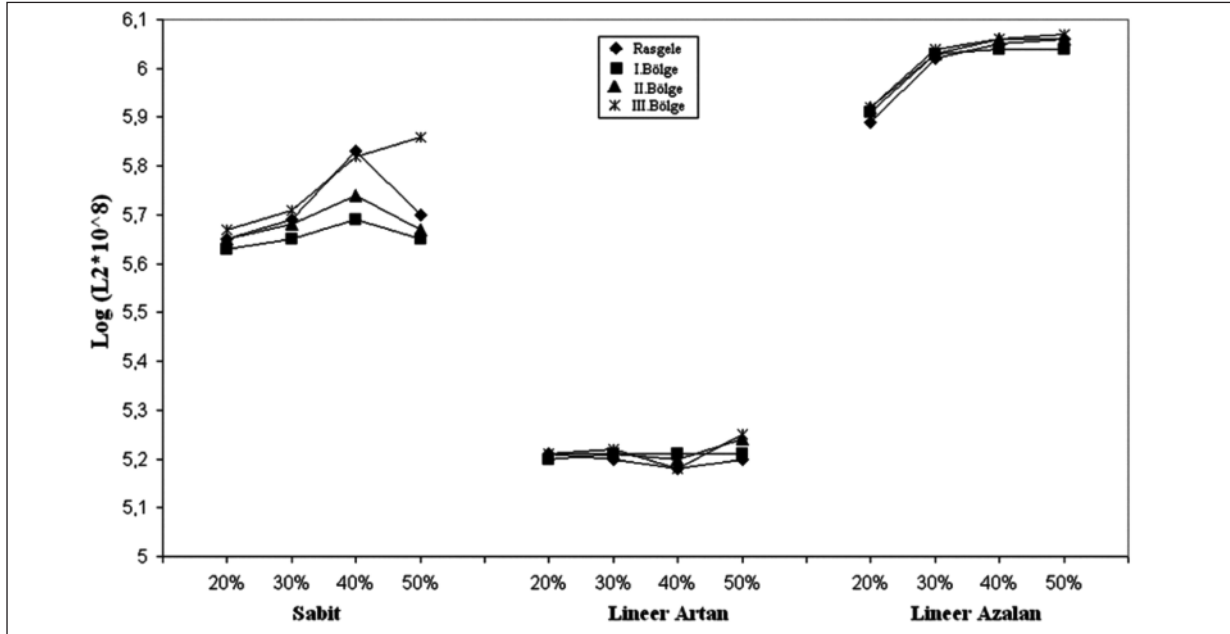
Şekil 2 ve 3'de verilen hata miktarları karşılaştırıldığında örnek genişliğinin 250 olduğu durumda hata oranlarında sabit ve lineer azalan fonksiyonlar için önemli bir değişim gözlenmezken lineer artan fonksiyon için azalış gözlenmiştir.

Simülasyonlarda kullanılan dağılımlar karşılaştırıldığında her iki örnek genişliği ve tüm sansür oranlarında en başarılı tahminler doğrusal artan hazard fonksiyonuna ait tahmininde ortaya çıkmıştır.

Tüm dağılımlarda sansürlü gözlem miktarındaki artışın tahminlerin başarısını genellikle olumsuz yönde etkilediği ortaya konmuştur. Bu durum özellikle örnek genişliği arttırıldığında daha net gözlenmektedir (Şekil 3).



ŞEKİL 2: N=100, K=1000 iken Hazard oranı tahminlerine ait tahmin hataları.



ŞEKİL 3: N=250, K=1000 iken Hazard oranı tahminlerine ait tahmin hataları.

Sansürlü gözlemlerin yerleşiminin Hazard oranı tahminleri üzerine bazı noktalarda istatistiksel olarak anlamlı düzeyde etki yaptığını göstermiştir. Hata miktarlarına bakıldığı zaman, rasgele olmayan sansürlü gözlemlerin yerleşimlerinin tahminleri üzerine etkisinin klinik olarak anlamlılık yaratacak düzeyde olmadığı görülmektedir.

TARTIŞMA

Bu çalışmada 3 farklı dağılım, 4 farklı sansür oranı, 4 farklı sansürlü gözlem yerleşim yeri ve iki farklı örnek genişliği kullanılarak oldukça detaylı bir simülasyon çalışması ortaya konmuştur. Bu çalışmanın literatürde şu ana kadar var olan çalışmalardan

en temel farklılığı sansürlü gözlemlerin miktarlarının yanı sıra veri seti içindeki yerleşimlerinin de hazard tahmini üzerine etkisinin araştırılmasıdır. Literatürde kernel tahmin yönteminin etkinliğini gösteren en detaylı simülasyon çalışması Hess ve ark.¹³ tarafından yapılmıştır. Bu çalışmada sansürlü gözlemlerin veri seti içinde tamamen rasgele olarak dağıldığı varsayılmıştır. Bizim çalışmamızda ise sansürlü gözlemler veri seti içinde 4 farklı şekilde yerleştirilmiştir.

Hazard oranlarının fonksiyonlarına ait tahminlerde Kernel düzgünleştirme yöntemi kullanılmıştır. Literatürde birçok Kernel fonksiyonu yer almaktadır. Etkinliklerine göre sıralandığında etkinliği en iyi Kernel fonksiyonunun Epanechnikov Kernel olması ve hesaplanma kolaylığı nedeni ile bu çalışmada Epanechnikov Kernel fonksiyonu kullanılmıştır.³

Kernel fonksiyonunun belirlenmesinin tahmin performansı üzerine büyük bir etkisi olmadığı asıl olarak bant genişliği belirleme yönteminin Kernel tahminlerinin performansını önemli düzeyde etkilediği bilinmektedir.^{9-12,14,15} Bu nedenle çalışmamızda R paket programının içinde yer alan birçok bant genişliği belirleme yöntemiyle elde edilen sonuçlar karşılaştırılmıştır. Kullanılan bant genişliği belirleme yöntemleri şunlardır: Yanlı çapraz geçerlilik (biased cross validation), yansız çapraz geçerlilik (unbiased cross validation). Ayrıca çalışma simülasyon çalışması olduğu ve verilerin türetildikleri fonksiyonlar bilindiği için gerçek fonksiyonla tahmin edilen fonksiyonlar arasındaki farkı minimum yapan bant genişliğini kullanacak şekilde bir algoritma yazılmıştır ve bant genişliklerinin performansları karşılaştırıldığında bu bant genişliği belirleme yönteminin kullanılmasına karar verilmiştir. Fakat simülasyon çalışmaları dışındaki gerçek veriler ile çalışılan durumda gerçek fonksiyonu bilmek mümkün olmadığı için bu yöntemin bant genişliği belirlenmesinde kullanılması mümkün değildir. Verinin gerçek dağılımı bilinmediğinde literatürden yararlanılarak en uygun bant genişliği belirleme yöntemine karar vermek gerekecektir.

Örnek genişliği arttırıldığında son bölgede oluşan tahmin hatalarının da azalmış olmasından

dolayı, III. Bölgede oluşan tahmin hatalarındaki kötüleşmenin nedeni olarak çalışmanın sonuna doğru izlemdeki birey sayısının azalması olarak değerlendirilmiştir. Çalışmanın sonuna doğru yüksek oranda sansürlü gözlem oluşmasının beklendiği bir klinik çalışmada örnek genişliğinin arttırılması bir çözüm olarak önerilebilir.

Simülasyon çalışmamızın sonuçları literatürdeki en detaylı simülasyon çalışması ile karşılaştırıldığında¹³ çalışmamızdaki Hazard oranı tahminlerinin performansının daha yüksek olduğu görülmüştür. Bunun nedeni her veri için ideal bir bant genişliği belirlememizdir. Hata oranları oldukça düşük olduğundan sağdan ya da soldan düzeltme yapmaya ihtiyaç duyulmamıştır.

Sansürlü gözlem miktarları ele alındığında, sansür miktarındaki artış hazard tahminlerini kötüleştirmektedir. Yapılan simülasyonların hepsinde 0.30 oranında sansürün varlığında tahminlerdeki hata oranlarının makul düzeyde olduğu fakat bu oran arttığında hataların belirgin bir şekilde yükseldiği görülmüştür.

Çalışmanın sonlarına doğru oluşan hataların yüksek olması hazard tahminlerinin izlemdeki gözlem sayısının azalmasından etkilendiğini göstermektedir. Ayrıca örnek genişliğini arttırmak hazard oranı tahminlerinin performansını etkilememiştir, bu bulguda Hess ve ark. bulgusu ile benzerdir. Her iki simülasyon çalışmasında da en kötü tahminler doğrusal azalan hazard fonksiyonun tahmininde ortaya çıkmıştır.

Sonuç olarak, rasgele olmayan sansürlü gözlemlerin varlığının Hazard oranı tahmininde kullanılan Kernel düzgünleştirme yöntemi üzerine yorumları klinik olarak etkileyecek düzeyde bir etkisinin olmadığı ortaya konmuştur. Bu sonuçta dikkat edilmesi gereken konu rasgele olmayan sansürlü gözlemlerin klinik ve hayvan deneylerinde birçok şekilde ortaya çıkabileceğidir. Yarışan bir riskin varlığı nedeni ile ya da bir prognostik faktörden dolayı ortaya çıkan sansürlerde rasgele olmayan sansürlerin özel bir halidir.^{15,16} Yarışan risk olduğu durumda oluşacak rasgele olmayan sansürün kendisinin de bir bilgi içermesi söz konusudur, çünkü eşlik eden hastalığın yaşam süresi üzerine

genel bir etkisi olacaktır.¹⁷ Bu durumda yarışan risklere özel istatistiksel yöntemlerin kullanılması önerilmektedir.¹⁸ Bu çalışmada oluşturulan rasgele olmayan sansürlü gözlemlerin yaşam süresine etkisi olmadığı varsayılmıştır.

Hazard oranı tahmini için elde edilen sonuçlar, rasgele olmayan sansürlü gözlemlerin Kaplan Meier tahminleri ve Logrank test sonuçlarına etkisi ile benzer çıkmıştır.¹⁹

SONUÇ

En genel sonuç olarak, sansürlü gözlemlerin miktarının artırılmasının hazard tahminlerinin başarısını genel olarak olumsuz yönde etkilediği görülmüştür. Sansür miktarları ile tahmin hataları arasındaki ilişki daha detaylı incelendiğinde, hayatta kalma analizi çalışmalarında 0,30 sansür oranının makul olacağı ve daha fazla sansürlü veri içeren verilerden yapılan tahminlerde tahmin başarılarının düşeceği sonucuna varılmıştır.

Çalışmanın sonlarına doğru oluşan hataların yüksek olması hazard tahminlerinin izlemdeki gözlem sayısının azalmasından olumsuz etkilendiği göstermiştir. Örneklem genişliğindeki değişimin tahminleri çok etkilemediği fakat tamamlanmış gözlem sayısının az olduğu bölgelerde tahminlerin kötüleştiği gözlenmiştir.

Sonuç olarak, sansürlü gözlemlerin veri seti içindeki yerleşiminin bazı tahminlerde tamamen rasgele oluşturulmuş sansür yerleşimine göre istatistiksel olarak anlamlı farklılıklar yarattığı gözlenmiştir. Fakat bu farklılıkların binde birlik düzeylerde olduğu dikkate alındığında sonuçların klinik olarak bir farklılık oluşturmayacağına karar verilmiştir.

KAYNAKLAR

- 1 Kleinbaum DG, Klein M. Introduction to survival analysis. In: Gail M, Krieger K, eds. *A Survival Analysis: A Self-Learning Text*. 2nd ed. New York: Springer; 2005. p.13-5.
- 2 Kaplan EI, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 1958;53(282):457-81.
- 3 Wang JL. Smoothing hazard rates. In: Armitage P, Colton T, eds. *Encyclopedia of Biostatistics*. 2nd ed. London: Wiley; 2005. p. 4986-7.
- 4 Ramlau-Hansen H. Smoothing counting process intensities by means of kernel functions. *Ann Statist* 1983;11(2):453-66.
- 5 Sheather SJ, Jones MC. A reliable data-based bandwidth selection method for kernel density estimation. *JR Statist Soc* 1991;53(3):683-90.
- 6 Sheiner LB, Beal SL, Dunne A, Lin X, Little RJ, Gelman A, et al. Analysis of Nonrandomly censored ordered categorical longitudinal data from analgesic trials. *J Am Stat Assoc* 1997; 92(440):1235-55.
- 7 Ganju J, Lakatos E, Rothe E. A simple way to estimate the median time and compare survival distributions in analgesic trials under informative censoring. *J Biopharm Stat* 1999; 9(4):683-93.
- 8 Tanner MA. Note on the variable kernel estimator of the hazard function from randomly censored data. *Ann Stat* 1983;11(3): 994-8.
- 9 Mugdadi AR, Ahmad IA. A bandwidth selection for kernel density estimation of functions of random variables. *Computational Statistics & Data Analysis* 2004;47(1):49-62.
- 10 Uzunogulları Ü, Wang LJ. A comparison of hazard rate estimators for left truncated and right censored data. *Biometrika* 1992;79(2): 297-310.
- 11 Müller HG, Wang JL. Hazard rate estimation under random censoring with varying kernels and bandwidths. *Biometrics* 1994;50(1):61-76.
- 12 Scott DW, Terrell GR. Biased and unbiased cross-validation in density estimation. *J Am Stat Assoc* 1987;82(400):1131-46.
- 13 Hess KR, Serachitopol DM, Brown BW. Hazard function estimators: a simulation study. *Stat Med* 1999;18(22):3075-88.
- 14 Bowman AW. An alternative method of cross-validation for the smoothing density estimates. *Biometrika* 1984;71(2):353-60.
- 15 Lagakos SW. General right censoring and its impact on the analysis of survival data. *Biometrics* 1979;35(1):139-56.
- 16 Wu Y, Furnary AP, Grunkemeier GL. Using the National Death Index to validate the non-informative censoring assumption of survival estimation. *Ann Thorac Surg* 2008;85(4): 1256-60.
- 17 Scharfstein D, Robins JM, Eddings W, Rotnitzky A. Inference in randomized studies with informative censoring and discrete time-to-event endpoints. *Biometrics* 2001;57(2):404-13.
- 18 Borgan Ø. Aalen-Johansen Estimator. In: Armitage P, Colton T. *Encyclopedia of Biostatistics*. 2nd ed. London: Wiley; 2005. p.1-5.
- 19 Kanik A, Kul S. Impact of non-random right censoring on Kaplan Meier estimates and logrank test results: a simulation study. *Turkiye Klinikleri J Biostat* 2010;2(2):82-8.