

Kesirli Polinomlar ile Cox Regresyon Modeli: Prostat Kanseri Veri Kümesi Üzerine Bir Uygulama

Cox Regression Model with Fractional Polynomials: An Application on Prostate Cancer Data Set

✉ Hazal DİNÇ^a, ✉ Nihal ATA TUTKUN^a

^aHacettepe Üniversitesi Fen Fakültesi, İstatistik Bölümü, Ankara, TÜRKİYE

ÖZET Amaç: Kesirli polinomlu regresyon modelleri, değişkenlerin doğrusal olmayan etkilerini modellemek için önerilmiştir. Değişkenlerin sürekli değişken olarak modele dahil edilmesi ya da kategorik biçime dönüştürülerek modele dahil edilmesi sonuçları değiştirebilir. Sürekli değişkenleri kategorik biçime dönüştürmek bilgi kaybına neden olmaktadır. Bu çalışmanın amacı kesirli polinomların Cox regresyon modelinde kullanımını incelemektir. **Gereç ve Yöntemler:** Prostat kanseri hastalığının 3. ve 4. aşamalarında olan ve 5 yıl boyunca izlenen 475 hastaya ait bir veri kümesine (Byar and Green, 1980) klasik ve kesirli polinomlu Cox regresyon modelleri uygulanmıştır. **Bulgular:** Tedavi türü ve evre değişkenleri kategorik ve yaş değişkeni sürekli kabul edilerek klasik Cox regresyon çözümü yapılmış, yaş değişkeninin ve evre değişkeninin istatistiksel olarak anlamlı olduğu ancak orantılı tehlikeler varsayımının sağlanmadığı görülmüştür. Yaş değişkeni kategorik olarak ele alındığında ise 81 yaş ve üzeri yaş grubunun ve evre değişkeninin anlamlı çıktığı görülmüştür. Bununla birlikte, orantılı tehlikeler varsayımı sağlanmadığından klasik Cox regresyon modelinin kullanılması da uygun değildir. Prostat kanseri veri kümesi için kesirli polinomlu Cox regresyon modeli incelendiğinde ise yaş değişkeninin yaşam süresi ile doğrusal olmayan bir ilişkiye sahip olduğu belirlenmiştir. Yaş ve evre değişkenlerinin istatistiksel olarak anlamlı olduğu ve kesirli polinomlu Cox regresyon modeli için orantılı tehlikeler varsayımının da sağlandığı görülmüştür. **Sonuç:** Kesirli polinomlu Cox regresyon modelinin prostat kanseri veri kümesi için daha uygun olduğu sonucuna ulaşılmıştır.

ABSTRACT Objective: The use of regression models with fractional polynomial was proposed to model nonlinear effects of covariates. The inclusion of covariates into the model as continuous or categorical can change the results. However, converting continuous covariates into categorical format causes a loss of information. In this case, models with fractional polynomial become an appropriate alternative. The aim of this study is examining the use of fractional polynomials in Cox regression model. **Material and Methods:** Classical Cox regression model and Cox regression model with fractional polynomial were applied to the data set of 475 patients (Byar and Green, 1980) who were in the 3rd and 4th stages of prostate cancer disease and followed for 5 years. **Results:** Classical Cox regression analysis was performed by including type of treatment and stage of disease as categorical and age as continuous and it was found that age and stage of disease were statistically significant, but the assumption of proportional hazards was violated. When age was taken as a categorical covariate, the 81 and more level of age and the stage of disease was significant. However, the use of classical Cox regression model is not correct since the proportional hazards assumption was violated. When the Cox regression model with fractional polynomials was run for prostate cancer data set, it was found that the age variable had a nonlinear relationship with the survival time. Age and stage of disease were statistically significant and the assumption of proportional hazards was provided. **Conclusion:** It was concluded that the Cox regression model with fractional polynomials is more suitable for the prostate cancer dataset.

Anahtar Kelimeler: Cox regresyon modeli; doğrusal olmama; kesirli polinomlar; orantılı tehlikeler varsayımı

Keywords: Cox regression model; nonlinearity; fractional polynomials; proportional hazards assumption

Açıklayıcı değişkenlerin yaşam süresi ile ilişkisini çözümlmek için literatürde birçok yaşam modeli mevcuttur. Yöntemler yarı-parametrik, parametrik olmayan ve parametrik yaklaşımları içermektedir. Yaşam fonk-

Correspondence: Nihal ATA TUTKUN

Hacettepe Üniversitesi Fen Fakültesi, İstatistik Bölümü, Ankara, TÜRKİYE/TURKEY

E-mail: nihalata@hacettepe.edu.tr

Peer review under responsibility of Türkiye Klinikleri Journal of Biostatistics.

Received: 16 Oct 2019 **Received in revised form:** 02 Jan 2020 **Accepted:** 22 Jan 2020 **Available online:** 27 Jan 2020

2146-8877 / Copyright © 2020 by Türkiye Klinikleri. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

siyonunun parametrik olmayan bir tahmin edicisi olarak Kaplan-Meier (KM) yöntemi, yaşam olasılıklarının zamanın bir fonksiyonu olarak tahmin etmek ve grafiklemek için kullanılır. Yarı parametrik bir yöntem olarak Cox regresyon modeli (CRM) ise yaşam çözümlemesinde en yaygın kullanılan model olma özelliğine sahiptir. Parametrik yaşam modelleri ise yaşam süresi dağılımının bilindiği durumlarda kullanılır ve üstel, Weibull, log-normal dağılım parametrik dağılımlara örnek olarak verilebilir. Bu modellerin herbirinde açıklayıcı değişkenlerin yaşam süresi ile doğrusal bir ilişkiye sahip olduğu varsayımı yer almaktadır. Modele dahil edilen açıklayıcı değişkenlerin etkilerinin doğrusal olmaması durumunda ise kesirli polinomların kullanılması önerilmektedir.²

Kesirli polinomlar, ilk kez Royston ve Altman (1994) tarafından regresyon modelinde sürekli açıklayıcı değişkenler ile sonuç değişkeni arasındaki ilişkinin modellenmesi için önerilmiştir.³ Sauerbrei ve diğ. (1999) tarafından meme kanseri verisinde semptom ve tanı verileri modellenmiştir.⁴ Berger ve diğ. (2000) mide kanseri verilerini kullanarak kesirli polinomlarda dinamik CRM'yi kullanarak model oluşturmuş ve klasik yöntemler ile sonuçları karşılaştırmışlardır.⁵ Berger ve diğ. (2003) mide kanseri verisi için CRM'yi kullanarak yaşam verilerindeki dinamik etkilerin modellenmesi için bir yöntem önermiş ve klasik yöntemler ile karşılaştırma yapmışlardır.⁶ Bellera ve diğ. (2010) meme kanseri veri kümesi için dinamik CRM'yi kullanarak kanserin yayılma belirtilerini incelemişlerdir.⁷ Buchholz ve Sauerbrei (2011) meme kanseri ve akciğer kanseri verilerini kullanarak çok değişkenli modellerde doğrusal olmayan ve zamanla değişen etkileri değerlendirmek için model oluşturma yöntemlerini karşılaştırmışlardır.⁸ Zhang ve ark. (2014) ilk laktat ve normalleşme süresinin tehlike fonksiyonu ile olan ilişkisini araştırırken kesirli polinomlu CRM'yi kullanmışlar ve en uygun modeli bulmaya çalışmışlardır.⁹ Buchholz ve diğ. (2014) Rotterdam akciğer kanseri verilerini kullanarak zamana bağlı değişkenlere ait yaşam eğrilerini kesirli polinomlarda dinamik CRM ile incelemişlerdir.¹⁰

Bu çalışmada, STATA istatistiksel paket programı kullanılarak literatürde yer alan prostat kanseri veri kümesine klasik CRM ve kesirli polinomlu CRM uygulanarak elde edilen sonuçlar yorumlanmıştır.¹¹

GEREÇ VE YÖNTEMLER

Araştırmacılar regresyon modellerinde çoğu zaman sonuçları etkileyebilecek birçok değişken ile karşı karşıya kalmaktadır.¹²

Model oluşturmanın zorluğu ise sonuç ile sürekli bir değişken arasındaki ilişkide doğrusal olmama sorunu ile nasıl başa çıkılacağıdır. Alternatif olarak, sürekli değişkenler iki veya daha fazla kategoriye ayrılarak kategorik değişkenlere dönüştürülebilir.

Sağlık bilimlerinde sürekli değişken ile sonuç değişkeni arasındaki etkileşim araştırılırken en bilinen yaklaşım tek bir kesim noktası kullanılarak iki grup oluşturmaktır. Kesim noktasının seçiminde, en tatmin edici sonucu veren bir değerin arayışı vardır. Tüm olası kesim noktaları denir ve istatistiksel önemi en üst düzeye çıkarak değer seçilir.^{13,14,15} Alternatif bir yaklaşım, değişkeni sürekli tutmak ve doğrusal olmama durumuna izin vermektir. Karesel veya kübik polinomlar kullanmak yerine, Royston ve Altman (1994) tarafından genel bir regresyon modeli önerilmiştir.¹⁶ Bu model kesirli polinom fonksiyonları kullanmaktadır ve kesirli polinomlu regresyon modeli olarak adlandırılır.

Kesirli Polinom (KP) yöntemi, açıklayıcı bir değişkenin model için önemli olup olmadığını ve fonksiyonel formunun belirlenmesine izin veren bir yöntemdir. Kesirli polinomlar doğrusallık sağlanmadığında sürekli değişkenlerin yapısını korumak için kullanılır.

Bir sonuç değişkenine, tek bir sürekli değişken X 'e ve bunları ilişkilendiren uygun bir regresyon modeline sahip olduğumuzu varsayalım. Başlangıç noktası doğrusal model olan $\beta_1 X$ 'dir. Çoğu zaman bu, ilişkinin yeterli bir açıklaması olacaktır, ancak diğer modeller, olası iyileştirmeler için araştırılmalıdır. Doğrusal modelin dö-

nüshüm formülü $\beta_1 X^p$ 'dir. Bu model değişik p değerleriyle kullanılmaktadır. Model, Royston ve Altman (1994) tarafından birinci dereceden kesirli polinom veya KP1 şeklinde adlandırılmış ve

$$KP1 = \beta_1 X^p \quad (1)$$

biçiminde ifade edilmiştir.³ p değerleri sınırlı bir S adı verilen diziden seçilir. Royston ve Altman (1994), X^0 'ın $\log(x)$ 'i ifade ettiği S kümesi için $\{-2,-1,-0.5,0,0.5,1,2,3\}$ değerlerini önermişlerdir. S kümesi, logaritmik, karekök vb. dönüşümleri içermez.¹⁶ Model uyumu için 8 tane p değeri denenir. En uygun model, en yüksek olasılığı veren değerdir. Doğrusal olmama, $p=1$ 'e karşı $p \neq 1$ 'in anlamlılık testidir. Bu test, sapma farkını serbestlik derecesi 1 olan ki-kare dağılımı ile karşılaştırarak gerçekleştirilmektedir. Eğer test anlamlı değilse doğrusal model kabul edilir, yoksa doğrusal olmayan kesirli polinom tercih edilir.

Polinomiyal regresyonda olduğu gibi, KP1 fonksiyonundan daha karmaşık olan KP2 fonksiyonuna geçilir. $\beta_1 X_1 + \beta_2 X_2$ denklemini kullanmak yerine, KP2 fonksiyonları S kümesinden alınan p_1 ve p_2 değerleriyle $\beta_1 X^{p_1} + \beta_2 X^{p_2}$ olarak tanımlanır. KP2 ise,

$$\begin{aligned} KP2 &= \beta_1 X^{p_1} + \beta_2 X^{p_2} & p_1 \neq p_2 \\ &= \beta_1 X^{p_1} + \beta_2 X^{p_2} \ln x & p_1 = p_2 = p \end{aligned} \quad (2)$$

biçimindedir.¹⁷ Eğer $p_2=p_1$ ise; Royston ve Altman (1994) tekrarlı üslü adımı verdikleri $\beta_1 X^{p_1} + \beta_2 X^{p_2} \ln x$ fonksiyonunu önermişlerdir.¹⁶ S kümesinden elde edilen 36 tane üs kombinasyonu arasından en yüksek olasılık değeri olarak tanımlanır. KP modelinin genel denklemi ise,

$$\begin{aligned} \varphi_m^*(x;p) &= \beta_0 + \varphi_m(x;p) = \sum_{j=1}^m \beta_j h_j(x), \quad j=1, \dots, m \\ h_j(x) &= X^{p_j} & p_j \neq p_{j-1} \\ &= h_{j-1} \log x & p_j = p_{j-1} \end{aligned} \quad (3)$$

biçimindedir.¹⁸

MODEL SEÇİMİNDE KULLANILACAK YÖNTEMLER

KP1 için 8 ve KP2 için 36 olmak üzere toplam 44 model veri kümesi için çalıştırılır ve bu modellerden en uygun olanının seçilmesi gerekmektedir. Bu amaçla, kapalı test yöntemi(RA2 yöntemi) ve ardışık yöntem olmak üzere iki yöntem önerilmiştir.

KAPALI TEST YÖNTEMİ

Bu algoritma Ambler ve Royston (2001) ve Sauerbrei ve Royston (2002) tarafından tanımlanmıştır. Bu algoritma RA2 olarak bilinir. Bu test her bir değişken için yaklaşık 1.tip hata değerini verir. Bu yöntem, aday modeller arasından, belirlenmiş bir minimumdan (yokluk modeli), test sonuçlarının düzenli bir sıralamasına göre öngörülen maksimuma (kp) kademeli olarak artmasına izin verir. Algoritma adımları aşağıdaki gibidir:

1. En iyi ikinci dereceden kesirli polinom modeli 4 serbestlik derecesinde alfa anlamlılık düzeyinde yokluk modeline göre test edilir. Eğer test anlamlı değilse x değişkeni önemli değildir ve bir sonraki aşamaya geçilmez, x değişkeni anlamlıysa bir sonraki aşamaya geçilir.
2. En iyi ikinci dereceden kesirli polinom modeli 3 serbestlik derecesinde alfa anlamlılık düzeyinde doğrusal modele göre test edilir. Eğer test anlamlı değilse bir sonraki aşamaya geçilmez yani son model doğrusal modeldir yoksa bir sonraki aşamaya geçilir.
3. En iyi ikinci dereceden kesirli polinom modeli 2 serbestlik derecesinde alfa anlamlılık düzeyinde en iyi birinci dereceden kesirli polinom modeline göre test edilir. Eğer test anlamlı değilse en iyi model birinci dereceden kesirli polinom modelidir, yoksa en iyi model ikinci dereceden kesirli polinom model olarak alınır.

Sırasıyla 1., 2. ve 3. adımdaki testler genel ilişkiyi, doğrusal olmamayı ve daha basit veya karmaşık kesirli polinom modelleri arasındaki seçimi inceler.

ARDIŞIK YÖNTEM

En iyi modeli seçmek için kullanılan ikinci alternatif algoritma olan ardışık yöntemin adımları aşağıdaki gibidir:

1. En iyi ikinci dereceden kesirli polinom modeli 2 serbestlik derecesinde alfa anlamlılık düzeyinde birinci dereceden kesirli polinom modeline göre test edilir. Eğer test anlamlıysa bir sonraki aşamaya geçilmez ve en iyi model ikinci dereceden kesirli polinom modeli olarak alınır.
2. En iyi birinci dereceden kesirli polinom modelini 1 serbestlik derecesinde alfa anlamlılık düzeyinde doğrusal modele göre test edilir. Eğer test anlamlıysa bir sonraki aşamaya geçilmez ve en iyi model birinci dereceden kesirli polinom olarak alınır. Yoksa bir sonraki aşamaya geçilir.
3. Doğrusal modeli alfa anlamlılık düzeyinde, 1 serbestlik derecesinde x 'in olmadığı modele göre test edilir. Eğer test anlamlıysa en iyi model doğrusal model olarak alınır. Yoksa x modele dahil edilmez.¹⁹

Doğru ilişki doğrusal model olduğunda 1. tip hata değeri nominal alfa değerini aşabilir. Bu yüzden prosedür basit olanlara kıyasla daha karmaşık modelleri tercih etmektedir. Amber ve Royston (2001)'nin yapmış olduğu simülasyon çalışmasının sonuçlarına dayanarak en uygun fonksiyon seçimi için kapalı test yöntemi (RA2 yöntemi) önerilir.^{11,20}

ÖLÇEKLEME VE ODAKLAMA

Dönüştürülmüş x^p değerlerinin hassas ölçeklendirilmiş ve ortalanmış (centred) olmasını sağlamak için sayısal ve kavramsal avantajlar vardır. Ölçeklendirme, model tahmininde belirsizlikleri ve zorlukları azaltır. Merkezleme (centring), CRM'de β_0 kesim noktasının veya temel tehlike fonksiyonunun anlamlı bir şekilde yorumlanmasını sağlar. Ölçekleme işlemi aşağıdaki gibi ifade edilebilir:

$$\text{Irange} = \log_{10}[\max(x) - \min(x)] \quad (4)$$

$$\text{ölçek} = 10^{\text{sign(Irange)} \cdot \text{int}(\text{Irange})} \quad (5)$$

$$x^* = x/\text{ölçek} \quad (6)$$

Burada $\text{int}(x)$, x 'in altındaki en yakın tam sayıyı gösterir. Miktar ölçeği 10'un üssü olan, bir tamsayıdır. X 'in KP dönüşümü daha sonra x^* gözlenen değerlerinin ortalamasına odaklanır. Örneğin, KP1 modeli için $\beta_0 + \beta_1 X^{p_1}$, STATA programı tarafından

$$\beta_0 + \beta_1^* \left(\left(\frac{x}{\text{ölçek}} \right)^p - \left(\frac{1}{n} \sum_{i=1}^n x_i^* \right)^p \right) \quad (7)$$

biçiminde uyarlanır.

İşlem sırası ilk olarak ölçekleme, ikincisi kesirli polinom üslerini tahmin etme ve son olarak merkezlemedir. İlk aşamada merkezleme farklı üslere yol açabilir ve bu yüzden ilk aşamada kullanılması tavsiye edilmemektedir. Üslerin tahmini ölçeklemeden etkilenmez. Ölçeğin herhangi bir pozitif değeri için aynı üsler $x/\text{ölçek}$ için bulunur. Ölçekleme, modelin doğruluğunu korumak için gereklidir.²¹

KESİRLİ POLİNOMLAR İLE ÇOK DEĞİŞKENLİ MODEL OLUŞTURMA

Model oluşturmada amaç uygun, yorumlanabilir çok değişkenli bir model türetmek ve daha güçlü tahmin ediciler elde etmektir. Sürekli değişkenler için doğru fonksiyonel biçimlerin belirlenmesi de önemlidir. Doğrusal

olmayan modeller için uygun modeli oluştururken değişken seçim yöntemi olarak geriye doğru seçim yöntemi de ileriye doğru değişken seçim yöntemi ile birlikte kullanılır. Sauerbrei ve Royston (1999) bunu çok değişkenli kesirli polinom (Multivariate Fractional Polynomial, MFP) olarak adlandırmışlardır.²² Çok değişkenli kesirli polinom, yalnızca regresyon modelleri oluşturma hakkında genel bilgi gerektirir. İki anlamlılık seviyesine ihtiyaç vardır: Geriye doğru seçim ile değişkenleri seçmek için α_1 ve ileriye doğru seçim yöntemi içindeki işlevlerin uyumunu karşılaştırmak için α_2 dir. Çok değişkenli kesirli polinom (α_1, α_2) notasyonu, Geriye doğru seçim için α_1 ve ileriye doğru seçim yöntemi için α_2 kullanılarak seçilen bir modeli gösterir. Ortak durum $\alpha_1 = \alpha_2 = \alpha$, çok değişkenli kesirli polinom (α) olarak kısaltılmıştır.

Çok değişkenli kesirli polinom algoritması, yani MFP (α_1, α_2) prosedürünün adımları aşağıdaki gibidir. Tüm anlamlılık testleri ki-kare olasılık oranını esas almaktadır.

1. Nominal p değerleri α_1, α_2 olarak seçilmiştir. $\alpha_1 = \alpha_2 = 0.05$ 'dir. Değerler değişkenler arasında değişebilir. Bir değişken için $\alpha_1=1$ alındığında değişken modele alınır. Sürekli bir değişken için $\alpha_2 = 1$ olduğunda kesirli polinom fonksiyonu için modele alınır.

2. Maksimum serbestlik dereceli kesirli polinom fonksiyonları seçilir. Örneğin;

serbestlik derecesi=2 ve serbestlik derecesi=1 sırasıyla KP2, KP1 ve doğrusal fonksiyonlar anlamına gelir.

3. Tahmin edicilerin modele alınma sırası, her bir tahmin edicinin modelden çıkarılması için p değerine göre belirlenir. En önemli tahmin edici, ilk sırada modele alınır. Prosedürün tüm döngülerinde tutulan x_1, \dots, x_k değişkenlerinin bu sırayla düzenlendiği varsayılır.

4. Döngü sayacı $c=0$ olarak başlatılsın.

5. Her döngü içindeki değişken sayacını başlatmak için $j = 1$ olsun.

6. Eğer x_j sürekli değişkense 7. adıma gidilir; değilse x_j değişkeni kategorik ya da ikili değişkendir. Kukla değişkenlerin ortak önemi α_1 anlamlılık düzeyinde test edilir. Diğer tüm değişkenler varsayım olarak modele dahil edilmiştir. Eğer x_j önemliyse modelde tutulur; değilse modelden çıkarılır. 8.adıma gidilir.

7. Fonksiyon seçim prosedürünün 1. adımı $\alpha = \alpha_1$ seviyesinde x_j 'ye uygulanır. Eğer x_j önemli değilse modelden çıkarılır; önemliyse fonksiyon seçim prosedürünün 2. ve 3. adımları $\alpha = \alpha_2$ seviyesinde, kesirli polinom veya doğrusal fonksiyonu seçmek için uygulanır. Diğer tüm değişkenler varsayım olarak modele dahil edilmiştir. p üslerinin seçildiği kesirli polinom durumunda, x_j dönüştürülmüş x_j^p olarak temsil edilir.

8. x_j 'nin modele dahil edilmesi veya modelden çıkarılması, bir sonraki döngüde ele alınana kadar geçerlidir.

9. $j=j+1$ olsun. Eğer $j \leq k$ ise, sonraki tahmin edici için 6. adıma geri dönülür; değilse 10. adıma gidilir.

10. $c=c+1$ olsun. c. döngüde adımlar tamamlanır. Eğer $c > c_{\max}$ ise dur. c_{\max} döngülerinde, algoritma başarısız olur. Aksi halde, dahil edilen değişkenlerin ve FP dönüşümlerinin $c-1$ döngüsünden c döngüsüne değişip değişmediği kontrol edilir. Eğer öyleyse, yeni bir döngü başlatmak için 5. adıma dönülür. Yoksa, döngü durdurulur ve mevcut model tahminleri kullanılır.²¹

Bu çalışmada, Byar ve Green (1980) tarafından incelenen ve hastalığının 3. ve 4. aşamalarında olan 475 hasta-ya ait prostat kanseri veri kümesi kesirli polinomlar ile Cox regresyon modelini uygulamak için kullanılmıştır.¹ Byar ve Green(1980)'in prostat kanseri verisi literatürde birçok çalışmada farklı amaçlar için kullanılmıştır. İlk olarak, Byar ve Green (1980) üstel regresyon modelini kullanarak tedavinin yaşam süresi üzerindeki etkisini incelemişlerdir.¹ Kay (1986) çalışmasında ise tedavi ile farklı nedenlerden ölüm arasındaki ilişkiyi yarışan riskleri kullanarak incelemiştir.²² Schemper ve Heinze (1997) prognostik faktör çalışmaları için önerdikleri koşullu olasılık yerine koyma yöntemini incelemek için bu veri kümesini kullanmıştır.²³ Chen vd. (1998) bağımlı yarışan risklerde birikimli insidans fonksiyonu için güven aralıkları ve bantlarını elde etmek için bir yöntem önermiş ve prostat kanseri veri kümesinde uygulamışlardır.²⁴ Hunt ve Jorgensen (1999) karma model

sınıflandırması için bu veri kümesini kullanmıştır.²⁵ Bu çalışmada ise prostat kanseri veri kümesine ilk kez kesirli polinomlu Cox regresyon modeli uygulanmıştır.

Çalışma Helsinki Deklerasyonu Prensipleri'ne uygun olarak yapılmıştır.

BULGULAR

Prostat kanseri veri kümesindeki hastalar 5 yıl boyunca izlenmiş ve 475 hastadan 338'inde (%71.1) başarısızlık ve 137'sinde (%28.8) durdurma gözlenmiştir. Hastaların ölümü başarısızlık olarak ele alınmıştır. Yaş, tedavi türü ve hastalığın evresi açıklayıcı değişkenler olarak ele alınmıştır. Yaş değişkeni sürekli bir değişken, diğer değişkenler ise kategorik değişken olarak modellere dahil edilmiştir. Hastaların yaş ortalaması 71.56 ± 0.318 (min=48, max=89), ortanca yaşam süresi 33.5 ± 2.137 ay ve 5 yıllık yaşam olasılığı %25.14 olarak bulunmuştur. Kategorik değişkenlere ait bilgiler Tablo 1'de verilmiştir.

Prostat kanseri veri kümesine ait KM yaşam eğrisi Şekil 1'de verilmiştir.

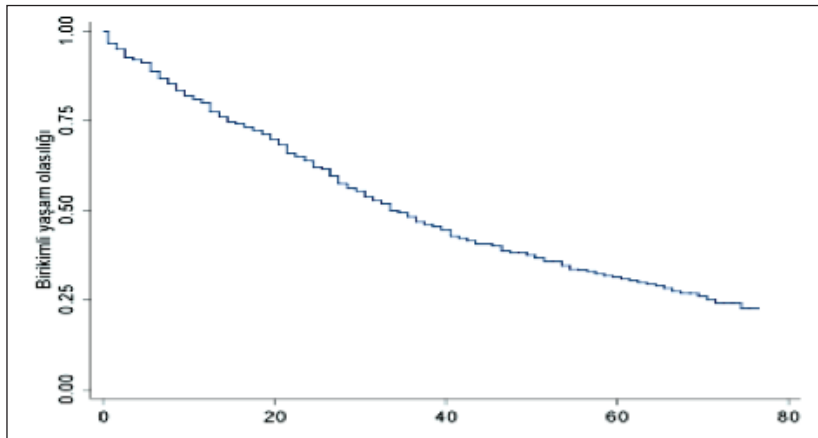
Çalışmada kullanılan her bir açıklayıcı değişken için KM yaşam eğrileri ise Şekil 2(a)-(c)'de verilmiştir.

Evre ve tedavi değişkenlerinde değişken düzeylerinin yaşam eğrileri birbirine çok yakın olarak elde edilmiştir. Yaş değişkeni için düzeyler arasında farklılığın daha fazla olduğu görülmektedir. Yaşı 65 yaş ve altı, 66-71 yaş arası ve 72-76 yaş arası olan hastaların yaşam olasılıklarının birbirine yakın olduğu ve 77-80 yaş arası ile 81 yaş ve üzerinde olan hastaların yaşam olasılıklarından daha yüksek olduğu görülmektedir (Şekil 2).

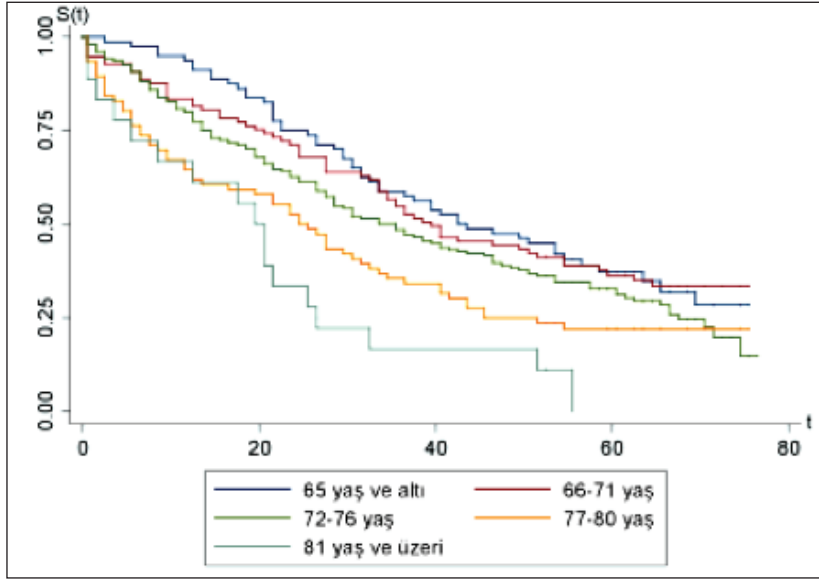
TABLO 1: Prostat kanseri veri kümesi için kullanılan değişkenler ve düzeyleri.

Değişken	Değişken Düzeyleri	n (%)	Başarısız Gözlemlerin Sayısı	Durdurulmuş Gözlemlerin Sayısı
Kategorik yaş*	1: 65 yaş ve altı	80 (%35)	52	28
	2: 66-71 yaş	97 (%35.1)	63	34
	3: 72-76 yaş	204 (%27.9)	147	57
	4: 77-80	76 (%22.4)	59	17
	5: 81 yaş ve üzeri	18 (%5.6)	17	1
Hastalığın evresi	Evre 3:Yerel uzak metastaz belirtisi olmayan	272 (%33.5)	181	91
	Evre 4:Uzak metastaz belirtisi olan	203 (%22.7)	157	46
Tedavi türü	0: Placebo	238 (%33.6)	158	80
	1: TRT E	237 (%24.1)	180	57

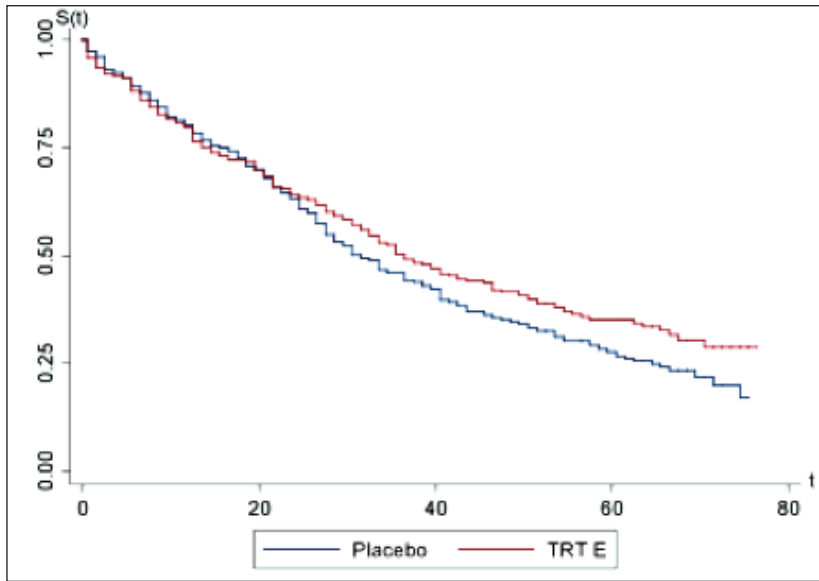
*: Yaş değişkeni kategorik olarak alındığında kullanılan sınıflandırma.



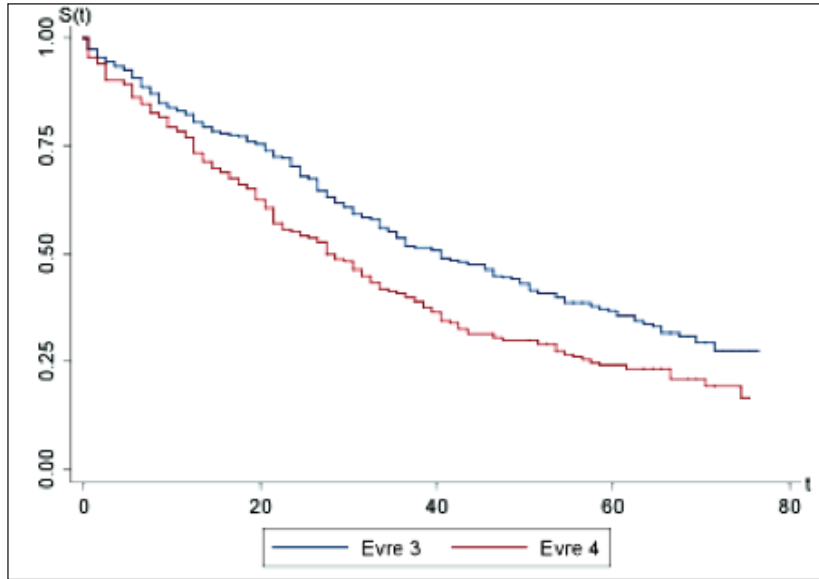
ŞEKİL 1: Prostat kanseri verisi için KM yaşam eğrisi.



ŞEKİL 2(a): Yaş değişkeni için KM eğrisi.



ŞEKİL 2(b): Tedavi değişkeni için KM eğrisi.



ŞEKİL 2(c): Evre değişkeni için KM eğrisi.

Prostat kanseri veri kümesi için log-rank testi sonuçları **Tablo 2**'de gibi elde edilmiştir. Yaşam olasılıkları açısından yaş ($p < 0.001$) ve evre ($p < 0.001$) değişkenlerinin tüm düzeyleri arasında anlamlı bir farklılık olduğu %95 güven düzeyinde söylenebilir (**Tablo 2**).

KLASİK COX REGRESYON MODELİ SONUÇLARI

Öncelikle yaş sürekli bir değişken olarak ele alınmış ve çok değişkenli CRM veri kümesine uygulanmıştır. Modelin uygunluğuna ilişkin p değeri $p < 0.001$ olarak elde edildiğinden, modelin %95 güven düzeyinde anlamlı olduğu sonucuna ulaşılmıştır. Bu modele ait AIC= 3781.429 ve BIC= 3793.919 olarak elde edilmiştir. Kurulan CRM'ye ait sonuçlar **Tablo 3**'de verilmiştir.

Tüm değişkenlerin birinci düzeyi referans kategorisi olarak alınmıştır. Yaş ($p < 0.001$) ve evre ($p=0.001$) değişkenleri önemli risk faktörleri olarak belirlenmiştir. CRM'den elde edilen sonuçlara göre, yaş değişkenindeki 1 birimlik artış ölüm riskini 1.031 ($e^{0.031}=1.031$) birim arttırmaktadır. Evre 4 (uzak metastaz belirtisi olan) aşamasında olan hastaların, evre 3 (yerel uzak metastaz belirtisi olmayan) aşamasında olan hastalara göre 1.429 ($e^{0.357}=1.429$) kat ölüm riskinin daha fazla riskli olduğu görülmüştür (**Tablo 3**).

TABLO 2: Prostat kanseri verisi için log-rank testi sonuçları.

Değişken	Ki-kare	p değeri
Yaş	24.176	$p < 0.001$
Evre	10.107	0.001
Tedavi türü	2.663	0.103

TABLO 3: Prostat kanseri veri kümesi için yaş değişkeni sürekli kabul edilerek oluşturulan çok değişkenli klasik CRM sonuçları.

Değişken	β	Standart hata	p değeri	Exp(β)	Exp(β) için %95 güven aralıkları
Yaş	0.031	0.009	$p < 0.001$	1.031	1.023, 1.039
Evre	0.357	0.109	0.001	1.429	1.281, 1.593
Tedavi türü	-0.163	0.109	0.135	0.849	0.761, 0.947

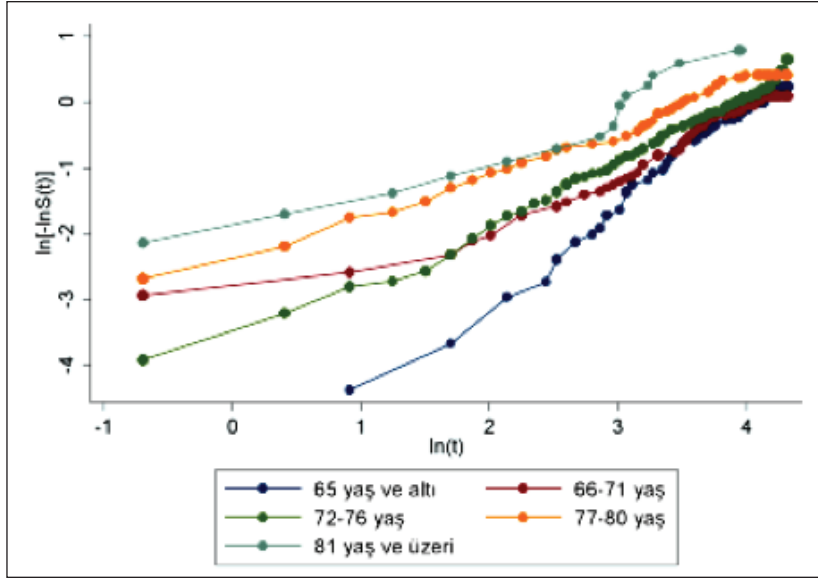
TABLO 4: Prostat kanseri veri kümesi için Schoenfeld artıkları yöntemi sonuçları.

Değişken	r^*	Ki-kare	Sd	p değeri
Yaş	-0.150	9.76	1	0.001
Evre	-0.043	0.65	1	0.420
Tedavi türü	-0.077	2.09	1	0.148

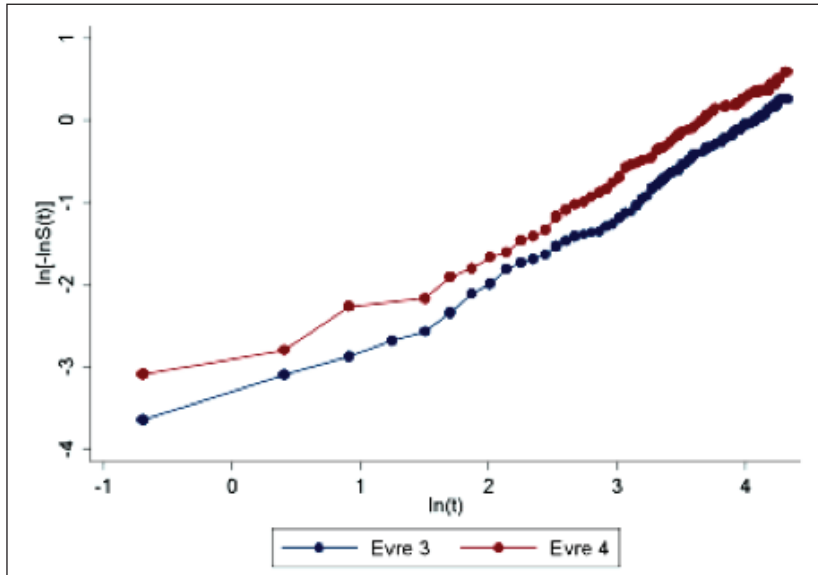
* r: Schoenfeld artıkları ile logt arasındaki Pearson çarpım-moment korelasyon katsayısı.

Çalışmada orantılı tehlikeler varsayımı (OT) Schoenfeld artıkları yöntemi ile incelenmiştir. Yaş değişkeni için $p < 0.05$ olduğundan, varsayımının sağlanmadığı sonucuna ulaşılmıştır (Tablo 4).

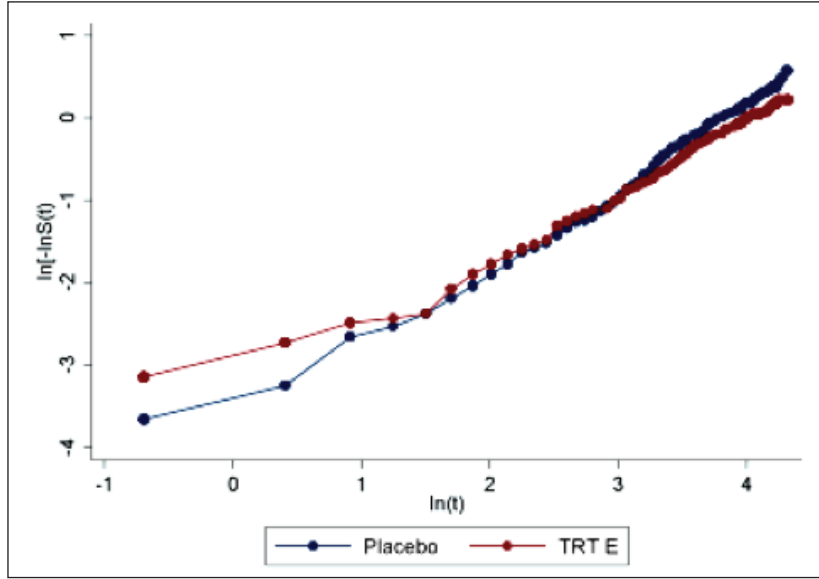
Çalışmada kullanılan prostat kanseri veri kümesine ait OT varsayımı grafiksel yöntemle de incelenmiş olup, sonuçlar Şekil 3'te verilmiştir.



ŞEKİL 3(a): Yaş değişkeni için log-log grafiği.



ŞEKİL 3(b): Evre değişkeni için log-log grafiği.



ŞEKİL 3(c): Tedavi değişkeni için log-log grafiği.

Değişkenlerin $\ln(t)$ 'ye karşı $-\ln(-\ln S(t))$ grafikleri çizilmiştir. Yaş ve tedavi değişkenlerinin düzeyleri çakıştığı için OT varsayımının sağlanmadığı görülmektedir (Şekil 3(a), Şekil 3(c)).

Klasik CRM için, ikinci olarak model sadece sürekli yaş değişkeni için kurulmuş ve %95 güven düzeyinde model anlamlı bulunmuştur ($p=0.0001$). Bu modele ait $AIC=3789.696$ ve $BIC=3793.859$ olarak elde edilmiştir.

Bu model için yaş değişkeninin anlamlı olduğu görülmüştür. Yaş değişkenindeki 1 birimlik artış, ölüm riskini 1.030 ($e^{0.033}=1.030$) birim arttırmaktadır (Tablo 5).

Üçüncü olarak, klasik CRM için model sadece kategorik yaş değişkeni için kurulmuş ve model %95 güven düzeyinde anlamlı bulunmuştur ($p < 0.001$). Bu CRM için $AIC=3790.376$ ve $BIC=3807.029$ olarak elde edilmiştir.

Sadece kategorik yaş değişkeni ele alınarak klasik CRM uygulandığında 77-80 yaş grubu ve 81 yaş ve üzeri yaş grubu anlamlı çıkmaktadır. 77-80 yaş arasındaki hastaların, 65 yaş ve altındaki hastalara göre 1.729 ($e^{0.548}=1.729$) kat ölüm riskinin daha fazla riskli olduğu söylenebilmektedir. 81 yaş ve üzeri yaş arasındaki hastaların, 65 yaş ve altındaki hastalara göre 2.989 ($e^{1.095}=2.989$) kat ölüm riskinin daha fazla riskli olduğu söylenebilmektedir (Tablo 6).

TABLO 5: Prostat kanseri veri kümesi için yaş değişkeni sürekli kabul edilerek oluşturulan tek değişkenli klasik CRM sonuçları.

Değişken	β	Standart hata	p değeri	Exp(β)	Exp(β) için %95 güven aralıkları
Yaş	0.033	0.089	$p < 0.001$	1.030	0.945, 1.129

TABLO 6: Prostat kanseri veri kümesi için yaş değişkeni kategorik kabul edilerek oluşturulan tek değişkenli klasik CRM sonuçları.

Değişken	β	Standart hata	p değeri	Exp(β)	Exp(β) için %95 güven aralıkları
Kategorik yaş			$p < 0.001$		
66- 71 yaş	0.054	0.197	0.773	1.055	0.866, 1.285
72-76 yaş	0.266	0.210	0.099	1.304	1.057, 1.609
77-80 yaş	0.548	0.329	0.004	1.729	1.244, 2.403
81 yaş ve üzeri	1.095	0.840	$p < 0.001$	2.989	1.290, 6.924

TABLO 7: Modele tüm değişkenler dahil edildiğinde yaş değişkeni kategorik kabul edilerek oluşturulan klasik CRM sonuçları.

Değişken	β	Standart hata	p değeri	Exp(β)	Exp(β) için %95 güven aralıkları
Kategorik yaş					
66-71 yaş	0.063	0.187	0.736	1.065	0.883, 1.284
72-76 yaş	0.250	0.161	0.121	1.284	1.093, 1.508
77-80 yaş	0.501	0.193	0.090	1.650	1.360, 2.001
81 yaş ve üzeri	1.113	0.284	p <0.001	3.043	2.291, 4.043
Evre	0.334	0.109	0.002	1.396	1.252, 1.557
Tedavi türü	-0.193	0.824	0.090	0.849	0.361, 1.879

* Değişkenler için referans düzeyi: yaş için "65 yaş ve altı", hastalığın evresi için "Evre 3", tedavi türü için "0: Placebo" olarak alınmıştır.

Klasik CRM için son olarak modele tüm değişkenler dahil edilerek ve yaş değişkeni kategorik kabul edilerek kurulmuş ve model %95 güven düzeyinde anlamlı bulunmuştur ($p <0.001$). Bu model için AIC= 3782.703 ve BIC= 3807.683 olarak elde edilmiştir.

81 yaş ve üzeri yaş grubunun anlamlı çıktığı görülmektedir. Ölüm riski üzerinde evre değişkeninin anlamlı bir etkisi olduğu görülmüştür (Tablo 7). Evre 4 (uzak metastaz belirtisi olan) aşamasında olan hastaların, evre 3 (yerel uzak metastaz belirtisi olmayan) aşamasında olan hastalara göre 1.396 ($e^{0.334}=1.396$) kat ölüm riskinin daha fazla riskli olduğu söylenebilir. 81 yaş ve üzerindeki hastaların, 65 yaş ve altındaki hastalara göre 3.043 kat ölüm riskinin daha fazla riskli olduğu görülmektedir (Tablo 7).

KESİRLİ POLİNOMLU COX REGRESYON MODELİ SONUÇLARI

Kesirli polinom yöntemi, açıklayıcı bir değişkenin model için önemli olup olmadığını ve fonksiyonel biçiminin belirlenmesine izin veren bir yöntemdir. Kesirli polinomlar doğrusallık sağlanmadığında sürekli değişkenlerin yapısını korumak için kullanılmaktadır. KP1 ve KP2 olmak üzere 2 modelden oluşmaktadır ve toplamda veri kümesi için 44 model denenmektedir. Polinomun derecesi olarak ifade edilen p değerleri $S = \{-2, -1, -0.5, 0, 0.5, 1, 2, 3\}$ kümesindeki değerlerinden oluşmaktadır. En uygun model en yüksek olasılığı veren modeldir. $p=1$ 'e karşı $p \neq 1$ 'in anlamlılık testi doğrusal olmamanın testidir. Eğer test anlamlı değilse doğrusal model kabul edilir, yoksa doğrusal olmayan kesirli polinomlu model tercih edilir.

Çalışmada öncelikle kesirli polinomlu CRM sadece yaş değişkeni kullanılarak kurulmuş ve kapalı test seçim yöntemi ile elde edilen sonuçlar Tablo 8 ve Tablo 9'da verilmiştir.

TABLO 8: Prostat kanseri veri kümesi için tek değişkenli kesirli polinomlu CRM için kapalı test seçim yöntemi sonuçları.

Model	Sapma	p derece	Adım
KP2	3778.586	3,3	1
KP1	3784.375	3	2
Doğrusal	3787.696	1	3
Yokluk	3802.495	-	-

TABLO 9: Prostat kanseri veri kümesi için tek değişkenli kesirli polinomlu CRM için kapalı test seçim yöntemi sonuçları.

Karşılaştırma	Sapma farkı	p değeri
KP2 ile yokluk	23.909	$p <0.001$
KP2 ile doğrusal	9.110	0.028<0.05
KP2 ile KP1	5.790	0.055>0.05

KP2 ile yokluk modeli karşılaştırıldığında, yaş değişkeninin anlamlı olduğu ve KP2 ile doğrusal model karşılaştırıldığında yaş değişkeninin bağımlı değişken ile doğrusal olmayan bir ilişkisi olduğu görülmüştür. KP2 ile KP1 karşılaştırıldığında ise test anlamlı çıkmamaktadır bu yüzden en iyi modelin KP1 modeli olduğu söylenebilir (Tablo 9). Tablodaki KP2 ve KP1'in sapma değerleri karşılaştırıldığında da en iyi kesirli polinomlu CRM olarak derecesi 3, sapması 3784.375 değeriyle KP2'nin sapma değerinden yüksek olmasından dolayı KP1 modeli olduğunu söyleyebiliriz (Tablo 8).

Tek değişkenli kesirli polinomlu CRM'ye ait sonuçlar Tablo 10'da verilmiştir.

KP2 modelleri için de en uygun model $p_1=p_2=3$ olarak elde edilmiştir. Yani,

$$KP(2) = \beta_1 X^3 + \beta_2 X^3 \ln x \text{ dir.}$$

İkinci olarak, veri kümesine çok değişkenli kesirli polinomlu uygulanmış ve kapalı test seçim yöntemi ile sonuçlar elde edilmiştir (Tablo 11 ve Tablo 12).

KP2 ile yokluk modeli karşılaştırıldığında, yaş değişkeninin anlamlı olduğu ve KP2 ile doğrusal model karşılaştırıldığında yaş değişkeni ile bağımlı değişken arasında doğrusal olmayan bir ilişkisi olduğu görülmüştür. KP2 ile KP1 karşılaştırıldığında ise test anlamlı çıkmamaktadır bu yüzden en iyi modelin KP1 modeli olduğu söylenebilir (Tablo 12). KP2 ve KP1'in sapma değerleri karşılaştırıldığında da en iyi kesirli polinomlu CRM olarak derecesi 3, sapması 3772.374 değeriyle KP2'nin sapma değerinden yüksek olmasından dolayı KP1 modeli olduğunu söyleyebiliriz (Tablo 11).

Prostat kanseri veri kümesi için çok değişkenli kesirli polinomlu CRM sonuçları Tablo 13'deki gibidir (Tablo 13).

KP2 modelleri için de en uygun model $p_1=p_2=3$ olarak elde edilmiştir. Yani,

$$KP(2) = \beta_1 X^3 + \beta_2 X^3 \ln x$$

biçimindedir. Çok değişkenli kesirli polinomlu CRM modelinde yaş ve evre değişkenlerinin anlamlı çıktığı görülmektedir. Evre 4 (uzak metastaz belirtisi olan) olan hastaların, evre 3 (yerel uzak metastaz belirtisi olma-

TABLO 10: Prostat kanseri veri kümesi için tek değişkenli kesirli polinomlu CRM sonuçları.

Değişken	β	Standart hata	p değeri	Exp(β)	Exp(β) için %95 güven aralıkları
Yaş_1	-0.038	0.016	0.017	0.962	0.947; 0.978
Yaş_2	0.017	0.007	0.011	1.017	1.010; 1.024

TABLO 11: Prostat kanseri veri kümesi için çok değişkenli kesirli polinomlu CRM için kapalı test seçim yöntemi sonuçları.

Model	Sapma	p derece	Adım
KP2	3766.934	3,3	1
KP1	3772.374	3	2
Doğrusal	3775.429	1	3
Yokluk	3789.494	-	-

TABLO 12: Prostat kanseri veri kümesi için çok değişkenli kesirli polinomlu CRM için kapalı test seçim yöntemi sonuçları.

Karşılaştırma	Sapma farkı	p değeri
KP2 ile yokluk	22.560	$p < 0.001$
KP2 ile doğrusal	8.495	$0.037 < 0.05$
KP2 ile KP1	5.440	$0.066 > 0.05$

TABLO 13: Prostat kanseri veri kümesi için çok değişkenli kesirli polinomlu CRM sonuçları.

Değişken	β	Standart hata	p değeri	Exp(β)	Exp(β) için %95 güven aralıkları
Yaş_1	-0.037	0.016	0.021	0.963	0.948, 0.979
Yaş_2	0.017	0.007	0.014	1.017	1.010, 1.024
Tedavi türü	-0.194	0.110	0.078	0.823	0.737, 0.919
Evre	0.332	0.109	0.003	1.393	1.249, 1.554

*Evre değişkeni için Evre 3 referans kategorisi olarak alınmıştır. Tedavi türü değişkeni için 0: Placebo referans kategorisi olarak alınmıştır.

TABLO 14: Prostat kanseri veri kümesi için kesirli polinomlu çok değişkenli CRM'ye ait Schoenfeld artıkları yöntemi sonuçları.

Değişken	r^*	Ki-kare	Sd	p değeri
Yaş_1	-0.063	1.31	1	0.251
Yaş_2	0.060	1.16	1	0.281
Evre	-0.057	1.11	1	0.291
Tedavi türü	-0.071	1.77	1	0.183

* r : Schoenfeld artıkları ile logt arasındaki Pearson çarpım-moment korelasyon katsayısı.

yan) aşamasında olan hastalara göre 1.393 ($e^{0.332}=1.393$) kat ölü riskinin daha fazla riskli olduğu söylenebilir (Tablo 13).

Prostat veri kümesi için çok değişkenli kesirli polinomlu CRM sonuçları elde edildikten sonra, OT varsayımı yeniden incelenmiş ve elde edilen sonuçlar Tablo 14'te verilmiştir. Tablo 14'ten, $p > 0.05$ olduğundan OT varsayımının sağlandığı görülmüştür.

Klasik CRM sonuçlarında prostat veri kümesi için OT varsayımı sağlanmadığından CRM sonuçlarının kullanılması ve yorumlanması uygun değildir. Ancak kesirli polinomlu CRM sonrası yaş değişkeninin bağımlı değişken ile doğrusal olmayan bir ilişkiye sahip olduğu belirlendiğinden ve tekrar Schoenfeld artıkları yöntemi sonuçlarına bakıldığında OT varsayımı sağlandığından kesirli polinomlu CRM'nin prostat veri kümesi için kullanılmasının daha uygun olduğu sonucuna varılmıştır (Tablo 14).

TARTIŞMA

Yaşam modelleri, açıklayıcı değişkenlerin yaşam süresi ile ilişkisini incelemek için tasarlanmaktadır. Yarı parametrik bir yaşam modeli olan Cox regresyon modeli yaşam çözümlenmesinde en çok kullanılan model olma özelliğine sahiptir. Modele dahil edilen açıklayıcı değişkenlerin etkilerinin doğrusal olmaması durumunda kesirli polinomlu Cox regresyon modelinin kullanımı önerilmiştir.

Uygulamalarda çoğunlukla doğrusal olmayan ilişkilerden kaçınmak için açıklayıcı değişkenler kategorik biçime dönüştürülerek modele dahil edilmektedir. Ancak, veriyi kategorik biçime dönüştürürken, değişken düzeylerinin nasıl belirleneceği konusunda da kararsızlıklar ortaya çıktığı için kesirli polinomlu modelleri kullanmak daha uygundur. Sürekli açıklayıcı değişkenin, kategorik biçime dönüştürülmeden modele dahil edilmesi ve bağımlı değişken ile doğrusal ilişkiye sahip olup olmadığının incelenmesi önerilen modelin uygunluğuna karar verebilmek için önemlidir.

Bu çalışmada prostat kanseri veri kümeleri kullanılmıştır. Modelde yer alan sürekli değişken dikkate alınarak tek değişkenli ve çok değişkenli klasik ve kesirli polinomlu CRM modelleri veri kümelerine uygulanmıştır. Ayrıca değişkenin anlamlılığı konusunda bir farklılık olup olmadığını görebilmek için de ilgili sürekli değişken kategorik biçime dönüştürülerek de klasik CRM'ye dahil edilmiştir.

Prostat kanseri veri kümesi için model kurarken ilk olarak, modele tüm değişkenler dahil edilerek sadece yaş değişkenini sürekli kabul ederek klasik Cox regresyon çözümlemesi yapılmıştır ve OT varsayımını sağlamadığı görülmüştür. Ancak kesirli polinomlu Cox regresyon modeli ile klasik Cox regresyon modeli sonuçları arasındaki farkı görebilmek için analize devam edilmiştir. Model sonucunda yaş ve evre değişkenlerinin anlamlı olduğu görülmüştür. İkinci olarak, sadece sürekli yaş değişkeni modele dahil edilerek klasik Cox regresyon çözümlemesi yapıldığında yaş değişkeninin anlamlı olduğu görülmektedir. Üçüncü olarak, sadece kategorik yaş değişkeni modele dahil edilerek klasik Cox regresyon çözümlemesi yapıldığında ise 77-80 yaş grubu ve 81 yaş ve üzeri yaş grubu anlamlı çıkmıştır. Dördüncü olarak, modele tüm değişkenler dahil edilerek ve yaş değişkeni kategorik değişken olarak ele alınıp klasik Cox regresyon çözümlemesi yapıldığında 81 yaş ve üzeri yaş grubunun ve evre değişkeninin anlamlı çıktığı görülmüştür.

Yaş değişkeninin bağımlı değişken ile doğrusal ilişkili olup olmadığını görmek için tek değişkenli ve çok değişkenli kesirli polinom CRM uygulanmıştır. Her iki modelde de yaş değişkeninin anlamlı olduğu ve bağımlı değişken ile doğrusal olmayan bir ilişkiye sahip olduğu sonucuna ulaşılmıştır. Ayrıca kesirli polinomlu CRM kurulduktan sonra, OT varsayımı yeniden incelenmiş ve varsayımın bu defa sağlandığı görülmüştür. Buna göre kesirli polinomlu CRM'nin bu veri kümesi için kullanılması ve yorumlanması daha doğru olduğu sonucuna ulaşılmıştır.

SONUÇ

Kesirli polinomlu Cox regresyon modelinin kullanıma ilişkin ulusal literatürde bir çalışmaya rastlanamamaktadır. Kesirli polinomlu Cox regresyon modeli kategorik biçime dönüştürme aşamasındaki bilgi kaybını engellemesi, doğrusal olmayan ilişkileri ve fonksiyonel biçimini belirlemesi nedeni ile uygulamacılar için oldukça kullanışlıdır. Ayrıca istatistiksel paket programları için de son yıllarda modele ait kodların geliştirilmesi ile birlikte kullanımı kolaylaştırmıştır.

Paket programlar kullanılarak elde edilen kesirli polinomlu modeller, modelde yer alan tek bir değişken için kesirli polinomların kullanımına izin vermektedir. Bundan sonraki çalışmalarda birden fazla sürekli değişken için modelin kullanılmasına izin veren yazılımlar geliştirilebilir.

Finansal Kaynak

Bu çalışma sırasında, yapılan araştırma konusu ile ilgili doğrudan bağlantısı bulunan herhangi bir ilaç firmasından, tıbbi alet, gereç ve malzeme sağlayan veya üreten bir firma veya herhangi bir ticari firmadan, çalışmanın değerlendirme sürecinde, çalışma ile ilgili verilecek kararı olumsuz etkileyebilecek maddi ve/veya manevi herhangi bir destek alınmamıştır.

Çıkar Çatışması

Yazarlar herhangi bir çıkar çatışması veya finansal destek bildirmemiştir.

Yazar Katkıları

Fikir/Kavram: Hazal Dinç, Nihal Ata Tutkun; **Tasarım:** Hazal Dinç, Nihal Ata Tutkun **Denetleme/Danışmanlık:** Hazal Dinç, Nihal Ata Tutkun; **Veri Toplama ve/veya İşleme:** Hazal Dinç, Nihal Ata Tutkun; **Analiz ve/veya Yorum:** Hazal Dinç, Nihal Ata Tutkun; **Kaynak Taraması:** Hazal Dinç, Nihal Ata Tutkun; **Makalenin Yazımı:** Hazal Dinç, Nihal Ata Tutkun; **Eleştirel İnceleme:** Hazal Dinç, Nihal Ata Tutkun

KAYNAKLAR

1. Byar DP, Green SB. The choice of treatment for cancer patients based on covariate information. Bull Cancer. 1980;67(4):477-90. PMID: 7013866.
2. Austin PC, Park-Wyllie LY, Juurlink DN. Using fractional polynomials to model the effect of cumulative duration of exposure on outcomes: applications to cohort and nested case, control designs. Pharmacoepidemiol Drug Saf. 2014;23(8):819-29. PMID: 24664670. [Crossref] [PubMed] [PMC]
3. Royston P, Altman DG. Regression using fractional polynomials of continuous covariates: parsimonious parametric modelling. Applied Statistics. 1994;43(3):429-67. [Crossref]

4. Royston P, Ambler G, Sauerbrei W. The use of fractional polynomials to model continuous risk variables in epidemiology. *Int J Epidemiol.* 1999;28(5):964-74. PMID: 10597998. [[Crossref](#)] [[PubMed](#)]
5. Berger, Ursula, Gerein, Pia, Ulm, Kurt, Schäfer, Juliane. On the use of Fractional Polynomials in Dynamic Cox Models. Collaborative Research Center 386, Discussion Paper 207; 2000 (<https://doi.org/10.5282/ubm/epub.1597>).
6. Berger U, Schäfer J, Ulm K. Dynamic Cox modelling based on fractional polynomials: time-variations in gastric cancer prognosis. *Stat Med.* 2003;22(7):1163-80. PMID: 12652560. [[Crossref](#)] [[PubMed](#)]
7. Bellera CA, MacGrogan G, Debled M, de Lara CT, Brouste V, Mathoulin Pélissier S. Variables with time-varying effects and the Cox model: some statistical concepts illustrated with a prognostic factor study in breast cancer. *BMC Med Res Methodol.* 2010;10(20):1471-2288. PMID: 20233435. [[Crossref](#)] [[PubMed](#)] [[PMC](#)]
8. Buchholz A, Sauerbrei W. Comparison of procedures to assess non-linear and time-varying effects in multivariable models for survival data. *Biom J.* 2011;53(2):308-31. PMID: 21328605. [[Crossref](#)] [[PubMed](#)]
9. Zhang Z, Chen K, Ni H, Fan H. Predictive value of lactate in unselected critically ill patients: an analysis using fractional polynomials. *J Thorac Dis.* 2014;6(7):995-1003. PMID: 25093098.]
10. Buchholz A, Sauerbrei W, Royston P. A measure for assessing functions of time-varying effects in survival analysis. *Open J Stat.* 2014;4(11):977-98. [[Crossref](#)]
11. Sauerbrei W, Royston P, Meier-Hirmer C, Benner A. Multivariable regression model building by using fractional polynomials: description of SAS, STATA, R programs. *Comput Stat Data Anal.* 2006;50(12):3464-85. [[Crossref](#)]
12. Miller AJ. Least-squares computations. In: Isham V, Keiding N, Louis T, Reid N, Tibshirani R, Tong H, eds. *Subset Selection in Regression*. 2nd ed. New York: Chapman & Hall; 1990. p.11-34. [[Crossref](#)] [[PMC](#)]
13. Altman DG, Lausen B, Sauerbrei W, Schumacher M. The dangers of using "optimal" cutpoints in the evaluation of prognostic factors. *J Natl Cancer Inst.* 1994;86(11):829-35. PMID: 8182763. [[Crossref](#)] [[PubMed](#)]
14. Miller R, Siegmund D. Maximally selected chi-square statistics. *Biometrics.* 1982;38(4):1011-6. [[Crossref](#)]
15. Lausen B, Schumacher M. Evaluating the effect of optimized cut-off values in the assessment of prognostic factors. *Comput Stat Data Anal.* 1996;21(3):307-26. [[Crossref](#)]
16. Ata Tutkun N. [Kesirli polinomlu lojistik regresyon modeli]. *Journal of Statisticians: Statistics and Actuarial Sciences.* 2015;8(2):27-35.
17. Royston P, Sauerbrei W. Fractional polynomials for one variable, MFP: multivariable model-building with fractional polynomials. In: Balding DJ, Cressie NAC, Fitzmaurice GM, Johnstone IM, Molenberghs G, Scott DW, et al, eds. *Multivariable Model-Building*. 1st ed. New York: John Wiley & Sons, Inc; 2008. p.72-4.
18. Ambler G, Royston P. Fractional polynomial model selection procedures: investigation of type I error rate. *J Stat Comput Simul.* 2001;69(1):89-108. [[Crossref](#)]
19. Royston P, Sauerbrei W. Fractional polynomials for one variable, MFP: multivariable model-building with fractional polynomials. In: Balding DJ, Cressie NAC, Fitzmaurice GM, Johnstone IM, Molenberghs G, Scott DW, et al, eds. *Multivariable Model-Building*. 1st ed. New York: John Wiley & Sons, Inc; 2008. p.84-5.
20. Royston P, Sauerbrei W. Building multivariable prognostic and diagnostic models: transformation of the predictors by using fractional polynomials. *Journal of the Royal Statistical Society Series A.* 1999;162(1):71-94. [[Crossref](#)]
21. Royston P, Sauerbrei W. Fractional polynomials for one variable, MFP: multivariable model-building with fractional polynomials. In: Balding DJ, Cressie NAC, Fitzmaurice GM, Johnstone IM, Molenberghs G, Scott DW, et al, eds. *Multivariable Model-Building*. 1st ed. New York: John Wiley & Sons, Inc; 2008. p.116-8.
22. Kay R. Treatment effects in competing-risks analysis of prostate cancer data. *Biometrics.* 1986;42(1):203-11. PMID: 3521753. [[Crossref](#)] [[PubMed](#)]
23. Schemper M, Heinze G. Probability imputation revisited for prognostic factor studies. *Stat Med.* 1997;16(1-3):73-80. PMID: 9004384. [[Crossref](#)]
24. Cheng SC, Fine JP, Wei LJ. Prediction of cumulative incidence function under the proportional hazards model. *Biometrics.* 1998;54(1):219-28. [[Crossref](#)] [[PubMed](#)]
25. Hunt L, Jorgensen M. Mixture model clustering using the multimix program. *Australia and New Zealand Journal of Statistics.* 1999;41(2):154-71. [[Crossref](#)]