



**T.C.
ADNAN MENDERES ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ
BİYOİSTATİSTİK ANABİLİM DALI
BİS-YL-2015-0002**

**SAĞKALIM ANALİZİNDE KANTİL REGRESYON VE
PARAMETRİK REGRESYON MODELLERİNİN
PERFORMANSLARININ KARŞILAŞTIRILMASI**

YÜKSEK LİSANS TEZİ

Can TÜRKİŞ

**DANIŞMAN
Doç. Dr. İmran KURT ÖMÜRLÜ**

AYDIN-2015

**T.C.
ADNAN MENDERES ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ
BİYOİSTATİSTİK ANABİLİM DALI
BİS-YL-2015-0002**

**SAĞKALIM ANALİZİNDE KANTİL REGRESYON VE
PARAMETRİK REGRESYON MODELLERİNİN
PERFORMANSLARININ KARŞILAŞTIRILMASI**

YÜKSEK LİSANS TEZİ

Can TÜRKİŞ

**DANIŞMAN
Doç. Dr. İmran KURT ÖMÜRLÜ**

AYDIN-2015

T.C.
ADNAN MENDERES ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜNE
AYDIN

Biyostatistik Anabilim Dalı Yüksek Lisans Programı öğrencisi Can TÜRKİŞ tarafından hazırlanan “Sağkalım Analizinde Kantil Regresyon ve Parametrik Regresyon Modellerinin Performanslarının Karşılaştırılması.” başlıklı tez, 30/07/2015 tarihinde yapılan savunma sonucunda aşağıda isimleri bulunan jüri üyelerince kabul edilmiştir.

Ünvanı, Adı ve Soyadı :

Üniversitesi :

İmzası:

1- Prof. Dr. Mevlüt TÜRE

Adnan Menderes Üniversitesi

2- Prof. Dr. Mehmet N. ORMAN

Ege Üniversitesi

3- Doç. Dr. İmran KURT ÖMÜRLÜ

Adnan Menderes Üniversitesi

Jüri üyeleri tarafından kabul edilen bu Yüksek Lisans Tezi Enstitü Yönetim Kurulunun 2015-19/I Sayılı kararıyla 12.08.2015 tarihinde onaylanmıştır.

Prof. Dr. Ahmet CEYLAN

Enstitü Müdürü

ÖNSÖZ

Sağkalım analizinde, bireylerin sağkalım sürelerini etkileyen faktörleri belirlemek için birçok yöntem bulunmaktadır. Parametrik Regresyon (PR) ve Kantil Regresyon (KR) modelleri de bu amaçla kullanılan yöntemlerdendir. PR modelleri, bireylerin sağkalım sürelerine ait dağılım koşuluna sahiptir. Genellikle kullanılan PR modelleri, üstel, Weibull, log-normal ve log-lojistik regresyon modelleridir. Eğer bireylerin sağkalım süreleri; üstel dağılıma sahip ise üstel regresyon modelini, Weibull dağılımına sahip ise Weibull regresyon modelini, log-normal dağılımına sahip ise log-normal regresyon modelini, log-lojistik dağılımına uygun ise log-lojistik regresyon modelini kullanmak uygun olacaktır. KR, klasik regresyon analizinin varsayımlarından olan hata terimlere ilişkin normal dağılıma uygunluk varsayımını ihmal eden robust bir regresyon tekniği olarak ortaya çıkmıştır. Koenker ve Basett (1978) tarafından önerilen kantil regresyon modeli, sağkalım analizinde yaşam sürelerini etkileyen faktörleri belirlemek amacıyla da kullanılan yöntemlerden biridir. Heterojen yapıdaki veri setleri için oldukça kullanışlıdır. Bireylerin yaşam sürelerini etkileyen faktörler, farklı kantil değerleri ele alınarak belirlenebilir. Her bir kantil değeri için hatayı minimum yapan parametre tahminlerini belirlemeye çalışır. Herhangi bir varsayıma sahip olmadığı için esnek bir yöntemdir.

Yapılan bu çalışmada KR yöntemi ile PR yöntemlerinin performansları, hem 1000 tekrarlı simülasyon çalışmasında hem de homojen ve heterojen yapıdaki gerçek sağkalım verileri üzerinde karşılaştırılmış ve minimum hataya sahip olan model belirlenmiştir.

İÇİNDEKİLER

| | Sayfa |
|---------------------------------------|--------------|
| KABUL VE ONAY | i |
| ÖNSÖZ | ii |
| İÇİNDEKİLER | iii |
| SİMGELER VE KISALTMALAR | v |
| ÇİZELGELER | vi |
| ŞEKİLLER | vii |
| EKLER | ix |
| 1. GİRİŞ | 1 |
| 1.1. Genel Bilgiler | 2 |
| 1.1.1. Sağkalım Analizi | 2 |
| 1.1.2. Sansürlü Veri Türleri | 3 |
| 1.1.2.1. Sağdan Sansürlü Veri | 3 |
| 1.1.2.2. Soldan Sansürlü Veri | 4 |
| 1.1.2.3. Aralıklı Sansürlü Veri | 5 |
| 1.1.3. Sağkalım Zamanı Fonksiyonları | 5 |
| 1.1.3.1. Sağkalım Fonksiyonu | 6 |
| 1.1.3.2. Olasılık Yoğunluk Fonksiyonu | 6 |
| 1.1.3.3. Hazard Fonksiyonu | 7 |
| 1.2. Parametrik Regresyon Modelleri | 8 |
| 1.2.1. Üstel Regresyon Modeli | 10 |
| 1.2.2. Weibull Regresyon Modeli | 11 |
| 1.2.3. Log-Normal Regresyon Modeli | 12 |
| 1.2.4. Log-Lojistik Regresyon Modeli | 14 |
| 1.3. Kantil Regresyon | 16 |

| | |
|---|----|
| 1.3.1. Kantil Kavramı | 16 |
| 1.3.2. Kantil Dağılım Fonksiyonu | 17 |
| 1.3.3. Kantil Yoğunluk Fonksiyonu | 18 |
| 1.3.4. Kantil Regresyon Modeli ve Tahmini | 18 |
| 1.3.5. Kantil Regresyonun Özellikleri | 19 |
| 1.3.6. Sağkalım Analizinde Kantil Regresyon | 20 |
| 1.3.7. Portnoy Tahmin Edicisi ile Parametre Tahmini | 21 |
| 1.4. Model Değerlendirme Ölçütleri | 22 |
| 1.4.1. Hata Kareler Ortalamasının Karekökü | 22 |
| 1.4.2. Ortalama Mutlak Hata | 22 |
| 2. GEREÇ VE YÖNTEM | 23 |
| 2.1. Kullanılan Programlar | 23 |
| 2.2. Simülasyon Uygulaması | 23 |
| 2.2.1. Algoritma | 23 |
| 2.3. Gerçek Veri Seti Uygulaması | 25 |
| 2.3.1. Meme Kanseri Veri Seti | 27 |
| 2.3.2. Gbcs Veri Seti | 28 |
| 3. BULGULAR | 29 |
| 3.1. Simülasyon Sonuçları | 29 |
| 3.2. Gerçek Uygulama Sonuçları | 33 |
| 3.2.1. Meme Kanseri Veri Seti | 33 |
| 3.2.2. Gbcs Veri Seti | 43 |
| 4. TARTIŞMA | 55 |
| 5.SONUÇ | 59 |
| ÖZET | 61 |
| SUMMARY | 62 |
| KAYNAKLAR | 63 |
| ÖZGEÇMİŞ | 67 |
| TEŞEKKÜR | 68 |

SİMGELER VE KISALTMALAR

| | | |
|--------------|---|-------------------------------------|
| n | : | Birim sayısı |
| t | : | Sağkalım süresi |
| τ | : | Kantil değeri |
| δ_i | : | Sansür durumu |
| $w_i(\tau)$ | : | Ağırlıklandırılmış τ değeri |
| $S(t)$ | : | Sağkalım fonksiyonu |
| $\hat{S}(t)$ | : | Sağkalım fonksiyonunun tahmini |
| $f(t)$ | : | Olasılık yoğunluk fonksiyonu |
| $h(t)$ | : | Hazard fonksiyonu |
| $h_0(t)$ | : | Temel hazard fonksiyonu |
| $Q(\tau)$ | : | Kantil dağılım fonksiyonu |
| $q(\tau)$ | : | Kantil yoğunluk fonksiyonu |
| KR | : | Kantil regresyon |
| PR | : | Parametrik regresyon |
| HO | : | Hazard oranı |
| KRM | : | Kantil regresyon modeli |
| HKOK: | : | Hata kareler ortalamasının karekökü |
| OMH | : | Ortalama mutlak hata |
| K-S | : | Kolmogorov-Smirnov |

ÇİZELGELER

| | Sayfa |
|---|--------------|
| Çizelge 1.1. Parametrik dağılımlar için temel hazard, olasılık yoğunluk ve sağkalım fonksiyonları | 16 |
| Çizelge 2.1. Meme kanseri hastalarına ilişkin nüksetme durumuna göre bağımsız değişkenlerin tanımlayıcı istatistikleri | 27 |
| Çizelge 2.2. Gbcs veri setindeki meme kanseri hastalarına ilişkin ölüm olayına göre bağımsız değişkenlerin tanımlayıcı istatistikleri | 28 |
| Çizelge 3.1. Değişen n değerlerine göre modellerin HKOK ve OMH değerleri | 29 |
| Çizelge 3.2. Meme kanseri veri seti için K-S test sonuçları | 33 |
| Çizelge 3.3. KR'nin farklı τ değerlerine karşılık gelen HKOK ve OMH değerleri | 37 |
| Çizelge 3.4. Meme kanseri veri seti için üstel regresyon sonuçları | 39 |
| Çizelge 3.5. Meme kanseri veri seti için Weibull regresyon sonuçları | 39 |
| Çizelge 3.6. Meme kanseri veri seti için log-normal regresyon sonuçları | 40 |
| Çizelge 3.7. Meme kanseri veri seti için log-lojistik regresyon sonuçları | 40 |
| Çizelge 3.8 Meme kanseri veri seti için $\tau=0,12$ değerindeki KR sonuçları | 41 |
| Çizelge 3.9. Modellerin HKOK ve OMH değerleri | 42 |
| Çizelge 3.10. Gbcs veri setinin K-S test sonuçları | 43 |
| Çizelge 3.11. KR'nin farklı τ değerlerine karşılık gelen HKOK ve OMH değerleri | 49 |
| Çizelge 3.12. Gbcs veri seti için üstel regresyon sonuçları | 51 |
| Çizelge 3.13. Gbcs veri seti için Weibull regresyon sonuçları | 51 |
| Çizelge 3.14. Gbcs veri seti için log-normal regresyon sonuçları | 52 |
| Çizelge 3.15. Gbcs veri seti için log-lojistik regresyon sonuçları | 52 |
| Çizelge 3.16. Gbcs veri seti için $\tau=0,45$ değerindeki KR sonuçları | 53 |
| Çizelge 3.17. Modellerin HKOK ve OMH değerleri | 54 |

ŞEKİLLER

| | Sayfa |
|---|--------------|
| Şekil 1.1. Sağdan sansürlü veri örneği | 4 |
| Şekil 1.2. Soldan sansürlü veri örneği | 4 |
| Şekil 1.3. Aralıklı sansürlü veri örneği | 5 |
| Şekil 1.4. Olasılık yoğunluk fonksiyonunun farklı iki gösterimi | 7 |
| Şekil 1.5. Üstel dağılımının λ parametresinin farklı değerlerindeki olasılık yoğunluk fonksiyon grafiği | 10 |
| Şekil 1.6. Weibull dağılımının $\lambda=1$ ve farklı α değerlerine göre olasılık yoğunluk fonksiyon grafiği | 11 |
| Şekil 1.7. Log-normal dağılımının $\mu=1$ ve farklı σ değerlerine göre olasılık yoğunluk fonksiyon grafiği | 13 |
| Şekil 1.8. Log-lojistik dağılımının $\alpha=0.2$ ve farklı λ değerlerine göre olasılık yoğunluk fonksiyon grafiği | 15 |
| Şekil 1.9. Kantil dağılım fonksiyon grafiği | 18 |
| Şekil 1.10. 100 birimlik sağkalım veri seti için P-P grafiği | 20 |
| Şekil 2.1. Simülasyon uygulaması için hazırlanan programın akış şeması | 25 |
| Şekil 2.2. Gerçek veri seti uygulaması için hazırlanan programın akış şeması | 26 |
| Şekil 3.1. $n=100$ olduğunda HKOK ve OMH değerlerinin modellere göre değişimi | 30 |
| Şekil 3.2. $n=250$ olduğunda HKOK ve OMH değerlerinin modellere göre değişimi | 30 |
| Şekil 3.3. $n=500$ olduğunda HKOK ve OMH değerlerinin modellere göre değişimi | 31 |
| Şekil 3.4. $n=750$ olduğunda HKOK ve OMH değerlerinin modellere göre değişimi | 31 |
| Şekil 3.5. $n=1000$ olduğunda HKOK ve OMH değerlerinin modellere göre değişimi | 32 |
| Şekil 3.6. Meme kanseri tanısı konan hastaların hastalıksız sağkalım süreleri için P-P grafiği | 33 |
| Şekil 3.7. Farklı τ değerleri için parametre tahminlerinin grafiksel gösterimi | 35 |

| | |
|---|----|
| Şekil 3.8. HKOK değerinin τ değerlerine göre değişim grafiği | 38 |
| Şekil 3.9. OMH değerinin τ değerlerine göre değişim grafiği | 38 |
| Şekil 3.10. HKOK ve OMH değerlerinin modellere göre değişimi | 42 |
| Şekil 3.11. Gbcs veri setindeki meme kanseri tanısı konan hastaların sağkalım süreleri için P-P grafiği | 43 |
| Şekil 3.12. Farklı τ değerleri için parametre tahminlerinin grafiksel gösterimi | 44 |
| Şekil 3.13. HKOK değerinin τ değerlerine göre değişim grafiği | 50 |
| Şekil 3.14. OMH değerinin τ değerlerine göre değişim grafiği | 50 |
| Şekil 3.15. HKOK ve OMH değerlerinin modellere göre değişimi | 54 |

EKLER

| | Sayfa |
|---|--------------|
| Ek 1. Meme kanseri veri setine ait KR'nin farklı τ deęerleri için parametre tahminleri | 69 |
| Ek 2. Gbcs veri setine ait KR'nin farklı τ deęerleri için parametre tahminleri | 77 |

1. GİRİŞ

Sağkalım analizi; çeşitli konularda risk etkenlerini, hastalığın seyrini etkileyen faktörleri, tedavi başarılarını ortaya koyabilmesi açısından önem taşıyan ve özellikle hasta bireylerin sağkalım sürelerine etki eden faktörleri belirlemek amacıyla sağlık alanında sıklıkla kullanılan yöntemlerdir (Ata ve ark 2008; Çelik 2011; Özdamar 2003; Şenocak 1998).

Sağkalım süresini etkileyen faktörleri belirlemek amacıyla kullanılan parametrik regresyon (PR) modelleri, sağkalım analizinde en yaygın kullanılan regresyon yöntemleridir. Sağkalım analizinde kullanılan regresyon modellerinin diğer regresyon modellerinden temel farkı, sansürlü veriyi de analize dahil etmeleridir. Sağkalım süreleri için belirli bir olasılık dağılımı varsayımı geçerli olduğunda, bu varsayıma dayalı tahminler daha kesindir. Parametre tahminlerinin standart hataları daha küçüktür. Sağkalım sürelerinin dağılımına göre genellikle kullanılan PR modelleri üstel, Weibull, log-normal ve log-lojistik modelleridir (Collett 1994; Efron 1977; Nardi ve Schemper 2003). Araştırmacıların genellikle karşılaştıkları temel problem, sağkalım süresinin dağılımının belirlenememesidir. Koenker ve Bassett (1978) tarafından geliştirilen ve Portnoy (2003) tarafından sağkalım analizine uyarlanan kantil regresyon (KR) yönteminde, sağkalım süresini etkileyen faktörleri belirlemek için herhangi bir dağılım varsayımı gerekmemektedir. KR'nin, kantillere bağlı olarak regresyon katsayılarını tahmin ettiği ve sapan değerlerin bulunduğu heterojen (sağkalım süresinin herhangi bir dağılıma uygun olmaması) yapıdaki veri setlerinde daha iyi tahminler verdiği bilinmektedir (Koenker ve Bassett 1978; Leng ve Tong 2013).

Son yıllarda araştırmacılar, KR modelinin yaşam analizinde kullanılmasına yönelik çalışmalar yapmışlardır. Lin ve Rodriguez (2013), primer biliyer siroz veri setini kullanarak 418 hastanın sağkalım süresi (gün) üzerine etki eden; yaş (yıl), albümin (g/dl), bilirubin (mg/dl), ödem durumu ve protrombin zamanı değişkenlerini ele alarak koşullu sağkalım fonksiyonlarının tahminlerinden ve sağkalım analizinde KR prosedüründe sansürlenmiş veriler için analiz metotlarından yararlanmışlardır. Peng ve Huang (2008), türetilmiş veriler üzerinde log-linear model ile KR modelinin performanslarını karşılaştırmışlardır. Kantillerin, sağkalım sürelerini karakterize etmede ortalamaya dayalı yöntemlere göre daha esnek ve güçlü bir yöntem olduğunu rapor etmişlerdir.

Bu çalışmanın amacı, homojen (sağkalım süresinin herhangi bir dağılıma uygun olması) ve heterojen yapıdaki veri setlerinde KR ile PR modellerinin sağkalım sürelerini tahmin etmedeki performanslarını karşılaştırmaktır. Modellerin performanslarının karşılaştırılması, hata kareler ortalamasının karekökü (HKOK) ve ortalama mutlak hata (OMH) kriterleri baz alınarak yapılacaktır. Bu amaçla, hem simülasyon uygulaması hem de gerçek veri setleri üzerinde sağkalım süresini minimum hata ile tahmin eden en uygun model belirlenecektir.

1.1. Genel Bilgiler

1.1.1. Sağkalım Analizi

Sağkalım süresi, tedaviye yanıt vermeyi ve hastalığın niteliğiyle ilişkili olan iyileşme süresini veya hastalığın gelişme süresini kapsar. Sağkalım analizi, hastalığa cevap verme olasılıklarını, hastalığa yakalanan kişinin aldığı tedavi türünün değişimine ve hastanın zamana bağlı olarak durumunda meydana gelen değişmelere neden olan değişkenlerin sağkalım süresi üzerine etkilerini, iyileşme veya ortalama yaşam süresini tahmin eder (Hosmer ve ark 2008; Kleinbaum ve Klein 2005; Lee ve Wang 2003; Özdamar 2003; Şenocak 1998).

Sağkalım analizinde temel amaç, gözlenen ölüm ya da yaşam süresi değişkenlerinin incelenmesidir. Dolayısıyla, bu değişkenin açık ve doğru bir şekilde tanımlanması gereklidir. İlgilenilen olaya göre farklı şekillerde ortaya çıkan bu değişken ölçülürken aşağıdaki kriterlere dikkat edilmelidir (Cox 1984; Lee ve Wang 2003):

- ✓ Başlangıç zamanı, her bir birey için şüpheye yer vermeyecek şekilde tanımlanmalıdır.
- ✓ Ölçülen zaman için ölçek olmalıdır (gün, ay, yıl gibi).
- ✓ Her bir birim için ölüm olayının gerçekleştiği zaman net olmalıdır.

Sağkalım analizinde, sağkalım sürelerinin tam olarak belirlenemediği durumlarda araştırmacılar sansür problemi ile karşılaşır. Sansürlü veri; zaman ve maliyet gibi birtakım sınırlamalar nedeni ile sağkalım süresinin kesin olarak bilinemediği, herhangi bir sebeple gözlenemediği verilerdir.

Bireylerin sağkalım sürelerinin sansürlü olarak kabul edilmesi için aşağıdaki durumlardan en az biri meydana gelmelidir:

Birey,

- Tedavi gördüğü süre içerisinde farklı bir sebepten dolayı ölmüş olabilir.
- Tedaviye başka bir hastanede veya başka bir şehirde devam etmek zorunda kalmış olabilir.
- Tedaviye cevap vermemiş olabilir.
- Tedavi süresi içerisinde başka bir hastalığa yakalanmış ve tedaviye ara vermiş olabilir (Çelik 2011; Özdamar 2003; Şenocak 1998; Yayla 2013).

Kleinbaum ve Klein (2005) göre, bireylerin sağkalım sürelerinin sansürlü olup olmadığına karar vermek için üç durum vardır.

- a) Birey gözlem esnasında ölebilir.
- b) Birey gözlemden çekilebilir. İlgilenilen olay dışında başka bir nedenden dolayı ölebilir veya uygulanan yöntemlerden istenmeyen sonuçlar alınabilir.
- c) Birey gözlem sonunda hala yaşıyor olabilir.

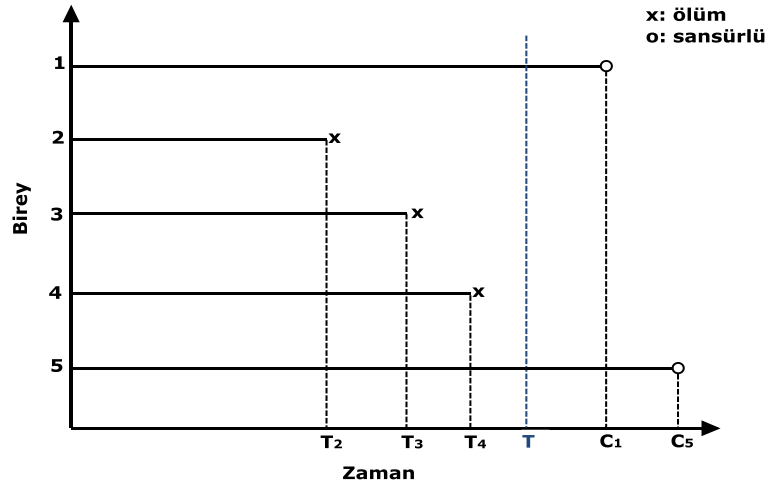
(a) durumunda, bireyin sağkalım süresi bilindiğinden sansürlü değildir. (b) durumunda, bireyin sağkalım süresi gözlemden çekilme zamanından itibaren sansürlüdür. (c) durumunda ise, bireyin sağkalım süresinin çalışmanın sonlandırılma zamanına kadar bilinmesine rağmen, gözlem sonrası hakkında herhangi bir bilgi olmadığından bu bireyin sağkalım süresi de sansürlüdür (Kleinbaum ve Klein 2005).

1.1.2. Sansürlü Veri Türleri

Sağkalım analizinde üç farklı sansürlü veri vardır. Bunlar sağdan sansürlü, soldan sansürlü ve aralıklı sansürlüdür.

1.1.2.1. Sağdan Sansürlü Veri

Başarısızlık olarak adlandırılan olay (ölüm), çalışma için belirlenen bir durma zamanına kadar gerçekleşmez ise, bireyin sağkalım süresinin uzunluğu çalışmanın durma zamanının sağ tarafına geçer (Şekil 1.1). Böyle bir durumda bireyin sağkalım süresi tam olarak bilinmeyecek ve birey gözleme alınamayacaktır. Yani bireyin sağkalım süresi sansürlenmiş olacaktır. Bu tip veriye “sağdan sansürlü veri” denir (Lawless 1983; Liu 2012).



Şekil 1.1. Sağdan sansürlü veri örneği

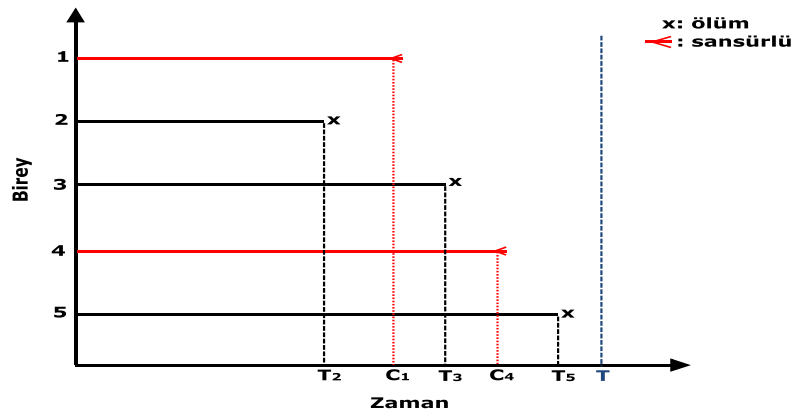
C_i ; sansürleme zamanı, T_i ; bireyin sağkalım süresi olmak üzere $T_i > C_i$ olduğunda bu bireyin sağkalım süresi sağdan sansürlüdür.

$$\delta_i = \begin{cases} 0, & T_i > C_i \\ 1, & T_i \leq C_i \end{cases} \quad i = 1, 2, \dots, n$$

Eğer $\delta_i=0$ ise birey sansürlüdür, $\delta_i=1$ ise bireyin sağkalım süresi gözlenmiştir (Nelson 1982).

1.1.2.2. Soldan Sansürlü Veri

Çalışmaya alınan bireylerin gerçek sağkalım süresi gözlemci tarafından belirlenmiş olan süreden daha az ise elde edilen veriler soldan sansürlüdür (Şekil 1.2). Soldan sansürlü verilerde sadece ölümün belirlenen bir zamandan önce olduğu bilinir (Lawless 1983; Liu 2012).



Şekil 1.2. Soldan sansürlü veri örneği

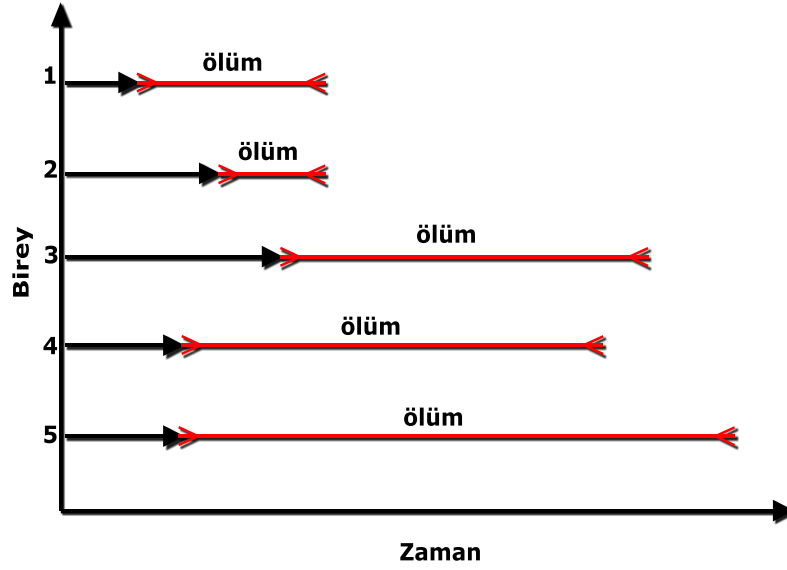
$T_i \leq C_i$ olduğunda bu bireyin sağkalım süresi soldan sansürlüdür.

$$\delta_i = \begin{cases} 0, & T_i \leq C_i \\ 1, & T_i > C_i \end{cases} \quad i = 1, 2, \dots, n$$

Eğer $\delta_i=0$ ise birey sansürlüdür, $\delta_i=1$ ise bireyin sağkalım süresi gözlenmiştir (Lawless 1983; Liu 2012).

1.1.2.3. Aralıklı Sansürlü Veri

Bireylerin belirli bir zaman aralığında başarısızlığa sahip olmaları olarak bilinir. Genellikle takip gerektiren olaylarda kullanılır. Çalışmaya konu olan olayın meydana gelme süresi, bir aralıkla ifade edilir (Şekil 1.3). Eğer aralıklı sansürlemesi, sağdan sansürlemenin genelleştirilmiş biçimi olarak ifade ediliyorsa, aralık $(0, C_i]$ olarak alınır. Eğer soldan sansürlemenin genelleştirilmiş biçimi olarak ifade ediliyorsa aralık $(C_i, \infty]$ olarak alınır (Nelson 1982).



Şekil 1.3. Aralıklı sansürlü veri örneği

1.1.3. Sağkalım Zamanı Fonksiyonları

Sağkalım zamanı, kesin bir olayda ölçülen verilerin zamanıdır ve T ile gösterilir. Başarısızlık, ölüm, hastalığın iyileşmesi vb. durumlar kesin olaylardır. Sağkalım verileri genellikle üç fonksiyon ile tanımlanır.

1. Sağkalım Fonksiyonu
2. Olasılık Yoğunluk Fonksiyonu
3. Hazard Fonksiyonu

Bu üç fonksiyon arasında matematiksel bir ilişki söz konusudur. Bu fonksiyonlardan herhangi biri verildiğinde diğer fonksiyonlar bulunabilir.

1.1.3.1. Sağkalım Fonksiyonu

Sağkalım fonksiyonu $S(t)$ olarak gösterilir. Bir bireyin t zamanından daha uzun sürede hayatta kalma olasılığını ifade eder.

$$\begin{aligned} S(t) &= P(\text{bireyin } t \text{ zamanından daha uzun yaşaması}) \\ &= P(T > t) \end{aligned}$$

t 'nin kümülatif olasılık yoğunluk fonksiyonu $F(t)$ ile gösterilirse sağkalım fonksiyonu;

$$\begin{aligned} S(t) &= 1 - P(\text{bireyin } t \text{ zamanından önce başarısız olması}) \\ &= 1 - F(t) \end{aligned}$$

$S(t)$, t zamanında sürekli, artmayan bir fonksiyondur.

$$S(t) = \begin{cases} 1, & t = 0 \\ 0, & t = \infty \end{cases}$$

$t=0$ zamanında sağkalım olasılığı 1'dir ve t sonsuza giderken sağkalım olasılığı 0'dır. Eğer sansürlü gözlemler bulunmuyor ise sağkalım fonksiyonu t 'den daha uzun hayatta kalan hastaların oranı olarak tahmin edilir (Kleinbaum ve Klein 2005; Lawless 1983; Lee ve Wang 2003):

$$\hat{S}(t) = \frac{\text{t zamanından daha uzun süre hayatta kalan hasta sayısı}}{\text{Toplam hasta sayısı}}$$

1.1.3.2. Olasılık Yoğunluk Fonksiyonu

Olasılık yoğunluk fonksiyonu ($f(t)$), $(t, t + \Delta t)$ aralığında başarısız olan bir kişinin olasılıksal limitidir veya kısa bir aralığa düşen başarısızlık oranını göstermektedir.

$$f(t) = \frac{\lim_{\Delta t \rightarrow 0} P((t, t + \Delta t) \text{ aralığında ölen bir birey})}{\Delta t}$$

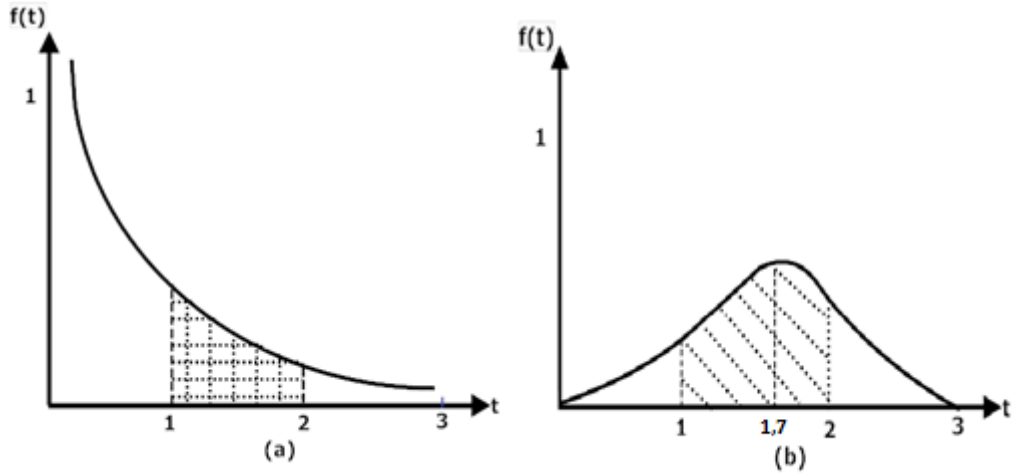
$f(t)$ fonksiyonunun iki temel özelliği vardır:

1. $f(t)$ negatif olmayan bir fonksiyondur. Tüm $t \geq 0$ için $f(t) \geq 0$, aksi durumda $f(t) = 0$ olmaktadır.
2. Yoğunluk eğrisiyle t eksenini arasında kalan alan 1'e eşittir. Eğer sansürlü veri yoksa olasılık yoğunluk fonksiyonu $\hat{f}(t)$ olarak tanımlanır:

$$\hat{f}(t) = \frac{t \text{ zamanıyla başlayan aralıkta ölen birey sayısı}}{(\text{toplam birey sayısı}) \times (\text{aralık genişliği})}$$

Herhangi bir zaman aralığına düşen bireylerin başarısızlık oranı ve başarısızlığın frekansları yoğunluk fonksiyonundan bulunabilir (Hosmer ve ark 2008; Kleinbaum ve Klein 2005; Lawless 1983; Lee ve Wang 2003).

Şekil 1.4.(a), bir bireyin çalışmanın başındaki yüksek başarısızlık oranını verir ve zaman arttıkça başarısızlık oranının azaldığını göstermektedir. Şekil 1.4.(b)'de ise, yüksek başarısızlık frekansı yaklaşık 1,7 olarak tespit edilmiştir. 1. ve 2. zaman biriminin arasına düşen kişilerin oranı, t eksenini ve olasılık yoğunluk fonksiyonu eğrisi altında kalan alana eşittir (Lee ve Wang 2003).



Şekil 1.4. Olasılık yoğunluk fonksiyonunun farklı iki gösterimi

1.1.3.3. Hazard Fonksiyonu

T sağkalım zamanının hazard fonksiyonu olan $h(t)$, koşullu başarısızlık oranını verir. Bu fonksiyon dar bir zaman aralığında devam eden başarısızlık olarak tanımlanır ve birey bu aralığın başında hayattadır.

$$h(t) = \frac{\lim_{\Delta t \rightarrow 0} P[t \text{ yaşındaki } (t, t + \Delta t) \text{ zaman aralığındaki başarısızlığı}]}{\Delta t}$$

Hazard fonksiyonu ayrıca olasılık yoğunluk ve sağkalım fonksiyonu ile de tanımlanabilir:

$$h(t) = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{S(t)}$$

Eğer sansürlü veri yok ise hazard fonksiyonu;

$$\hat{h}(t) = \frac{t \text{ zamanda başlayan } (t, t + \Delta t) \text{ zaman aralığında ölen birey sayısı}}{t \text{ zamanında sağkalan birey sayısı}}$$

şeklinde tanımlanır (Hosmer ve ark 2008; Kleinbaum ve Klein 2005; Lawless 1983).

Hazard fonksiyonu artabilir, azalabilir, sabit kalabilir veya daha karmaşık bir durum gösterebilir.

Kümülatif hazard fonksiyonu;

$$H(t) = \int_0^t h(x) dx \\ = -\log S(t)$$

şeklinde tanımlanır. Böylece $t=0$ olduğunda $S(t)=1$ ve $H(t)=0$, $t=\infty$ iken ise $S(t)=0$ ve $H(t)=\infty$ olacaktır (Lee ve Wang 2003).

Hazard fonksiyonunun incelenmesinin araştırmacıya sağladığı yararlar aşağıdaki gibi sıralanabilir (Cox 1984):

1. t yaşına kadar yaşamını sürdürdüğü bilinen bir bireyin var olan başarısızlık riski hazard fonksiyonu ile dikkate alınır.
2. Hazard fonksiyonu yardımı ile farklı gruplar kesin bir şekilde karşılaştırılabilir.
3. Hazard fonksiyonunu temel alan modellerin kullanılması, sansürleme veya farklı tipte başarısızlığın bulunduğu durumlarda daha uygundur.
4. Başarısızlık zamanlarının bir üstel dağılıma sahip olduğu varsayılarak, farklı grupların karşılaştırılması hazard fonksiyonu ile daha kolay yapılır.
5. Hazard “tek başarısızlık” içeren sistemler için özel bir form oluşturur.

1.2. Parametrik Regresyon Modelleri

PR modellerinde, sağkalım süresine ait dağılımın biliniyor olmasından dolayı olasılık yoğunluk, hazard ve sağkalım fonksiyonları kesin olarak tanımlanmaktadır. Bu fonksiyonların tanımlanması, tahmin edilen sağkalım eğrisinin tüm veriler hakkında bilgi vermesini ve parametre tahmini yapılmasını sağlar (Avcı ve Bekiroğlu 2014).

PR modelleri ile en çok olabilirlik tahmin yöntemi kullanılarak parametrelerin tahmini gerçekleştirilir. Modellerden elde edilen parametre tahminleri kullanılarak minimum hata ile sağkalım süreleri tahmin edilir. Hatalar, gözlenen ve tahmin edilen zaman değerleri arasındaki fark olarak hesaplanır (Hosmer ve ark 2008).

Sağkalım süresini etkileyen bağımsız değişkenleri PR modelleri ile modellemek için iki yaklaşım vardır:

Birinci yaklaşım, klasik doğrusal regresyon yaklaşımına benzemektedir. Bu yaklaşımda, sağkalım süresinin doğal logaritması ($Y=\ln(T)$) modellenir. $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_p)$ olmak üzere, Y için doğrusal model,

$$Y = \mu + \mathbf{X}\boldsymbol{\beta}' + \varepsilon$$

biçimindedir. Burada $\boldsymbol{\beta}' = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_p)$ regresyon katsayılar vektörü ve ε hata terimidir. Hata teriminin dağılımı için en sık kullanılan dağılımlar, standart normal dağılım, uç değer dağılımı ve lojistik dağılımdır. Bu yaklaşım için kullanılan PR modelleri, hızlandırılmış başarısızlık zamanı (Accelerated Failure Time, AFT) modelleri olarak da adlandırılmaktadır. Ancak bu yaklaşımda hata terimi için dağılım varsayımı gerektiğinden birinci yaklaşımın kullanım alanları sınırlıdır. Bu durumda ikinci yaklaşımın kullanılması daha uygun olur.

Bağımsız değişkenlerin sağkalım fonksiyonu üzerindeki etkilerini modellemek için kullanılan ikinci yaklaşımda, bağımsız değişkenlerin bir fonksiyonu olan koşullu hazard fonksiyonu kullanılır. Koşullu hazard fonksiyonu:

$$h(t|\mathbf{x})=h_0(t)\exp(\mathbf{x}'\boldsymbol{\beta})$$

biçiminde tanımlanır (Avcı ve Bekiroğlu 2014; Hosmer ve ark 2008; London 1997).

PR modeli için koşullu hazard fonksiyonu belirlenirken, sağkalım süresi hangi dağılıma uygun ise o dağılıma ait temel hazard fonksiyonu kullanılmalıdır. En sık kullanılan PR modelleri üstel, Weibull, log-normal ve log-lojistik dağılımlarına dayanmaktadır.

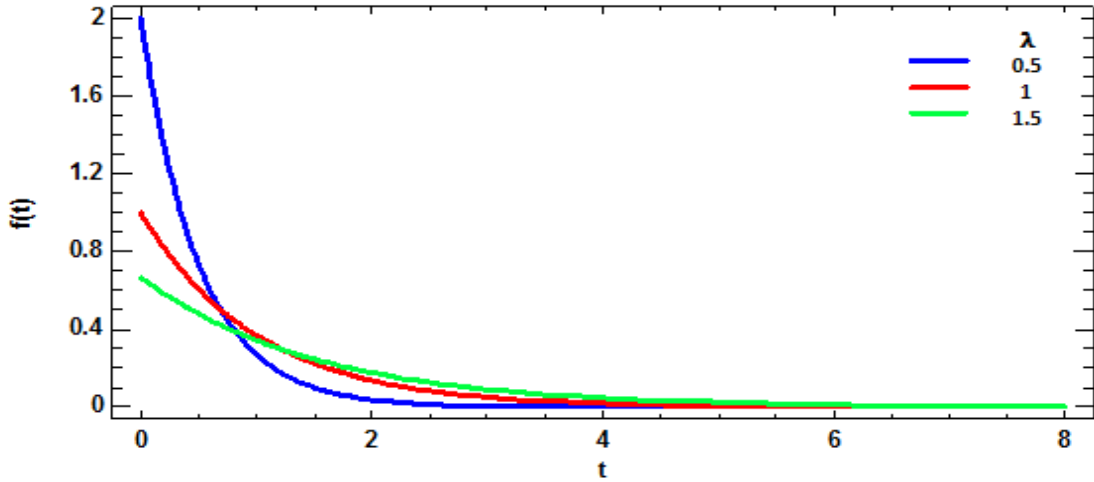
1.2.1. Üstel Regresyon Modeli

Üstel regresyon modeli, sağkalım analizinde kullanılan en temel ve en basit PR modelidir (Kleinbaum ve Klein 2005).

Üstel dağılımın tek parametresi, λ ölçek parametresidir. Bu parametre sabit hazard oranı olarak tanımlanır. Yüksek λ değerleri, yüksek risk ve kısa sağkalım süresini verirken; düşük λ değerleri, düşük risk ve uzun sağkalım süresini verir (Hosmer ve ark 2008).

Üstel regresyon modelinde; $t = (t_1, t_2, \dots, t_n)'$ sağkalım sürelerinin bağımsız, homojen olduğu ve her birinin λ parametresi ile üstel dağılıma sahip olduğu varsayılır. Olasılık yoğunluk fonksiyonu (Şekil 1.5);

$$f(t_i|\lambda) = \lambda \exp(-\lambda t_i)$$



Şekil 1.5. Üstel dağılımının λ parametresinin farklı değerlerindeki olasılık yoğunluk fonksiyon grafiği

Hazard fonksiyonu;

$$h(t) = \lambda$$

Sağkalım fonksiyonu ise ;

$$S(t_i|\lambda) = \exp(-\lambda t_i)$$

şeklinde tanımlanır (Hosmer ve ark 2008; Kleinbaum ve Klein 2005).

λ parametresi için üstel sağkalım modelinin olabilirlik fonksiyonu;

$$L(\lambda|t) = \prod_{i=1}^n f(t_i|\lambda)^{v_i} S(t_i|\lambda)^{1-v_i}$$

$$= \lambda \sum_{i=1}^n v_i \exp(-\lambda \sum_{i=1}^n t_i)$$

şeklinde tanımlanır. Burada v_i , gösterge değişkenidir. Eğer $t_i > T_i$ ise, i . veri sansürlenmiş veri;

$$v_i = \begin{cases} 1, & t_i \leq T_i \\ 0, & t_i > T_i \end{cases}$$

olarak tanımlanır (İbrahim ve ark 2001).

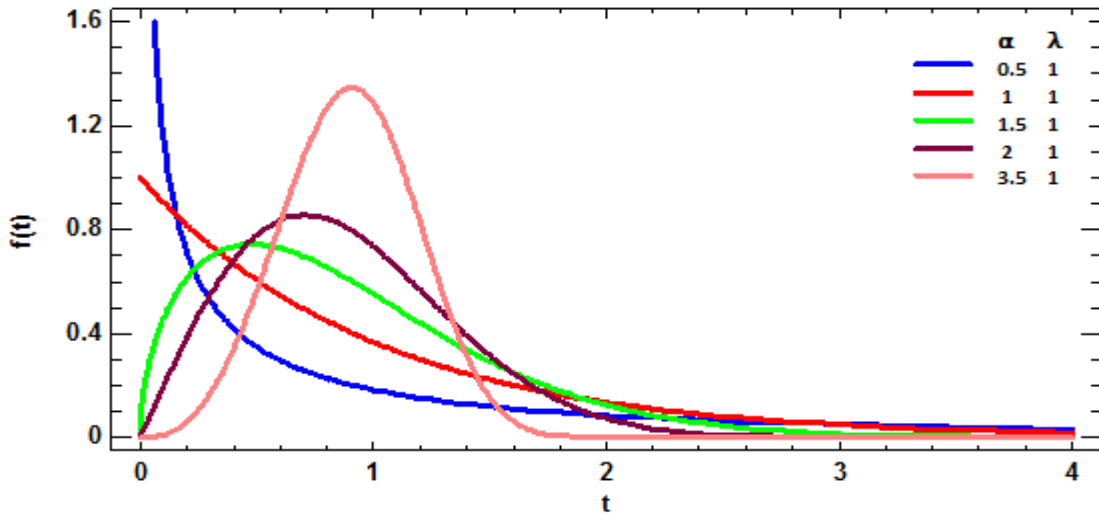
1.2.2. Weibull Regresyon Modeli

Weibull regresyon modeli, sağkalım verilerinin analizinde sık kullanılan parametrik sağkalım modellerinden biridir. Weibull regresyon modeli, üstel regresyon modelinin geliştirilmiş halidir. α şekil parametresi, λ ölçek parametresi olmak üzere iki parametreye sahiptir (Hosmer ve ark 2008).

Weibull dağılımının olasılık yoğunluk fonksiyonu;

$$f(t|\lambda, \alpha) = \lambda \alpha t^{\alpha-1} \exp(-\lambda t^\alpha); \quad \lambda > 0, \alpha > 0$$

şeklinde tanımlanır. $\lambda=1$ ve α parametresinin farklı değerlerinde olasılık yoğunluk fonksiyonunun değişimi Şekil 1.6'da verilmiştir.



Şekil 1.6. Weibull dağılımının $\lambda=1$ ve farklı α değerlerine göre olasılık yoğunluk fonksiyon grafiği

Weibull sağkalım modelinin hazard fonksiyonu;

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

α , hazard fonksiyonunun şeklini belirler. Eğer $\alpha > 1$ ise, sağkalım süresi arttıkça hazard oranı da artar; $\alpha = 1$ ise, hazard oranı sabittir; $\alpha < 1$ ise, hazard oranı zamanla azalır. α parametresinin dağılımda yer alması, Weibull regresyon modelini üstel regresyon modelinden daha esnek kılar (Kleinbaum ve Klein 2005).

Weibull sağkalım fonksiyonu;

$$S(t|\lambda, \alpha) = \exp(-\lambda t^\alpha)$$

olarak ifade edilir (Hosmer ve ark 2008; Ibrahim ve ark 2001).

λ ve α parametreleri için Weibull sağkalım modelinin olabilirlik fonksiyonu;

$$\begin{aligned} L(\lambda, \alpha|t) &= \prod_{i=1}^n f(t_i|\lambda, \alpha)^{v_i} S(t_i|\lambda, \alpha)^{1-v_i} \\ &= \lambda^{\sum_{i=1}^n v_i} \exp\left\{\alpha \left(\sum_{i=1}^n v_i\right) + \sum_{i=1}^n (v_i(\lambda - 1) \log(t_i) - t_i^\lambda \exp(\alpha))\right\} \end{aligned}$$

ve Weibull sağkalım modeli;

$$\log T_i = a_0 + \sum_{k=1}^p a_k X_{ki} + \sigma \varepsilon = \mu_i + \sigma \varepsilon_i$$

şeklinde tanımlanır. Burada T , $\lambda_i = \exp\left(-\frac{\mu_i}{\sigma}\right)$ ve $\alpha = \frac{1}{\sigma}$ parametrelili Weibull dağılımına sahiptir (Lee ve Wang 2003).

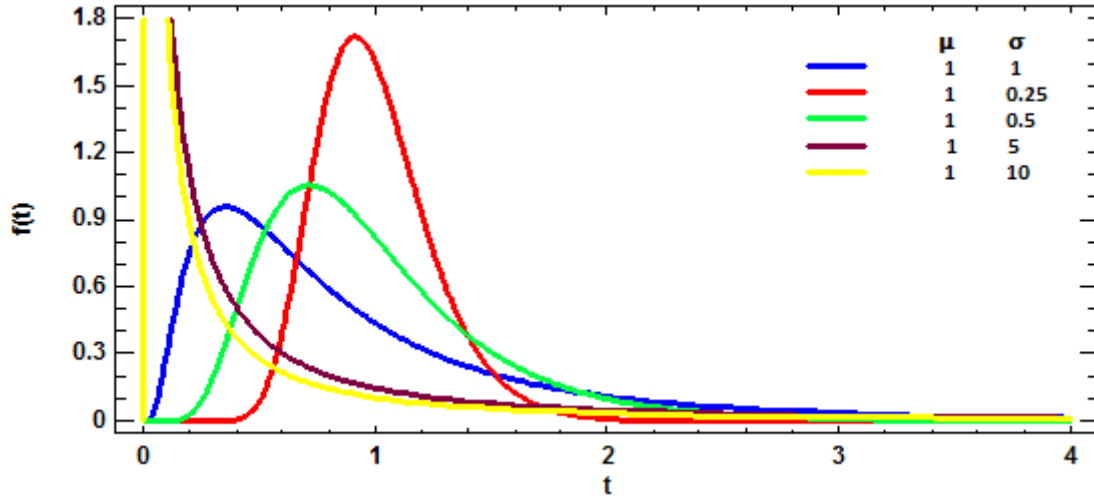
1.2.3. Log-Normal Regresyon Modeli

Log-normal regresyon modelinde, sağkalım sürelerinin log-normal dağıldığı varsayılır (Ibrahim ve ark 2001).

$\log_e T$, μ ortalamalı ve σ^2 varyanslı normal dağılıma sahiptir. T değişkeni, logaritmik normal dağılıma sahiptir ve $T \sim \Lambda(\mu, \sigma^2)$ şeklinde gösterilir. Hazard fonksiyonu başta hızla artar, medyanı geçtikten sonra sonsuza giderken 0'a yaklaşır. Yani logaritmik normal dağılım, ilk başta artan daha sonra azalan hazard oranına sahip sağkalım süreleri için daha uygundur (Hosmer ve ark 2008; Kleinbaum ve Klein 2005).

Log-normal dağılımının olasılık yoğunluk fonksiyonu (Şekil 1.7);

$$f(t_i|\mu, \sigma) = \frac{1}{t_i\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2} (\log_e t_i - \mu)^2\right\}$$



Şekil 1.7. Log-normal dağılımının $\mu=1$ ve farklı σ değerlerine göre olasılık yoğunluk fonksiyon grafiği

Hazard fonksiyonu;

$$h(t) = \frac{f(t)}{S(t)} = \frac{\frac{1}{t_i\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2} (\log_e at_i)^2\right\}}{1 - \Phi\left(\frac{\log_e at_i}{\sigma}\right)}$$

şeklinde tanımlanır (Hosmer ve ark 2008; Lee ve Wang 2003).

Sağkalım fonksiyonu;

$$S(t_i|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_t^{\infty} \frac{1}{x} \exp\left[-\frac{1}{2\sigma^2} (\log_e x - \mu)^2\right] dx$$

şeklinde tanımlanır.

$a = \exp(-\mu)$ dönüşümü yapılarak ve μ yerine $-\mu = \log_e a$ ifadesi yukarıdaki eşitlikte yazılarak;

$$f(t_i|\mu, \sigma) = \frac{1}{t_i\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2} (\log_e at_i)^2\right\}$$

$$S(t_i|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_t^{\infty} \exp\left[-\frac{1}{2\sigma^2} (\log_e ax)^2\right] \frac{dx}{x}$$

$$S(t_i|\mu, \sigma) = 1 - \Phi\left(\frac{\log_e at_i}{\sigma}\right)$$

eşitlikleri elde edilir.

μ ve σ parametreleri için Log-normal sağkalım modelinin olabirlilik fonksiyonu;

$$\begin{aligned} L(\mu, \sigma|t) &= \prod_{i=1}^n f(t_i|\mu, \sigma)^{v_i} S(t_i|\mu, \sigma)^{1-v_i} \\ &= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^{\sum_{i=1}^n v_i} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n v_i (\log t_i - \mu)^2\right\} \prod_{i=1}^n \left(\frac{1}{t_i}\right)^{v_i} \left(1 - \Phi\left(\frac{\log t_i - \mu}{\sigma}\right)\right)^{1-v_i} \end{aligned}$$

şeklinde tanımlanır (Ibrahim ve ark 2001).

Log-normal sağkalım modeli;

$$\log T_i = a_0 + \sum_{k=1}^p a_k X_{ki} + \sigma \varepsilon_i = \mu_i + \sigma \varepsilon_i$$

şeklinde oluşturulur (Hosmer ve ark 2008; Lee ve Wang 2003).

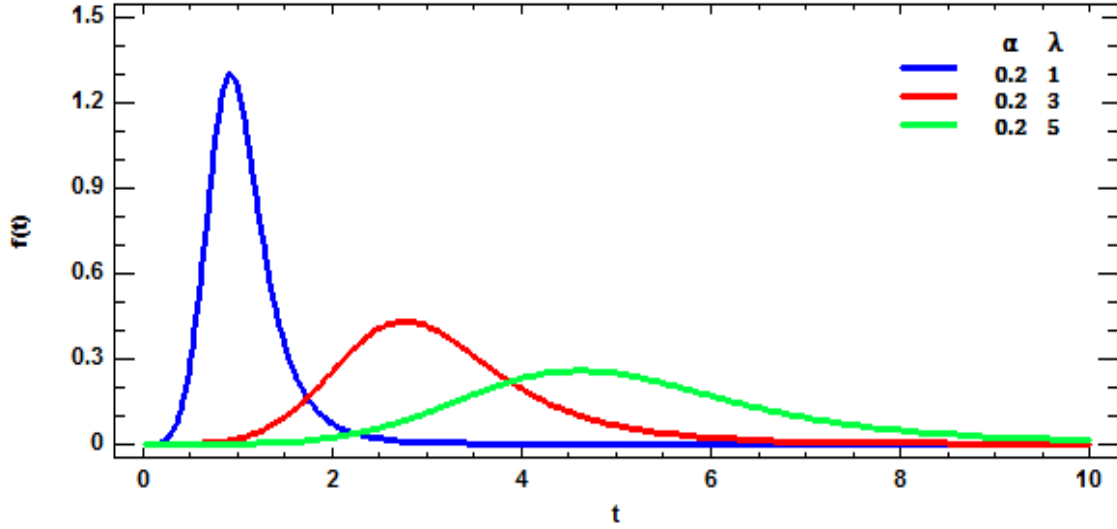
1.2.4. Log-Lojistik Regresyon Modeli

Log-lojistik regresyon modeli, sağkalım süresi değişkeninin log-lojistik dağılıma uygun olduğu durumda kullanılır. Eğer $\log(T)$ lojistik dağılıma sahip ise T sağkalım süresi log-lojistik dağılır. Log-lojistik dağılımının, λ ölçek ve α şekil parametresi olmak üzere iki parametresi vardır (Kleinbaum ve Klein 2005).

Log-lojistik dağılımının olasılık yoğunluk fonksiyonu (Şekil 1.8);

$$f(t_i|\lambda, \alpha) = \frac{\alpha t^{\alpha-1} \lambda}{[1 + \lambda t^\alpha]^2}$$

şeklinde tanımlanır.



Şekil 1.8. Log-lojistik dağılımının $\alpha=0.2$ ve farklı λ değerlerine göre olasılık yoğunluk fonksiyon grafiği

Hazard fonksiyonu;

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

şeklinde tanımlanır. Burada $\alpha \leq 1$ ise hazard oranı sifira yaklaşır. Eğer $\alpha > 1$ ise hazard oranı maksimum noktaya ulaşır daha sonra sifira yaklaşır (Hosmer ve ark 2008; Lee ve Wang 2003).

Sağkalım fonksiyonu;

$$S(t) = \frac{1}{1 + \lambda t^\alpha} ; \quad \alpha > 0, \quad \lambda > 0$$

şeklinde, λ ve α parametreleri için log-lojistik sağkalım modelinin olabilirlik fonksiyonu ise;

$$\begin{aligned} L(\lambda, \alpha | t) &= \prod_{i=1}^n f(t_i | \lambda, \alpha)^{v_i} S(t_i | \lambda, \alpha)^{1-v_i} \\ &= \prod_{i=1}^n \left[\frac{\alpha t^{\alpha-1} \lambda}{[1 + \lambda t^\alpha]^2} \right]^{v_i} \left[\frac{1}{1 + \lambda t^\alpha} \right]^{1-v_i} \end{aligned}$$

şeklinde tanımlanır (Hosmer ve ark 2008; Ibrahim ve ark 2001).

PR modellerine ait temel hazard fonksiyonları $h_0(t)$, olasılık yoğunluk fonksiyonları $f(t)$ ve sağkalım fonksiyonları $S(t)$ Çizelge 1.1’de verilmiştir.

Çizelge 1.1. Parametrik dağılımlar için temel hazard, olasılık yoğunluk ve sağkalım fonksiyonları

| Dağılım | Parametre | $h_0(t)$ | $f(t)$ | $S(t)$ |
|--------------|-------------------------------|---|--|--|
| Üstel | $\lambda > 0$ | λ | $\lambda \exp(-\lambda t)$ | $\exp(-\lambda t)$ |
| Weibull | $\alpha > 0$ $\lambda > 0$ | $\lambda \alpha t^{\alpha-1}$ | $\lambda \alpha t^{\alpha-1} \exp(-\lambda t^\alpha)$ | $\exp(-\lambda t^\alpha)$ |
| Log-normal | $\mu \geq 0$ $\sigma > 0$ | $\frac{f(t)}{S(t)}$ | $\frac{\exp\left[-\frac{1}{2}\left(\frac{\ln(t) - \mu}{\sigma}\right)^2\right]}{t\sigma\sqrt{2\pi}}$ | $1 - \Phi\left[\frac{\ln(t) - \mu}{\sigma}\right]$ |
| Log-lojistik | $\alpha > 0$ $\lambda > 0$ | $\frac{\alpha t^{\alpha-1}\lambda}{1 + \lambda t^\alpha}$ | $\frac{\alpha t^{\alpha-1}\lambda}{[1 + \lambda t^\alpha]^2}$ | $\frac{1}{1 + \lambda t^\alpha}$ |

Φ : Standart normal dağılımın kümülatif dağılım fonksiyonu

1.3. Kantil Regresyon

KR, koşullu kantil fonksiyonların modelini tahmin etmek amacıyla geliştirilen ve diğer regresyon yöntemlerine göre sağkalım süresi hakkında daha ayrıntılı bilgi veren istatistiksel bir yöntemdir (Koenker ve Bassett 1978). Günümüzde klasik en küçük kareler yöntemi, koşullu ortalamaya bağlı iken; KR yöntemi ise koşullu kantil fonksiyonuna bağlıdır (Koenker 2005).

Değişkenler arasındaki ilişkiyi açıklamak için PR modelleri kullanıldığında, modelin fonksiyonel şeklinin doğru seçilmesi, verinin homojen yapıda olması, sapan değerlerin olmaması önemlidir. Bu nedenlerden dolayı PR modelleri, sağkalım süresini etkileyen faktörlerin belirlenmesinde kısıtlılıklara sahiptir. Bu kısıtlılıkları gidermek için daha esnek bir yöntem olan KR modelini kullanmak daha uygundur.

1.3.1. Kantil Kavramı

Kantil değeri, değişkenin dağılımında yer alan ve dağılımı kendisinden büyük olanlar ve kendisinden küçük olanlar diye ikiye bölen herhangi bir değerdir. Bir seriye ait gözlem değerleri küçükten büyüğe doğru sıralandığında serinin toplam frekansını 2, 4, 10 ve 100 eşit parçaya ayıran değerler genel olarak kantil (eşbölen) olarak adlandırılmaktadır. Seriyi iki eşit parçaya bölmek için hesaplanan değerlere medyan, dört eşit parçaya bölmek için hesaplanan değerlere kartil, on eşit parçaya bölmek için hesaplanan değerlere desil, yüz eşit parçaya bölmek için hesaplanan değerlere santil adı verilmektedir (Turanlı ve

Gürüş 2005). Burada r bir tamsayıyı göstermektedir ve h , 1 ile $(r-1)$ arasında bütün tamsayıları alabilen değerlerdir. r 'nin çeşitli değerleri için kantil grupları Çizelge 1.2'de gösterilmiştir.

Çizelge 1.2. Çeşitli r değerlerine göre kantil grupları

| r | $Q_{\frac{h}{r}}$ | Kantil Grupları |
|-----|---------------------|-----------------|
| 2 | $Q_{\frac{h}{2}}$ | Medyan |
| 4 | $Q_{\frac{h}{4}}$ | Dörttebirlikler |
| 10 | $Q_{\frac{h}{10}}$ | Ondabirlikler |
| 100 | $Q_{\frac{h}{100}}$ | Yüzdebirlikler |

h tamsayısının r 'ye bağlı olarak aldığı değerler, çeşitli kantil gruplarındaki kantillerin isimlendirilmesini (yedinci ondabirlik, onüçüncü yüzdebirlik gibi) sağladığı gibi, alabileceği farklı değerler de her kantil grubu için hesaplanabilecek ortalama sayısını gösterir. Örneğin $r=10$ için h , 1'den 9'a kadar dokuz değer alabildiğine göre ondabirlik grup için 9 tane kantil hesaplanır. Dağılım için yalnızca bir ortanca ($Q_{1/2}$), üç tane dörttebirlik ($Q_{1/4}, Q_{2/4}, Q_{3/4}$), dokuz tane ondabirlik ($Q_{1/10}, Q_{2/10}, \dots, Q_{8/10}, Q_{9/10}$) ve doksandokuz tane yüzdebirlik ($Q_{1/100}, Q_{2/100}, \dots, Q_{99/100}$) değer hesaplanır (Çömlekçi 1994).

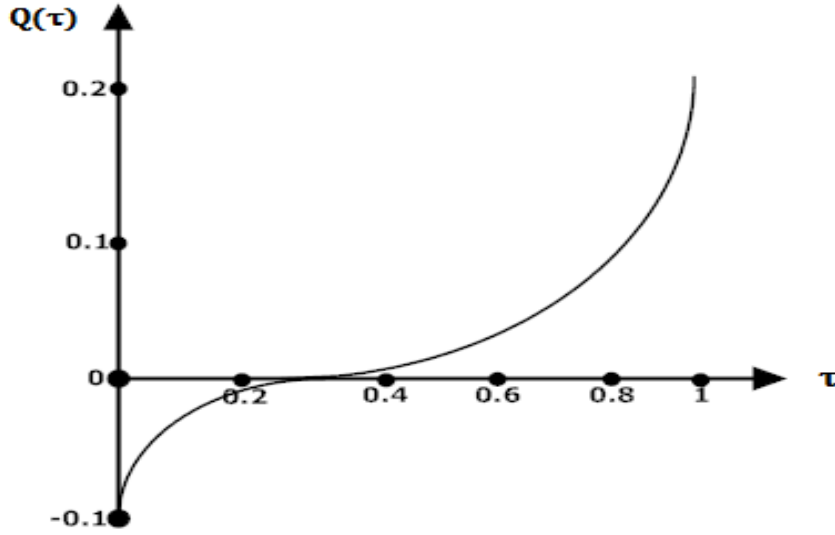
1.3.2. Kantil Dağılım Fonksiyonu

Kantil fonksiyonu sürekli bir fonksiyon olduğu için $F(x_{\tau}) = \tau$ olarak tanımlanır ve r . kantil dağılım fonksiyonu QF veya $Q(\tau)$ ile gösterilir.

x_{τ} 'nin değeri, toplumun τ .kantili olarak adlandırılır. $x_{\tau} = Q(\tau)$ fonksiyonu, τ .kantilleri, τ 'nin bir fonksiyonu olarak ifade edilir ve kantil fonksiyonu olarak adlandırılır. $Q(\tau)$, τ 'nin tüm olasılıkları için ($0 \leq \tau \leq 1$) kantil değerini verir. Kantil fonksiyonu ve dağılım fonksiyonu, herhangi (x, τ) ikilisi için $x = Q(\tau)$ ve $\tau = F(x)$

şeklinde yazılır. Bu fonksiyonlar sürekli, artan fonksiyonlardır (Koenker ve Hallock 2001) (Şekil 1.9).

$$Q(\tau) = F^{-1}(\tau) \text{ ve } F(x) = Q^{-1}(x)$$



Şekil 1.9. Kantil dağılım fonksiyon grafiği

Dağılımları modelleyebilmek için kantil fonksiyonu kullanılır. x verildiğinde y 'nin τ .kantilleri,

$$Q_y(\tau|x) = \beta x + \varepsilon(\tau)$$

olarak gösterilir. Burada $\varepsilon(\tau)$, hata terimini ifade etmektedir ve bu fonksiyona y 'nin x üzerindeki “koşullu kantil fonksiyonu” denir (Koenker ve Bassett 1978; Parzen 1979).

1.3.3. Kantil Yoğunluk Fonksiyonu

Dağılımları modelleyebilmek için kantil dağılım fonksiyonunun türevi alınarak kantil yoğunluk fonksiyonu elde edilir ve

$$q(\tau) = \frac{dQ(\tau)}{d\tau}$$

şeklinde tanımlanır. $Q(\tau)$ azalmayan bir fonksiyon olduğu için eğimi, $q(\tau)$ negatif değildir (Buhai 2004; Koenker ve Bassett 1978; Koenker ve Hallock 2001).

1.3.4. Kantil Regresyon Modeli ve Tahmini

Kantil fonksiyonu ile tanımlanan kantil regresyon modeli (KRM), bağımlı değişken ile bağımsız değişkenlerin arasındaki ilişkiyi modeller. KRM'de, hata terimlerinin ve

bağımsız değişkenlerin dağılımına ve varyansların homojen olup olmamasına ilişkin herhangi bir varsayım bulunmamaktadır. KRM, bağımsız değişkenlerin bağımlı değişkenlerin üzerindeki etkisini tahmin eder ve kantillere bağlı olarak regresyon katsayılarını belirler (Baur ve ark 2004; Hao ve Naiman 2007; Koenker 2005; Leng ve Tong 2013).

KRM, doğrusal regresyon modelinin bir uzantısıdır. Bağımsız değişkenlerdeki değişim, doğrusal regresyon modelinde bağımlı değişkenin koşullu ortalamasını değiştirirken, KR'de koşullu kantil değerini değiştirir (Hao ve Naiman 2007).

KRM'de tahmin ediciler doğrusal programlama problemi olarak formüle edilir. Hataların iki parçalı doğrusal amaç fonksiyonu optimize edilir ve simpleks yöntem kullanılarak sayısal değerler hesaplanır. F dağılımına sahip olan Y bağımlı değişkeninin τ . regresyon kantili, rastgele örneklem $\{y_i, i = 1, \dots, n\}$, $\{x_i, i = 1, \dots, n\}$, \mathbf{b} tahmin edilecek katsayı vektörü ve $e_i = y_i - \mathbf{x}_i \mathbf{b}$ hata değeri olmak üzere;

$$\min_{\mathbf{b} \in \mathbb{R}} \left[\sum_{i \in \{i: y_i \geq \mathbf{x}_i' \mathbf{b}\}} \tau |y_i - \mathbf{x}_i' \mathbf{b}| + \sum_{i \in \{i: y_i < \mathbf{x}_i' \mathbf{b}\}} (1 - \tau) |y_i - \mathbf{x}_i' \mathbf{b}| \right], \quad 0 \leq \tau \leq 1$$

ifadesinin minimize edilmesiyle hesaplanır (Buhai 2004; Koenker ve Bassett 1978; Koenker ve Hallock 2001).

τ . kantilin hata teriminin sıfır olduğu varsayılır;

$$Q_{\varepsilon_i(\tau)}(\tau | x_{i1}, \dots, x_{iK}) = 0$$

Bu özellikten, y 'nin τ . koşullu kantili;

$$Q_{y_i}(\tau | x_{i1}, \dots, x_{iK}) = \beta_0(\tau) + \sum_{k=1}^K \beta_k(\tau) x_{ik} + \varepsilon_i(\tau)$$

biçiminde tanımlanır (Baur ve ark 2004; Koenker 2008; Koenker ve Bassett 1978).

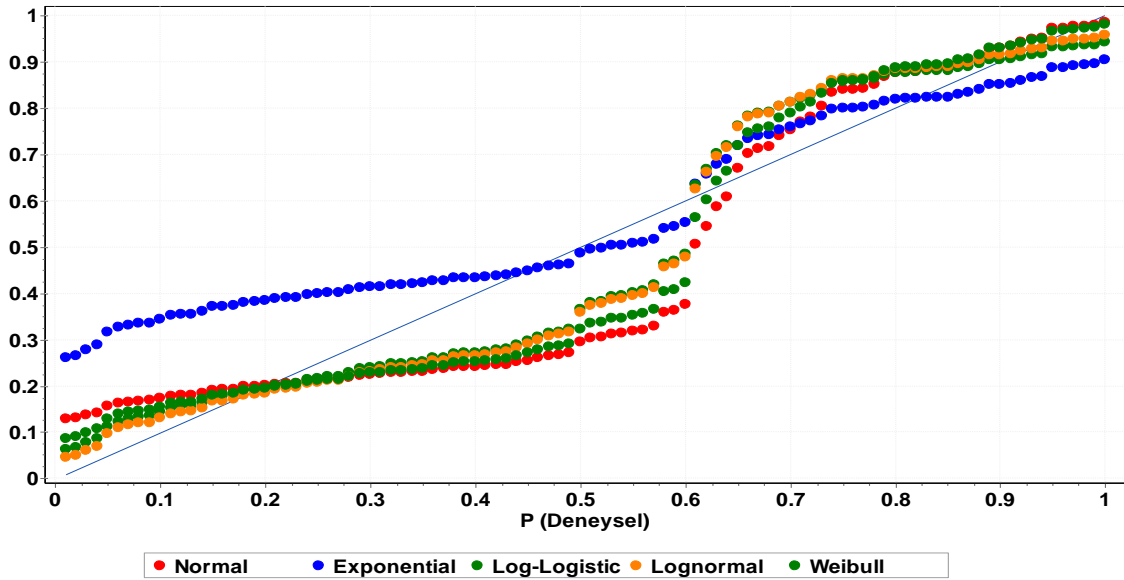
1.3.5. Kantil Regresyonun Özellikleri

1. Kantiller, bağımlı değişkendeki sapan değerlere karşı duyarlı değildir. Bu durum, KR yönteminin robust bir yöntem olduğunun göstergesidir.
2. Hata terimi normal dağılmadığında KR tahmin edicileri çok daha etkindir.
3. KR, değişen varyansın belirlenmesine olanak sağlar.

4. KR, heterojen yapıdaki veri setlerinin analizinde oldukça kullanışlıdır. (Koenker 2005; Koenker ve Bassett 1978; Lin ve Rodriguez 2013).

1.3.6. Sağkalım Analizinde Kantil Regresyon

Kliniksel çalışmalarda KR, toplumdaki hastaların sağkalım sürelerini ve hastalığın seyrini etkileyen faktörleri belirlemeye yardımcı olur. Bireylerin sağkalım süresinin kantillerle modellenmesi, sağkalım süresinin ortak değişkenlerden ne kadar etkilendiğinin belirlenmesini sağlar. PR modelleri ile karşılaştırıldığında KR'nin iki önemli avantajı vardır. (1) Sağkalım süresinin dağılımı için herhangi bir varsayım yoktur ve her bir kantil değeri için regresyon modeli oluşturulur. (2) Heterojen yapıdaki veri setlerinin analizinde kullanışlıdır (Lin ve Rodriguez 2013). Heterojen veri, popülasyonun en az iki farklı karakteristiğe sahip iki alt popülasyondan oluşmasıdır (Erişoğlu ve ark 2011). Örneğin; Marmara bölgesinde, akciğer kanseri teşhisi konulmuş 40 hastanın sağkalım süreleri, ortalaması 150 ve standart sapması 30 olan normal dağılıma sahip olsun. Doğu bölgesinde, akciğer kanseri teşhisi konulmuş 60 hastanın sağkalım süreleri, ortalaması 50 ve standart sapması 10 olan normal dağılıma sahip olsun. İki bölgeden alınan 40 ve 60 birimlik örneklemeleri bir arada kullanarak oluşturulan 100 birimlik sağkalım verisi normal dağılıma uygun değildir (Şekil 1.10).



Şekil 1.10. 100 birimlik sağkalım veri seti için P-P grafiği

Sansürlü KRM'de, klasik regresyon yönteminin parametre tahmin edicileri doğrudan kullanılamaz. Bu nedenle, parametre nokta tahmin yöntemlerinden biri olan Portnoy yöntemi kullanılabilir. Bu yöntem ile her bir kantil değerinin katsayı tahmini

yapıldığından kullanıcıya esnek bir modelleme imkanı sunar (Peng ve Huang 2008; Portnoy 2003; Wang ve ark 2013; Ying ve ark 1995).

Kantillere bağlı olarak sağkalım modeli;

$$Q_{\log(t|x)}(\tau|x) = \mathbf{x}'\boldsymbol{\beta}(\tau)$$

veya

$$Q_{t|x}(\tau|x) = e^{\mathbf{x}'\boldsymbol{\beta}(\tau)}$$

şeklinde oluşturulur.

Hazard fonksiyonu;

$$h(Q_{t|x}(\tau|x)) = \lim_{\Delta\tau \rightarrow 0} \frac{\frac{\Delta\tau}{1-\tau} \Delta Q_{t|x}(\tau|x)}{1-\tau}$$

Sağkalım fonksiyonu;

$$Q_{t|x}(\tau|x) = S_0^{-1} \left((1-\tau)e^{\frac{1}{\mathbf{x}'\boldsymbol{\beta}}} \right)$$

şeklinde tanımlanır (Koenker 2005).

1.3.7. Portnoy Tahmin Edicisi ile Parametre Tahmini

Sağkalım verilerinde sansür probleminin KR yöntemi ile çözülmesi için Portnoy tarafından bir algoritma geliştirilmiştir. KRM kurulurken, tek örneklem durumunda ortaya çıkmayan birçok belirsizlik bulunmaktadır. Bunlardan en önemlisi, KR süreci tarafından tahmin edilmiş sansürlü gözlemlerin olasılığıdır ve bu olasılık negatif hataya sahiptir. Parametre tahmin yöntemi, negatif hatalardan sonra gelen τ değerleri için sıfır hatalı en uygun kurala dönüşür. Tek örneklem olduğu durumlarda Kaplan-Meier tahmin edicisinin monotonluğu nedeniyle bu durum gerçekleşmez. Ancak p boyutlu regresyon modeli kurulduğunda, monotonluk koşulunun zayıflığı nedeniyle en uygun kural aşağıdaki ağırlıklandırma ile elde edilir (Koenker 2008).

$$w_i(\tau) = \frac{\frac{1}{n} - \tau_i}{1 - \tau_i}$$

Portnoy, tahmin edilemeyen model için $\tau \in [0,1]$ aralığında değişen sansürlü gözlemlerin fazla olmasından kaynaklanan karışıklıkları incelemiş ve bu sorunu ortadan

kaldırmak için etkili bir yöntem tanımlamıştır. τ değerini sıfır civarından başlatarak, her adımda ağırlıklandırılmış KR probleminin,

$$\min \sum_{i \in K} \rho_{\tau}(Y_i - \mathbf{x}'_i \boldsymbol{\beta}) + \sum_{i \in K} [w_i(\tau) \rho_{\tau}(Y_i - \mathbf{x}'_i \boldsymbol{\beta}) + (1 - w_i(\tau)) \rho_{\tau}(Y_{\infty} - \mathbf{x}'_i \boldsymbol{\beta})]$$

en uygun çözümü için τ aralıklarının uzunluğunu ölçmek mümkündür. Her bir τ değeri için parametre tahminleri yapılır ve aralığın üst sınırından tekrar çözülür. Geriye sadece pozitif hatalı, tekrar ağırlıklandırma yapılmamış sansürlü gözlemler kaldığı durumda iterasyon sonlanır (Koenker ve Hallock 2001; Portnoy 2003).

1.4. Model Değerlendirme Ölçütleri

Bu tez çalışmasında modellerin performanslarını karşılaştırmak için iki kriter kullanılmıştır. Bu kriterler HKOK ve OMH'dir.

1.4.1. Hata Kareler Ortalamasının Karekökü

HKOK, bir model tarafından tahmin edilen değerler ile gerçekte gözlenen değerler arasındaki farkın bir ölçütüdür. HKOK, uç değerlere karşı hassasiyet gösterir. HKOK ölçütüne göre modellerin performansları incelenirken HKOK değeri en küçük olan model en uygun model olarak kabul edilir. Basitleştirmek için n birim için hesaplanan hata e_i ($i = 1, 2, \dots, n$) olduğunda HKOK ölçütü,

$$\text{HKOK} = \sqrt{\frac{\sum_{i=1}^n e_i^2}{n}} = \sqrt{\frac{\sum_{i=1}^n (t_{\text{gözlenen}} - t_{\text{tahmin}})^2}{n}}$$

formülü ile hesaplanır (Chai ve Draxler 2014; Sayegh ve ark 2014).

1.4.2. Ortalama Mutlak Hata

OMH, model değerlendirmede yaygın olarak kullanılan başka bir ölçüttür. OMH, gerçek değer ile tahmin edilmiş değer arasındaki farkın mutlak değeridir. OMH ölçütüne göre modellerin performansları değerlendirilirken, OMH değeri en küçük olan model en uygun model olarak kabul edilir. n birim için hesaplanan hata e_i ($i = 1, 2, \dots, n$) olduğunda OMH,

$$\text{OMH} = \frac{\sum_{i=1}^n |e_i|}{n} = \frac{\sum_{i=1}^n |t_{\text{gözlenen}} - t_{\text{tahmin}}|}{n}$$

formülü ile hesaplanır (Chai ve Draxler 2014; Sayegh ve ark 2014; Willmott ve Matsuura 2005).

2. GEREÇ VE YÖNTEM

Bu çalışmanın uygulama bölümü iki farklı aşamadan oluşmuştur. Birinci aşamasında, yöntemlerin performanslarının karşılaştırılması amacıyla simülasyon ile türetilmiş heterojen yapıdaki veriler üzerinde KR ve PR yöntemleri uygulanmış ve yöntemlerin HKOK ve OMH değerleri hesaplanmıştır. Bu kriterler dikkate alınarak regresyon modellerinin performansları karşılaştırılmıştır.

İkinci aşamada, yöntemlerin kullanılabilirliğini göstermek amacıyla, 1998-2007 yılları arasında meme kanserli hastaların hastalısız sağkalım sürelerine etki eden risk faktörlerini içeren veri seti ve herkesin ücretsiz erişimine açık olan, 1984-1992 yıllarına ilişkin, sağkalım süresinin heterojen yapıdaki German Breast Cancer Study (gbcs) adlı veri seti kullanılmıştır. Parametre tahminleri hesaplanmış ve risk faktörleri değerlendirilmiştir.

2.1. Kullanılan Programlar

Analizler, R 3.0.2 programında gerçekleştirildi. Sağkalım veri setlerinin türetilmesinde R programlama dilinden yararlanıldı ve *quantreg*, *survival*, *SparseM*, *graphics*, *base*, *stats*, *HydroGOF*, *zoo* ve *foreign* paketleri kullanıldı. Sağkalım veri setinin heterojen yapıda olup olmadığının belirlenmesi için EasyFit 5.6 trial programından yararlanıldı.

2.2. Simülasyon Uygulaması

Sağkalım sürelerine ilişkin veri setlerinde kullanılan KR ve PR yöntemlerinin karşılaştırılması 1000 tekrarlı simülasyon çalışması ile yapıldı. Karşılaştırmalar, HKOK ve OMH ölçütleri kullanılarak gerçekleştirildi.

2.2.1. Algoritma

1. Örneklem büyüklüğü için sabit bir değer atandı ($n=100, 250, 500, 750, 1000$).
2. Tahmin edilecek olan simülasyon parametreleri $\beta_1 = 2$ ve $\beta_2 = -1$ olarak alındı.
3. X_1 ve X_2 değişkenleri için $N(0,1)$ dağılımında veri türetimi yapıldı.
4. $\lambda_t = \exp(\beta_1 * X_1 + \beta_2 * X_2)$ fonksiyonu yardımı ile heterojen veri seti türetmek için üstel dağılımın ve Weibull dağılımının ölçek parametreleri hesaplandı.

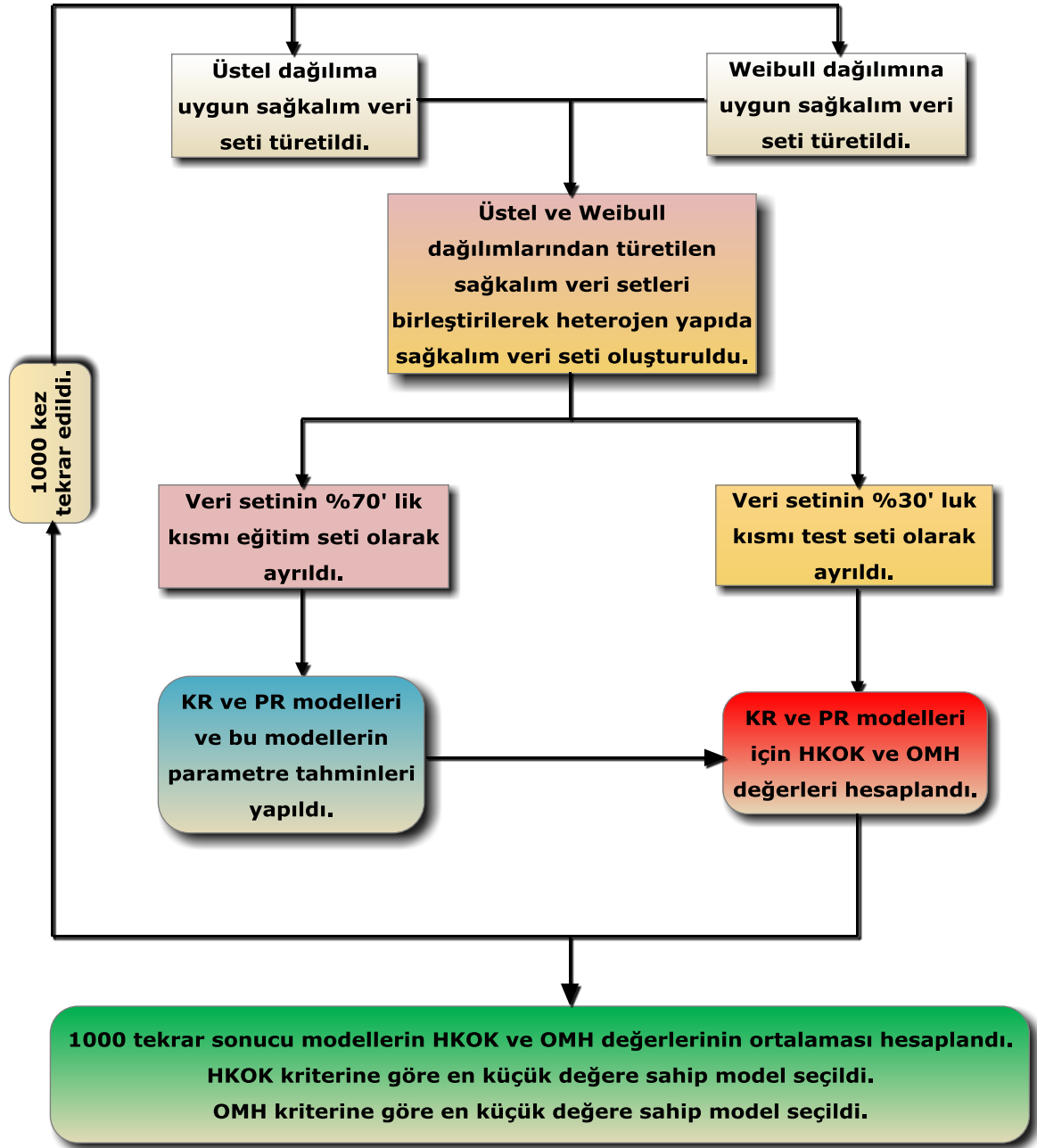
5. $T_1 \sim \text{Üstel}(\lambda_t)$ ve $T_2 \sim \text{Weibull}(\alpha = 10, \lambda = \lambda_t)$ dağılımlara sahip sağkalım süreleri türetildi ve gerçek sağkalım süresine yakın olması için bu değerler 100 ile çarpıldı.
6. $C_1 \sim \text{Üstel}(\lambda = 1)$ ve $C_2 \sim \text{Weibull}(\alpha = 10, \lambda = 1)$ dağılımlarına sahip sansürlü gözlemleri belirlemek için sağkalım süreleri türetildi ve bu değerler 100 ile çarpıldı.
7. Üstel ve Weibull dağılım gösteren sağkalım sürelerini belirlemek için $\min(T, C)$ kriteri kullanıldı.
8. Birimin sansürlü olup olmadığına karar vermek için;

$$\delta_i = \begin{cases} 1, & T_i \leq C_i \\ 0, & T_i > C_i \end{cases} \quad i: 1, 2, \dots, n$$

kriteri kullanıldı.

9. Hem üstel hem de Weibull dağılımından türetilen sağkalım süresi, sağkalım durumu ve X_1, X_2 değişkenleri birleştirilerek heterojen yapıda veri seti oluşturuldu.
10. Veri setinin %70'lik kısmı eğitim, %30'luk kısmı ise test seti olarak ayrıldı.
11. Eğitim veri setinden KR ve PR modelleri için parametre tahminleri yapıldı ve regresyon modelleri oluşturuldu.
12. Eğitim veri setinden elde edilen regresyon modellerinin parametreleri kullanılarak bağımsız bir veri seti olan test setinde modellerin geçerliliği denendi. HKOK ve OMH değerleri hesaplandı.
13. 3.-12. adımlar 1000 kez tekrar edildi.
14. 1000 tekrar sonucunda elde edilen HKOK ve OMH değerlerinin ortalaması alındı.
15. Ortalama HKOK ve OMH değerleri en küçük olan model en uygun model olarak seçildi

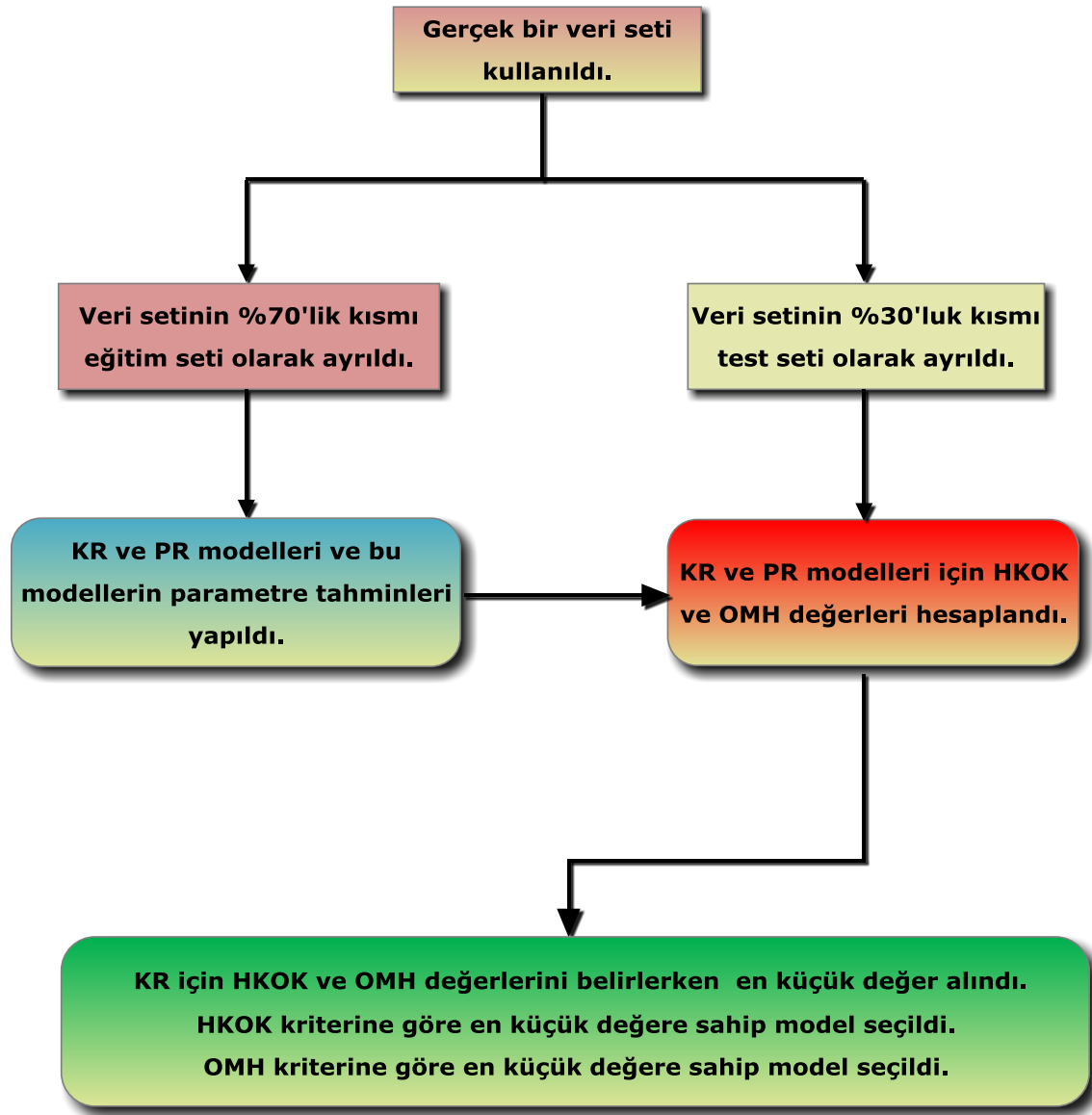
Bu algoritmanın akış şeması Şekil 2.1'de verilmiştir.



Şekil 2.1. Simülasyon uygulaması için hazırlanan programın akış şeması

2.3. Gerçek Veri Seti Uygulaması

Çalışmada meme kanseri hastalarına ilişkin homojen ve heterojen yapıda gerçek veri setleri kullanıldı. Bu verilerin analizi, Şekil 2.2’de hazırlanan programın akış şemasına göre gerçekleştirildi.



Şekil 2.2. Gerçek veri seti uygulaması için hazırlanan programın akış şeması

Her iki veri setindeki nicel değişkenlerin normal dağılıma uygunluğu Kolmogorov-Smirnov (K-S) testi ile incelendi. Nicel değişkenler normal dağılıma uygun olmadığından meme kanseri verisi için nüksetmenin olduğu ve olmadığı gruplara; gbcs veri seti için ölüm olayının gerçekleştiği ve gerçekleşmediği gruplara göre karşılaştırılmasında Mann-Whitney U testi kullanıldı ve tanımlayıcı istatistikler medyan (%25-%75 persantil) şeklinde gösterildi. Kategorik yapıdaki değişkenlerin karşılaştırılmasında ise Ki-kare analizi kullanıldı ve tanımlayıcı istatistikler frekans (%) şeklinde gösterildi (Çizelge 2.1 ve Çizelge 2.2).

Bu veri setleri için KR ve PR modellerinin parametre tahminleri yapıldı ve modeller oluşturuldu. Oluşturulan modellerin HKOK ve OMH değerlerini hesaplamak için

sağkalım sürelerinin tahminleri hesaplandı. En küçük HKOK ve OMH değerine sahip olan model, bireylerin sağkalım sürelerini tahmin etmek için oluşturulan en uygun regresyon modeli olarak kabul edildi.

2.3.1. Meme Kanseri Veri Seti

Çalışmada kullanılan, 1998-2007 yılları arasında 485 meme kanserli hastanın hastalıksız sağkalım sürelerine etki eden risk faktörlerini içeren veri seti “Bayesgil Yaşam Analizi ve Cox Regresyon Yaşam Analizinin Türetilmiş ve Gerçek Veri Setlerinde Uygulanması” adlı çalışmadan alınarak analizler gerçekleştirilmiştir (Ömürlü ve ark 2009).

Bireylerin hastalıksız sağkalım süresini etkileyen risk faktörleri olarak yaş, aksiller lenf nodu tutulumu, tümör büyüklüğü, progesteron reseptör durumu, kapsül dışı lenf nodu tutulumu ve hormon terapisi değişkenleri alındı.

Çizelge 2.1. Meme kanseri hastalarına ilişkin nüksetme durumuna göre bağımsız değişkenlerin tanımlayıcı istatistikleri

| Bağımsız Değişkenler | | Nüksetme | | p |
|--------------------------------|----------------------|----------------|----------------|--------|
| | | Yok (n=345) | Var (n=140) | |
| Yaş (yıl) | | 49(43-58) | 52(43-62) | 0,890 |
| Tümör Büyüklüğü (cm) | | 3(2-4) | 3,5(2,35-5) | <0,001 |
| Progesteron Reseptör Durumu | Negatif | 85(24,6) | 59(42,1) | <0,001 |
| | Pozitif | 260(75,4) | 81(57,9) | |
| Aksiller Lenf Nodu Tutulumu | Negatif | 165(47,8) | 31(22,1) | <0,001 |
| | Pozitif Lenf Nodu1-3 | 94(27,2) | 34(24,3) | |
| | Pozitif Lenf Nodu≥4 | 86(24,9) | 75(53,6) | |
| Kapsül Dışı Lenf Nodu Tutulumu | Negatif | 249(72,2) | 65(46,4) | <0,001 |
| | Pozitif | 96(27,8) | 75(53,6) | |
| Hormon Terapisi | Yok | 66(19,1) | 55(39,3) | <0,001 |
| | Var | 279(80,9) | 85(60,7) | |

Çizelge 2.1 incelendiğinde, meme kanserli hastalarda kanserin nüksetme durumu üzerinde tümör büyüklüğü, progesteron reseptör durumu, aksiller lenf nodu tutulumu, kapsül dışı lenf nodu tutulumu ve hormon terapisi değişkenlerinin önemli etkisi olduğu belirlendi (p<0,001).

2.3.2. Gbcs Veri Seti

Erişime açık, 1984-1992 yılları arasında 686 hastanın heterojen yapıda sağkalım sürelerini içeren gbcs veri seti kullanıldı. Gbcs veri setine ftp://ftp.wiley.com/public/sci_tech_med/survival adresinden ulaşıldı (Sauerbrei ve ark 1999).

Sağkalım sürelerini etkileyen risk faktörleri olarak; yaş, hormon terapisi, tümör büyüklüğü, tümör evresi, progesteron reseptör sayısı, östrojen reseptörlerin sayısı, menopoz durumu ve nodüllerin sayısı değişkenleri alındı.

Çizelge 2.2. Gbcs veri setindeki meme kanseri hastalarına ilişkin ölüm olayına göre bağımsız değişkenlerin tanımlayıcı istatistikleri

| Bağımsız Değişkenler | | Ölüm | | p |
|-----------------------------|-----|----------------|----------------|--------|
| | | Yok (n=515) | Var (n=171) | |
| Yaş (yıl) | | 53(46-61) | 53(47-62) | 0,591 |
| Tümör Büyüklüğü (mm) | | 25(20-35) | 30(23-39) | <0,001 |
| Progesteron Reseptör Sayısı | | 55(11-164) | 10(1-32) | <0,001 |
| Östrojen Reseptör Sayısı | | 44(11-128) | 14(2-67) | <0,001 |
| Nodül Sayısı | | 3(1-5) | 5(3-11) | <0,001 |
| Menopoz Durumu | Var | 223(43,3) | 67(39,2) | 0,345 |
| | Yok | 292(56,7) | 104(60,8) | |
| Hormon Terapisi | Var | 325(63,1) | 115(67,3) | 0,327 |
| | Yok | 190(36,9) | 56(32,7) | |
| Tümör evresi | 1 | 75(14,6) | 6(3,5) | <0,001 |
| | 2 | 337(65,4) | 107(62,6) | |
| | 3 | 103(20) | 58(33,9) | |

Çizelge 2.2 incelendiğinde, gbcs veri setindeki meme kanserli hastaların ölüm olayı üzerinde tümör büyüklüğü, progesteron reseptör sayısı, östrojen reseptör sayısı, nodül sayısı ve tümör evresi değişkenlerinin önemli etkisi olduğu belirlendi ($p<0,001$).

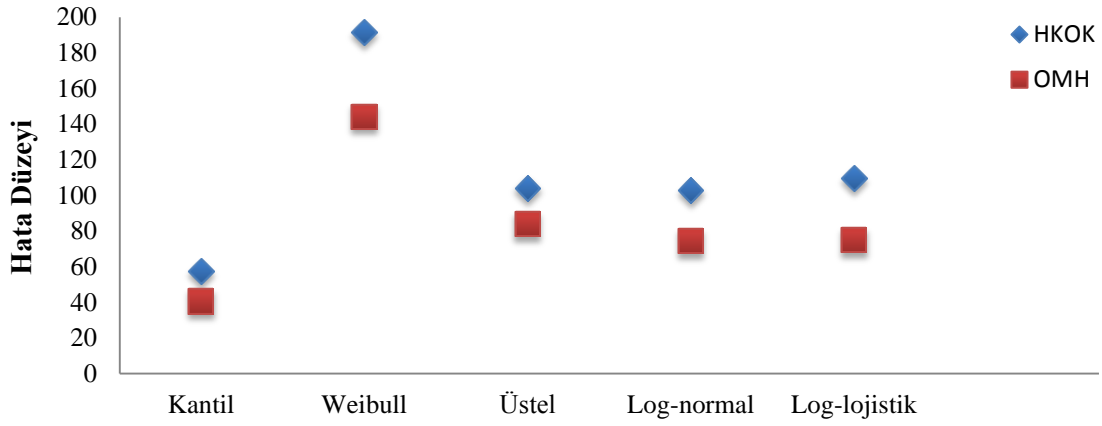
3. BULGULAR

3.1. Simülasyon Sonuçları

Algoritma kullanılarak yapılan 1000 tekrarlı simülasyon çalışması sonunda regresyon modellerinin karşılaştırılmasında kullanılan HKOK ve OMH ölçütlerinden hesaplanan hata değerlerinin, n=100, 250, 500, 750, 1000 olduğu durumlara göre değişimleri Çizelge 3.1’de verilmiştir. Farklı örneklem büyüklüklerine göre regresyon modellerinin hataları hesaplandığında her iki kritere göre de minimum hata değerlerinin KR modelinde olduğu gözlemlendi. Farklı örneklem büyüklüklerine göre simülasyon sonucunda elde edilen regresyon modellerine ilişkin HKOK ve OMH değerlerinin KR ve PR modellerine göre değişimi Şekil 3.1–Şekil 3.5’de gösterilmiştir.

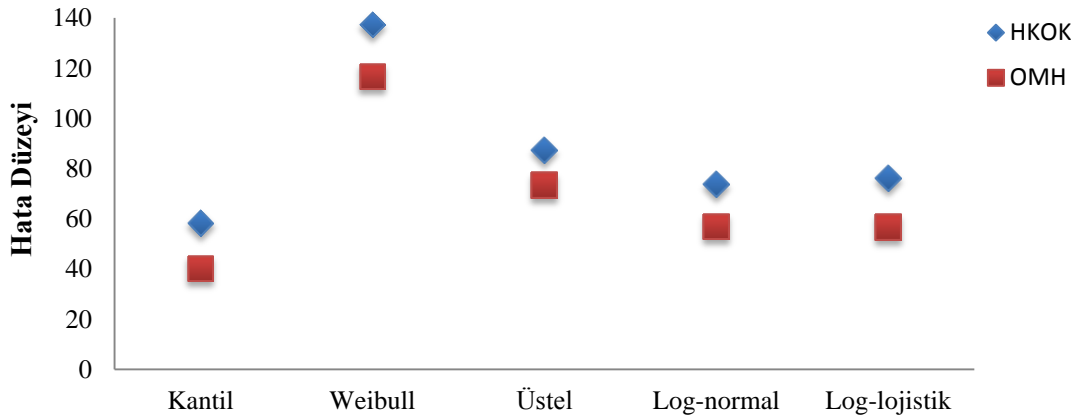
Çizelge 3.1. Değişen n değerlerine göre modellerin HKOK ve OMH değerleri

| n | MODELLER | HKOK | OMH |
|------|--------------|--------------|--------------|
| 100 | Kantil | 57,50 | 40,63 |
| | Weibull | 191,51 | 144,17 |
| | Üstel | 104,19 | 83,86 |
| | Log-normal | 102,90 | 74,35 |
| | Log-lojistik | 109,64 | 74,84 |
| 250 | Kantil | 58,32 | 40,28 |
| | Weibull | 137,40 | 116,63 |
| | Üstel | 87,47 | 73,34 |
| | Log-normal | 73,73 | 56,99 |
| | Log-lojistik | 76,15 | 56,69 |
| 500 | Kantil | 58,80 | 39,94 |
| | Weibull | 124,66 | 109,35 |
| | Üstel | 83,66 | 70,80 |
| | Log-normal | 65,82 | 51,41 |
| | Log-lojistik | 68,38 | 51,25 |
| 750 | Kantil | 59,40 | 39,86 |
| | Weibull | 121,06 | 107,78 |
| | Üstel | 82,65 | 70,19 |
| | Log-normal | 63,74 | 49,74 |
| | Log-lojistik | 65,86 | 49,36 |
| 1000 | Kantil | 59,66 | 39,93 |
| | Weibull | 119,58 | 106,60 |
| | Üstel | 82,15 | 69,69 |
| | Log-normal | 62,86 | 48,91 |
| | Log-lojistik | 65,27 | 48,64 |



Şekil 3.1. n=100 olduğunda HKOK ve OMH değerlerinin modellere göre değişimi

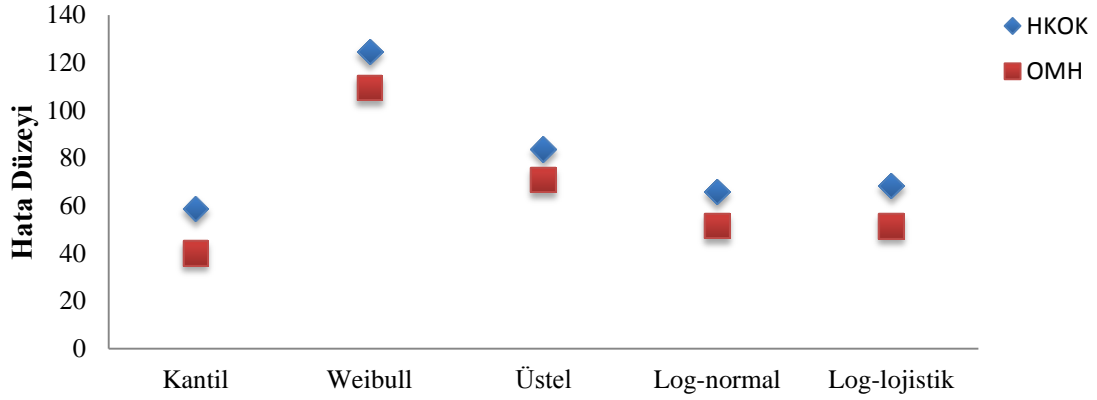
n=100 olduğu durumda, HKOK kriterine göre minimum hata değeri 57,50 olarak KR modelinden, maksimum hata değeri 191,51 olarak Weibull regresyon modelinden elde edildi. OMH kriterine göre ise minimum hata değeri 40,63 olarak KR modelinden, maksimum hata değeri 144,17 ile Weibull regresyon modelinden elde edildi. PR modelleri arasında her iki kritere göre minimum hataya sahip modelin log-normal regresyon modeli, maksimum hataya sahip modelin Weibull regresyon modeli olduğu bulundu (Şekil 3.1).



Şekil 3.2. n=250 olduğunda HKOK ve OMH değerlerinin modellere göre değişimi

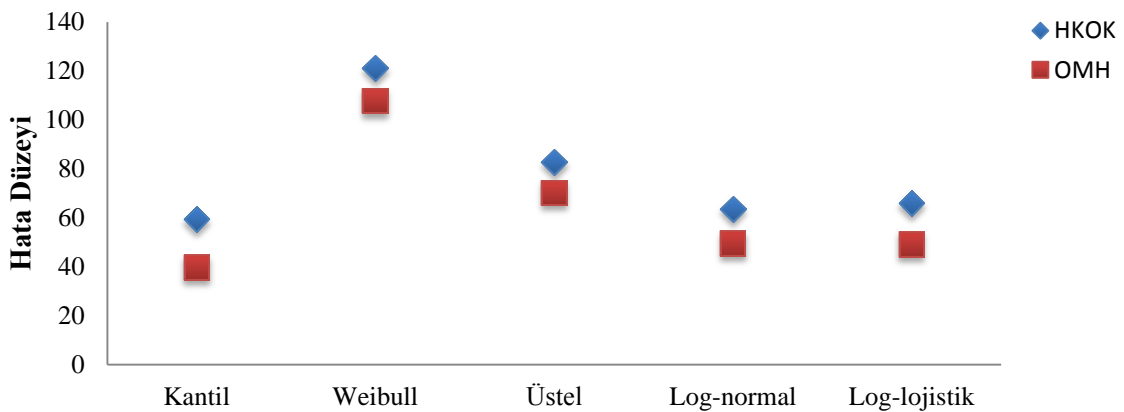
n=250 olduğu durumda, HKOK kriterine göre minimum hata değeri 58,32 olarak KR modelinden, maksimum hata değeri 137,40 olan Weibull regresyon modelinden elde edildi. OMH kriterine göre ise minimum hata değeri 40,28 olarak KR modelinden, maksimum hata değeri 116,63 olan Weibull regresyon modelinden elde edildi. PR modelleri arasında HKOK kriterine göre minimum hataya sahip modelin log-normal regresyon modeli, maksimum hataya sahip modelin Weibull regresyon modeli olduğu bulundu. OMH kriterine göre minimum hataya sahip PR modeli log-lojistik regresyon

modeli, maksimum hataya sahip modelin Weibull regresyon modeli olduğu bulundu (Şekil 3.2).



Şekil 3.3. n=500 olduğunda HKOK ve OMH değerlerinin modellere göre değişimi

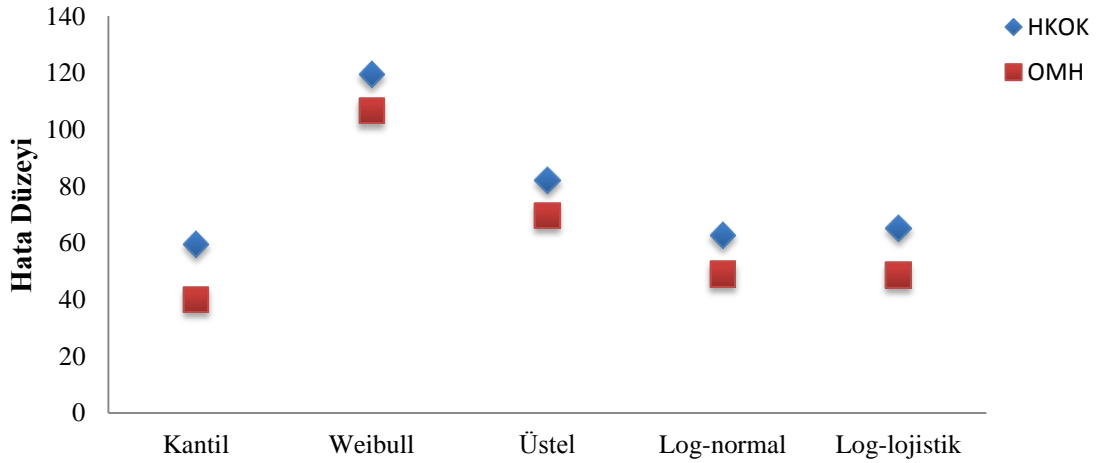
n=500 olduğu durumda, HKOK kriterine göre minimum hata değeri 58,80 olan KR modelinden, maksimum hata değeri 124,66 olan Weibull regresyon modelinden elde edildi. OMH kriterine göre ise minimum hata değeri 39,94 olan KR modelinden, maksimum hata değeri 109,35 olan Weibull regresyon modelinden elde edildi. PR modelleri arasında HKOK kriterine göre minimum hataya sahip modelin log-normal regresyon modeli, maksimum hataya sahip modelin Weibull regresyon modeli olduğu bulundu. OMH kriterine göre minimum hataya sahip PR modelinin log-lojistik regresyon modeli, maksimum hataya sahip modelin Weibull regresyon modeli olduğu bulundu (Şekil 3.3).



Şekil 3.4. n=750 olduğunda HKOK ve OMH değerlerinin modellere göre değişimi

n=750 olduğu durumda, HKOK kriterine göre minimum hata değeri 59,40 olarak KR modelinden, maksimum hata değeri 121,06 olarak Weibull regresyon modelinden elde

edildi. OMH kriterine göre ise minimum hata değeri 39,86 olarak KR modelinden, maksimum hata değeri 107,78 olan Weibull regresyon modelinden elde edildi. PR modelleri arasında HKOK kriterine göre minimum hataya sahip modelin 63,74 değeri ile log-normal regresyon modeli, maksimum hataya sahip modelin Weibull regresyon modeli olduğu bulundu. OMH kriterine göre minimum hataya sahip PR modelinin 49,36 değeri ile log-lojistik regresyon modeli, maksimum hataya sahip modelin Weibull regresyon modeli olduğu bulundu (Şekil 3.4).



Şekil 3.5. n=1000 olduğunda HKOK ve OMH değerlerinin modellere göre değişimi

n=1000 olduğu durumda, HKOK kriterine göre minimum hata değeri 59,66 olarak KR modelinden, maksimum hata değeri 119,58 olarak Weibull regresyon modelinden elde edildi. OMH kriterine göre ise minimum hata değeri 39,93 olarak KR modelinden, maksimum hata değeri 106,60 olarak Weibull regresyon modelinden elde edildi. PR modelleri arasında HKOK kriterine göre minimum hataya sahip model 62,86 değeri ile log-normal regresyon modeli, maksimum hataya sahip modelin Weibull regresyon modeli olduğu bulundu. OMH kriterine göre minimum hataya sahip PR modelinin 48,64 değeri ile log-lojistik regresyon modeli, maksimum hataya sahip modelin Weibull regresyon modeli olduğu bulundu (Şekil 3.5).

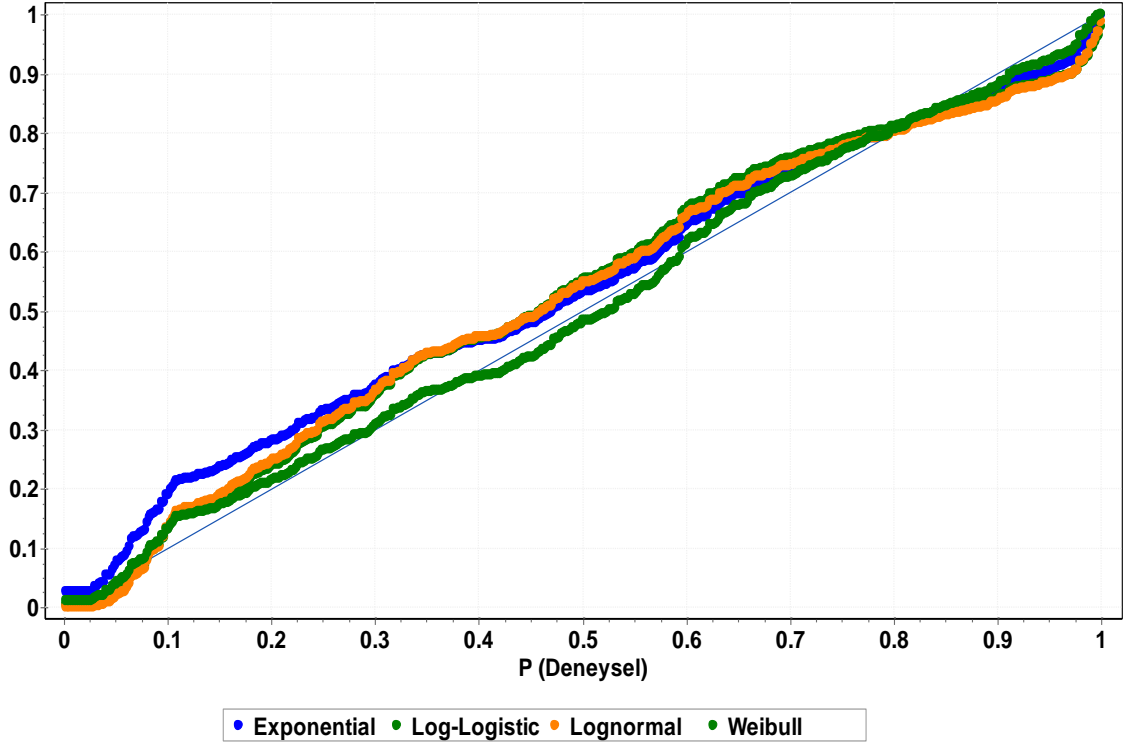
3.2. Gerçek Uygulama Sonuçları

3.2.1. Meme Kanseri Veri Seti

Meme kanseri veri setinin heterojen yapıda olup olmadığına karar vermek için K-S test istatistiğinin p değerleri Çizelge 3.2’de verilmiştir. Hastaların hastalıksız sağkalım süreleri için P-P grafiği Şekil 3.6’da gösterilmiştir.

Çizelge 3.2. Meme kanseri veri seti için K-S test sonuçları

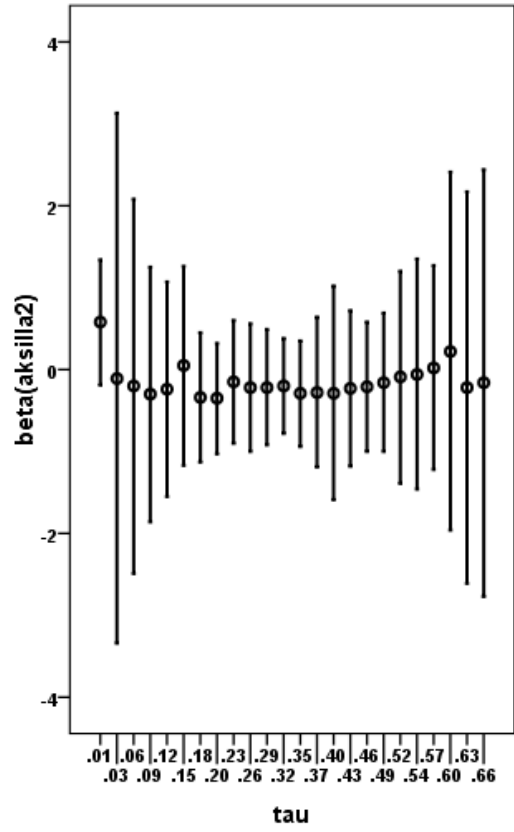
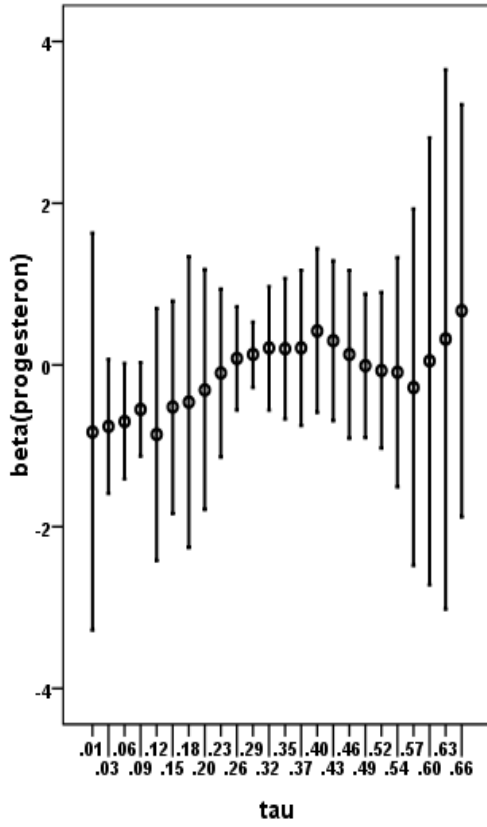
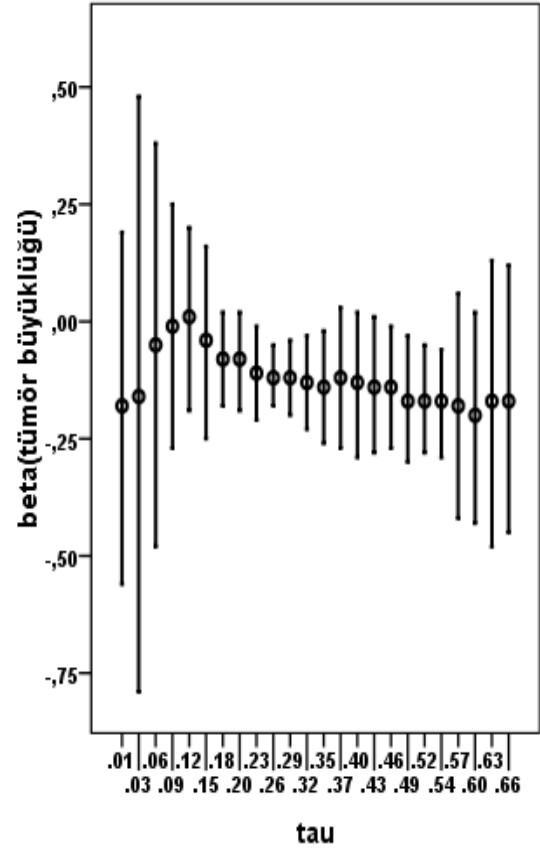
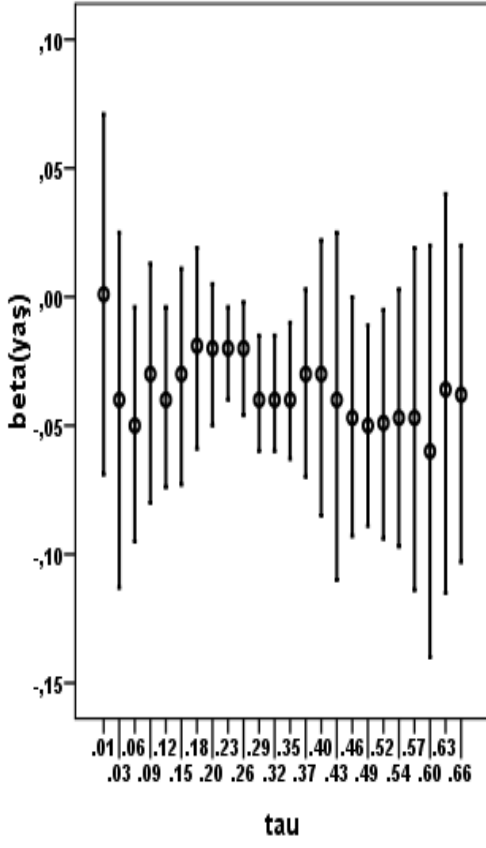
| Dağılımlar | K-S p |
|--------------|--------|
| Weibull | 0,264 |
| Üstel | <0,001 |
| Log-normal | 0,003 |
| Log-lojistik | 0,005 |



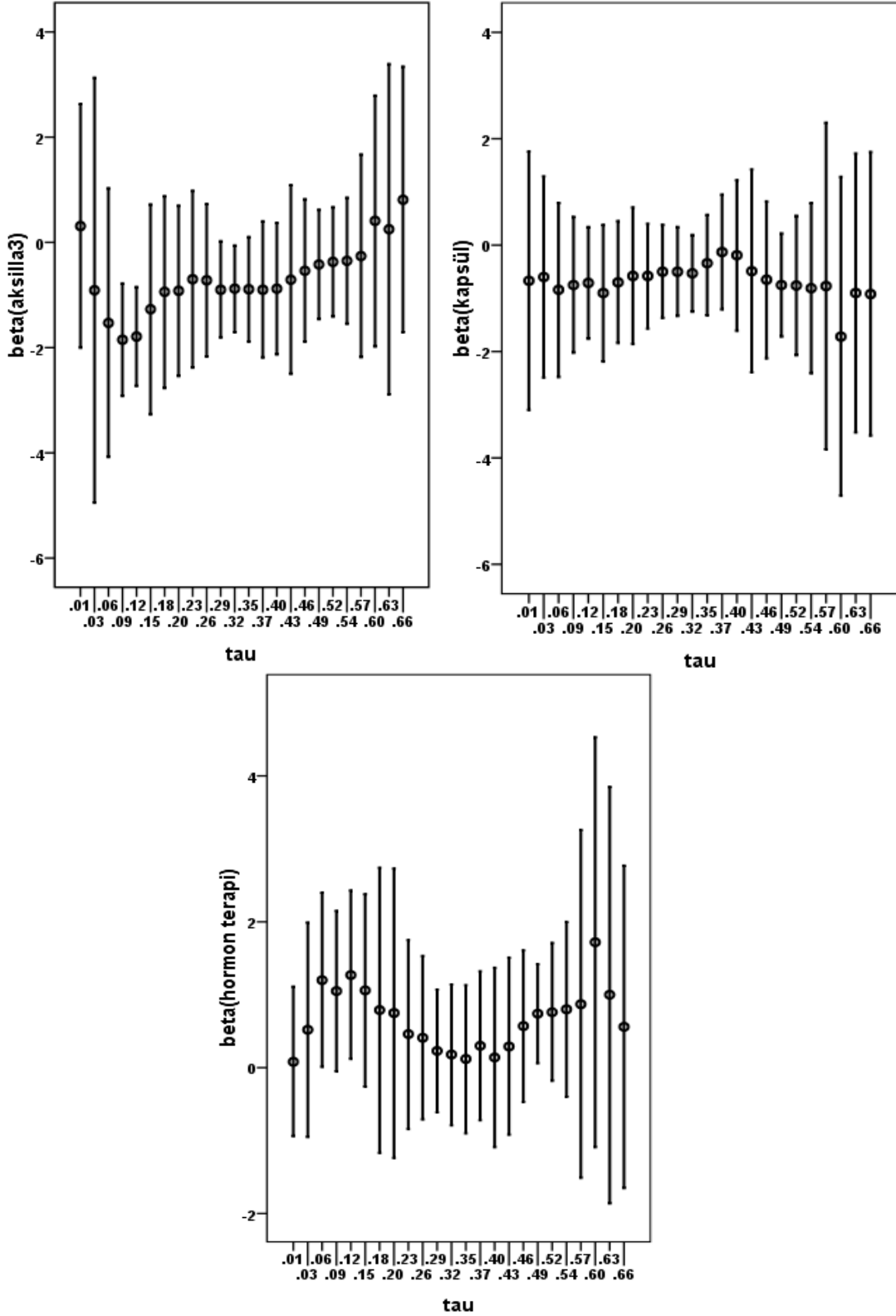
Şekil 3.6. Meme kanseri tanısı konan hastaların hastalıksız sağkalım süreleri için P-P grafiği

Meme kanseri tanısı konan hastaların hastaliksız sađkalım sürelerinin homojen yapıda ($p>0,05$) olduđu; üstel, log-normal ve log-lojistik dağılımlarına uygun olmadığı ($p<0,05$) bulundu (Şekil 3.6).

Meme kanseri veri setine uygulanan KR yönteminin farklı τ değerlerine göre parametre tahminleri Ek 1’de verilmiştir. Parametre tahminlerinin değışimi Şekil 3.7’de gösterilmiştir.



Şekil 3.7. Farklı τ değerleri için parametre tahminlerinin grafiksel gösterimi



Şekil 3.7. Farklı τ değerleri için parametre tahminlerinin grafiksel gösterimi (devam)

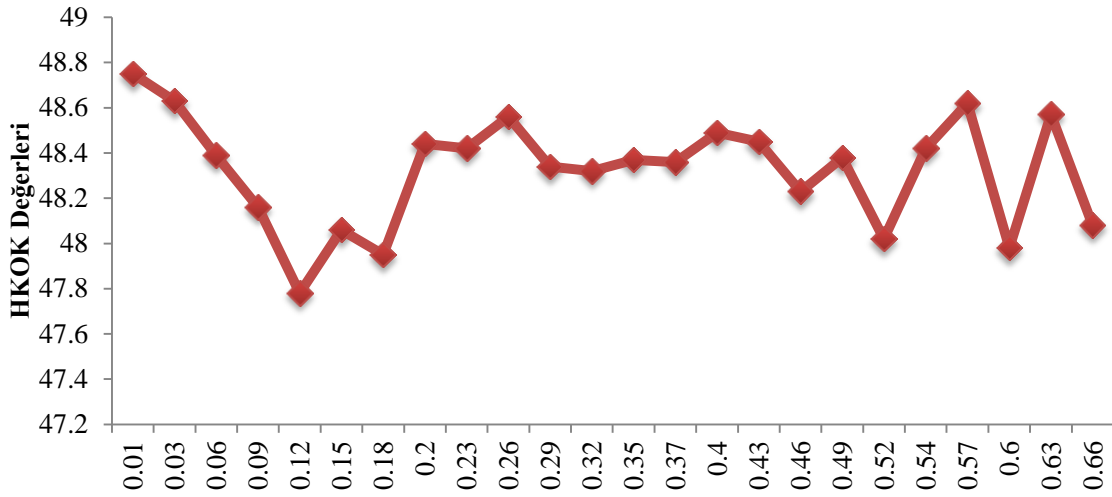
τ değerleri için HKOK ve OMH kriterine göre hata değerleri hesaplandı. Minimum hata $\tau=0,12$ değerinde elde edildi. HKOK kriterine göre minimum hatanın 47,78 olduğu;

OMH kriterine göre minimum hatanın 34,90 olduğu bulundu (Çizelge 3.3). PR modelleri ile karşılaştırma yapılırken $\tau=0,12$ için KR sonuçları kullanıldı.

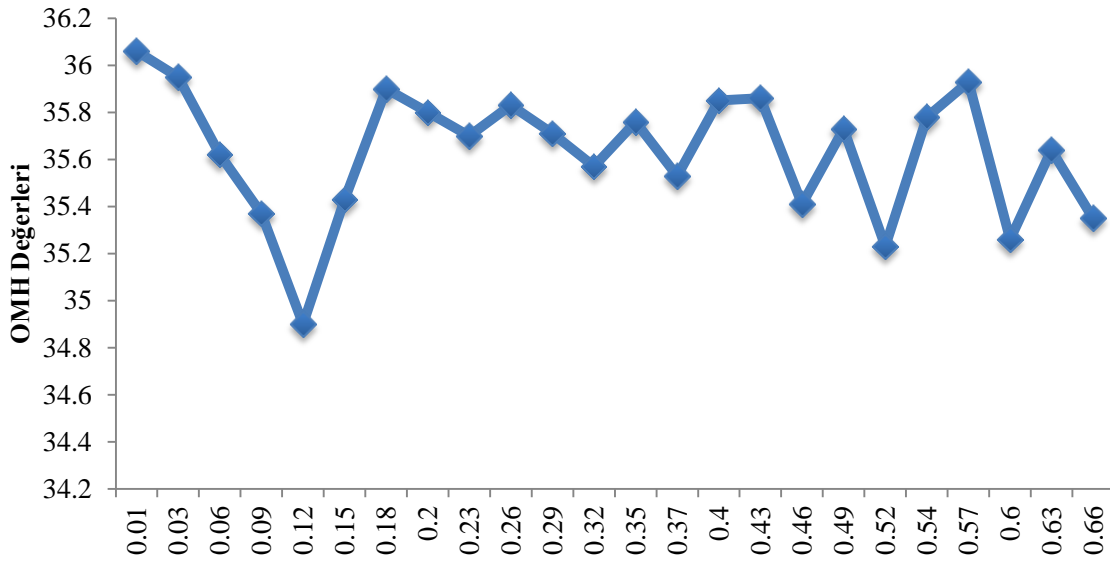
Çizelge 3.3. KR'nin farklı τ değerlerine karşılık gelen HKOK ve OMH değerleri

| τ | HKOK | OMH |
|-------------|--------------|--------------|
| 0,01 | 48,75 | 36,06 |
| 0,03 | 48,63 | 35,95 |
| 0,06 | 48,39 | 35,62 |
| 0,09 | 48,16 | 35,37 |
| 0,12 | 47,78 | 34,90 |
| 0,15 | 48,06 | 35,43 |
| 0,18 | 47,95 | 35,90 |
| 0,20 | 48,44 | 35,80 |
| 0,23 | 48,42 | 35,70 |
| 0,26 | 48,56 | 35,83 |
| 0,29 | 48,34 | 35,71 |
| 0,32 | 48,32 | 35,57 |
| 0,35 | 48,37 | 35,76 |
| 0,37 | 48,36 | 35,53 |
| 0,40 | 48,49 | 35,85 |
| 0,43 | 48,45 | 35,86 |
| 0,46 | 48,23 | 35,41 |
| 0,49 | 48,38 | 35,73 |
| 0,52 | 48,02 | 35,23 |
| 0,54 | 48,42 | 35,78 |
| 0,57 | 48,62 | 35,93 |
| 0,60 | 47,98 | 35,26 |
| 0,63 | 48,57 | 35,64 |
| 0,66 | 48,08 | 35,35 |

HKOK ve OMH değerlerinin τ değerlerine göre değişimi birbirine çok yakın dalgalanmalar göstermiştir (Şekil 3.8, Şekil 3.9).



Şekil 3.8. HKOK değerinin τ değerlerine göre değişim grafiği



Şekil 3.9. OMH değerinin τ değerlerine göre değişim grafiği

Meme kanseri veri setine uygulanan PR yöntemleri sonuçları Çizelge 3.4 - Çizelge 3.7’de verilmiştir. KR analizi sonucunda, minimum hata değerine sahip modelin $\tau=0,12$ ’de olduğu bulundu. Analiz sonuçları Çizelge 3.8’da verildi.

Çizelge 3.4. Meme kanseri veri seti için üstel regresyon sonuçları

| | Bağımsız Değişkenler | | $\hat{\beta}$ | $\hat{\sigma}_{\beta}$ | z | p | HO |
|-------|-----------------------------------|---------------------|---------------|------------------------|--------|------------------|-------|
| Üstel | Yaş (yıl) | | -0,0299 | 0,0086 | -3,474 | <0,001 | 0,970 |
| | Tümör Büyüklüğü (cm) | | -0,1205 | 0,0453 | -2,660 | 0,007 | 0,886 |
| | Progesteron Reseptör Durumu | Negatif | 0,0336 | 0,2786 | 0,121 | 0,904 | 1,034 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | -0,0952 | 0,3326 | -0,286 | 0,775 | 0,909 |
| | | Pozitif ≥ 4 | -0,5020 | 0,3480 | -1,443 | 0,149 | 0,605 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,5819 | 0,2788 | -2,087 | 0,036 | 0,558 |
| | Hormon Terapisi | Var | 0,4441 | 0,2823 | 1,573 | 0,116 | 1,559 |

Çizelge 3.5. Meme kanseri veri seti için Weibull regresyon sonuçları

| | Bağımsız Değişkenler | | $\hat{\beta}$ | $\hat{\sigma}_{\beta}$ | z | p | HO |
|---------|-----------------------------------|---------------------|---------------|------------------------|--------|--------------|-------|
| Weibull | Yaş (yıl) | | -0,0342 | 0,0106 | -3,239 | 0,001 | 0,966 |
| | Tümör Büyüklüğü (cm) | | -0,1330 | 0,0541 | -2,459 | 0,013 | 0,875 |
| | Progesteron Reseptör Durumu | Negatif | 0,0393 | 0,3335 | 0,118 | 0,906 | 1,040 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | -0,1380 | 0,3998 | -0,345 | 0,730 | 0,871 |
| | | Pozitif ≥ 4 | -0,6129 | 0,4227 | -1,450 | 0,147 | 0,542 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,6423 | 0,3377 | -1,902 | 0,057 | 0,526 |
| | Hormon Terapisi | Var | 0,5542 | 0,3437 | 1,613 | 0,107 | 1,741 |

Çizelge 3.6. Meme kanseri veri seti için log-normal regresyon sonuçları

| | Bağımsız Değişkenler | | $\hat{\beta}$ | $\hat{\sigma}_{\beta}$ | z | p | HO |
|------------|-----------------------------------|---------------------|---------------|------------------------|--------|--------------|-------|
| Log-normal | Yaş (yıl) | | -0,0382 | 0,0126 | -3,024 | 0,002 | 0,963 |
| | Tümör Büyüklüğü (cm) | | -0,1155 | 0,0672 | -1,719 | 0,085 | 0,891 |
| | Progesteron Reseptör Durumu | Negatif | -0,2787 | 0,4050 | -0,688 | 0,491 | 0,757 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | -0,1218 | 0,4340 | -0,281 | 0,779 | 0,885 |
| | | Pozitif ≥ 4 | -0,8755 | 0,4862 | -1,800 | 0,071 | 0,417 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,6640 | 0,4025 | -1,650 | 0,099 | 0,515 |
| | Hormon Terapisi | Var | 0,7988 | 0,4149 | 1,925 | 0,054 | 2,223 |

Çizelge 3.7. Meme kanseri veri seti için log-lojistik regresyon sonuçları

| | Bağımsız Değişkenler | | $\hat{\beta}$ | $\hat{\sigma}_{\beta}$ | z | p | HO |
|--------------|--------------------------------------|---------------------|---------------|------------------------|--------|--------------|-------|
| Log-lojistik | Yaş (yıl) | | -0,0364 | 0,0115 | -3,151 | 0,001 | 0,964 |
| | Tümör Büyüklüğü (cm) | | -0,1264 | 0,0588 | -2,150 | 0,031 | 0,881 |
| | Progesteron Reseptör urumu | Negatif | -0,1179 | 0,3817 | -0,309 | 0,757 | 0,889 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | -0,2134 | 0,4068 | -0,525 | 0,600 | 0,808 |
| | | Pozitif ≥ 4 | -0,8397 | 0,4508 | -1,863 | 0,062 | 0,431 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,5730 | 0,3699 | -1,549 | 0,121 | 0,564 |
| | Hormon Terapisi | Var | 0,6894 | 0,3930 | 0,492 | 0,079 | 1,992 |

Çizelge 3.8. Meme kanseri veri seti için $\tau=0,12$ değerindeki KR sonuçları

| Bağımsız Değişkenler | | $\hat{\beta}$ | $\hat{\sigma}_{\hat{\beta}}$ | t | p | HO |
|--------------------------------|------------------|---------------|------------------------------|---------|------------------|-------|
| Yaş (yıl) | | -0,0395 | 0,0148 | -2,6645 | 0,007 | 0,673 |
| Tümör Büyüklüğü (cm) | | 0,0055 | 0,0729 | 0,0758 | 0,939 | 1,005 |
| Progesteron Reseptör Durumu | Negatif | -0,8567 | 0,7204 | -1,1891 | 0,234 | 0,424 |
| Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | -0,2357 | 0,4915 | -0,4795 | 0,631 | 0,790 |
| | Pozitif ≥ 4 | -1,7936 | 0,3970 | -4,5177 | <0,001 | 0,166 |
| Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,7108 | 0,5632 | -1,2619 | 0,206 | 0,491 |
| Hormon Terapisi | Var | 1,2730 | 0,4622 | 2,7540 | 0,005 | 3,571 |

PR yöntemlerinden üstel regresyon sonucunda progesteron reseptör durumu, aksiller lenf nodu tutulumu ve hormon terapisinin bireylerin hastalıksız sağkalım süreleri üzerinde etkili olmadığı; yaş, tümör büyüklüğü ve kapsül dışı lenf nodu tutulumunun ise etkili olduğu bulundu (Çizelge 3.4).

Weibull regresyon sonucunda progesteron reseptör durumu, aksiller lenf nodu tutulumu, hormon terapisi ve kapsül dışı lenf nodu tutulumunun sağkalım üzerinde etkili olmadığı; yaş ve tümör büyüklüğünün ise etkili olduğu bulundu (Çizelge 3.5).

Log-normal regresyon sonucunda, bireylerin sadece yaşlarının hastalıksız sağkalım süreleri üzerinde etkili olduğu bulundu. Log-normal regresyon modelinde tümör büyüklüğü, progesteron reseptör durumu, aksiller lenf nodu tutulumu, kapsül dışı lenf nodu tutulumu ve hormon terapisinin etkili olmadığı bulundu ($p>0,05$) (Çizelge 3.6).

Log-lojistik regresyon sonucunda progesteron reseptör durumu, aksiller lenf nodu tutulumu, hormon terapisi ve kapsül dışı lenf nodu tutulumunun hastalıksız sağkalım süreleri üzerinde etkili olmadığı; yaş ve tümör büyüklüğünün ise etkili olduğu ve log-lojistik regresyon modelinde olması gerektiği tespit edildi (Çizelge 3.7).

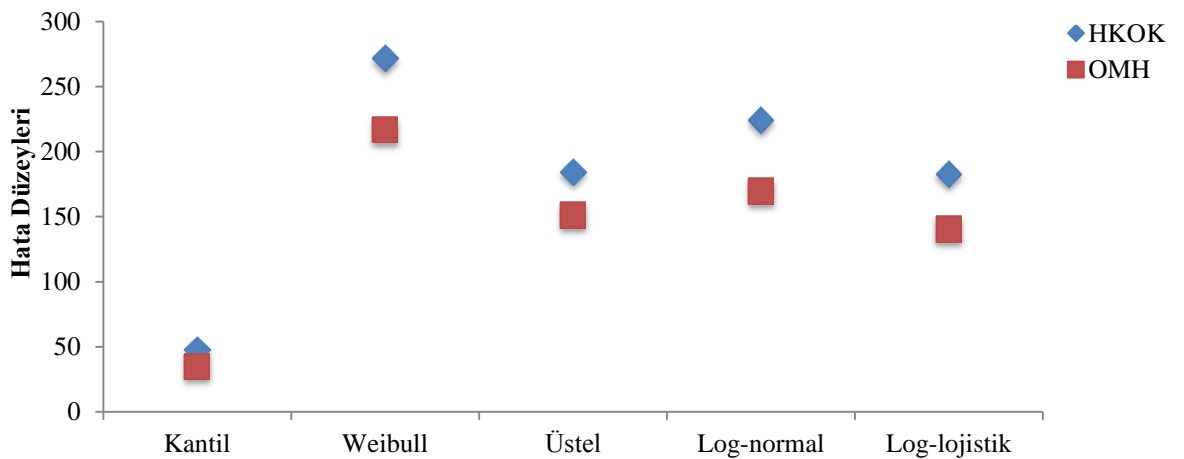
KR yöntemi sonucunda yaş ($p<0,01$), hormon terapisi ($p<0,01$) ve aksiller lenf nodu tutulumunun ($p<0,001$) hastaların hastalıksız sağkalım süreleri üzerinde etkili olduğu

bulundu. Tümör büyüklüğü, progesteron reseptör durumu ve kapsül dışı lenf nodu tutulumunun ise hastalısız sağkalım süreleri üzerinde etkili olmadığı bulundu ($p>0,05$) (Çizelge 3.8).

Meme kanseri teşhisi konan hastalardan oluşan homojen yapıdaki veri setine KR ve PR yöntemleri uygulandığında hastaların hastalısız sağkalım sürelerini minimum hata ile tahmin eden yöntemin KR olduğu bulundu. HKOK kriterine göre karşılaştırma yapıldığında 47,78 değeri ile en uygun modelin KR olduğu tespit edildi. PR modellerinden 182,93 değeri ile minimum hataya sahip olan modelin log-lojistik regresyon modeli olduğu bulundu. OMH kriterine göre karşılaştırma yapıldığında 34,90 hata değeri ile en uygun modelin KR olduğu tespit edildi. PR modellerinden minimum OMH değerine sahip modelin 140,41 ile log-lojistik regresyon modeli olduğu bulundu (Çizelge 3.9, Şekil 3.10).

Çizelge 3.9. Modellerin HKOK ve OMH değerleri

| MODELLER | HKOK | OMH |
|--------------|--------|--------|
| Kantil | 47,78 | 34,90 |
| Weibull | 271,89 | 216,71 |
| Üstel | 184,45 | 151,19 |
| Log-normal | 224,21 | 169,68 |
| Log-lojistik | 182,93 | 140,41 |



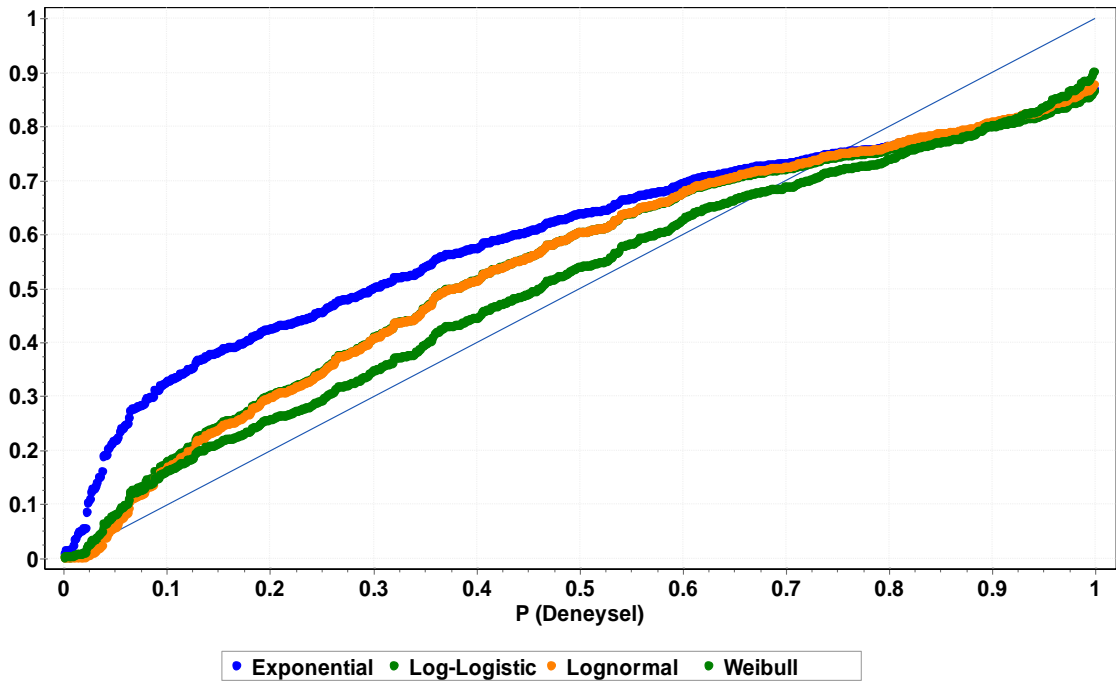
Şekil 3.10. HKOK ve OMH değerlerinin modellere göre değişimi

3.2.2. Gbcs Veri Seti

Gbcs veri setinin heterojen yapıda olup olmadığına karar vermek için K-S test istatistiğinden yararlanılmıştır ve p değerleri verilmiştir. Gbcs veri setindeki bireylerin sağkalım süreleri Weibull, üstel, log-normal ve log-lojistik dağılım göstermediğinden ($p < 0,001$) sağkalım sürelerinin heterojen yapıda olduğu kabul edildi (Çizelge 3.10, Şekil 3.11).

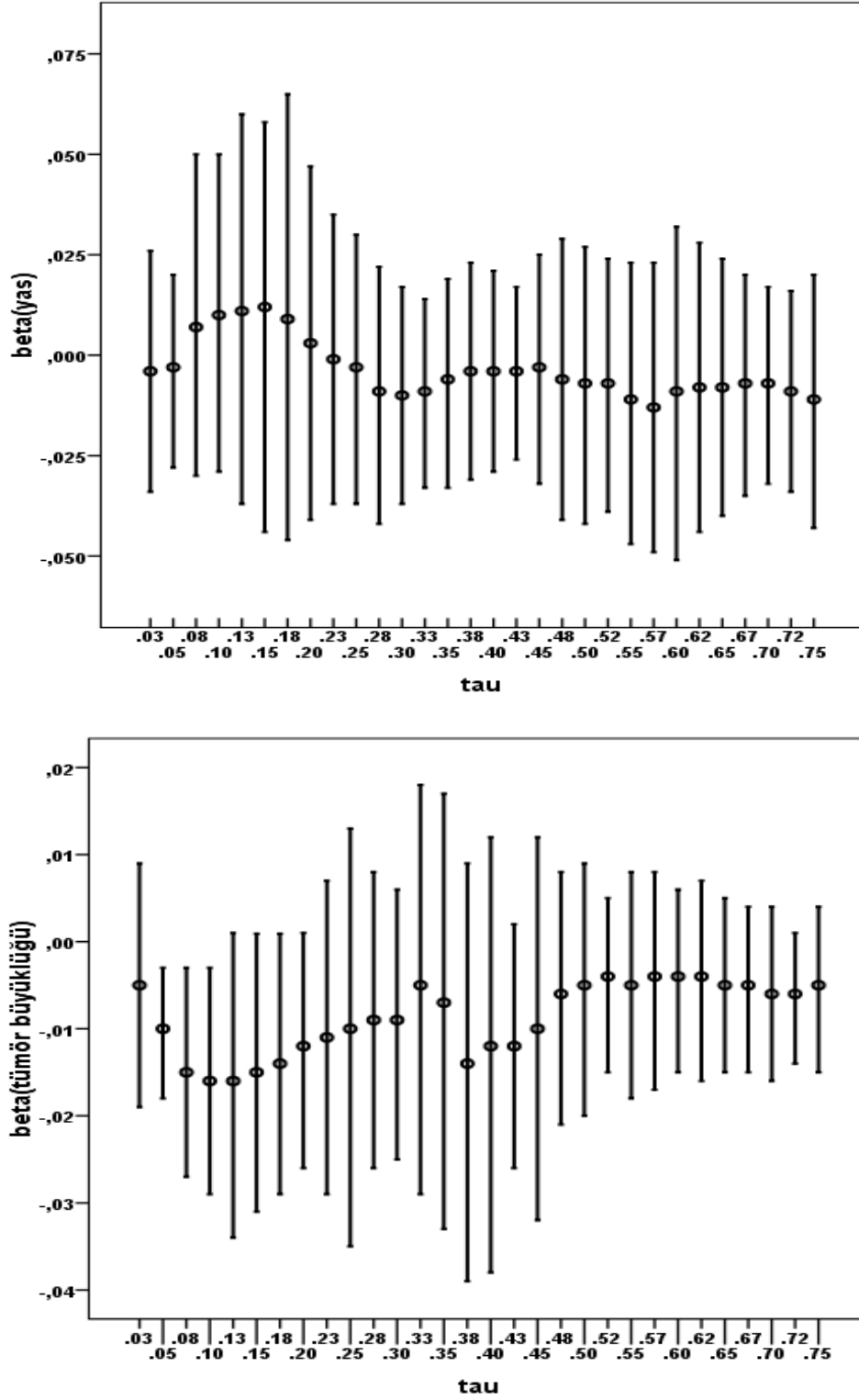
Çizelge 3.10. Gbcs veri setinin K-S test sonuçları

| Dağılımlar | K-S p |
|--------------|--------|
| Weibull | <0,001 |
| Üstel | <0,001 |
| Log-normal | <0,001 |
| Log-lojistik | <0,001 |

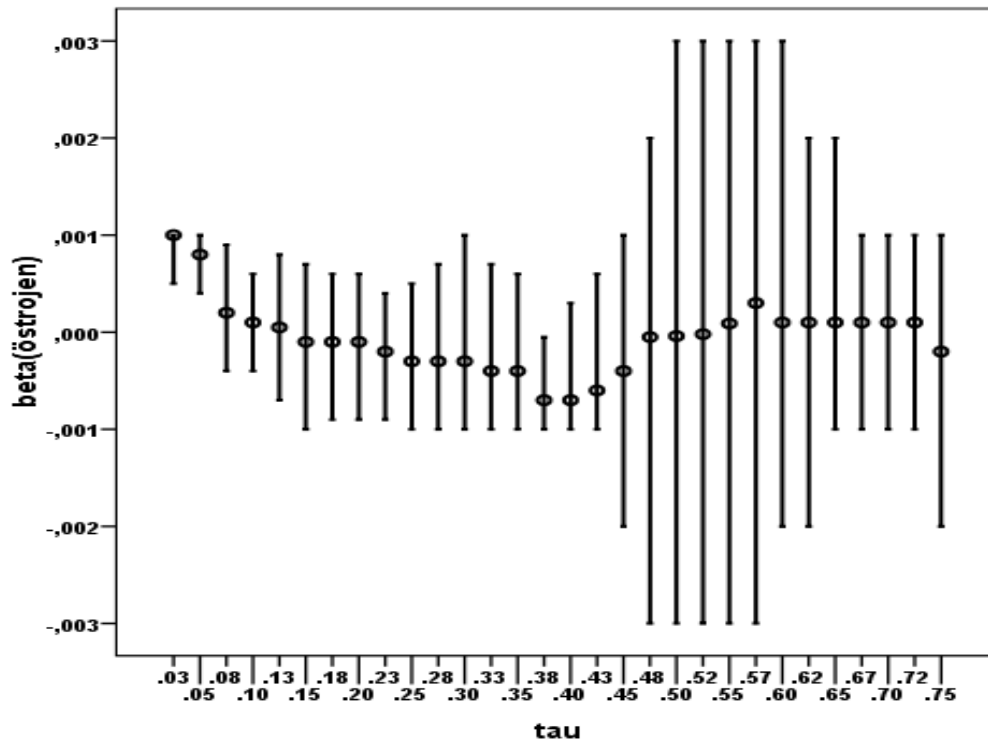
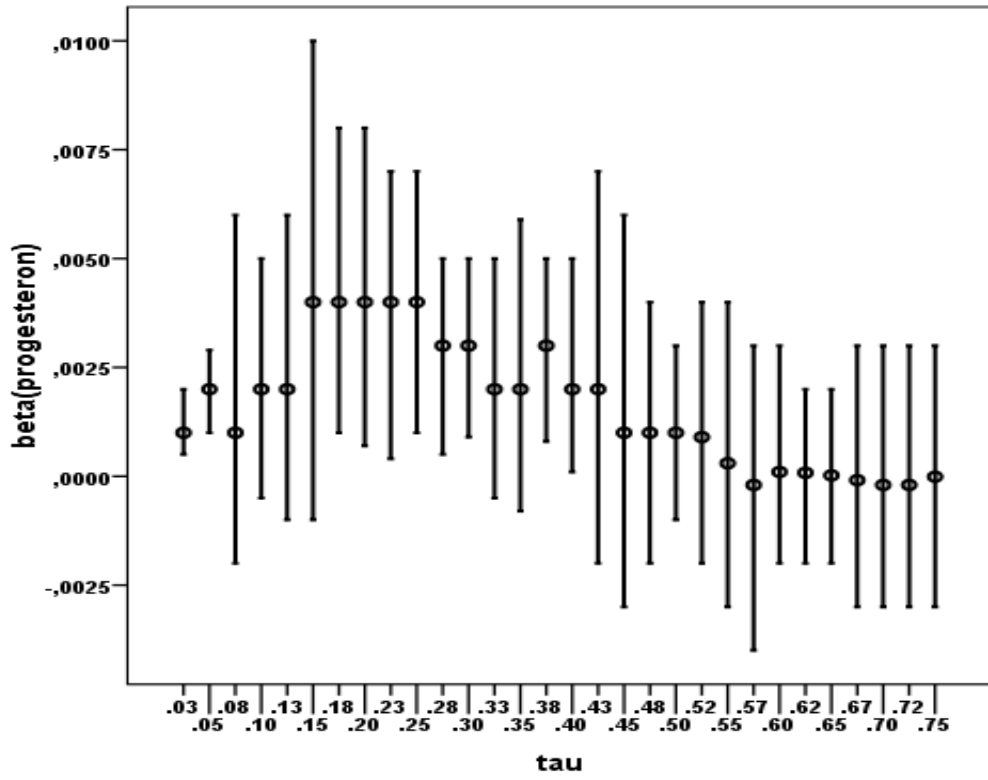


Şekil 3.11. Gbcs veri setindeki meme kanseri tanısı konan hastaların sağkalım süreleri için P-P grafiği

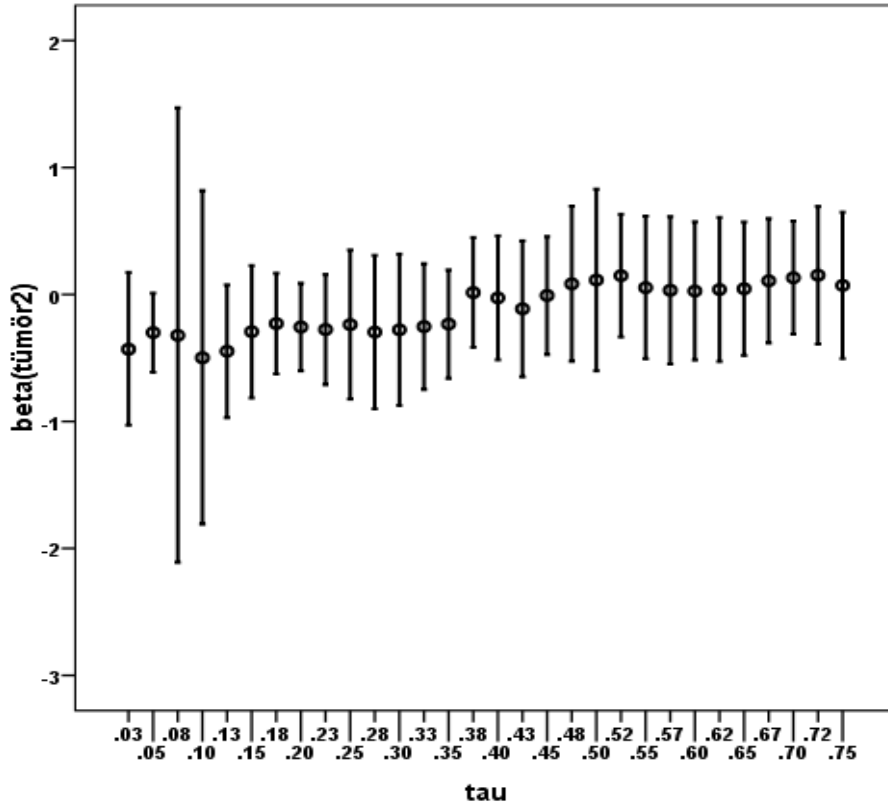
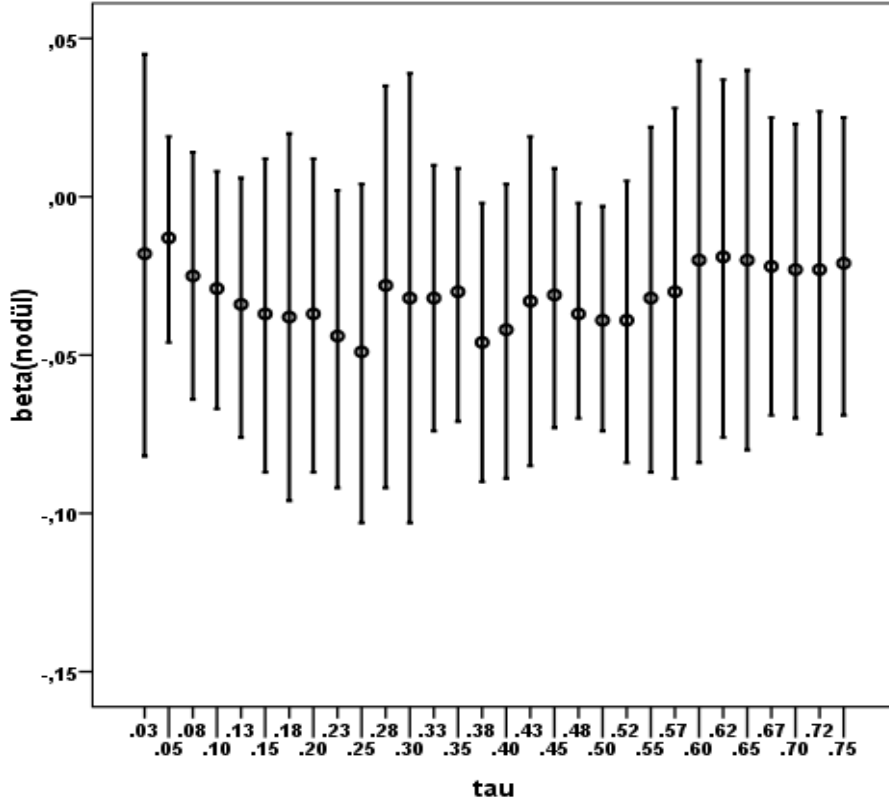
Gbcs veri setinde her bir τ değeri için KR analizi sonuçları Ek 2’de verilmiştir. KR yönteminden her bir τ değeri için parametre tahminlerinin grafikleri Şekil 3.12’de gösterilmiştir.



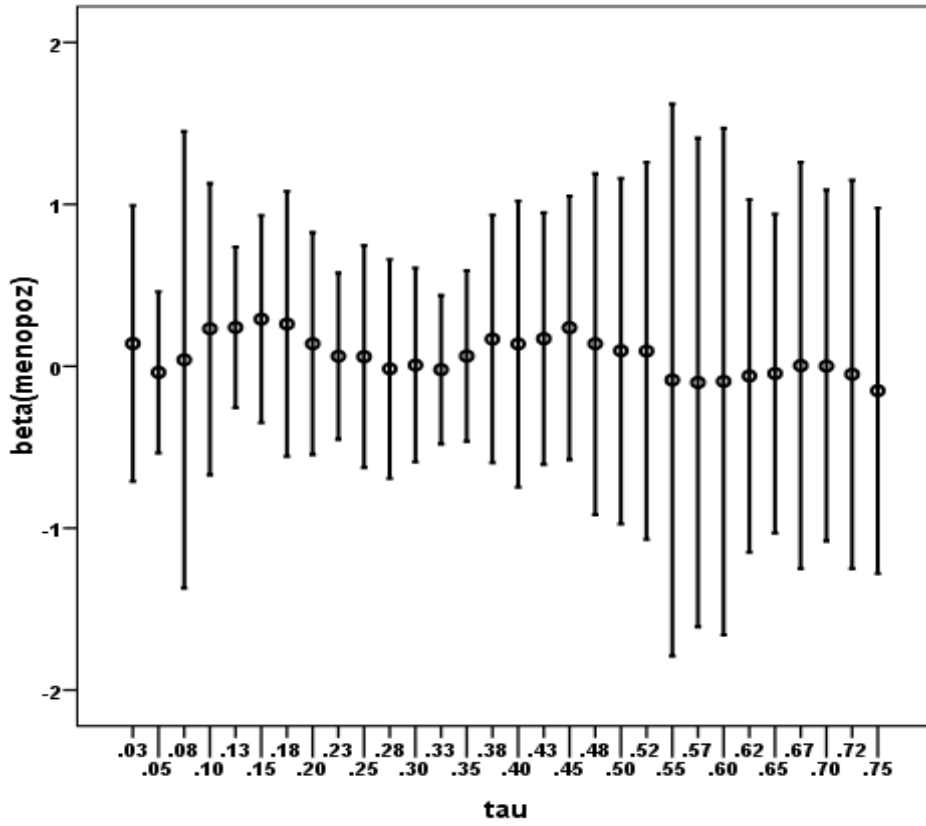
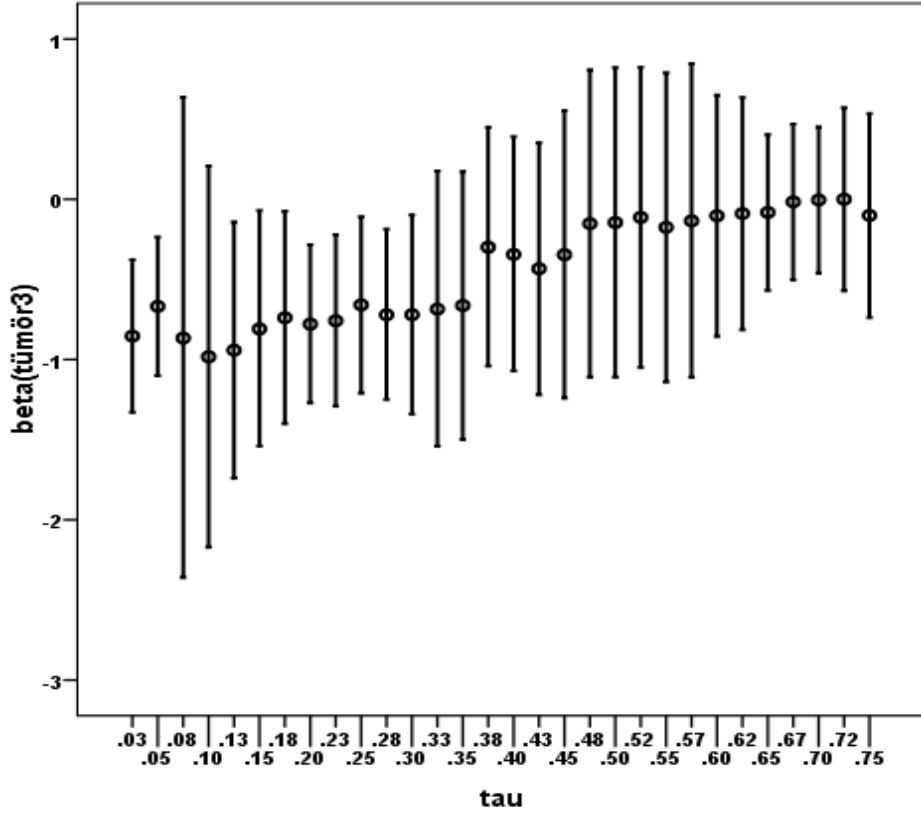
Şekil 3.12. Farklı τ değerleri için parametre tahminlerinin grafiksel gösterimi



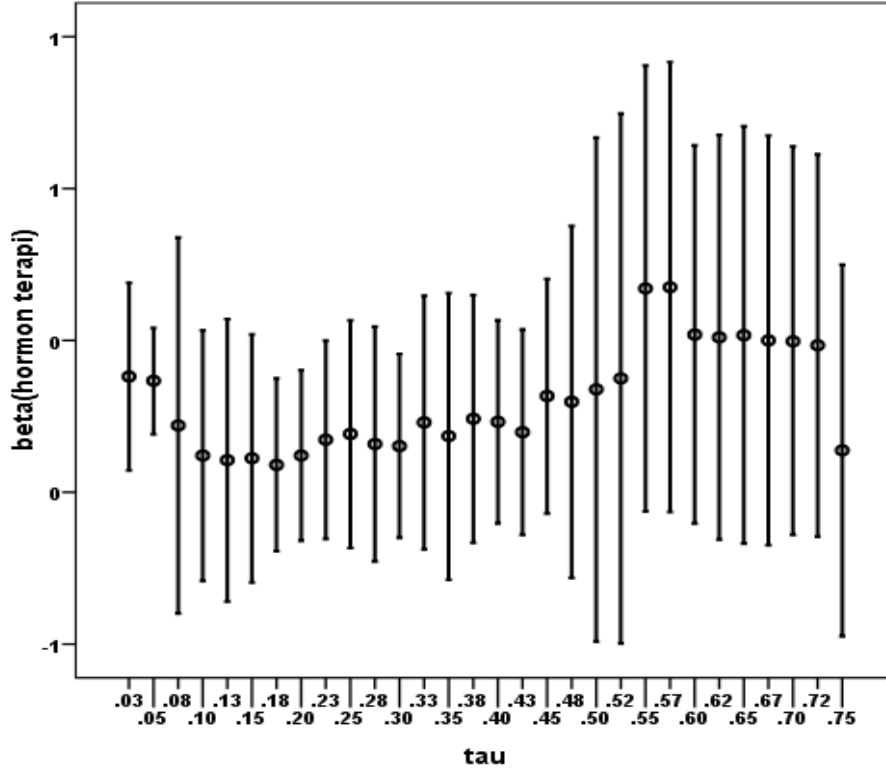
Şekil 3.12. Farklı τ değerleri için parametre tahminlerinin grafiksel gösterimi (devam)



Şekil 3.12. Farklı τ değerleri için parametre tahminlerinin grafiksel gösterimi (devam)



Şekil 3.12. Farklı τ değerleri için parametre tahminlerinin grafiksel gösterimi (devam)



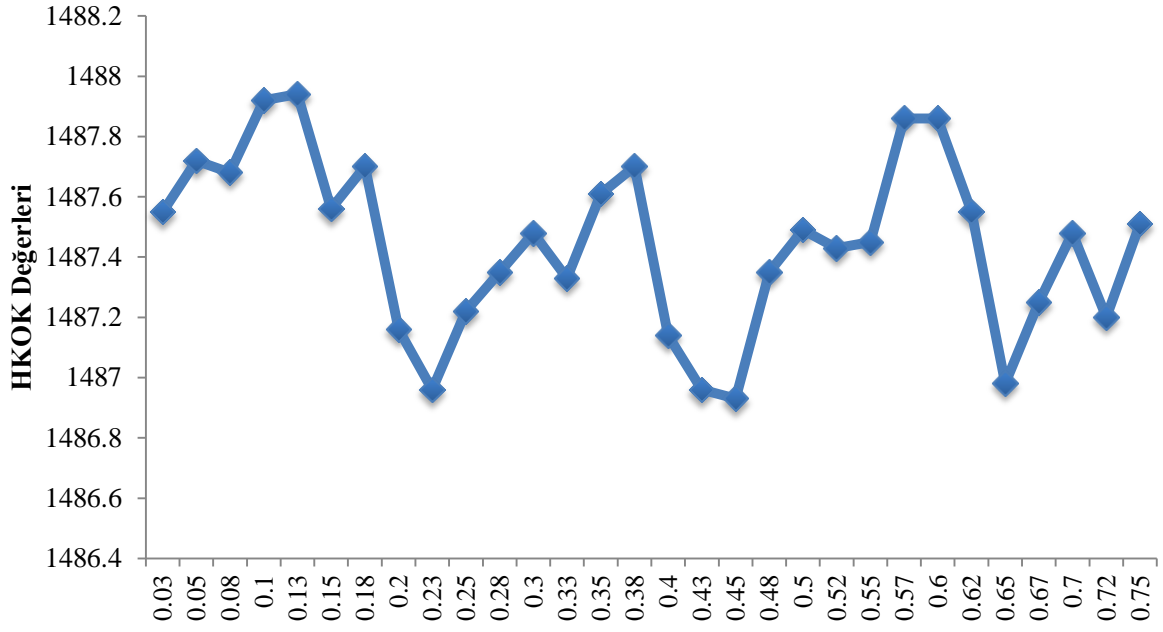
Şekil 3.12. Farklı τ değerleri için parametre tahminlerinin grafiksel gösterimi (devam)

Her bir τ için tahmin edilen modellerin HKOK ve OMH değerleri Çizelge 3.11'de verilmiştir. Farklı τ değerlerine göre parametre tahminleri yapıldı ve HKOK ve OMH değerlerine göre en uygun modelin KR'nin $\tau=0,45$ değerinde olduğu belirlendi. $\tau=0,45$ değerinde HKOK minimum değeri 1486,93 olarak; OMH minimum değeri 1351,44 olarak hesaplanmıştır. τ değerlerinin değişmesi ile HKOK ve OMH değerleri benzer dalgalanmalar göstermiştir (Şekil 3.13, Şekil 3.14).

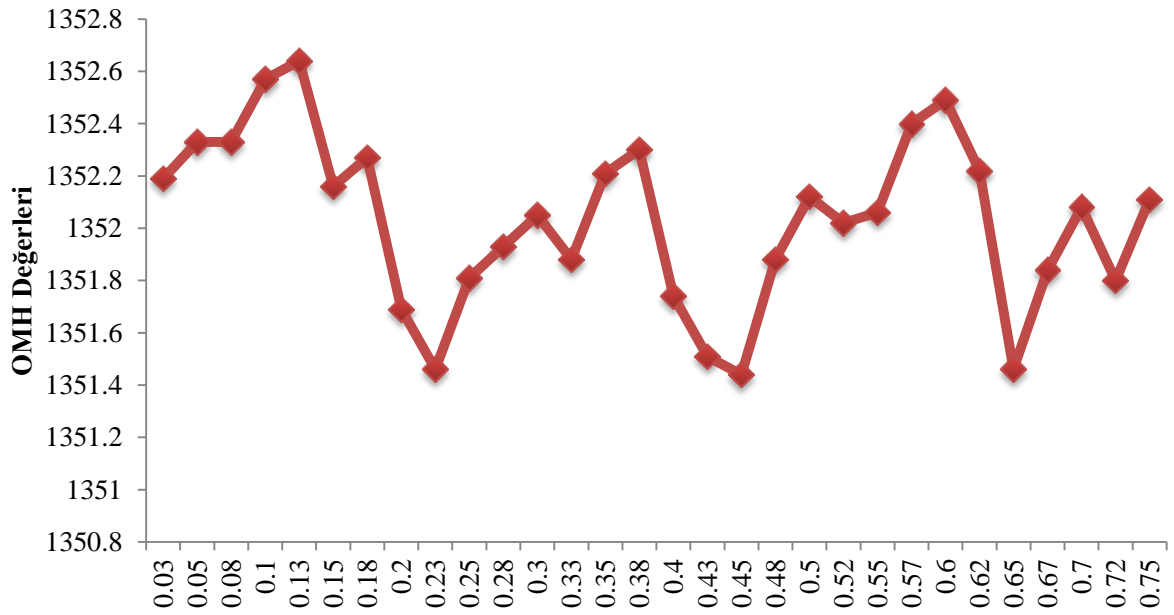
Çizelge 3.11. KR'nin farklı τ değerlerine karşılık gelen HKOK ve OMH değerleri

| τ | HKOK | OMH |
|--------|---------|---------|
| 0,03 | 1487,55 | 1352,19 |
| 0,05 | 1487,72 | 1352,33 |
| 0,08 | 1487,68 | 1352,33 |
| 0,10 | 1487,92 | 1352,57 |
| 0,13 | 1487,94 | 1352,64 |
| 0,15 | 1487,56 | 1352,16 |
| 0,18 | 1487,70 | 1352,27 |
| 0,20 | 1487,16 | 1351,69 |
| 0,23 | 1486,96 | 1351,46 |
| 0,25 | 1487,22 | 1351,81 |
| 0,28 | 1487,35 | 1351,93 |
| 0,30 | 1487,48 | 1352,05 |
| 0,33 | 1487,33 | 1351,88 |
| 0,35 | 1487,61 | 1352,21 |
| 0,38 | 1487,70 | 1352,30 |

| τ | HKOK | OMH |
|-------------|----------------|----------------|
| 0,40 | 1487,14 | 1351,74 |
| 0,43 | 1486,96 | 1351,51 |
| 0,45 | 1486,93 | 1351,44 |
| 0,48 | 1487,35 | 1351,88 |
| 0,50 | 1487,49 | 1352,12 |
| 0,52 | 1487,43 | 1352,02 |
| 0,55 | 1487,45 | 1352,06 |
| 0,57 | 1487,86 | 1352,40 |
| 0,60 | 1487,86 | 1352,49 |
| 0,62 | 1487,55 | 1352,22 |
| 0,65 | 1486,98 | 1351,46 |
| 0,67 | 1487,25 | 1351,84 |
| 0,70 | 1487,48 | 1352,08 |
| 0,72 | 1487,20 | 1351,80 |
| 0,75 | 1487,51 | 1352,11 |



Şekil 3.13. HKOK değerinin τ değerlerine göre değişim grafiği



Şekil 3.14. OMH değerinin τ değerlerine göre değişim grafiği

Gbcs veri setine uygulanan PR yöntemlerinin sonuçları Çizelge 3.12-Çizelge 3.15’de verilmiştir. KR analizi sonunda, minimum hata değerine sahip model $\tau=0,45$ için bulundu ve analiz sonuçları Çizelge 3.16’da verildi.

Çizelge 3.12. Gbcs veri seti için üstel regresyon sonuçları

| | Bağımsız Değişkenler | | $\hat{\beta}$ | $\hat{\sigma}_{\beta}$ | z | p | HO |
|-----------------|-----------------------------|--------|---------------|------------------------|---------|------------------|-------|
| Üstel | Yaş (yıl) | | -0,0022 | 0,0149 | -0,1537 | 0,878 | 0,997 |
| | Tümör Büyüklüğü (mm) | | -0,0116 | 0,0058 | -1,9939 | 0,046 | 0,998 |
| | Progesteron Reseptör Sayısı | | 0,0045 | 0,0013 | 3,3614 | <0,001 | 1,004 |
| | Östrojen Reseptör Sayısı | | 0,00002 | 0,0006 | 0,0398 | 0,968 | 1,000 |
| | Nodül Sayısı | | -0,0510 | 0,0130 | -3,9069 | <0,001 | 0,950 |
| | Tümör Evresi | 2 | -0,5220 | 0,0058 | -1,2067 | 0,228 | 0,593 |
| | | 3 | -0,9680 | 0,4590 | -2,1096 | 0,034 | 0,379 |
| | Menopoz Durumu | Var | -0,2590 | 0,3167 | -0,8187 | 0,413 | 0,771 |
| Hormon Terapisi | Var | 0,3820 | 0,2137 | 1,7872 | 0,073 | 1,465 | |

Çizelge 3.13. Gbcs veri seti için Weibull regresyon sonuçları

| | Bağımsız Değişkenler | | $\hat{\beta}$ | $\hat{\sigma}_{\beta}$ | z | p | HO |
|-----------------|-----------------------------|--------|---------------|------------------------|---------------|------------------|-------|
| Weibull | Yaş (yıl) | | -0,0037 | 0,0088 | -0,421 | 0,674 | 0,996 |
| | Tümör Büyüklüğü (mm) | | -0,0072 | 0,0034 | -2,086 | 0,037 | 0,992 |
| | Progesteron Reseptör Sayısı | | -0,0027 | 0,0008 | 3,282 | 0,001 | 0,997 |
| | Östrojen Reseptör Sayısı | | -0,00005 | 0,0003 | 0,143 | 0,886 | 0,999 |
| | Nodül Sayısı | | -0,0339 | 0,0080 | -4,227 | <0,001 | 0,966 |
| | Tümör Evresi | 2 | -0,3130 | 0,2592 | -1,209 | 0,227 | 0,731 |
| | | 3 | -0,6550 | 0,2766 | -2,369 | 0,017 | 0,519 |
| | Menopoz Durumu | Var | -0,1440 | 0,1902 | -0,758 | 0,449 | 0,865 |
| Hormon Terapisi | Var | 0,2790 | 0,1274 | 2,192 | 0,0284 | 1,321 | |

Çizelge 3.14. Gbcs veri seti için log-normal regresyon sonuçları

| | Bağımsız Değişkenler | | $\hat{\beta}$ | $\hat{\sigma}_{\beta}$ | z | p | HO |
|-----------------|-----------------------------|--------|---------------|------------------------|--------------|------------------|--------|
| Log-normal | Yaş (yıl) | | -0,0015 | 0,0099 | -0,159 | 0,874 | 0,998 |
| | Tümör Büyüklüğü (mm) | | -0,0082 | 0,0042 | -1,930 | 0,053 | 0,991 |
| | Progesteron Reseptör Sayısı | | 0,0026 | 0,0006 | 3,779 | <0,001 | 1,002 |
| | Östrojen Reseptör Sayısı | | 0,0001 | 0,0004 | 0,257 | 0,797 | 1,000 |
| | Nodül Sayısı | | -0,0389 | 0,0109 | -3,551 | <0,001 | 0,961 |
| | Tümör Evresi | 2 | -0,3351 | 0,2389 | -1,402 | 0,161 | 0,7152 |
| | | 3 | -0,7735 | 0,2621 | -2,951 | 0,003 | 0,461 |
| | Menopoz Durumu | Var | -0,1655 | 0,2046 | -0,809 | 0,419 | 0,847 |
| Hormon Terapisi | Var | 0,3051 | 0,1386 | 2,202 | 0,027 | 1,356 | |

Çizelge 3.15. Gbcs veri seti için log-lojistik regresyon sonuçları

| | Bağımsız Değişkenler | | $\hat{\beta}$ | $\hat{\sigma}_{\beta}$ | z | p | HO |
|-----------------|-----------------------------|--------|---------------|------------------------|--------------|------------------|-------|
| Log-lojistik | Yaş (yıl) | | -0,0018 | 0,0096 | -0,189 | 0,85 | 0,998 |
| | Tümör Büyüklüğü (mm) | | -0,0084 | 0,0040 | -2,071 | 0,038 | 0,991 |
| | Progesteron Reseptör Sayısı | | 0,0028 | 0,0007 | 3,509 | <0,001 | 1,002 |
| | Östrojen Reseptör Sayısı | | 0,00002 | 0,0004 | 0,049 | 0,961 | 1,000 |
| | Nodül Sayısı | | -0,0368 | 0,0101 | -3,629 | <0,001 | 0,963 |
| | Tümör Evresi | 2 | -0,3110 | 0,2457 | -1,265 | 0,206 | 0,732 |
| | | 3 | -0,7010 | 0,2678 | -2,618 | 0,008 | 0,496 |
| | Menopoz Durumu | Var | -0,1630 | 0,2025 | -,804 | 0,421 | 0,849 |
| Hormon Terapisi | Var | 0,3020 | 0,1344 | 2,243 | 0,024 | 1,352 | |

Çizelge 3.16. Gbcs veri seti için $\tau=0,45$ değerindeki KR sonuçları

| Bağımsız Değişkenler | | $\hat{\beta}$ | $\hat{\sigma}_{\hat{\beta}}$ | t | p | HO |
|-----------------------------|-----|---------------|------------------------------|---------|--------------|-------|
| Yaş (yıl) | | -0,0031 | 0,0178 | -0,1766 | 0,859 | 0,996 |
| Tümör Büyüklüğü (mm) | | -0,0100 | 0,0084 | -1,1912 | 0,233 | 0,990 |
| Progesteron Reseptör Sayısı | | 0,0016 | 0,0026 | 0,6193 | 0,535 | 1,001 |
| Östrojen Reseptör Sayısı | | -0,0004 | 0,0006 | -0,7499 | 0,453 | 0,999 |
| Nodül Sayısı | | -0,0319 | 0,0140 | -2,2725 | 0,023 | 0,968 |
| Tümör Evresi | 2 | -0,0082 | 0,2800 | -0,0293 | 0,976 | 0,991 |
| | 3 | -0,3468 | 0,3408 | -1,0174 | 0,308 | 0,706 |
| Menopoz Durumu | Var | 0,2387 | 0,4120 | 0,5793 | 0,562 | 1,269 |
| Hormon Terapisi | Var | -0,1839 | 0,2416 | -0,7611 | 0,446 | 0,832 |

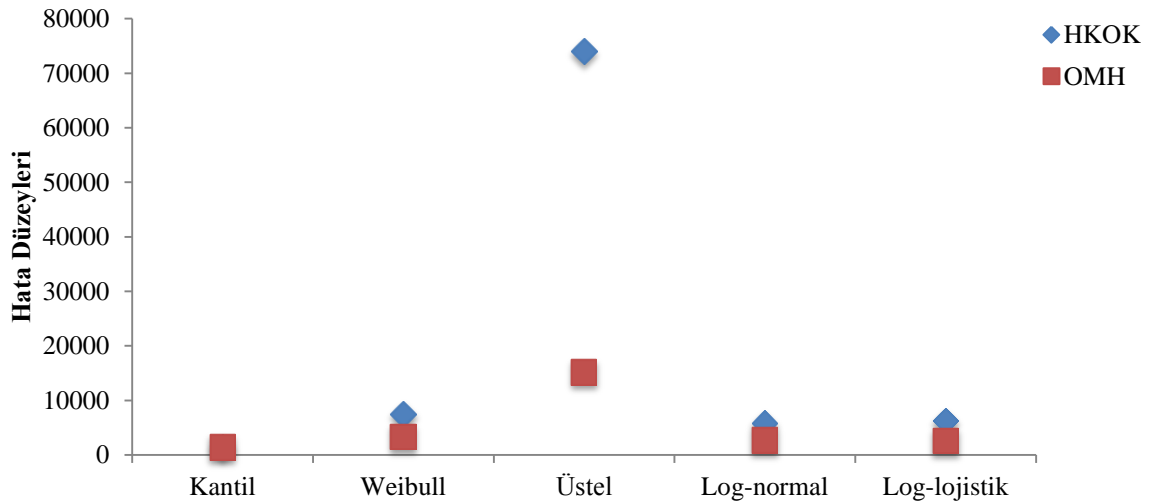
Üstel regresyon ve Weibull regresyon sonuçlarına göre bireylerin sağkalım süreleri üzerinde tümör büyüklüğünün, progesteron reseptör sayısının, nodül sayısının ve tümör evresinin etkili olduğu bulundu (Çizelge 3.12, Çizelge 3.13). Log-normal regresyon sonucunda progesteron reseptör sayısının, nodül sayısının, tümör evresinin ve hormon terapisinin sağkalım süresi üzerinde anlamlı etkiye sahip olduğu bulundu (Çizelge 3.14). Log-lojistik regresyon sonucunda ise tümör büyüklüğünün, nodül sayısının, tümör evresinin ve hormon terapisinin sağkalım süresi üzerinde etkili olduğu bulundu (Çizelge 3.15).

KR yöntemine ($\tau=0,45$ için) göre meme kanseri tanısı konulan hastaların yaş, tümör büyüklüğü, progesteron reseptör sayısı, tümör evresi, menopoz durumu, östrojen reseptör sayısı ve hormon terapisi sağkalım süresi üzerinde etkili olmadığı ($p>0,05$); nodül sayısının ise sağkalım süresi üzerinde etkili olduğu ($p<0,05$) bulundu (Çizelge 3.16).

Meme kanseri teşhisi konan hastalardan oluşan heterojen yapıda olan gbcs veri setine KR ve PR yöntemleri uygulandığında hastaların sağkalım sürelerini minimum hata ile tahmin eden yöntemin KR olduğu bulundu. HKOK kriterine göre minimum tahmini hata değeri 1486,93 ve OMH kriterine göre minimum tahmini hata değeri 1351,44 olan KR yönteminden elde edildi. PR modelleri içerisinde HKOK kriterine göre minimum hataya sahip modelin log-normal regresyon modeli olduğu; OMH kriterine göre hesaplanan minimum tahmini hataya sahip modelin log-lojistik regresyon modeli olduğu bulundu. Ayrıca tüm modeller içerisinde HKOK ve OMH kriterlerine göre hesaplanan hatanın maksimum olduğu modelin üstel regresyon modeli olduğu belirlendi (Çizelge 3.17). HKOK ve OMH değerlerinin KR ve PR modellerine göre hata düzeylerinin dalgalanmaları benzerlik göstermektedir (Şekil 3.15).

Çizelge 3.17. Modellerin HKOK ve OMH değerleri

| MODELLER | HKOK | OMH |
|--------------|----------------|----------------|
| Kantil | 1486,93 | 1351,44 |
| Weibull | 7407,38 | 3286,29 |
| Üstel | 74020,21 | 15136,53 |
| Log-normal | 5796,72 | 2633,56 |
| Log-lojistik | 6246,93 | 2543,32 |



Şekil 3.15. HKOK ve OMH değerlerinin modellere göre değişimi

4. TARTIŞMA

Sağkalım modellerinin tahmin performanslarının karşılaştırılmasına yönelik çok sayıda çalışma yapılmıştır. Fakat sağkalım verilerinin tahmininde KR yönteminin kullanıldığı çalışma sayısı azdır. Bu çalışmada, KR yöntemi ile bireylerin sağkalım süreleri üzerine etki eden risk faktörleri belirlenerek PR yöntemleriyle karşılaştırılması ve minimum hata ile sağkalım sürelerini tahmin eden en uygun regresyon modelinin belirlenmesi hedeflendi.

Literatür incelendiğinde sağkalım modellerinin tahmininde kullanılan KR ile EKK yöntemlerinin karşılaştırıldığı çalışmalar dikkat çekmektedir: Altındağ (2010) çalışmasında, 1999-2006 yılları arasında Konya iline ait aylık toplam güneşlenme süresi ve aylık ortalama bulutluluk süresi verilerini kullanarak KR yöntemi ile EKK yöntemini karşılaştırmıştır. Hata kareler ortalaması (HKO) kriterine göre, minimum hata tahmin değeri 1337,109 olarak EKK yönteminde elde etmiştir. Ortalama mutlak sapma (OMS) kriterine göre, minimum tahmini hata değerini 28,5 olarak KR yönteminde ($\tau=0.50$ için) elde etmiştir. Ayrıca çalışmasında, 2000-2009 yılları arasında aylık ortalama hava sıcaklığını etkileyen faktörleri (hava basıncı (hPa), buhar basıncı (hPa), nisbi nem (%) ve rüzgar hızı (m/sec)) alarak, HKO kriterine göre minimum tahmin hata değerini 0,938 olarak EKK yönteminden elde etmiştir. OMS kriterine göre, minimum tahmin hata değerini, 0,726 olarak KR yönteminden ($\tau=0.50$ için) elde etmiştir (Altındağ 2010). Keskin (2012), Dünya Bankası'nın veri tabanından alınan, 2005-2009 yılları arasında Türkiye'nin de içinde bulunduğu Ekonomik Kalkınma ve İşbirliği Örgütü üyesi 25 ülkeye ait, bağımlı değişkeni beklenen yaşam süresi olan ve 30 bağımsız değişken içeren veri setini kullanarak KR ve EKK yöntemlerinin performanslarını karşılaştırmıştır. Beklenen yaşam süresi çarpık bir dağılıma sahip olduğu için KR yönteminin kullanıldığını rapor etmiştir. Sonuç olarak beklenen yaşam süresini minimum hata ile tahmin eden en iyi yöntemin KR yöntemi ($\tau=0.05$ için) olduğunu bildirmiştir (Keskin 2012). Çelik (2011), TÜİK'in yapmış olduğu 2011 Hanehalkı İşgücü Anketi Mikro veri setini kullanarak KR ve EKK yöntemlerini karşılaştırmıştır. EKK'nın daha genel sonuçlar gösterdiğini, KR'nin ise daha açıklayıcı ve daha belirgin sonuçlar gösterdiğini tespit etmiştir (Çelik 2011). Köse ve ark. (2014), 2-18 yaş grubu sağlıklı çocukların kan basıncı normal değer eğrilerinin polinomial ve KR modelleriyle belirlenmesini amaçlamışlardır. Çalışmalarında kız ve erkek çocukların sistolik ve diyastolik kan basınçlarının KR yönteminde ($\tau=0.5$ için) birbirine yakın olduğu, hipertansiyon tanı değerinin KR yönteminde ($\tau=0.9$ ve $\tau=0.95$ için) elde edilen tahminlerin

polinomiyal regresyona göre daha yüksek olduğu bulgularına ulaşmışlardır (Köse ve ark 2014).

Sağkalım verileri ile ilgili çalışmalar incelendiğinde PR modellerinin yaygın olarak kullanıldığı dikkat çekmektedir. PR modellerinin iyi bir tahmin gerçekleştirilmesi ancak sağkalım sürelerine ilişkin dağılımın doğru bir şekilde tespit edilmesiyle sağlanır. Literatür incelendiğinde: Raman ve Venkatesan (2012)'nin yaptıkları çalışmada 522 meme kanserli hastanın sağkalım süresi üzerinde evre, düşük sayısı, çocuk sayısı, kemoterapi, radyoterapi, ailedeki kanser hikayesi ve menapoz durumu değişkenlerinin etkilerinin incelenmesi amaçlanmış ve 3 farklı PR modeli (Weibull, üstel ve log-normal) uygulanmıştır. İncelenen faktörlerden; evre ve kemoterapi değişkenlerinin sağkalım süresi üzerine etkisinin önemli olduğu saptanmıştır (Raman ve Venkatesan 2012). Hayat ve ark. (2010), meme kanserine yakalanan hastaların sağkalım süresinin belirlenmesinde sadece yaş faktörünü ele alarak beş farklı PR modeli (log-normal, log-lojistik, Weibull, Gamma ve Gompertz) kullanmışlardır. Sağkalım süreleri, Gompertz dağılımına uygun olduğu için Gompertz sağkalım modeliyle yaşın sağkalım süresi üzerine etkisi olup olmadığı araştırılmış ve istatistiksel olarak anlamlı etkiye sahip olmadığı tespit edilmiştir (Hayat ve ark 2010). Chapman ve ark. (2006), 415 meme kanserli hastanın sağkalım süresine yaş, ağırlık, tümör büyüklüğü, histolojik tip, östrojen reseptör durumu, progesteron reseptör durumu, tedavi türü, lenf nodu tutulumu faktörlerinin etkilerini saptamada log-normal sağkalım modeli ile Cox modelini karşılaştırmışlardır. Sonuç olarak alınan faktörlerin sağkalım süresi üzerindeki etkisinin belirlenmesinde log-normal regresyon modelinin daha avantajlı olduğunu tespit etmişlerdir (Chapman ve ark 2006). Avcı ve Bekiroğlu (2014), meme kanseri tanısı konan hastaların sağkalım sürelerine etki eden faktörleri belirlemek için üstel, Weibull, log-lojistik ve log-normal PR modellerini kullanmışlardır. Sağkalım sürelerinin dağılımının log-normal dağılıma uygun olduğunu tespit etmişlerdir. Log-normal regresyon modeline göre, tümör büyüklüğü, TNM evresi IIA, TNM evresi IIB, östrojen reseptör durumu ve aksiller lenf nodu değişkenlerinin sağkalım süresine etkilerini istatistiksel olarak anlamlı bulmuşlardır. PR modelleri, $-2\log L$ ve Akaike bilgi kriteri değerlerine göre karşılaştırma yapıldığında en düşük değerlerin log-normal sağkalım modelinde olduğu bulunmuştur (Avcı ve Bekiroğlu 2014). Ata ve ark. (2008), 236 akciğer kanseri hastaya ait veri setini kullanarak Cox regresyon ve PR modellerini incelemiştir. Hastalar, ameliyat olduktan sonra hastalıklarının ilk kez nüks etmesine kadar geçen süre boyunca izlenmiş ve sağkalım sürelerini etkileyen faktörlerin (yaş, sigara tüketimi,

genişletilmiş rezeksiyon, boyut, invazyon ve patolojik evre) belirlenmesi amaçlanmıştır. Hastalara ait sağkalım sürelerinin dağılımının log-lojistik dağılıma uygun olduğu bulunmuş ve log-lojistik regresyon modeline göre, sigara tüketimi, boyut ve patolojik evre değişkenlerinin sağkalım süreleri üzerinde istatistiksel olarak anlamlı etkiye sahip olduğu tespit edilmiştir. Sağkalım süreleri bir dağılıma uygunluk gösterdiğinde, PR modellerinin daha doğru sonuçlar vereceğini rapor etmişlerdir (Ata ve ark 2008).

Sağkalım sürelerinin heterojen yapıda olduğu veri setleri için, KR yönteminin oldukça avantajlı olduğu ileri sürülmektedir. KR yöntemi, hata dağılımı hakkında herhangi bir varsayıma ihtiyaç duymadığı gibi uç değerlerden etkilenmemektedir. Bu durum da KR yönteminin, robust bir yöntem olduğunu göstermektedir (Koenker 2005; Koenker ve Bassett 1978; Lin ve Rodriguez 2013). Bu nedenle çalışmamızda, heterojen yapıda veri setleri türetilerek KR ve PR yöntemlerinin performansları karşılaştırılmıştır.

Ülkemizde, sağkalım analizinde KR yönteminin kullanıldığı herhangi bir çalışma bulunmamaktadır. Uluslararası çalışmalarda ise bu yöntemin sağkalım analizinde kullanımı yaygın değildir. Lin ve Rodriguez (2013), primer biliyer siroz veri setini kullanarak 418 hastanın sağkalım süresi (gün) üzerine etki eden faktörleri (yaş (yıl), albümin (g/dl), bilirubin (mg/dl), ödem durumu ve protrombin) belirlemeye çalışmışlardır. KR ve PR yöntemlerini kullanarak parametre tahminlerini gerçekleştirmişlerdir. PR yönteminde, sadece protrombin değerlerinin sağkalım süreleri üzerinde etkili olduğunu bulmuşlardır. KR yönteminde ise protrombin değerinin sağkalım süresi üzerinde etkisinin farklı kantil değerlerine göre değiştiğini tespit etmişlerdir. $\tau=0.1, 0.2, 0.3, 0.4$ için protrombini istatistiksel olarak anlamlı olduğu; $\tau=0.5, 0.6, 0.7$ için sağkalım süresi üzerine etkisinin olmadığı bulgusuna ulaşmışlardır. Düşük kantil değerlerinde değişkenlerin anlamlılık seviyelerinin daha düşük olduğunu rapor etmişlerdir (Lin ve Rodriguez 2013). Literatürde sağdan sansürlü sağkalım verileri ele alınarak sağkalım süreleri tahmin edilmiştir (Peng ve Huang 2008; Portnoy 2003; Powell 1984). Koenker (2008) çalışmasında, sansürlü sağkalım verileri için parametre tahmin edicilerinin performanslarını karşılaştırmıştır. Farklı örneklem büyüklüklerine göre yapılan çalışmada OMH ve HKOK kriterlerine göre minimum hataya sahip tahmin edicinin Portnoy ve Peng-Huang olduğunu rapor etmiştir.

Klasik regresyon yöntemi ile KR yönteminin performansları karşılaştırıldığında KR yönteminin yapılan literatür taramasında daha iyi sonuçlar verdiği görülmektedir.

Çalışmamızda da KR yönteminin sağkalım verilerinde uygulanabilirliği ve PR yöntemlerine göre minimum hata ile sağkalım sürelerini tahmin eden bir yöntem olduğu gösterilmiştir.

5. SONUÇ

KR ve PR yöntemleri, farklı büyüklüklerdeki örneklerde gerçekleştirilen simülasyon uygulaması ve sağkalım sürelerinin hem homojen ve hem de heterojen yapıdaki meme kanserli hastalara ilişkin veri setleri uygulamasıyla karşılaştırıldı.

Farklı büyüklükteki örneklerde gerçekleştirilen simülasyon uygulamasında, HKOK ve OMH kriterlerine göre minimum hata değerleri KR modelinde gözlenirken, maksimum hata değerleri PR yöntemlerinden Weibull regresyon modelinde gözlemlendi. PR modelleri kendi içinde HKOK ve OMH kriterlerine göre değerlendirildiğinde ise minimum hata değerlerinin örneklem büyüklüklerine göre değişkenlik gösterdiği belirlendi. $n=100$ olduğu durumda, her iki kritere göre minimum hata değerlerine sahip olan modelin log-normal regresyon modeli olduğu bulundu. $n=250, 500, 750$ olduğu durumlarda ise, HKOK kriterine göre minimum hataya sahip olan modelin log-normal regresyon modeli, OMH kriterine göre minimum hataya sahip olan PR modelinin ise log-lojistik regresyon modeli olduğu belirlendi. $n=1000$ olduğu durumda, her iki kritere göre minimum hata değerinin log-lojistik regresyon modelinde olduğu gözlemlendi.

Homojen yapıdaki 485 meme kanserli hastanın hastalısız sağkalım sürelerine ilişkin uygulamada, KR ve PR yöntemlerinin performansları karşılaştırıldığında, her iki kritere göre minimum hata değerleri KR modelinde ($\tau=0,12$ için) gözlemlendi. Bireylerin hastalısız sağkalım süreleri Weibull dağılımına uygun olduğu halde Weibull regresyon modelinin, sağkalım süresini maksimum hata ile tahmin ettiği gözlemlendi. KR analizi sonucunda yaş ($HO=0,673, p<0,05$) ve aksiller lenf nodu tutulumu (pozitif lenf nodu ≥ 4 için $HO=0,166, p<0,001$) değişkenlerinin hastalısız sağkalım süresi üzerinde önemli etkisinin olduğu belirlendi. Weibull regresyon sonucunda, hastalısız sağkalım süresi üzerinde yaş ($HO=0,966, p=0,001$) ve tümör büyüklüğü ($HO=0,875, p<0,05$) değişkenlerinin etkisinin önemli olduğu bulgusuna ulaşıldı. Üstel regresyon sonucunda yaş ($HO=0,970, p<0,001$), tümör büyüklüğü ($HO=0,886, p<0,01$) ve kapsül dışı lenf nodu tutulumunun (pozitif için $HO=0,558, p<0,05$) hastalısız sağkalım süresi üzerinde önemli etkisi olduğu bulundu. Log-normal regresyon sonucunda, bireylerin sadece yaşının ($HO=0,963, p<0,01$) hastalısız sağkalım süresi üzerinde etkili olduğu bulgusuna ulaşıldı. Log-lojistik regresyon sonucunda ise yaş ($HO=0,964, p<0,001$) ve tümör büyüklüğünün ($HO=0,881, p<0,05$) hastalısız sağkalım süresi üzerinde etkili olduğu belirlendi. Bu

sonuçlara göre yaş değişkeninin, yöntemlerin hepsinde hastaliksız sağkalım süresi üzerinde etkili olduğu bulundu.

Heterojen yapıdaki 686 birimlik gbcs veri seti uygulamasında, her iki kritere göre KR ve PR yöntemlerinin performansları karşılaştırıldığında, minimum hata değerlerinin KR modelinde ($\tau=0,45$ için) olduğu gözlemlendi. KR sonucuna göre bireylerin sağkalım süreleri üzerinde sadece nodül sayısının ($HO=0,968$, $p<0,05$) etkili olduğu belirlendi. Üstel regresyon sonucunda tümör büyüklüğünün ($HO=0,998$, $p<0,05$), progesteron reseptör sayısının ($HO=1,004$, $p<0,001$), nodül sayısının ($HO=0,950$, $p<0,001$) ve tümör evresinin (3.evresi için $HO=0,379$, $p<0,05$) sağkalım süresi üzerinde önemli etkisi olduğu belirlendi. Weibull regresyon sonucunda tümör büyüklüğünün ($HO=0,992$, $p<0,05$), progesteron reseptör sayısının ($HO=1,004$, $p<0,001$), nodül sayısının ($HO=0,966$, $p<0,001$) ve tümör evresinin (3.evresi için $HO=0,519$, $p<0,05$) sağkalım süresi üzerinde önemli etkisi olduğu belirlendi. Log-normal regresyon sonucunda progesteron reseptör sayısının ($HO=1,002$, $p<0,001$), nodül sayısının ($HO=0,961$, $p<0,001$), tümör evresinin (3.evresi için $HO=0,461$, $p<0,05$) ve hormon terapisinin (var için $HO=1,356$, $p<0,05$) sağkalım süresi üzerinde önemli etkisi olduğu belirlendi. Log-lojistik regresyon sonucunda ise tümör büyüklüğünün ($HO=0,991$, $p<0,05$), nodül sayısının ($HO=0,963$, $p<0,001$), tümör evresinin (3.evresi için $HO=0,496$, $p<0,01$) ve hormon terapisinin (var için $HO=1,352$, $p<0,05$) sağkalım süresi üzerinde etkili olduğu; yaş, progesteron reseptör sayısı, östrojen reseptör sayısı ve menopoiz durumunun bireylerin sağkalım süreleri üzerinde anlamlı etkisi olmadığı belirlendi. Bu sonuçlara göre KR ve PR yöntemlerinin her ikisinde de sağkalım süresi üzerinde etkisi olan değişkenin nodül sayısı olduğu belirlendi.

Bu çalışmada, her bir kantil değeri için ayrı ayrı model oluşturan; araştırmacıya, oluşturulan bu modeller arasından minimum hatalı olanı seçme imkanı sunan ve heterojen yapıdaki veri setlerinde kullanışlı olan KR yönteminin, sağkalım süresini PR yöntemlerine göre daha az hata ile tahmin eden bir yöntem olduğu sonucuna ulaşıldı. Bundan sonra yapılacak çalışmalarda, sağkalım analizi yöntemlerinin kullanılacağı tüm veri setleri üzerinde (örneğin; hastanın hastaneden çıkış süresi, bireylere uygulanan aşının bağışıklık sistemine olan etkisinin araştırılması) KR yöntemi de güçlü bir alternatif olarak görülmelidir. Bu çalışmanın, sağkalım verileri ile çalışacak araştırmacılara kılavuz olacağını düşünmekteyiz.

ÖZET

Sağkalım Analizinde Kantil Regresyon ve Parametrik Regresyon Modellerinin Performanslarının Karşılaştırılması. Bu çalışmada, farklı örneklem büyüklüklerine göre simülasyon ile türetilen sağkalım veri setleri ile homojen, heterojen yapıdaki sağkalım verilerinde KR ile PR modellerinin sağkalım sürelerini tahmin etmedeki performanslarının karşılaştırılması ve minimum hataya sahip modelin belirlenmesi hedeflendi.

Ayrıca çalışmanın uygulama aşaması için R yazılım programı kullanılarak simülasyon uygulaması için bir algoritma oluşturuldu. Bu algoritma yardımıyla, heterojen yapıda türetilmiş veri setleri ve gerçek veri setleri üzerinde hem PR modellerinin hem de KR modelinin sağkalım analizindeki performansları, HKOK ve OMH kriterlerine göre karşılaştırıldı.

Yapılan simülasyon çalışması sonunda, farklı örneklem büyüklüklerinde KR modelinin hatalarının, her iki kritere göre minimum olduğu gözlemlendi. PR modelleri arasında, HKOK kriterine göre minimum hata değerine sahip olan modelin log-normal regresyon; OMH kriterine göre minimum hata değerine sahip olan modelin log-lojistik regresyon; her iki kritere göre maksimum hata değerine sahip olan modelin Weibull regresyon modeli olduğu belirlendi.

Homojen ve heterojen yapıdaki meme kanseri veri setlerinin uygulamasında, KR yönteminin performansının PR yöntemlerinden daha iyi olduğu gözlemlendi. Homojen yapıdaki meme kanseri veri seti ve heterojen yapıdaki gbcs veri seti için HKOK ve OMH kriterlerine göre minimum hatanın KR modelinde olduğu belirlendi. PR modelleri arasında, homojen yapıdaki meme kanseri veri setinde HKOK ve OMH kriterleri bakımından minimum hataya sahip olan modelin log-lojistik regresyon modeli olduğu; heterojen yapıdaki gbcs veri setinde ise her iki kritere göre minimum hataya sahip olan modellerin sırasıyla log-normal ve log-lojistik regresyon modelleri olduğu sonucuna ulaşıldı. Sonuç olarak, sağkalım analizinde KR yöntemi, hem homojen hem de heterojen yapıdaki veri setleri için PR yöntemlerine göre daha iyi performans göstermektedir.

Anahtar Kelimeler: Kantil regresyon, Parametrik regresyon, Sağkalım analizi, Simülasyon.

SUMMARY

Comparison of the performances of Quantile Regression and Parametric Regression Models in Survival Analysis. In this study, it is aimed to compare the performances of QR and PR models in predicting survival time in survival dataset created by a simulation according to different sample sizes and the survival dataset which have homogeneous, heterogeneous structure and is aimed to determine a model with minimum error.

Besides, an algorithm is created by using R software for simulation practice in the application step of the study. By means of this algorithm; on real data sets and the data sets produced in heterogeneous structure, both PR models and QR models are compared in terms of the performance in survival analysis according to RMSE and MAE criteria.

In the end of the simulation practice, it is determined that the errors of the QR model are minimum according to both two criteria and in different sample sizes. Among PR models, it is found that the model with minimum error according to RMSE is the log-normal regression; the model with minimum error according to MAE is log-logistic regression and the model with maximum error according to both two criteria is Weibull regression model.

In the application of homogeneous and heterogeneous breast cancer data sets; it is found that the performance of the QR method is better than PR models. For the homogeneous breast cancer data set and heterogeneous gbcs data set; according to RMSE and MAE criteria, the minimum error is found in QR model. Among PR models and according to RMSE and MAE criteria; in homogeneous breast cancer data set, it is found that the model with minimum error is log-logistic; in heterogeneous gbcs data set, it is found that the models with minimum error are log-normal and log-logistic, respectively. As a result; for both homogeneous and heterogeneous data sets, QR method has better performance than PR methods in survival analysis.

Keywords: Parametric regression, Quantile regression, Simulation, Survival analysis.

KAYNAKLAR

Ata N, Karasoy D, Sözer MT. Orantısız hazardlar için parametrik ve yarı parametrik yaşam modelleri. İstatistikçiler Dergisi: İstatistik ve Aktüerya 2008; 1(3): 125-134.

Altındağ İ. Quantile regresyon ve bir uygulama. Yüksek Lisans Tezi. Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya, Türkiye. 2010.

Avcı E, Bekiroğlu N. Parametrik sağkalım modellerinin sağkalım fonksiyonlarının tanıtıl grafikleri yardımıyla karşılaştırılması. Türkiye Klinikleri Journal of Biostatistics 2014; 6(1): 33-41.

Baur D, Saisana M, Schulze N. Modelling the effects of meteorological variables on ozone concentration-a quantile regression approach. Atmospheric Environment 2004; 38(28): 4689-4699.

Buhai S. Quantile regression: overview and selected applications. Ad-Astra-The Young Romanian Scientists' Journal 2004; 4: 1-17.

Chai T, Draxler RR. Root mean square error (RMSE) or mean absolute error (MAE) – Arguments against avoiding RMSE in the literature. Geoscientific Model Development 2014; 7(3): 1247-1250.

Chapman JAW, Lickley HLA, Trudeau ME, Hanna WM, Kahn HJ, Murray D, Pritchard KI. Ascertaining prognosis for breast cancer in node-negative patients with innovative survival analysis. The Breast Journal 2006; 12(1): 37-47.

Collett D. Modelling survival data in medical research. 2nd Ed. Florida: CRC Press; 2003. p. 1-408.

Cox D. Present position and potential developments: some personal views: design of experiments and regression. Journal of the Royal Statistical Society 1984; 147(2): 306-315.

Çelik MY Biyoistatistik bilimsel araştırma SPSS. 1. Baskı. Yusuf Çelik Yayınevi; 2011. p. 399-412.

Çömlekçi N. Deney tasarımı ilke ve teknikleri. 2. Baskı. İstanbul: Bilim Teknik Yayınevi; 1994. p. 136-137.

Efron B. The efficiency of Cox's likelihood function for censored data. Journal of the American Statistical Association 1977; 72(359): 557-565.

- Erişoğlu Ü, Erişoğlu M, Erol H. A mixture model of two different distributions approach to the analysis of heterogeneous survival data. *International Journal of Computational and Mathematical Sciences* 2011; (5): 75-79.
- Hao L, Naiman DQ. *Quantile regression*. 1st Ed. California: Sage publication Inc; 2007. p. 7-55.
- Hayat EA, Suner A, Burak U, Dursun Ö, Orman MN, Kitapçioğlu G. Comparison of five survival models: Breast cancer registry data from ege university cancer research center. *Türkiye Klinikleri Journal of Medical Sciences* 2010; 30(5): 1665-1674.
- Hosmer D, Lemeshow S, May S. *Applied survival analysis: Regression modelling of time to event data*. 2nd Ed. New York: John Wiley & Sons 2008. p. 1-286.
- Ibrahim JG, Chen MH, Sinha D. *Parametric models*. New York: Springer; 2001. p. 30-46.
- Keskin B. Sağlam bir çıkarsama yöntemi: Kantil regresyon. Yüksek Lisans Tezi. Muğla Sıtkı Koçman Üniversitesi Fen Bilimleri Enstitüsü, Muğla, Türkiye. 2012.
- Kleinbaum D, Klein M. *Survival analysis. A self-learning approach*. 2nd Ed. New York: Springer; 2005. p.1-331.
- Koenker R. *Quantile regression*. 1st Ed. New York: Cambridge University Press; 2005. p. 349.
- Koenker R. Censored quantile regression redux. *Journal of Statistical Software* 2008; 27(6): 1-25.
- Koenker R, Bassett Jr G. Regression quantiles. *Econometrica: Journal of the Econometric Society* 1978; 46(1): 33-50.
- Koenker R, Hallock K. Quantile regression: An introduction. *Journal of Economic Perspectives* 2001; 15(4): 43-56.
- Köse T, Keskinöglu P, Özgür S, Keskinöglu P, Orman MN. Sağlıklı çocuklarda kan basıncı eğrilerinin polinomial ve kantil regresyon modelleriyle belirlenmesi. XVI. Ulusal Biyoistatistik Kongresi. 10-12 Eylül 2014, Side Antalya; 2014. p. 91.
- Lawless J. *Statistical methods in reliability*. *Technometrics* 1983; 25(4): 305-316.
- Lee ET, Wang J. *Statistical methods for survival data analysis*. 3rd Ed. John Wiley & Sons; 2003; p. 256-298.

- Leng C, Tong X. A quantile regression estimator for censored data. *Bernoulli* 2013; 19(1): 344-361.
- Lin G, Rodriguez RN. Using the quantlife procedure for survival analysis. SAS Institute Inc 2013; 421.
- Liu X. *Survival analysis: models and applications*. 1st Ed. West Sussex: John Wiley & Sons; 2012. p. 93- 142.
- London D. *Survival models and their estimation*. 3rd Ed. Winsted, Conn: Actex Publications; 1997. p. 177-215.
- Nardi A, Schemper M. Comparing Cox and parametric models in clinical studies. *Statistics in Medicine* 2003; 22(23): 3597-3610.
- Nelson W. *Applied life data analysis*. New York: John Wiley & Sons; 1982. p. 584.
- Ömürlü İK, Özdamar K, Türe M. Comparison of Bayesian survival analysis and Cox regression analysis in simulated and breast cancer data sets. *Expert Systems with Applications* 2009; 36(8): 11341-11346.
- Özdamar K. *SPSS ile Biyoistatistik*. 5. Baskı. Eskişehir: Kaan Kitabevi; 2003. p. 473-502.
- Parzen E. Nonparametric statistical data modeling. *Journal of the American Statistical Association* 1979; 74(365): 105-121.
- Peng L, Huang Y. Survival analysis with quantile regression models. *Journal of the American Statistical Association* 2008; 103(482): 637-649.
- Portnoy S. Censored regression quantiles. *Journal of the American Statistical Association* 2003; 98(464): 1001-1012.
- Powell JL. Least absolute deviations estimation for the censored regression model. *Journal of Econometrics* 1984; 25(3): 303-325.
- Raman T, Venkatesan P. Accelerated failure time frailty model in survival analysis. *International Journal of Science and Technology* 2012; 2(2): 65-90.
- Sauerbrei W, Royston P, Bojar H, Schmoor C, Schumacher M. Modelling the effects of standart prognostic factors in node-positive breast cancer. *British Journal of Cancer* 1999; 79(11/12): 1752-1760. Erişim: ftp://ftp.wiley.com/public/sci_tech_med/survival. Erişim tarihi: 27 Ekim 2014.

Sayegh AS, Munir S, Habeebullah TM. Comparing the performance of statistical models for predicting PM10 concentrations. *Aerosol and Air Quality Research* 2014; 14(3): 653-665.

Şenocak M. Biyoistatistik. 1. Baskı. İstanbul: İstanbul Üniversitesi; 1998. p. 253-267.

Turanlı M, Güriş S. Temel istatistik. 2. Baskı. İstanbul: Der Yayınları; 2005. p. 785.

Wang HJ, Zhou J, Li Y. Variable selection for censored quantile regression. *Statistica Sinica* 2013; 23(1): 145-167.

Willmott CJ, Matsuura K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate research* 2005; 30(1): 79-82.

Yayla ME. Yaşam analizleri ve Cox regresyon modeli. www.academia.edu. Kasım 2013, Erişim Tarihi: 19 Nisan 2015.

Ying Z, Jung S, Wei L. Survival analysis with median regression models. *Journal of the American Statistical Association* 1995; 90(429): 178-184.

ÖZGEÇMİŞ

1990 yılında Muğla'nın Yatağan ilçesinde doğdu. İlk, orta ve lise öğrenimini Yatağan'da tamamladıktan sonra 2009 yılında Karadeniz Teknik Üniversitesi Fen Edebiyat Fakültesi'nde İstatistik ve Bilgisayar Bilimleri bölümünü okumaya hak kazandı ve 2013 yılında mezun oldu. Mezuniyetinden sonra 2013 yılında Adnan Menderes Üniversitesi Sağlık Bilimleri Enstitüsü Biyoistatistik Anabilim Dalı'nda yüksek lisans programına başladı.

TEŐEKKÜR

Yüksek lisans öğrenimim süresince ve tezimi tamamlama aşaması boyunca yardımlarıyla kendilerini yanımda hissettiğim değerli hocalarıma, aileme ve çalışma arkadaşlarıma teşekkürlerimi sunmak istiyorum:

Tez çalışmamın tüm aşamalarında desteklerini hiçbir zaman esirgemeyen, bilgi ve yardımlarıyla yanımda olan danışmanım Sayın Doç. Dr. İmran KURT ÖMÜRLÜ' ye,

Tez çalışmam boyunca değerli bilgilerinden yararlandığım ve tecrübesi ile katkıda bulunan Sayın Prof. Dr. Mevlüt TÜRE' ye,

Tez yazım aşamasında verdikleri moral ve destek sayesinde tezimi bitirmeme yardımcı olan Biyoistatistik Anabilim Dalı'ndaki çalışma arkadaşlarıma,

Fizîken yanımda olmadıkları zamanlarda bile aldığım her nefeste varlıklarını hissettiğim, maddi-manevi desteklerini hiçbir zaman esirgemeyen ve haklarını hiçbir zaman ödeyemeyeceğim annem Hülya TÜRKİŐ; babam Akın TÜRKİŐ ve ablam Filiz TÜRKİŐ'e sonsuz teşekkürlerimi sunarım.

EKLER

Ek 1. Meme kanseri veri setine ait KR'nin farklı τ değerleri için parametre tahminleri

| τ | Parametre | | $\hat{\beta}$ | $\hat{\sigma}_{\beta}$ | t | p |
|-------------|--------------------------------|------------------|---------------|------------------------|---------|--------------|
| 0,01 | Yaş (yıl) | | 0,0010 | 0,0393 | 0,0270 | 0,978 |
| | Tümör Büyüklüğü (cm) | | -0,1836 | 0,4083 | 0,2695 | 0,652 |
| | Progesteron Reseptör Durumu | Negatif | -0,8269 | 0,7034 | -1,1755 | 0,239 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | 0,5763 | 0,5366 | 1,0738 | 0,282 |
| | | Pozitif ≥ 4 | 0,3143 | 1,1660 | 0,2695 | 0,787 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,6700 | 0,9649 | -0,6943 | 0,487 |
| | Hormon Terapisi | Var | 0,0836 | 0,5394 | 0,1550 | 0,876 |
| 0,03 | Yaş (yıl) | | -0,0440 | 0,0352 | -1,2497 | 0,211 |
| | Tümör Büyüklüğü (cm) | | -0,1564 | 0,3112 | -0,5027 | 0,615 |
| | Progesteron Reseptör Durumu | Negatif | -0,7591 | 0,5482 | -1,3846 | 0,166 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | -0,1065 | 1,6495 | -0,0645 | 0,948 |
| | | Pozitif ≥ 4 | -0,9081 | 1,6059 | -0,5654 | 0,571 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,5992 | 1,2892 | -0,4648 | 0,642 |
| | Hormon Terapisi | Var | 0,5191 | 0,7833 | 0,6627 | 0,507 |
| 0,06 | Yaş (yıl) | | -0,0498 | 0,0255 | -1,9516 | 0,050 |
| | Tümör Büyüklüğü (cm) | | -0,0483 | 0,2204 | -0,2194 | 0,826 |
| | Progesteron Reseptör Durumu | Negatif | -0,6976 | 0,3717 | -1,8764 | 0,060 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | -0,2041 | 1,0622 | -0,1922 | 0,847 |
| | | Pozitif ≥ 4 | -1,5258 | 1,0657 | -1,4318 | 0,152 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,8435 | 0,8166 | -1,0329 | 0,301 |
| | Hormon Terapisi | Var | 1,2017 | 0,5664 | 2,1217 | 0,033 |

Ek 1. Meme kanseri veri setine ait KR'nin farklı τ deęerleri için parametre tahminleri
(devam)

| τ | Parametre | | $\hat{\beta}$ | $\hat{\sigma}_{\beta}$ | t | p |
|-----------------|--------------------------------|------------------|---------------|------------------------|--------------|------------------|
| 0,09 | Yaş (yıl) | | -0,0327 | 0,0220 | -1,4883 | 0,136 |
| | Tümör Büyüklüğü (cm) | | -0,0090 | 0,1458 | -0,0621 | 0,950 |
| | Progesteron Reseptör Durumu | Negatif | -0,5485 | 0,2936 | -1,8678 | 0,061 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | -0,3046 | 0,7558 | -0,4029 | 0,686 |
| | | Pozitif ≥ 4 | -1,8502 | 0,7023 | -2,6345 | 0,008 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,7476 | 0,6893 | -1,0845 | 0,278 |
| Hormon Terapisi | Var | 1,0503 | 0,5108 | 2,0559 | 0,039 | |
| 0,12 | Yaş (yıl) | | -0,0395 | 0,0148 | -2,6645 | 0,007 |
| | Tümör Büyüklüğü (cm) | | 0,0055 | 0,0729 | 0,0758 | 0,939 |
| | Progesteron Reseptör Durumu | Negatif | -0,8567 | 0,7204 | -1,1891 | 0,234 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | -0,2357 | 0,4915 | -0,4795 | 0,631 |
| | | Pozitif ≥ 4 | -1,7936 | 0,3970 | -4,5177 | <0,001 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,7108 | 0,5632 | -1,2619 | 0,206 |
| Hormon Terapisi | Var | 1,2730 | 0,4622 | 2,7540 | 0,005 | |
| 0,15 | Yaş (yıl) | | -0,0308 | 0,0241 | -1,2793 | 0,200 |
| | Tümör Büyüklüğü (cm) | | -0,0430 | 0,0866 | -0,4974 | 0,618 |
| | Progesteron Reseptör Durumu | Negatif | -0,5224 | 0,5963 | -0,8761 | 0,380 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | 0,0473 | 0,6036 | 0,0784 | 0,937 |
| | | Pozitif ≥ 4 | -1,2735 | 1,0434 | -1,2204 | 0,222 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,9046 | 0,6728 | -1,3445 | 0,178 |
| Hormon Terapisi | Var | 1,0594 | 0,4910 | 2,1574 | 0,030 | |

Ek 1. Meme kanseri veri setine ait KR'nin farklı τ deęerleri için parametre tahminleri (devam)

| τ | Parametre | | $\hat{\beta}$ | $\hat{\sigma}_{\beta}$ | t | p |
|-----------------|--------------------------------|------------------|---------------|------------------------|---------|--------------|
| 0,18 | Yaş (yıl) | | -0,0196 | 0,0189 | -1,0389 | 0,298 |
| | Tümör Büyüklüğü (cm) | | -0,0816 | 0,0659 | -1,2368 | 0,216 |
| | Progesteron Reseptör Durumu | Negatif | -0,4599 | 0,8695 | -0,5289 | 0,596 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | -0,3415 | 0,5670 | -0,6022 | 0,547 |
| | | Pozitif ≥ 4 | -0,9446 | 0,8634 | -1,0939 | 0,273 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,6953 | 0,6001 | -1,1586 | 0,246 |
| Hormon Terapisi | Var | 0,7878 | 1,0507 | 0,7498 | 0,453 | |
| 0,20 | Yaş (yıl) | | -0,0222 | 0,0147 | -1,5070 | 0,131 |
| | Tümör Büyüklüğü (cm) | | -0,0842 | 0,0589 | -1,4290 | 0,153 |
| | Progesteron Reseptör Durumu | Negatif | -0,3058 | 0,8213 | -0,3723 | 0,709 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | -0,3518 | 0,4860 | -0,7238 | 0,469 |
| | | Pozitif ≥ 4 | -0,9186 | 0,7706 | -1,1919 | 0,233 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,5788 | 0,5145 | -1,1248 | 0,260 |
| Hormon Terapisi | Var | 0,7453 | 0,9339 | 0,7979 | 0,424 | |
| 0,23 | Yaş (yıl) | | -0,0223 | 0,0119 | -1,8710 | 0,061 |
| | Tümör Büyüklüğü (cm) | | -0,1107 | 0,0476 | -2,3246 | 0,020 |
| | Progesteron Reseptör Durumu | Negatif | -0,0989 | 0,4898 | -0,2020 | 0,839 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | -0,1506 | 0,4938 | -0,3050 | 0,760 |
| | | Pozitif ≥ 4 | -0,7013 | 0,6557 | -1,0695 | 0,284 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,5820 | 0,4273 | -1,3619 | 0,173 |
| Hormon Terapisi | Var | 0,4576 | 0,6598 | 0,6934 | 0,488 | |

Ek 1. Meme kanseri veri setine ait KR'nin farklı τ deęerleri için parametre tahminleri
(devam)

| τ | Parametre | | $\hat{\beta}$ | $\hat{\sigma}_{\beta}$ | t | p |
|-----------------|--------------------------------|------------------|---------------|------------------------|---------|------------------|
| 0,26 | Yaş (yıl) | | -0,0247 | 0,0143 | -1,7290 | 0,083 |
| | Tümör Büyüklüğü (cm) | | -0,1179 | 0,0295 | -3,9884 | <0,001 |
| | Progesteron Reseptör Durumu | Negatif | 0,0813 | 0,2372 | 0,3429 | 0,731 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | -0,2182 | 0,4678 | -0,4665 | 0,640 |
| | | Pozitif ≥ 4 | -0,7186 | 0,6286 | -1,1432 | 0,252 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,4967 | 0,4038 | -1,2300 | 0,218 |
| Hormon Terapisi | Var | 0,4119 | 0,4919 | 0,8373 | 0,402 | |
| 0,29 | Yaş (yıl) | | -0,0381 | 0,0119 | -3,1949 | 0,001 |
| | Tümör Büyüklüğü (cm) | | -0,1200 | 0,0242 | -4,9495 | <0,001 |
| | Progesteron Reseptör Durumu | Negatif | 0,1250 | 0,2462 | 0,5077 | 0,611 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | -0,2156 | 0,3212 | -0,6714 | 0,501 |
| | | Pozitif ≥ 4 | -0,8958 | 0,3579 | -2,5032 | 0,012 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,4953 | 0,4106 | -1,2061 | 0,227 |
| Hormon Terapisi | Var | 0,2316 | 0,4145 | 0,5586 | 0,576 | |
| 0,32 | Yaş (yıl) | | -0,0384 | 0,0124 | -3,0858 | 0,002 |
| | Tümör Büyüklüğü (cm) | | -0,1278 | 0,0274 | -4,6601 | <0,001 |
| | Progesteron Reseptör Durumu | Negatif | 0,2059 | 0,4217 | 0,4883 | 0,625 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | -0,1983 | 0,2805 | -0,7072 | 0,479 |
| | | Pozitif ≥ 4 | -0,8827 | 0,3032 | -2,9112 | 0,003 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,5295 | 0,3831 | -1,3820 | 0,166 |
| Hormon Terapisi | Var | 0,1755 | 0,4137 | 0,4242 | 0,671 | |

Ek 1. Meme kanseri veri setine ait KR'nin farklı τ deęerleri için parametre tahminleri (devam)

| τ | Parametre | | $\hat{\beta}$ | $\hat{\sigma}_{\beta}$ | t | p |
|-----------------|--------------------------------|------------------|---------------|------------------------|---------|------------------|
| 0,35 | Yaş (yıl) | | -0,0372 | 0,0137 | -2,7015 | 0,006 |
| | Tümör Büyüklüğü (cm) | | -0,1410 | 0,0391 | -3,6051 | <0,001 |
| | Progesteron Reseptör Durumu | Negatif | 0,2020 | 0,4815 | 0,4194 | 0,674 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | -0,2919 | 0,3392 | -0,8605 | 0,389 |
| | | Pozitif ≥ 4 | -0,8948 | 0,5019 | -1,7827 | 0,074 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,3779 | 0,4113 | -0,9187 | 0,358 |
| Hormon Terapisi | Var | 0,1179 | 0,5474 | 0,2153 | 0,829 | |
| 0,37 | Yaş (yıl) | | -0,0341 | 0,0187 | -1,8159 | 0,069 |
| | Tümör Büyüklüğü (cm) | | -0,1218 | 0,0650 | -1,8740 | 0,060 |
| | Progesteron Reseptör Durumu | Negatif | 0,2101 | 0,5195 | 0,4044 | 0,685 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | -0,2752 | 0,4174 | -0,6594 | 0,509 |
| | | Pozitif ≥ 4 | -0,8955 | 0,5759 | -1,5548 | 0,119 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,1324 | 0,4892 | -0,2708 | 0,786 |
| Hormon Terapisi | Var | 0,2991 | 0,6379 | 0,4689 | 0,639 | |
| 0,40 | Yaş (yıl) | | -0,0313 | 0,0278 | -1,1235 | 0,261 |
| | Tümör Büyüklüğü (cm) | | -0,1349 | 0,0889 | -1,5173 | 0,129 |
| | Progesteron Reseptör Durumu | Negatif | 0,4237 | 0,5417 | 0,7822 | 0,434 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | -0,2856 | 0,5242 | -0,5447 | 0,585 |
| | | Pozitif ≥ 4 | -0,8792 | 0,8203 | -1,0717 | 0,283 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,1922 | 0,6328 | -0,3038 | 0,761 |
| Hormon Terapisi | Var | 0,1404 | 0,6874 | 0,2043 | 0,838 | |

Ek 1. Meme kanseri veri setine ait KR'nin farklı τ deęerleri için parametre tahminleri
(devam)

| τ | Parametre | | $\hat{\beta}$ | $\hat{\sigma}_{\beta}$ | t | p |
|-------------|--------------------------------|------------------|---------------|------------------------|---------|--------------|
| 0,43 | Yaş (yıl) | | -0,0429 | 0,0300 | -1,4293 | 0,152 |
| | Tümör Büyüklüğü (cm) | | -0,1378 | 0,0858 | -1,6064 | 0,108 |
| | Progesteron Reseptör Durumu | Negatif | 0,2966 | 0,5847 | 0,5072 | 0,611 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | -0,2313 | 0,5736 | -0,4033 | 0,686 |
| | | Pozitif ≥ 4 | -0,7054 | 1,006 | -0,7007 | 0,483 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,4852 | 0,9298 | -0,5219 | 0,601 |
| | Hormon Terapisi | Var | 0,2948 | 0,9042 | 0,3261 | 0,744 |
| 0,46 | Yaş (yıl) | | -0,0470 | 0,0218 | -2,1514 | 0,031 |
| | Tümör Büyüklüğü (cm) | | -0,1382 | 0,0894 | -1,5462 | 0,122 |
| | Progesteron Reseptör Durumu | Negatif | 0,1303 | 0,4616 | 0,2822 | 0,777 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | -0,2111 | 0,6272 | -0,3366 | 0,736 |
| | | Pozitif ≥ 4 | -0,5365 | 0,8691 | -0,6172 | 0,537 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,6547 | 0,6863 | -0,9539 | 0,340 |
| | Hormon Terapisi | Var | 0,5693 | 0,5496 | 1,0358 | 0,300 |
| 0,49 | Yaş (yıl) | | -0,0505 | 0,0183 | -2,7638 | 0,005 |
| | Tümör Büyüklüğü (cm) | | -0,1654 | 0,0831 | -1,9885 | 0,046 |
| | Progesteron Reseptör Durumu | Negatif | -0,0098 | 0,4380 | -0,0225 | 0,981 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | -0,1577 | 0,5626 | -0,2803 | 0,779 |
| | | Pozitif ≥ 4 | -0,4185 | 0,6383 | -0,6557 | 0,512 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,7535 | 0,5173 | -1,4565 | 0,145 |
| | Hormon Terapisi | Var | 0,7435 | 0,3856 | 1,9281 | 0,053 |

Ek 1. Meme kanseri veri setine ait KR'nin farklı τ deęerleri için parametre tahminleri
(devam)

| τ | Parametre | | $\hat{\beta}$ | $\hat{\sigma}_{\beta}$ | t | p |
|-----------------|--------------------------------|------------------|---------------|------------------------|---------|--------------|
| 0,52 | Yaş (yıl) | | -0,0499 | 0,0225 | -2,2178 | 0,026 |
| | Tümör Büyüklüğü (cm) | | -0,1674 | 0,0919 | -1,8217 | 0,068 |
| | Progesteron Reseptör Durumu | Negatif | -0,0674 | 0,5921 | -0,1138 | 0,909 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | -0,0920 | 0,7081 | -0,1299 | 0,896 |
| | | Pozitif ≥ 4 | -0,3708 | 0,6645 | -0,5581 | 0,576 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,7614 | 0,5862 | -1,2989 | 0,193 |
| Hormon Terapisi | Var | 0,7647 | 0,4958 | 1,5423 | 0,122 | |
| 0,54 | Yaş (yıl) | | -0,0471 | 0,0256 | -1,8361 | 0,066 |
| | Tümör Büyüklüğü (cm) | | -0,1739 | 0,0958 | -1,8158 | 0,069 |
| | Progesteron Reseptör Durumu | Negatif | -0,0882 | 0,6558 | -0,1346 | 0,892 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | -0,0560 | 0,7225 | -0,0776 | 0,938 |
| | | Pozitif ≥ 4 | -0,3495 | 0,6274 | -0,5570 | 0,577 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,8116 | 0,7922 | -1,0243 | 0,305 |
| Hormon Terapisi | Var | 0,8016 | 0,6121 | 1,3095 | 0,190 | |
| 0,57 | Yaş (yıl) | | -0,0477 | 0,0329 | -1,4471 | 0,147 |
| | Tümör Büyüklüğü (cm) | | -0,1829 | 0,0844 | -2,1650 | 0,030 |
| | Progesteron Reseptör Durumu | Negatif | -0,2781 | 0,8669 | -0,3208 | 0,748 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | 0,0213 | 0,7494 | 0,0285 | 0,977 |
| | | Pozitif ≥ 4 | -0,2558 | 0,8092 | -0,3161 | 0,751 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,7715 | 0,9592 | -0,8042 | 0,421 |
| Hormon Terapisi | Var | 0,8728 | 0,7545 | 1,1567 | 0,247 | |

Ek 1. Meme kanseri veri setine ait KR'nin farklı τ deęerleri için parametre tahminleri
(devam)

| τ | Parametre | | $\hat{\beta}$ | $\hat{\sigma}_{\beta}$ | t | p |
|-----------------|--------------------------------|------------------|---------------|------------------------|---------|--------------|
| 0,60 | Yaş (yıl) | | -0,0606 | 0,0370 | -1,6378 | 0,101 |
| | Tümör Büyüklüğü (cm) | | -0,2032 | 0,0995 | -2,0422 | 0,041 |
| | Progesteron Reseptör Durumu | Negatif | 0,0476 | 1,1734 | 0,0405 | 0,967 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | 0,2223 | 0,7740 | 0,2873 | 0,773 |
| | | Pozitif ≥ 4 | 0,4051 | 1,2929 | 0,3133 | 0,754 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -1,7163 | 1,2573 | -1,3650 | 0,172 |
| Hormon Terapisi | Var | 1,7219 | 1,2393 | 1,3893 | 0,164 | |
| 0,63 | Yaş (yıl) | | -0,0363 | 0,0333 | -1,0890 | 0,276 |
| | Tümör Büyüklüğü (cm) | | -0,1726 | 0,1331 | -1,2960 | 0,194 |
| | Progesteron Reseptör Durumu | Negatif | 0,3172 | 1,3406 | 0,2366 | 0,812 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | -0,2176 | 1,1013 | -0,1975 | 0,843 |
| | | Pozitif ≥ 4 | 0,2496 | 1,6144 | 0,1546 | 0,877 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,9006 | 1,5773 | -0,5709 | 0,568 |
| Hormon Terapisi | Var | 0,9971 | 1,3782 | 0,723 | 0,469 | |
| 0,66 | Yaş (yıl) | | -0,0388 | 0,0315 | -1,2305 | 0,218 |
| | Tümör Büyüklüğü (cm) | | -0,1674 | 0,0991 | -1,6896 | 0,091 |
| | Progesteron Reseptör Durumu | Negatif | 0,6703 | 1,3349 | 0,5021 | 0,615 |
| | Aksiller Lenf Nodu Tutulumu | Pozitif 1-3 | -0,1646 | 0,8551 | -0,1925 | 0,847 |
| | | Pozitif ≥ 4 | 0,8126 | 1,3613 | 0,5969 | 0,550 |
| | Kapsül Dışı Lenf Nodu Tutulumu | Pozitif | -0,9169 | 1,2071 | -0,7595 | 0,447 |
| Hormon Terapisi | Var | 0,5611 | 1,1938 | 0,4700 | 0,638 | |

Ek 2. Gbcs veri setine ait KR'nin farklı τ deęerleri için parametre tahminleri

| τ | Bağımsız Deęişkenler | | $\hat{\beta}$ | $\hat{\sigma}_{\beta}$ | t | p |
|-------------|-----------------------------|-----|---------------|------------------------|---------|------------------|
| 0,03 | Yaş (yıl) | | -0,040 | 0,0134 | -0,3007 | 0,763 |
| | Tümör Büyüklüğü (mm) | | -0,0051 | 0,0080 | -0,6352 | 0,525 |
| | Progesteron Reseptör Sayısı | | 0,0016 | 0,006 | 2,3255 | 0,020 |
| | Östrojen Reseptör Sayısı | | 0,0010 | 0,0002 | 4,1176 | <0,001 |
| | Nodül Sayısı | | -0,0185 | 0,0361 | -0,5120 | 0,608 |
| | Tümör Evresi | 2 | -0,4314 | 0,2647 | -1,6300 | 0,103 |
| | | 3 | -0,8548 | 0,2209 | -3,8695 | <0,001 |
| | Menopoz Durumu | Var | 0,1408 | 0,3482 | 0,4043 | 0,685 |
| | Hormon Terapisi | Var | -0,1194 | 0,1637 | -0,7294 | 0,465 |
| 0,05 | Yaş (yıl) | | -0,0036 | 0,0098 | -0,3744 | 0,708 |
| | Tümör Büyüklüğü (mm) | | -0,0108 | 0,0043 | -2,4911 | 0,012 |
| | Progesteron Reseptör Sayısı | | 0,0021 | 0,0005 | 3,7866 | <0,001 |
| | Östrojen Reseptör Sayısı | | 0,0008 | 0,0001 | 5,3873 | <0,001 |
| | Nodül Sayısı | | -0,0137 | 0,0223 | -0,6179 | 0,536 |
| | Tümör Evresi | 2 | -0,3004 | 0,1124 | -2,6719 | 0,007 |
| | | 3 | -0,6684 | 0,2099 | -3,1839 | 0,001 |
| | Menopoz Durumu | Var | -0,0380 | 0,2840 | -0,1338 | 0,893 |
| | Hormon Terapisi | Var | -0,1338 | 0,0966 | -1,3846 | 0,166 |
| 0,08 | Yaş (yıl) | | 0,0076 | 0,0251 | 0,3035 | 0,761 |
| | Tümör Büyüklüğü (mm) | | -0,0155 | 0,0078 | -1,9955 | 0,045 |
| | Progesteron Reseptör Sayısı | | 0,0017 | 0,0022 | 0,7715 | 0,440 |
| | Östrojen Reseptör Sayısı | | 0,0002 | 0,0003 | 0,7096 | 0,477 |
| | Nodül Sayısı | | -0,0251 | 0,0192 | -1,3104 | 0,190 |
| | Tümör Evresi | 2 | -0,3220 | 0,9540 | -0,3376 | 0,735 |
| | | 3 | -0,8662 | 0,8346 | -1,0378 | 0,299 |
| | Menopoz Durumu | Var | 0,0403 | 0,6786 | 0,0595 | 0,952 |
| | Hormon Terapisi | Var | -0,2803 | 0,2638 | -1,0623 | 0,288 |

Ek 2. Gbcs veri setine ait KR'nin farklı τ deęerleri için parametre tahminleri (devam)

| τ | Parametre | $\hat{\beta}$ | $\hat{\sigma}_{\beta}$ | t | p | |
|-------------|-----------------------------|---------------|------------------------|---------|--------------|--------------|
| 0,10 | Yaş (yıl) | 0,0102 | 0,0202 | 0,5041 | 0,614 | |
| | Tümör Büyüklüğü (mm) | -0,0162 | 0,0065 | -2,4929 | 0,012 | |
| | Progesteron Reseptör Sayısı | 0,0023 | 0,0017 | 1,3922 | 0,163 | |
| | Östrojen Reseptör Sayısı | 0,0001 | 0,0002 | 0,3416 | 0,732 | |
| | Nodül Sayısı | -0,0292 | 0,0208 | -1,4027 | 0,160 | |
| | Tümör Evresi | 2 | -0,4985 | 0,6473 | -0,7702 | 0,441 |
| | | 3 | -0,9836 | 0,7268 | -1,3533 | 0,175 |
| | Menopoz Durumu | Var | 0,2326 | 0,5052 | 0,4604 | 0,645 |
| | Hormon Terapisi | Var | -0,3796 | 0,2265 | -1,6753 | 0,093 |
| 0,15 | Yaş (yıl) | 0,0121 | 0,0217 | 0,5606 | 0,575 | |
| | Tümör Büyüklüğü (mm) | -0,0154 | 0,0075 | -2,0435 | 0,041 | |
| | Progesteron Reseptör Sayısı | 0,0045 | 0,0030 | 1,4863 | 0,137 | |
| | Östrojen Reseptör Sayısı | -0,0001 | 0,0003 | -0,4396 | 0,660 | |
| | Nodül Sayısı | -0,0370 | 0,0258 | -1,4331 | 0,151 | |
| | Tümör Evresi | 2 | -0,2939 | 0,2661 | -1,1042 | 0,269 |
| | | 3 | -0,8098 | 0,4086 | -1,9814 | 0,047 |
| | Menopoz Durumu | Var | 0,2911 | 0,4357 | 0,6683 | 0,503 |
| | Hormon Terapisi | Var | -0,3886 | 0,2411 | -1,6112 | 0,107 |
| 0,18 | Yaş (yıl) | 0,0094 | 0,0253 | 0,3721 | 0,709 | |
| | Tümör Büyüklüğü (mm) | -0,0143 | 0,0054 | -2,6114 | 0,009 | |
| | Progesteron Reseptör Sayısı | 0,0049 | 0,0020 | 2,4508 | 0,014 | |
| | Östrojen Reseptör Sayısı | -0,0001 | 0,0004 | -0,2937 | 0,768 | |
| | Nodül Sayısı | -0,0380 | 0,0262 | -1,4470 | 0,147 | |
| | Tümör Evresi | 2 | -0,2296 | 0,2529 | -0,9077 | 0,364 |
| | | 3 | -0,7390 | 0,3020 | -2,4468 | 0,014 |
| | Menopoz Durumu | Var | 0,2613 | 0,4462 | 0,5857 | 0,558 |
| | Hormon Terapisi | Var | -0,4104 | 0,1899 | -2,1606 | 0,030 |

Ek 2. Gbcs veri setine ait KR'nin farklı τ deęerleri için parametre tahminleri (devam)

| τ | Bağımsız Deęişkenler | | $\hat{\beta}$ | $\hat{\sigma}_{\beta}$ | t | p |
|-------------|-----------------------------|-----|---------------|------------------------|---------|--------------|
| 0,20 | Yaş (yıl) | | 0,0031 | 0,0211 | 0,1486 | 0,881 |
| | Tümör Büyüklüğü (mm) | | -0,0123 | 0,0053 | -2,2941 | 0,021 |
| | Progesteron Reseptör Sayısı | | 0,0047 | 0,0024 | 1,9051 | 0,056 |
| | Östrojen Reseptör Sayısı | | -0,0001 | 0,0003 | -0,3470 | 0,728 |
| | Nodül Sayısı | | -0,0375 | 0,0246 | -1,5223 | 0,127 |
| | Tümör Evresi | 2 | -0,2564 | 0,2144 | -1,1955 | 0,231 |
| | | 3 | -0,7799 | 0,3081 | -2,5312 | 0,011 |
| | Menopoz Durumu | Var | 0,1397 | 0,3907 | 0,3575 | 0,720 |
| | Hormon Terapisi | Var | -0,3793 | 0,1681 | -2,2562 | 0,024 |
| 0,23 | Yaş (yıl) | | -0,0011 | 0,0131 | -0,0883 | 0,929 |
| | Tümör Büyüklüğü (mm) | | -0,0110 | 0,0098 | -1,1256 | 0,260 |
| | Progesteron Reseptör Sayısı | | 0,0040 | 0,0018 | 2,2505 | 0,024 |
| | Östrojen Reseptör Sayısı | | -0,0002 | 0,0003 | -0,6186 | 0,536 |
| | Nodül Sayısı | | -0,0446 | 0,0273 | -1,6353 | 0,101 |
| | Tümör Evresi | 2 | -0,2769 | 0,1832 | -1,5111 | 0,130 |
| | | 3 | -0,7582 | 0,2346 | -3,2310 | 0,001 |
| | Menopoz Durumu | Var | 0,0624 | 0,3145 | 0,1984 | 0,842 |
| | Hormon Terapisi | Var | -0,3273 | 0,1457 | -2,2469 | 0,024 |
| 0,25 | Yaş (yıl) | | -0,0038 | 0,0133 | -0,2860 | 0,774 |
| | Tümör Büyüklüğü (mm) | | -0,0109 | 0,0120 | -0,9094 | 0,363 |
| | Progesteron Reseptör Sayısı | | 0,0042 | 0,0021 | 1,9956 | 0,045 |
| | Östrojen Reseptör Sayısı | | -0,0003 | 0,0004 | -0,7967 | 0,425 |
| | Nodül Sayısı | | -0,0492 | 0,0366 | -1,3441 | 0,178 |
| | Tümör Evresi | 2 | -0,2372 | 0,1832 | -1,2943 | 0,195 |
| | | 3 | -0,6599 | 0,2757 | -2,3935 | 0,016 |
| | Menopoz Durumu | Var | 0,0599 | 0,3912 | 0,1531 | 0,878 |
| | Hormon Terapisi | Var | -0,3087 | 0,1730 | -1,7846 | 0,074 |

Ek 2. Gbcs veri setine ait KR'nin farklı τ deęerleri için parametre tahminleri (devam)

| τ | Bağımsız Deęişkenler | | $\hat{\beta}$ | $\hat{\sigma}_{\beta}$ | t | p |
|-----------------|-----------------------------|---------|---------------|------------------------|---------|--------------|
| 0,28 | Yaş (yıl) | | -0,0099 | 0,0179 | -0,5557 | 0,578 |
| | Tümör Büyüklüğü (mm) | | -0,0091 | 0,0120 | -0,7621 | 0,445 |
| | Progesteron Reseptör Sayısı | | 0,0031 | 0,0017 | 1,7703 | 0,076 |
| | Östrojen Reseptör Sayısı | | -0,0003 | 0,0005 | -0,6049 | 0,545 |
| | Nodül Sayısı | | -0,0288 | 0,0457 | -0,6314 | 0,527 |
| | Tümör Evresi | 2 | -0,2960 | 0,2505 | -1,1816 | 0,237 |
| | | 3 | -0,7219 | 0,3137 | -2,3009 | 0,021 |
| | Menopoz Durumu | Var | -0,0168 | 0,3775 | -0,0446 | 0,964 |
| Hormon Terapisi | Var | -0,3414 | 0,1902 | -1,7952 | 0,072 | |
| 0,30 | Yaş (yıl) | | -0,0100 | 0,0163 | -0,6170 | 0,537 |
| | Tümör Büyüklüğü (mm) | | -0,0093 | 0,0090 | -1,0283 | 0,303 |
| | Progesteron Reseptör Sayısı | | 0,0032 | 0,0015 | 2,1226 | 0,033 |
| | Östrojen Reseptör Sayısı | | -0,0003 | 0,0003 | -1,0782 | 0,280 |
| | Nodül Sayısı | | -0,0322 | 0,0430 | -0,7509 | 0,452 |
| | Tümör Evresi | 2 | -0,2786 | 0,3003 | -0,9277 | 0,353 |
| | | 3 | -0,7209 | 0,3366 | -2,1416 | 0,032 |
| | Menopoz Durumu | Var | 0,0078 | 0,3237 | 0,0242 | 0,980 |
| Hormon Terapisi | Var | -0,3481 | 0,2173 | -1,6018 | 0,109 | |
| 0,35 | Yaş (yıl) | | -0,0069 | 0,0138 | -0,5043 | 0,613 |
| | Tümör Büyüklüğü (mm) | | -0,0076 | 0,0151 | -0,5029 | 0,615 |
| | Progesteron Reseptör Sayısı | | 0,0025 | 0,0018 | 1,3984 | 0,161 |
| | Östrojen Reseptör Sayısı | | -0,0004 | 0,0003 | -1,3630 | 0,172 |
| | Nodül Sayısı | | -0,0308 | 0,0190 | -1,6201 | 0,105 |
| | Tümör Evresi | 2 | -0,2336 | 0,2845 | -0,8213 | 0,411 |
| | | 3 | -0,6638 | 0,4198 | -1,5810 | 0,113 |
| | Menopoz Durumu | Var | 0,0632 | 0,2782 | 0,2272 | 0,820 |
| Hormon Terapisi | Var | -0,3159 | 0,3160 | -0,9995 | 0,317 | |

Ek 2. Gbcs veri setine ait KR'nin farklı τ deęerleri için parametre tahminleri (devam)

| τ | Bağımsız Deęişkenler | | $\hat{\beta}$ | $\hat{\sigma}_{\beta}$ | t | p |
|--------|-----------------------------|-----|---------------|------------------------|---------|--------------|
| 0,38 | Yaş (yıl) | | -0,0040 | 0,0114 | -0,3537 | 0,723 |
| | Tümör Büyüklüğü (mm) | | -0,0146 | 0,0126 | -1,1626 | 0,244 |
| | Progesteron Reseptör Sayısı | | 0,0030 | 0,0012 | 2,4816 | 0,013 |
| | Östrojen Reseptör Sayısı | | -0,0007 | 0,0004 | -1,6694 | 0,095 |
| | Nodül Sayısı | | -0,0461 | 0,0236 | -1,9475 | 0,051 |
| | Tümör Evresi | 2 | 0,0158 | 0,3055 | 0,0517 | 0,958 |
| | | 3 | -0,2994 | 0,3762 | -0,7959 | 0,426 |
| | Menopoz Durumu | Var | 0,1688 | 0,3743 | 0,4510 | 0,651 |
| | Hormon Terapisi | Var | -0,2584 | 0,2870 | -0,9003 | 0,367 |
| 0,40 | Yaş (yıl) | | -0,0041 | 0,01136 | -0,3667 | 0,713 |
| | Tümör Büyüklüğü (mm) | | -0,0129 | 0,0129 | -0,9967 | 0,318 |
| | Progesteron Reseptör Sayısı | | 0,0027 | 0,0015 | 1,7885 | 0,073 |
| | Östrojen Reseptör Sayısı | | -0,0007 | 0,0004 | -1,6071 | 0,108 |
| | Nodül Sayısı | | -0,0424 | 0,0253 | -1,6743 | 0,094 |
| | Tümör Evresi | 2 | -0,0270 | 0,2716 | -0,0995 | 0,920 |
| | | 3 | -0,3445 | 0,3348 | -1,0290 | 0,303 |
| | Menopoz Durumu | Var | 0,1388 | 0,4571 | 0,3175 | 0,750 |
| | Hormon Terapisi | Var | -0,2683 | 0,2378 | -1,1278 | 0,259 |
| 0,43 | Yaş (yıl) | | -0,0048 | 0,0138 | -0,3485 | 0,727 |
| | Tümör Büyüklüğü (mm) | | -0,0121 | 0,0130 | -0,9311 | 0,351 |
| | Progesteron Reseptör Sayısı | | 0,0023 | 0,0020 | 1,1733 | 0,240 |
| | Östrojen Reseptör Sayısı | | -0,0006 | 0,0007 | -0,8111 | 0,417 |
| | Nodül Sayısı | | -0,0332 | 0,0243 | -1,3651 | 0,172 |
| | Tümör Evresi | 2 | -0,1130 | 0,2105 | -0,5365 | 0,591 |
| | | 3 | -0,4337 | 0,2920 | -1,4853 | 0,137 |
| | Menopoz Durumu | Var | 0,1705 | 0,4059 | 0,4202 | 0,674 |
| | Hormon Terapisi | Var | -0,3022 | 0,1811 | -1,6685 | 0,095 |

Ek 2. Gbcs veri setine ait KR'nin farklı τ deęerleri için parametre tahminleri (devam)

| τ | Bağımsız Deęişkenler | | $\hat{\beta}$ | $\hat{\sigma}_{\beta}$ | t | p |
|-------------|-----------------------------|-----|---------------|------------------------|---------|--------------|
| 0,45 | Yaş (yıl) | | -0,0031 | 0,0178 | -0,1766 | 0,859 |
| | Tümör Büyüklüğü (mm) | | -0,0100 | 0,0084 | -1,1912 | 0,233 |
| | Progesteron Reseptör Sayısı | | 0,0016 | 0,0026 | 0,6193 | 0,535 |
| | Östrojen Reseptör Sayısı | | -0,0004 | 0,0006 | -0,7499 | 0,453 |
| | Nodül Sayısı | | -0,0319 | 0,0140 | -2,2725 | 0,023 |
| | Tümör Evresi | 2 | -0,0082 | 0,2800 | -0,0293 | 0,976 |
| | | 3 | -0,3468 | 0,3408 | -1,0174 | 0,308 |
| | Menopoz Durumu | Var | 0,2387 | 0,4120 | 0,5793 | 0,562 |
| | Hormon Terapisi | Var | -0,1839 | 0,2416 | -0,7611 | 0,446 |
| 0,48 | Yaş (yıl) | | -0,0061 | 0,0221 | -0,2760 | 0,782 |
| | Tümör Büyüklüğü (mm) | | -0,0063 | 0,0094 | -0,6706 | 0,502 |
| | Progesteron Reseptör Sayısı | | 0,0011 | 0,0019 | 0,6111 | 0,541 |
| | Östrojen Reseptör Sayısı | | -0,0001 | 0,0016 | -0,0302 | 0,975 |
| | Nodül Sayısı | | -0,0370 | 0,0157 | -2,3567 | 0,018 |
| | Tümör Evresi | 2 | 0,0844 | 0,3504 | 0,2410 | 0,809 |
| | | 3 | -0,1526 | 0,4305 | -0,3545 | 0,722 |
| | Menopoz Durumu | Var | 0,1390 | 0,4506 | 0,3085 | 0,757 |
| | Hormon Terapisi | Var | -0,2024 | 0,2761 | -0,7332 | 0,463 |
| 0,50 | Yaş (yıl) | | -0,0073 | 0,0191 | -0,3810 | 0,703 |
| | Tümör Büyüklüğü (mm) | | -0,0051 | 0,0061 | -0,8399 | 0,400 |
| | Progesteron Reseptör Sayısı | | 0,0010 | 0,0017 | 0,6255 | 0,531 |
| | Östrojen Reseptör Sayısı | | -0,0001 | 0,0014 | -0,0292 | 0,976 |
| | Nodül Sayısı | | -0,0390 | 0,0119 | -3,2651 | 0,001 |
| | Tümör Evresi | 2 | 0,1140 | 0,3712 | 0,3070 | 0,758 |
| | | 3 | -0,1456 | 0,4309 | -0,3379 | 0,735 |
| | Menopoz Durumu | Var | 0,0962 | 0,3750 | 0,2565 | 0,797 |
| | Hormon Terapisi | Var | -0,1616 | 0,4171 | -0,3874 | 0,698 |

Ek 2. Gbcs veri setine ait KR'nin farklı τ deęerleri için parametre tahminleri (devam)

| τ | Bağımsız Deęişkenler | | $\hat{\beta}$ | $\hat{\sigma}_{\beta}$ | t | p |
|--------|-----------------------------|-----|---------------|------------------------|---------|-------|
| 0,52 | Yaş (yıl) | | -0,0075 | 0,0178 | -0,4198 | 0,674 |
| | Tümör Büyüklüğü (mm) | | -0,0048 | 0,0043 | -1,1209 | 0,262 |
| | Progesteron Reseptör Sayısı | | 0,0009 | 0,0012 | 0,7794 | 0,435 |
| | Östrojen Reseptör Sayısı | | -0,0001 | 0,0016 | -0,0092 | 0,992 |
| | Nodül Sayısı | | -0,0395 | 0,0199 | -1,9832 | 0,047 |
| | Tümör Evresi | 2 | 0,1481 | 0,3986 | 0,3717 | 0,710 |
| | | 3 | -0,1132 | 0,4038 | -0,2804 | 0,779 |
| | Menopoz Durumu | Var | 0,0945 | 0,5127 | 0,1844 | 0,853 |
| | Hormon Terapisi | Var | -0,1252 | 0,3949 | -0,3170 | 0,751 |
| 0,55 | Yaş (yıl) | | -0,0115 | 0,0242 | -0,4775 | 0,632 |
| | Tümör Büyüklüğü (mm) | | -0,0051 | 0,0053 | -0,9675 | 0,333 |
| | Progesteron Reseptör Sayısı | | 0,0003 | 0,0016 | 0,1789 | 0,857 |
| | Östrojen Reseptör Sayısı | | 0,0001 | 0,0018 | 0,0489 | 0,960 |
| | Nodül Sayısı | | -0,0326 | 0,0307 | -1,0619 | 0,288 |
| | Tümör Evresi | 2 | 0,0549 | 0,3806 | 0,1442 | 0,885 |
| | | 3 | -0,1755 | 0,3861 | -0,4546 | 0,649 |
| | Menopoz Durumu | Var | -0,0856 | 0,8353 | -0,1025 | 0,918 |
| | Hormon Terapisi | Var | 0,1712 | 0,3363 | 0,5091 | 0,610 |
| 0,57 | Yaş (yıl) | | -0,0130 | 0,0245 | -0,5299 | 0,596 |
| | Tümör Büyüklüğü (mm) | | -0,0049 | 0,0053 | -0,9191 | 0,358 |
| | Progesteron Reseptör Sayısı | | -0,0001 | 0,0015 | -0,0154 | 0,987 |
| | Östrojen Reseptör Sayısı | | 0,0003 | 0,0015 | 0,2224 | 0,823 |
| | Nodül Sayısı | | -0,0303 | 0,0372 | -0,8154 | 0,414 |
| | Tümör Evresi | 2 | 0,0335 | 0,3575 | 0,0937 | 0,925 |
| | | 3 | -0,1357 | 0,4297 | -0,3160 | 0,752 |
| | Menopoz Durumu | Var | -0,1003 | 0,7481 | -0,1340 | 0,893 |
| | Hormon Terapisi | Var | 0,1756 | 0,3157 | 0,5564 | 0,577 |

Ek 2. Gbcs veri setine ait KR'nin farklı τ deęerleri için parametre tahminleri (devam)

| τ | Bağımsız Deęişkenler | | $\hat{\beta}$ | $\hat{\sigma}_{\beta}$ | t | p |
|--------|-----------------------------|-----|---------------|------------------------|---------|-------|
| 0,60 | Yaş (yıl) | | -0,0097 | 0,0229 | -0,4269 | 0,669 |
| | Tümör Büyüklüğü (mm) | | -0,0043 | 0,0044 | -0,9800 | 0,327 |
| | Progesteron Reseptör Sayısı | | 0,0001 | 0,0015 | 0,0868 | 0,930 |
| | Östrojen Reseptör Sayısı | | 0,0001 | 0,0011 | 0,1307 | 0,895 |
| | Nodül Sayısı | | -0,0201 | 0,0350 | -0,5734 | 0,566 |
| | Tümör Evresi | 2 | 0,0279 | 0,2907 | 0,0961 | 0,923 |
| | | 3 | -0,1032 | 0,3484 | -0,2963 | 0,766 |
| | Menopoz Durumu | Var | -0,0931 | 0,6870 | -0,1355 | 0,892 |
| | Hormon Terapisi | Var | 0,0195 | 0,3031 | 0,0643 | 0,948 |
| 0,62 | Yaş (yıl) | | -0,0083 | 0,0197 | -0,4238 | 0,671 |
| | Tümör Büyüklüğü (mm) | | -0,0048 | 0,0046 | -1,0235 | 0,306 |
| | Progesteron Reseptör Sayısı | | 0,0001 | 0,0013 | 0,0563 | 0,955 |
| | Östrojen Reseptör Sayısı | | 0,0001 | 0,0011 | 0,0815 | 0,935 |
| | Nodül Sayısı | | -0,0193 | 0,0259 | -0,7445 | 0,456 |
| | Tümör Evresi | 2 | 0,0393 | 0,2896 | 0,1359 | 0,891 |
| | | 3 | -0,0896 | 0,3172 | -0,2826 | 0,777 |
| | Menopoz Durumu | Var | -0,0606 | 0,5538 | -0,1094 | 0,912 |
| | Hormon Terapisi | Var | 0,0101 | 0,2809 | 0,0359 | 0,971 |
| 0,65 | Yaş (yıl) | | -0,0081 | 0,0136 | -0,5946 | 0,552 |
| | Tümör Büyüklüğü (mm) | | -0,0050 | 0,0043 | -1,1520 | 0,249 |
| | Progesteron Reseptör Sayısı | | 0,0001 | 0,0008 | 0,0264 | 0,978 |
| | Östrojen Reseptör Sayısı | | 0,0001 | 0,0009 | 0,1124 | 0,910 |
| | Nodül Sayısı | | -0,0200 | 0,0240 | -0,8339 | 0,404 |
| | Tümör Evresi | 2 | 0,0457 | 0,2290 | 0,1996 | 0,841 |
| | | 3 | -0,0828 | 0,2679 | -0,3093 | 0,757 |
| | Menopoz Durumu | Var | -0,0451 | 0,5180 | -0,0871 | 0,930 |
| | Hormon Terapisi | Var | 0,0177 | 0,2637 | 0,0674 | 0,946 |

Ek 2. Gbcs veri setine ait KR'nin farklı τ değerleri için parametre tahminleri (devam)

| τ | Bağımsız Değişkenler | $\hat{\beta}$ | $\hat{\sigma}_{\beta}$ | t | p | |
|-----------------|-----------------------------|---------------|------------------------|---------|---------|-------|
| 0,67 | Yaş (yıl) | -0,0072 | 0,0164 | -0,4430 | 0,657 | |
| | Tümör Büyüklüğü (mm) | -0,0056 | 0,0043 | -1,3003 | 0,193 | |
| | Progesteron Reseptör Sayısı | -0,0001 | 0,0008 | -0,1039 | 0,917 | |
| | Östrojen Reseptör Sayısı | -0,0001 | 0,0008 | 0,1491 | 0,881 | |
| | Nodül Sayısı | -0,0220 | 0,0239 | -0,9180 | 0,358 | |
| | Tümör Evresi | 2 | 0,1082 | 0,2180 | 0,4965 | 0,619 |
| | | 3 | -0,0168 | 0,3184 | -0,0530 | 0,957 |
| | Menopoz Durumu | Var | 0,0041 | 0,5067 | 0,0082 | 0,993 |
| Hormon Terapisi | Var | 0,0001 | 0,2557 | 0,0004 | 0,999 | |
| 0,70 | Yaş (yıl) | -0,0076 | 0,0184 | -0,4125 | 0,679 | |
| | Tümör Büyüklüğü (mm) | -0,0062 | 0,0039 | -1,5892 | 0,112 | |
| | Progesteron Reseptör Sayısı | -0,0002 | 0,0009 | -0,2730 | 0,784 | |
| | Östrojen Reseptör Sayısı | 0,0001 | 0,0007 | 0,1849 | 0,853 | |
| | Nodül Sayısı | -0,0235 | 0,0197 | -1,1890 | 0,234 | |
| | Tümör Evresi | 2 | 0,1321 | 0,2445 | 0,5406 | 0,588 |
| | | 3 | -0,0047 | 0,3017 | -0,0157 | 0,987 |
| | Menopoz Durumu | Var | 0,0012 | 0,5010 | 0,0024 | 0,998 |
| Hormon Terapisi | Var | -0,0031 | 0,2865 | -0,0111 | 0,991 | |
| 0,72 | Yaş (yıl) | -0,0090 | 0,0171 | -0,5276 | 0,597 | |
| | Tümör Büyüklüğü (mm) | -0,0062 | 0,0042 | -1,4828 | 0,138 | |
| | Progesteron Reseptör Sayısı | -0,0002 | 0,0008 | -0,2440 | 0,807 | |
| | Östrojen Reseptör Sayısı | 0,0001 | 0,0007 | 0,1624 | 0,870 | |
| | Nodül Sayısı | -0,0237 | 0,0186 | -1,2711 | 0,203 | |
| | Tümör Evresi | 2 | 0,1522 | 0,2484 | 0,6129 | 0,539 |
| | | 3 | 0,0004 | 0,3137 | 0,0013 | 0,998 |
| | Menopoz Durumu | Var | -0,0495 | 0,5001 | -0,0989 | 0,921 |
| Hormon Terapisi | Var | -0,0164 | 0,3024 | -0,0544 | 0,956 | |

Ek 2. Gbcs veri setine ait KR'nin farklı τ deęerleri için parametre tahminleri (devam)

| τ | Bağımsız Deęişkenler | $\hat{\beta}$ | $\hat{\sigma}_{\beta}$ | t | p | |
|-------------|-----------------------------|---------------|------------------------|---------|---------|-------|
| 0,75 | Yaş (yıl) | -0,0115 | 0,0189 | -0,6074 | 0,543 | |
| | Tümör Büyüklüğü (mm) | -0,0055 | 0,0045 | -1,2138 | 0,224 | |
| | Progesteron Reseptör Sayısı | -0,0001 | 0,0008 | -0,0072 | 0,994 | |
| | Östrojen Reseptör Sayısı | -0,0002 | 0,0008 | -0,2391 | 0,810 | |
| | Nodül Sayısı | -0,0218 | 0,0177 | -1,2346 | 0,216 | |
| | Tümör Evresi | 2 | 0,0718 | 0,2658 | 0,2702 | 0,786 |
| | | 3 | -0,1016 | 0,3045 | -0,3338 | 0,738 |
| | Menopoz Durumu | Var | -0,1519 | 0,5061 | -0,3001 | 0,764 |
| | Hormon Terapisi | Var | -0,3624 | 0,3281 | -1,1046 | 0,269 |