



HACETTEPE  
ÜNİVERSİTESİ

**İST493- SEMİNER**

**Tekrarlayan Olaylarda Sağkalım Analizi Araştırma Raporu**

**Prof. Dr. NİHAL ATA TUTKUN**

**İREM TARIM – 2210329047**

# ***Tekrarlayan Olaylarda Sağkalım Analizi***

## **1. Sağkalım analizi nedir?**

Sağkalım analizi, tıp, epidemiyoloji, halk sağlığı ve daha birçok alanda, belirli bir olayın gerçekleşmesine kadar geçen süreyi incelemek amacıyla kullanılan güçlü istatistiksel yöntemlerden oluşan bir yaklaşımdır. Bu analiz türü, genellikle bireylerin yalnızca tek bir olay yaşadığı varsayımına dayanır; örneğin ölüm, hastalık tanısı ya da hastalığın nüksetmesi gibi tekil sonuçlara odaklanır. Ancak gerçek yaşam koşullarında yürütülen birçok klinik ve gözlemsel çalışmada bireylerin zaman içinde benzer olayları birden fazla kez yaşamaları mümkündür. Bu tür olaylar literatürde "tekrarlayan olaylar" ya da "reküren olaylar" olarak adlandırılır. Bu durum, geleneksel sağkalım analizlerinin öngördüğü yapıdan sapma anlamına gelir ve dolayısıyla analiz stratejilerinde önemli metodolojik değişiklikleri gerekli kılar. Tekrarlayan olayların analizi, yalnızca olayların gerçekleşip gerçekleşmediğini değil, aynı zamanda bu olayların zaman içindeki düzenini ve birey içindeki tekrar yapısını da dikkate alan özel modellere ihtiyaç duyar. Bu sayede, elde edilen sonuçlar daha doğru, anlamlı ve klinik olarak yorumlanabilir hale gelir.

## **2. Dikkat edilmesi gerekenler**

Klasik sağkalım analizlerinde sıkça kullanılan Cox orantısal hazard modeli, birey başına yalnızca tek bir olayın meydana geldiği durumlar için tasarlanmıştır. Ancak tekrarlayan olayların söz konusu olduğu analizlerde bu modelin temel varsayımları geçerliliğini yitirir. Öncelikle, aynı bireyde meydana gelen ardışık olaylar genellikle birbirinden bağımsız değildir. Örneğin, bir hastanın geçirdiği ilk enfeksiyon, ikinci enfeksiyonun riskini artırabilir veya azaltabilir. Bu durum birey içi bağımlılık yaratır ve klasik Cox modeli bu korelasyonu dikkate almaz. Ayrıca, her tekrarlayan olayın meydana geldiği zaman farklıdır ve her biri kendi başına bir risk dönemi oluşturur. Dolayısıyla olaylar yalnızca varlığıyla değil, zaman içindeki konumlarıyla birlikte ele alınmalıdır. Üstelik, aynı bireydeki olayların risk düzeyleri de farklılık gösterebilir; ilk olay, sonraki olayların hem şiddetini hem de gerçekleşme süresini etkileyebilir. Bu nedenle analizde hangi zaman ölçüsünün kullanılacağı önemlidir. Olaylar arası geçen sürenin mi yoksa ilk olaydan itibaren geçen toplam sürenin mi esas alınacağı, araştırma sorusuna göre belirlenmelidir. Tekrarlayan olayları doğru şekilde analiz edebilmek için birey içi bağımlılıkları dikkate alan sağlam (robust) varyans tahmincilerinin kullanılması önerilir.

### 3. Veri Seti ve Yapısı

Bu çalışmada kullanılan veri seti, tekrarlayan mesane kanseri tümörlerine ilişkin olarak Byar (1980) ile Wei, Lin ve Weissfeld (1989) tarafından yayımlanan bir araştırmadan elde edilen bulgulara dayanmaktadır.

Veri seti, transüretal cerrahi eksizyon (TUR) sonrasında mesane tümörünün tekrarlamasına odaklanmaktadır. Çalışma kapsamında her biri 64 aya kadar takip edilen ilk 15 hastanın verileri analiz edilmiştir. Her hastanın izlem süresi boyunca geçirdiği tümör tekrarı olayları ve bu olaylara ilişkin zaman bilgileri, çok satırlı kayıt yapısıyla veri setine işlenmiştir. Böylelikle aynı birey için birden fazla olay gözlemi yapılabilmektedir. Bu yapı, Counting Process (CP) formatına uygun olacak şekilde düzenlenmiştir.

```
> bladder <- read.csv("C:/Users/iremt/OneDrive/Desktop/seminer/bladder.csv", header = TRUE, sep = ";")
> bladder[12:20, ]
  id int event start stop tx num size
12 10  2     1    12   16  1  1    1
13 10  3     0    16   18  1  1    1
14 11  1     0     0   23  1  3    3
15 12  1     1     0   10  1  1    3
16 12  2     1    10   15  1  1    3
17 12  3     0    15   23  1  1    3
18 13  1     1     0     3  1  1    1
19 13  2     1     3   16  1  1    1
20 13  3     1    16   23  1  1    1
```

#### 3.1 Veri Yapısı ve Değişkenler

Veri setinde her bireyin her bir olay (ya da olaysız izlem) periyoduna karşılık gelen bir gözlem satırı yer almaktadır. Örneğin, üç olay geçiren bir bireyin verileri üç gözlem (ve bir takip satırı) olarak temsil edilir. Her gözlemde, olayın gerçekleşip gerçekleşmediği, zaman aralığı (başlangıç ve bitiş ayı), tedavi durumu ve sabit kovaryatlar yer almaktadır.

Değişken	Açıklama
İd	Bireyin kimlik numarası (her hasta için eşsiz)
İnt	Olay sırası (1, 2, 3 ...) – veri zaman sırasına göre yapılandırılmıştır
Start	Belirli zaman periyodunun başlangıç zamanı (ay cinsinden)
Stop	Belirli zaman periyodunun bitiş zamanı (ay cinsinden)

Event	Belirli periyotta olay olup olmadığı (1 = olay var, 0 = yok)
Tx	Tedavi durumu (0 = plasebo, 1 = tiotepa ile tedavi)
Num	Başlangıçta tespit edilen tümör sayısı (zamandan bağımsız)
Size	Başlangıç tümör büyüklüğü (cm cinsinden; zamandan bağımsız)

Bu yapıda start–stop zaman aralıkları, Counting Process modelleme çerçevesine doğrudan uyumludur. Zaman yapısının bu şekilde tanımlanması, klasik Kaplan-Meier veya standart Cox modellerinin aksine, bireylerin olay yaşamaya tekrar tekrar aday olduğu durumlar için uygundur.

### 3.2 Analiz Edilen Olay ve İlgili Değişkenleri

Bu analizde ilgi duyulan olay, mesane tümörünün tekrarlamasıdır. Her tümör tekrarı hastanede cerrahi olarak tedavi edilmiştir. Bu nedenle bir olayın meydana gelmesi, bir sonraki gözlem periyodunun başlangıcı olarak yeni bir risk periyodu yaratmıştır.

Çalışmanın asıl maruz kalma (exposure) değişkeni ilaç tedavisi (tx) olup, iki düzeye sahiptir:

- 0 = plasebo
- 1 = tiotepa (ilaçla tedavi)

#### 4. Counting Process (CP) Yaklaşımı ile Cox Regresyon Modeli

Bu analizde, bireylerin zaman içinde birden fazla olay yaşaması olasılığı göz önünde bulundurularak, tekrarlayan olaylar için Counting Process (CP) yaklaşımı ile Cox regresyon modeli uygulanmıştır. Bu yaklaşımda, her bireyin her olay aralığı ayrı bir gözlem satırı olarak değerlendirilmiş ve zaman ölçümü bireyin çalışmaya dahil olduğu andan itibaren hesaplanmıştır.

Veri seti, uygun şekilde (start, stop] formatına dönüştürülmüş ve her bireyin tekrarlayan olaylarının zaman bilgisi bu yapıya uygun olarak tanımlanmıştır. Analiz öncesinde stop = start olan herhangi bir satır olmadığı kontrol edilmiş ve buna yönelik düzenlemeler yapılmıştır.

```
> library(dplyr)
> library(survival)
> bladder[bladder$stop <= bladder$start, ]
[1] id      int    event start stop  tx    num    size
<0 rows> (or 0-length row.names)
> Y=Surv(bladder$start,bladder$stop,bladder$event==1)
> coxph(Y ~ tx + num + size + cluster(id), data = bladder)
Call:
coxph(formula = Y ~ tx + num + size, data = bladder, cluster = id)

      coef exp(coef) se(coef) robust se      z      p
tx    0.99927   2.71630  0.39830   0.40250  2.483 0.013
num  -0.11354   0.89267  0.14211   0.10990 -1.033 0.302
size -0.05286   0.94851  0.13209   0.12478 -0.424 0.672

Likelihood ratio test=6.61 on 3 df, p=0.08553
n= 52, number of events= 31
```

Modelde bağımlı değişken, olayın meydana gelip gelmediğini ve bu olayın meydana geldiği zaman aralığını göstermektedir. cluster(id) ifadesi, aynı bireyde birden fazla gözlem olduğu için robust standart hataların kullanılmasını sağlar.

Tedavi etkisi (tx), istatistiksel olarak anlamlıdır ( $p = 0.013$ ). İlaç ile tedavi edilen bireylerde olay riski plaseboya göre yaklaşık 2.72 kat daha yüksektir. Bu sonuç, tedavinin tümör tekrarı üzerine etkili olabileceğini gösterir.

Başlangıç tümör sayısı ve boyutu değişkenleri anlamlı bulunmamıştır ( $p > 0.05$ ). Bu, bu iki değişkenin olay süresi üzerinde etkili olmadığını veya etkinin sınırlı olduğunu gösterir.

Modelin genel anlamlılığı, likelihood ratio testi ile test edilmiş ve 3 serbestlik derecesi ile hesaplanan istatistik değeri 6.61 olarak bulunmuştur ( $p = 0.085$ ).

## 5. Stratified Counting Process (SCP) Yaklaşımı ile Cox Regresyon Modeli

Bu analizde, Stratified Counting Process (SCP) yaklaşımı ile genişletilmiş Cox modeli uygulanmıştır. SCP modeli, olay sırasına göre her gözlem periyodunu farklı bir katman olarak tanımlar. Böylece 1. olay, 2. olay, 3. olay gibi ardışık olaylar birbirinden ayrıştırılarak farklı risk yapıları altında değerlendirilmiş olur.

```
> bladder<- bladder%>%
+   group_by(id) %>%
+   mutate(interval = row_number()) %>%
+   ungroup()
> coxph(Y ~ tx + num + size + strata(interval) + cluster(id), data = bladder)
Call:
coxph(formula = Y ~ tx + num + size + strata(interval), data = bladder,
      cluster = id)

      coef exp(coef) se(coef) robust se      z      p
tx      1.04205    2.83503  0.44318    0.34662  3.006 0.00264
num     -0.02755    0.97282  0.13395    0.07486 -0.368 0.71285
size     0.02655    1.02690  0.14905    0.12134  0.219 0.82681

Likelihood ratio test=5.48 on 3 df, p=0.1396
n= 52, number of events= 31
```

Burada interval, her birey için olay sırasını belirtir ve stratifikasyon değişkeni olarak modele dahil edilmiştir. Ayrıca, cluster(id) kullanılarak bir bireyin birden fazla olay yaşaması durumunda gözlemler arası bağımlılık için robust standart hatalar hesaplanmıştır.

Modelde tx değişkeni istatistiksel olarak anlamlı çıkmıştır ( $p = 0.0026$ ) ve hazard oranı (HR) 2.835 olarak tahmin edilmiştir. Bu, tedavi uygulanan bireylerin olay yaşama riskinin plasebo grubuna göre yaklaşık 2,8 kat daha yüksek olduğunu göstermektedir. Bu etki, CP modeline göre daha güçlüdür (CP modelinde  $HR \approx 2.72$ ,  $p = 0.013$ ).

Başlangıçtaki tümör sayısı ve büyüklüğü bu modelde anlamlı bulunmamıştır ( $p > 0.70$ ). Bu da SCP modeli kullanıldığında bu faktörlerin tekrarlanan olaylar üzerindeki etkisinin zayıf olabileceğini göstermektedir.

Modelin genel anlamlılık testi (LR test)  $p = 0.14$  ile istatistiksel olarak anlamlılık sınırının biraz dışındadır.

## 6. Etkileşimli Stratified Cox Modeli (tx × interval)

Tedavi etkisinin olay sırasına (birinci olay ve ikinci olay) göre değişip değişmediğini değerlendirmek amacıyla etkileşimli bir Cox regresyon modeli kurulmuştur. Bu modelde, tedavi değişkeni (tx) ile olay sırasını belirten faktör değişkeni (interval\_f) arasında etkileşim terimi (tx:interval\_f) modele dahil edilmiştir. Amaç, tedavi etkisinin her olay için farklı olup olmadığını test etmektir.

```
> bladder_sub <- bladder %>% filter(interval <= 2)
> # Stratification değişkeni faktöre çevrilir
> bladder_sub$interval_f <- as.factor(bladder_sub$interval)
> # tx * interval etkileşimli model kurulması
> cox_interaction <- coxph(Surv(start, stop, event) ~ tx * interval_f + num + size + cluster(id),
+ data = bladder_sub)
> # Model çıktısının gösterilmesi
> summary(cox_interaction)
Call:
coxph(formula = Surv(start, stop, event) ~ tx + interval_f +
      num + size + tx:interval_f, data = bladder_sub, cluster = id)

n= 39, number of events= 24

              coef exp(coef) se(coef) robust se         z Pr(>|z|)
tx             0.8014    2.2285  0.6240   0.5598    1.432   0.152
interval_f2    0.1897    1.2089  0.6371   0.5907    0.321   0.748
num            -0.0483    0.9528  0.1526   0.1101   -0.439   0.661
size            0.0221    1.0224  0.1670   0.1653    0.134   0.894
tx:interval_f2  0.9417    2.5643  1.0324   0.9864    0.955   0.340

              exp(coef) exp(-coef) lower .95 upper .95
tx             2.2285      0.4487    0.7440     6.676
interval_f2    1.2089      0.8272    0.3798     3.848
num             0.9528      1.0495    0.7679     1.182
size            1.0223      0.9781    0.7395     1.413
tx:interval_f2  2.5643      0.3900    0.3710    17.726

Concordance= 0.639 (se = 0.057 )
Likelihood ratio test= 8.91 on 5 df,  p=0.1
Wald test               = 31.79 on 5 df,  p=7e-06
Score (logrank) test = 12.62 on 5 df,  p=0.03, Robust = 7.74 p=0.2
```

Veri setinden yalnızca birinci ve ikinci olaylara ait gözlemler seçilmiştir. İnterval değişkeni faktör tipine dönüştürülerek etkileşim modeline uygun hale getirilmiştir. Modelde, bağımlı değişken olarak hayatta kalma süresi (başlangıç ve bitiş zamanlarıyla belirtilen sayma süreci formatında) kullanılmıştır. Bağımsız değişkenler olarak tedavi durumu (tx), başlangıçtaki tümör sayısı (num), tümör boyutu (size) ve tx \* interval\_f etkileşimi modele dahil edilmiştir. Birey bazında tekrarlayan gözlemler bulunduğu için cluster(id) ifadesi kullanılarak robust (empirical) standart hatalarla analiz yapılmıştır.

İlk olayda tedavinin etkisi istatistiksel olarak anlamlı bulunmamıştır (HR = 2.23; p = 0.152). İkinci olayda ise tedavi etkisi, ilk olaya göre daha yüksek görünmektedir. İkinci olayda toplam tedavi etkisi yaklaşık 5.71 olarak hesaplanmıştır (HR = 2.23 × 2.56), ancak bu fark da istatistiksel olarak anlamlı değildir (etkileşim terimi için p = 0.340). Model genel anlamda anlamlılık sınırında kalmıştır (LR testi p = 0.10), buna rağmen Wald testi anlamlı sonuç vermiştir (p < 0.001). Ayrıca, num ve size değişkenleri her iki olayda da anlamlı çıkmamıştır.

Bu sonuçlar, tedavi etkisinin olayın türüne göre değişebileceğini göstermektedir. Özellikle ikinci olayda tedavi etkisi daha belirgin hale gelmiştir. Bu farklılığı test etmek amacıyla kullanılan Etkileşimli Stratified Cox modeli, ikinci olayda tedavi etkisinin daha güçlü olabileceğini düşündürmektedir.

## 7. Gap Time Yaklaşımı ile Stratified Cox Modeli

Bu analizde, tekrarlayan olaylar arasındaki süreyi dikkate alarak oluşturulmuş. Gap Time yaklaşımı kullanılmıştır. Gap Time modeli, her olayın başlangıcında zaman ölçümünü sıfırlayarak olaylar arası geçen süreyi analiz eder.

```
> bladder$start2=0
> bladder$stop2 <- bladder$stop - bladder$start
> data.frame(id,event,start,stop,start2,stop2)[12:20, ]
   id event start stop start2 stop2
12 10     1    12   16      0     4
13 10     0    16   18      0     2
14 11     0     0   23      0    23
15 12     1     0   10      0    10
16 12     1    10   15      0     5
17 12     0    15   23      0     8
18 13     1     0    3      0     3
19 13     1     3   16      0    13
20 13     1    16   23      0     7
> Y2=Surv(bladder$start2,bladder$stop2,bladder$event)
> coxph(Y2 ~ tx + num + size + strata(interval) + cluster(id), data = bladder)
Call:
coxph(formula = Y2 ~ tx + num + size + strata(interval), data = bladder,
      cluster = id)

      coef exp(coef) se(coef) robust se         z        p
tx    0.48529   1.62464  0.43153   0.39475   1.229 0.219
num   -0.11184   0.89418  0.13958   0.07238  -1.545 0.122
size   0.01846   1.01863  0.14010   0.12279   0.150 0.881

Likelihood ratio test=2.11 on 3 df, p=0.5494
n= 52, number of events= 31
```

- start2 sabit olarak 0 tanımlanmış (her olay için zaman sıfırdan başlar),
- stop2 değeri, her olay için bir önceki olaydan itibaren geçen süreyi ifade eder (Gap Time),
- interval değişkeni olay sırasına göre stratifikasyon için kullanılmış,
- cluster(id) ile birey içi korelasyon dikkate alınmıştır (robust standart hata).

Modelde tedavi değişkeninin hazard oranı 1.625 olup istatistiksel olarak anlamlı değildir (p = 0.219). Bu sonuç, SCP modelinde anlamlı bulunan tedavi etkisinin, Gap Time yaklaşımında zayıfladığını göstermektedir.



Tümör sayısı (num),  $z=-1.545$ ,  $p=0.122$  ile sınırda anlamlılığa yakındır. Tümör büyüklüğü ise anlamlılık taşımamaktadır ( $p=0.881$ ).

Modelin genel anlamlılık testi (LR test) sonuçlarına göre model anlamlı değildir ( $p = 0.549$ ). Bu, Gap Time yaklaşımının tüm değişkenler açısından daha zayıf tahmin gücüne sahip olduğunu gösterir.

## 8. Marginal Yaklaşım ile Cox Modeli

Marginal model, tekrarlayan olayların analizinde birey içindeki olaylar arasında doğrudan bir sıra veya yapısal bağımlılık varsaymadan, her olayın ortalama etkisini değerlendirmeyi amaçlayan bir yaklaşımdır. Bu modelde, birey birden fazla olay yaşayabilir, ancak bu olaylar birbirinden bağımsız kabul edilmez; aksine, robust standart kullanılarak birey içi korelasyonlar hesaba katılır. Modelin ana amacı, tüm olayları birlikte değerlendirerek popülasyon düzeyinde ortalama tedavi etkisini incelemektir.

```
> cox_marginal <- coxph(Surv(start, stop, event) ~ tx + num + size + cluster(id), data = bladder)
> summary(cox_marginal)
Call:
coxph(formula = Surv(start, stop, event) ~ tx + num + size, data = bladder,
      cluster = id)

n= 52, number of events= 31

      coef exp(coef) se(coef) robust se      z Pr(>|z|)
tx    0.99927  2.71630  0.39830  0.40250  2.483  0.013 *
num   -0.11354  0.89267  0.14211  0.10990 -1.033  0.302
size  -0.05286  0.94851  0.13209  0.12478 -0.424  0.672
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

      exp(coef) exp(-coef) lower .95 upper .95
tx      2.7163    0.3681    1.2342    5.978
num      0.8927    1.1202    0.7197    1.107
size     0.9485    1.0543    0.7427    1.211

Concordance= 0.603 (se = 0.065 )
Likelihood ratio test= 6.61 on 3 df,  p=0.09
Wald test            = 9.09 on 3 df,  p=0.03
Score (logrank) test = 7.31 on 3 df,  p=0.06,  Robust = 4.81 p=0.2

(Note: the likelihood ratio and score tests assume independence of
      observations within a cluster, the Wald and robust score tests do not).
```

- Zaman yapısı klasik (start, stop) biçiminde tutulmuştur.
- cluster(id) ifadesi ile robust varyans tahmini yapılmıştır,
- Tedavi (tx), başlangıç tümör sayısı (num) ve tümör büyüklüğü (size) bağımsız değişkenler olarak modele dahil edilmiştir.

Tedavi etkisi istatistiksel olarak anlamlıdır ( $p = 0.013$ ), ve tahmin edilen hazard oranı 2.72'dir. Bu, tedavi grubunda olan bireylerin olay yaşama riskinin plasebo grubuna göre yaklaşık 2.7 kat daha yüksek olduğu anlamına gelir.

Her iki değişken de bu modelde istatistiksel olarak anlamlı bulunmamıştır. LR testi  $p = 0.09$  ile sınıra yakın anlamlılık göstermektedir. Wald testi anlamlıdır ( $p = 0.03$ ), ancak Score testine göre daha iyimser sonuç verir.

Marginal model ile elde edilen bulgular, tedavi değişkeninin tekrarlayan olaylar üzerinde istatistiksel olarak anlamlı bir ortalama etkisi olduğunu göstermektedir.

## 9. Paylaşılan Kırılğanlık (Shared Frailty) Cox Modeli

Tekrarlayan olay verilerinde aynı bireyden gelen birden fazla gözlemin birbirine bağımlı olması yaygın bir durumdur. Geleneksel Cox modelleri, birey içi bu bağımlılığı göz ardı eder. Ancak bu bağımlılığın doğrudan modele katılması, varyansların daha güvenilir tahmin edilmesini ve modelin gerçek hayattaki süreci daha doğru temsil etmesini sağlar.

```
> # Shared Frailty Modeli
> cox_frailty <- coxph(Surv(start, stop, event) ~ tx + num + size + frailty(id), data = bladder)
> summary(cox_frailty)
Call:
coxph(formula = Surv(start, stop, event) ~ tx + num + size +
      frailty(id), data = bladder)

n= 52, number of events= 31

      coef      se(coef)  se2    Chisq Df p
tx      0.99927 0.3983    0.3983 6.29 1 0.012
num     -0.11354 0.1421    0.1421 0.64 1 0.420
size    -0.05286 0.1321    0.1321 0.16 1 0.690
frailty(id)                0.00 0 0.940

      exp(coef) exp(-coef) lower .95 upper .95
tx      2.7163    0.3681    1.2444    5.929
num     0.8927    1.1202    0.6757    1.179
size    0.9485    1.0543    0.7322    1.229

Iterations: 6 outer, 22 Newton-Raphson
Variance of random effect= 5e-07 I-likelihood = -85.2
Degrees of freedom for terms= 1 1 1 0
Concordance= 0.615 (se = 0.065 )
Likelihood ratio test= 6.61 on 3 df, p=0.09
```

- start, stop, event değişkenleriyle olay süresi temsil edilmiş,
- tx (tedavi), num (başlangıçtaki tümör sayısı), size (tümör büyüklüğü) açıklayıcı değişkenler olarak modele dahil edilmiştir,
- frailty(id) ifadesiyle her birey için rastgele bir etkene izin verilmiştir.

Tedavi değişkeni (tx) istatistiksel olarak anlamlı bulunmuştur ( $p = 0.012$ ). Tedavi alan bireylerin olay yaşama riski tedavi almayanlara göre yaklaşık 2.72 kat daha yüksek çıkmıştır.

Başlangıç tümör sayısı (num) ve tümör büyüklüğü (size) değişkenleri için modelde anlamlılık gözlenmemiştir. Bu da bu iki değişkenin olay süresi üzerindeki etkisinin belirsiz veya zayıf olduğunu düşündürmektedir.

Modelde frailty(id) bileşeninin varyansı yaklaşık sıfıra ( $5e-07$ ) yakın bulunmuştur. Bu, aynı bireydeki tekrar eden gözlemler arasında anlamlı bir korelasyon olmadığını ya da bu korelasyonun model tarafından açıklanabilecek düzeyde olmadığını göstermektedir.

Paylaşılan kırılgenlik modeli, CP ve Marginal yaklaşımlarla karşılaştırıldığında daha gelişmiş bir yapı sunar. Özellikle gözlemler arasında korelasyonun yüksek olduğu durumlarda kullanılmalıdır. Bu analizde frailty varyansı düşük çıktığından dolayı modelden elde edilen sonuçlar, daha basit modellerdeki sonuçlarla büyük ölçüde benzerlik göstermektedir.

## 10. Olay Bazlı (Interval-Spesifik) Cox Modelleri ile Analiz

Stratified Cox modellerinde olayların sırası göz önüne alınabilirken, olay bazlı analizlerde her olay için ayrı bir model kurulması, tedavi ve diğer faktörlerin etkilerinin olaylara göre nasıl değiştiğini daha ayrıntılı şekilde değerlendirmeye olanak tanır.

Bu bölümde her biri için ayrı model kurularak, birinci olay ve ikinci olay için tedavi ve diğer değişkenlerin etkisi analiz edilmiştir.

### 10.1 Birinci Olay (interval == 1)

```
> cox_1 <- coxph(Surv(start, stop, event) ~ tx + num + size + cluster(id),
+               data = bladder %>% filter(interval == 1))
> summary(cox_1)
Call:
coxph(formula = Surv(start, stop, event) ~ tx + num + size, data = bladder %>%
      filter(interval == 1), cluster = id)

n= 25, number of events= 14

              coef exp(coef) se(coef) robust se      z Pr(>|z|)
tx      0.62451    1.86734  0.63125   0.60350  1.035   0.301
num     0.15217    1.16436  0.21354   0.29048  0.524   0.600
size    0.07575    1.07870  0.20221   0.19346  0.392   0.695

      exp(coef) exp(-coef) lower .95 upper .95
tx           1.867    0.5355   0.5722    6.094
num           1.164    0.8588   0.6589    2.058
size           1.079    0.9270   0.7383    1.576

Concordance= 0.635 (se = 0.082 )
Likelihood ratio test= 1.71 on 3 df,  p=0.6
Wald test               = 2.03 on 3 df,  p=0.6
Score (logrank) test = 1.88 on 3 df,  p=0.6, Robust = 2.35 p=0.5
```

Modelde hiçbir değişken anlamlı çıkmamıştır. Özellikle tedavi etkisi ( $HR = 1.867$ ,  $p = 0.301$ ), istatistiksel olarak anlamlı değildir. Bu, tedavinin birinci olay üzerinde belirgin bir etkisinin olmadığını göstermektedir.

## 10.2 İkinci Olay (interval == 2)

```
> cox_2 <- coxph(Surv(start, stop, event) ~ tx + num + size + cluster(id),
+               data = bladder %>% filter(interval == 2))
> summary(cox_2)
Call:
coxph(formula = Surv(start, stop, event) ~ tx + num + size, data = bladder %>%
      filter(interval == 2), cluster = id)

n= 14, number of events= 10

      coef exp(coef) se(coef) robust se      z Pr(>|z|)
tx      1.70642    5.50923  1.08177    1.36517  1.250   0.211
num     -0.18659    0.82978  0.26968    0.18774 -0.994   0.320
size     0.01589    1.01602  0.34588    0.34971  0.045   0.964

      exp(coef) exp(-coef) lower .95 upper .95
tx           5.5092    0.1815    0.3794    80.006
num           0.8298    1.2051    0.5743     1.199
size           1.0160    0.9842    0.5120     2.016

Concordance= 0.702 (se = 0.138 )
Likelihood ratio test= 6.04 on 3 df,  p=0.1
Wald test               = 8.96 on 3 df,  p=0.03
Score (logrank) test = 6.25 on 3 df,  p=0.1, Robust = 8.21 p=0.04
```

Burada tedavi değişkeni istatistiksel olarak anlamlılığa yaklaşmaktadır, özellikle Wald testi ve Robust score testi ile anlamlılık sınırına girmektedir ( $p = 0.04$ ). Bu durum, tedavi etkisinin ikinci olay üzerinde daha güçlü olduğunu göstermektedir ( $HR = 5.51$ ). Yani tedavi uygulanan bireylerde ikinci olayın meydana gelme riski, plasebo grubuna göre anlamlı düzeyde artmıştır.

## 11. SONUÇ VE GENEL DEĞERLENDİRME

Bu çalışmada, tekrarlayan olayların sağkalım analizi için önerilen farklı modelleme yaklaşımları kullanılarak, mesane kanseri verisi üzerinden ayrıntılı analizler gerçekleştirilmiştir. Amaç, tedavi etkisinin (tx) tekrarlayan olaylar üzerindeki istatistiksel ve klinik anlamlılığını çeşitli modelleme perspektiflerinden incelemek, aynı zamanda Kleinbaum & Klein (2012)'nin "*Recurrent Event Survival Analysis*" başlıklı 8. bölümünde sunulan yaklaşımları uygulamalı olarak değerlendirmektir.

Tedavi etkisi (tx), temel CP ve Marginal modellerde istatistiksel olarak anlamlıdır ve tedavinin olay yaşama riskini yaklaşık 2.7 kat artırdığı gözlenmiştir. Stratifikasyon uygulandığında (SCP), tedavi etkisi daha da güçlenmiş ve p-değeri daha da düşmüştür. Bu durum, olay sıralamasının dikkate alınmasının sonuçları etkileyebileceğini göstermektedir.

Gap Time modeli, zaman yapılandırmasının değiştirilmesiyle tedavi etkisinin anlamlılığını yitirdiğini göstermektedir. Bu da model seçimlerinin klinik yorumları değiştirebileceğini ortaya koyar. Sadece ilk olaylara dayalı analizlerde tedavi etkisi anlamlı bulunmazken, ikinci olaylar üzerinde yapılan analizlerde tedavi etkisi yüksek çıkmış ancak küçük örneklem nedeniyle anlamlılığa ulaşmamıştır.

Interaction modeli de benzer şekilde, tedavi etkisinin olay sırasına göre değiştiğini varsaymasına rağmen istatistiksel olarak güçlü bir bulgu sunmamıştır. Shared frailty modeli, birey içi bağımlılığı doğrudan modele dahil ederek yine anlamlı bir tedavi etkisi ortaya koymuştur. Ancak bireyler arası varyans çok düşük bulunmuş ve kırılabilirliğin anlamlı katkısı olmamıştır.

### **4.3. Klinik ve Yöntemsel Çıkarımlar**

Tedavi, özellikle ikinci olay gibi daha ciddi evrelerde etkili olabilir. Bu durum, Kleinbaum & Klein (2012)'de de vurgulandığı gibi ilk ve ikinci olayların doğasının farklı olabileceğini düşündürmektedir. Model tercihi, yalnızca istatistiksel uygunluk açısından değil, klinik bağlamda hangi soruya cevap aranacağına göre yapılmalıdır.

Farklı modellerle yapılan bu analizler, tedavinin bazı durumlarda koruyucu olabileceğini, özellikle olay sayısı arttıkça etkisinin daha belirgin hale gelebileceğini göstermektedir. Ancak bu etki tüm modellerde tutarlı değildir. Bu nedenle, model seçiminin araştırmanın amacı ve veri yapısı doğrultusunda dikkatle yapılması büyük önem taşımaktadır.

## 12. Kaynakça

<https://www.kaggle.com/datasets/emmanueljegou/counting-process-format-for-survival-data>

Kleinbaum, D. G., & Klein, M. (2012). *Survival Analysis: A Self-Learning Text* (3rd ed.). New York: Springer.

Byar, D. P. (1980). The Veterans Administration Study of Chemoprophylaxis for Recurrence of Superficial Bladder Cancer: Comparisons of Placebo, Pyridoxine, and Thiotepa. *Cancer*, 45(S10), 542–548.

Wei, L. J., Lin, D. Y., & Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *Journal of the American Statistical Association*, 84(408), 1065–1073.