SME0821 - Análise de Sobrevivência - Atividade II

Francisco Rosa Dias de Miranda - 4402962

maio 2022

Exercício 1

Considerando o modelo com função de risco dada por:

$$h(t) = \delta t + \frac{\theta}{1 + \beta t}, \quad t > 0$$

onde $\beta \geq 0, \delta \geq 0$ e $\theta \geq 0$.

Para verificar o comportamento de h, primeiramente vamos obter $h'(t) = \frac{\partial h}{\partial t}$

$$h'(t) = \frac{\partial}{\partial t}\left[\delta t + \frac{\theta}{1 + \beta t}\right] = \delta + \theta \, \frac{\partial}{\partial t}\left((1 + \beta t)^{-1}\right) = \delta + \theta(-(1 + \beta t)^{-2} \, \frac{\partial}{\partial t}\,(1 + \beta t)) = \delta - \frac{\theta \beta}{(1 + \beta t)^2}$$

- (a) $\theta = 0 \Rightarrow \; h$ é crescente

Se $\theta = 0$ então

$$h'(t) = \delta - \frac{0\beta}{(1 + \beta t)^2} = \delta$$

Como $\delta > 0$, h é crescente.

- (b) $\delta = 0 \Rightarrow \; h$ é decrescente

Se $\delta = 0 \Rightarrow \; h$ então

$$h'(t) = 0 - \frac{\theta \beta}{(1 + \beta t)^2} = -\theta \beta (1 + \beta t)^{-2} < 0$$

Pois $t > 0, \; \beta \geq 0, e \; \theta \geq 0$.

- (c) $\beta = \delta = 0 \Rightarrow \; h$ é constante

Substituindo os valores, temos que

$$h'(t) = 0 - \frac{0\theta}{(1+0t)^2} = 0$$

Portanto, $h(t)$ é constante.

- (d) Quando $\delta \geq \beta\theta$, h é crescente

$$h'(t) = \delta - \frac{\theta\beta}{(1+\beta t)^2} = \frac{\delta(1+\beta t)^2 - \theta\beta}{(1+\beta t)^2}$$

Portanto, $h(t)$ é constante.

- (f)

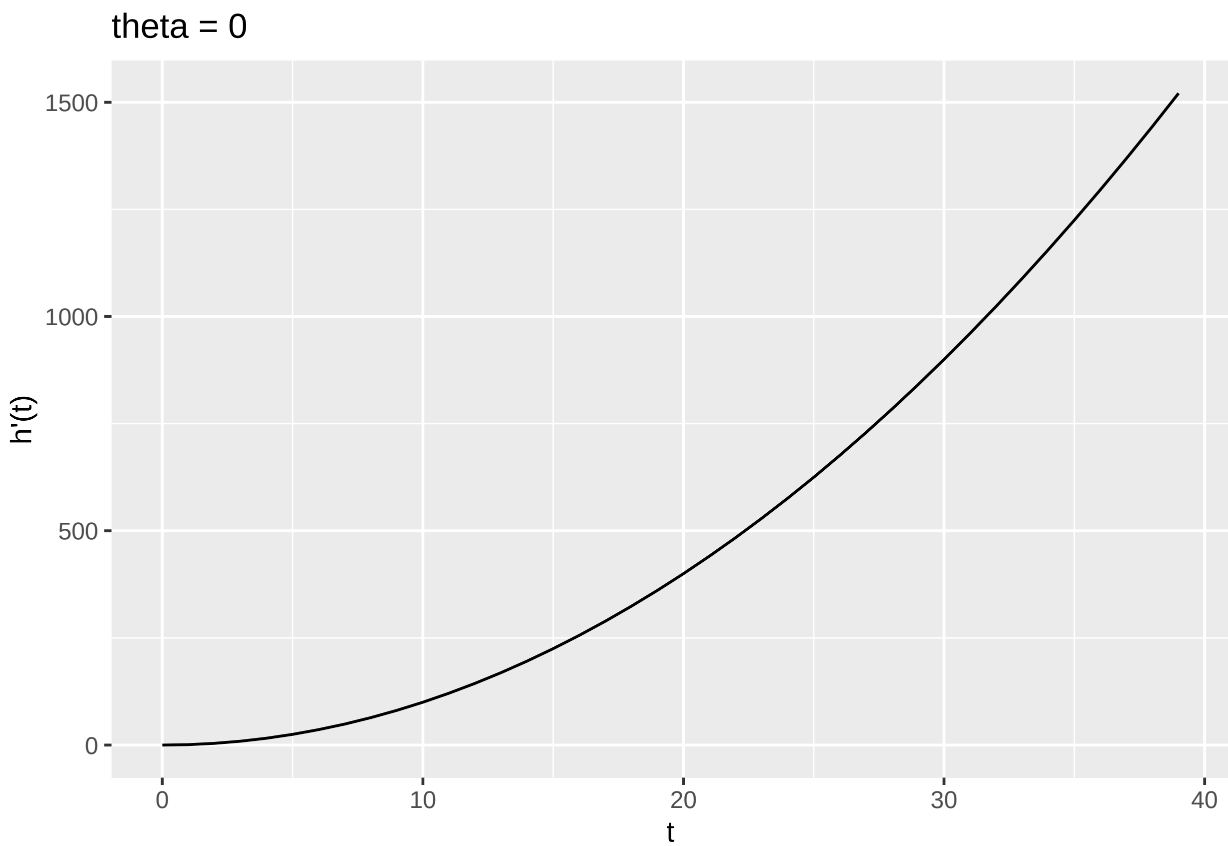
```
library(tidyverse)
library(survival)
library(survminer)
library(KMsurv)
library(biostat3)
```

```
h <- function(t, delta, theta, beta){
  delta*t + theta/(1 + beta*t)
}

h_prime <- function(t, delta, theta, beta){
  delta - (theta*beta)/(1+beta*t)^2
}
```

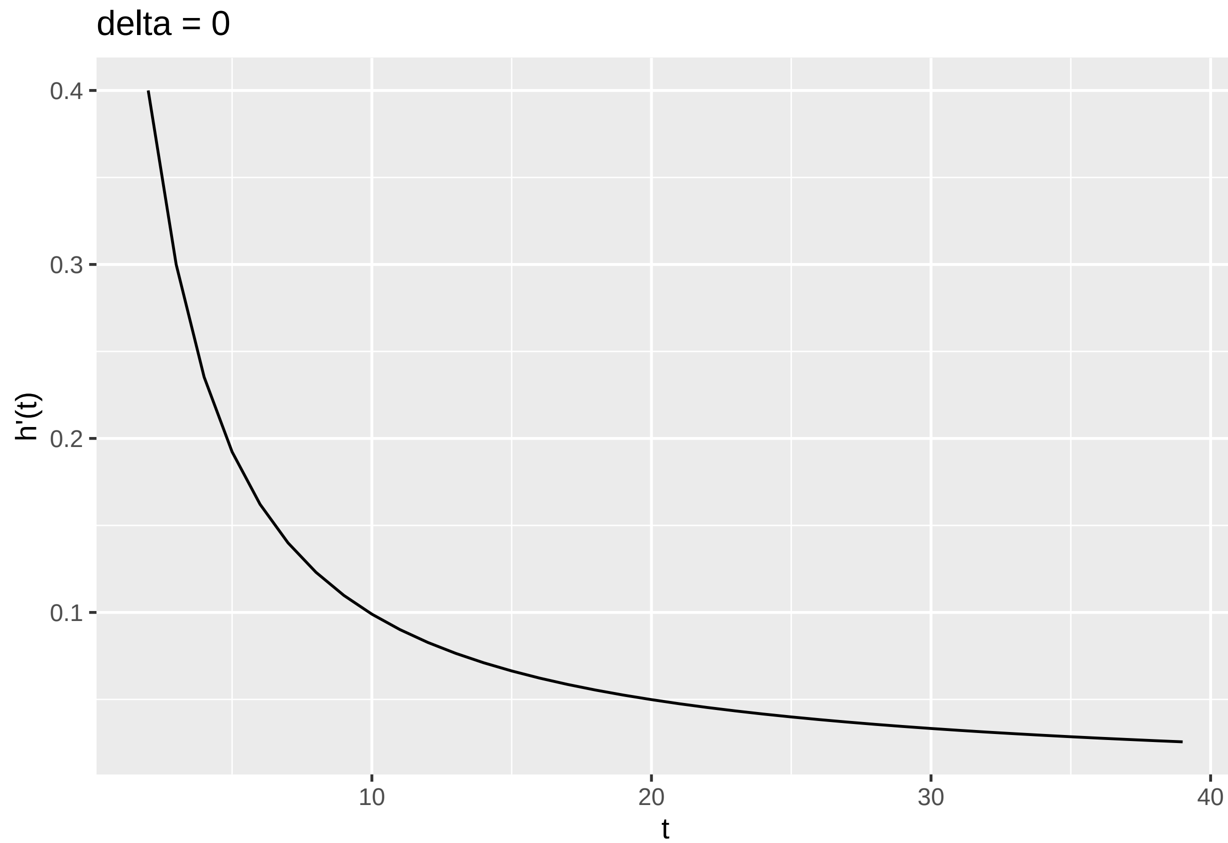
```
t <- seq(0, 39)
delta <-  seq(0,39)
theta <- seq(0,39)
beta <- seq(0,39)

tibble(t, delta, theta, beta) %>% mutate(h = h(t,delta,theta,beta))
```

```
## # A tibble: 40 x 5
##         t delta theta  beta     h
##     <int> <int> <int> <int> <dbl>
##  1      0     0     0     0   0
##  2      1     1     1     1   1.5
##  3      2     2     2     2   4.4
##  4      3     3     3     3   9.3
##  5      4     4     4     4  16.2
##  6      5     5     5     5  25.2
##  7      6     6     6     6  36.2
##  8      7     7     7     7  49.1
##  9      8     8     8     8  64.1
## 10      9     9     9     9  81.1
## # ... with 30 more rows
```

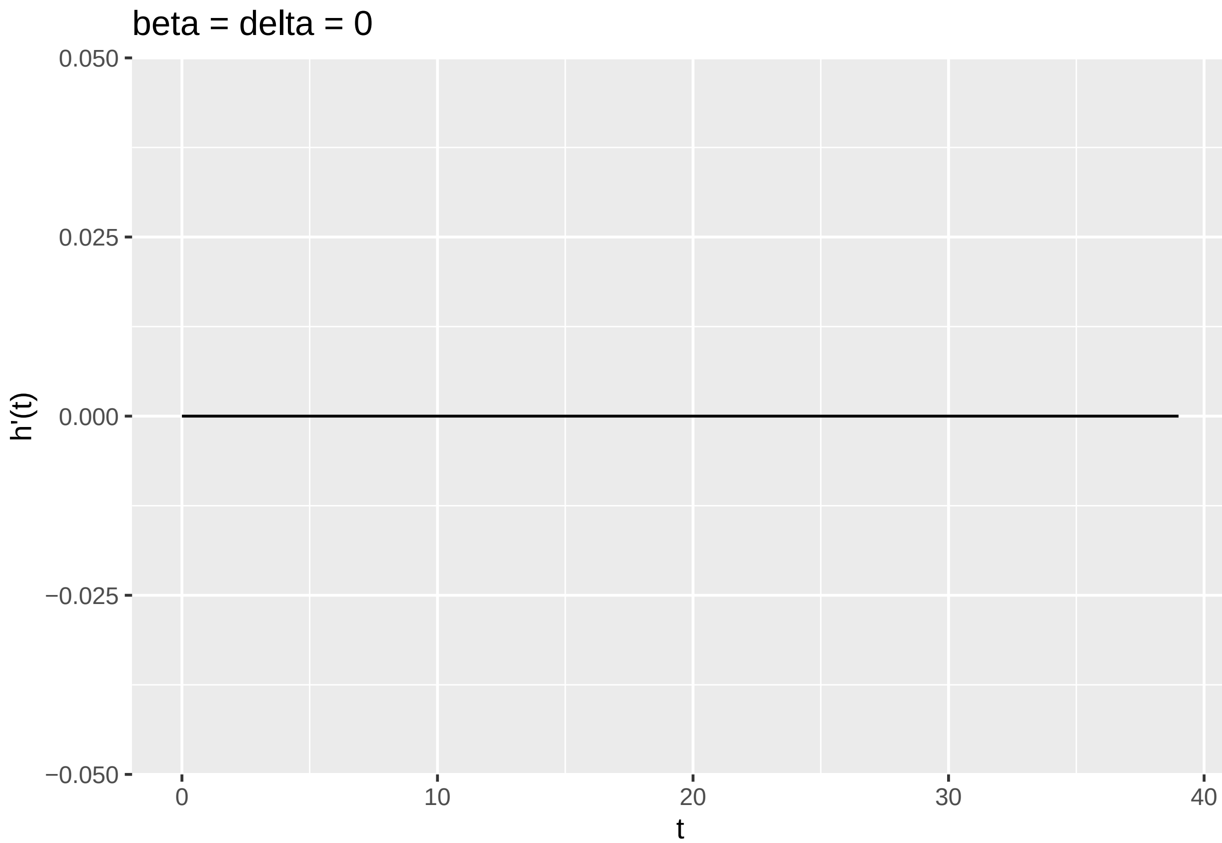
```
tibble(t, delta, beta) %>% mutate(h = h(t,delta,0,beta)) %>%
  ggplot() +
  geom_line(aes(x = t, y = h )) +
  labs(title = expression("theta = 0"),
       y = "h'(t)")
```



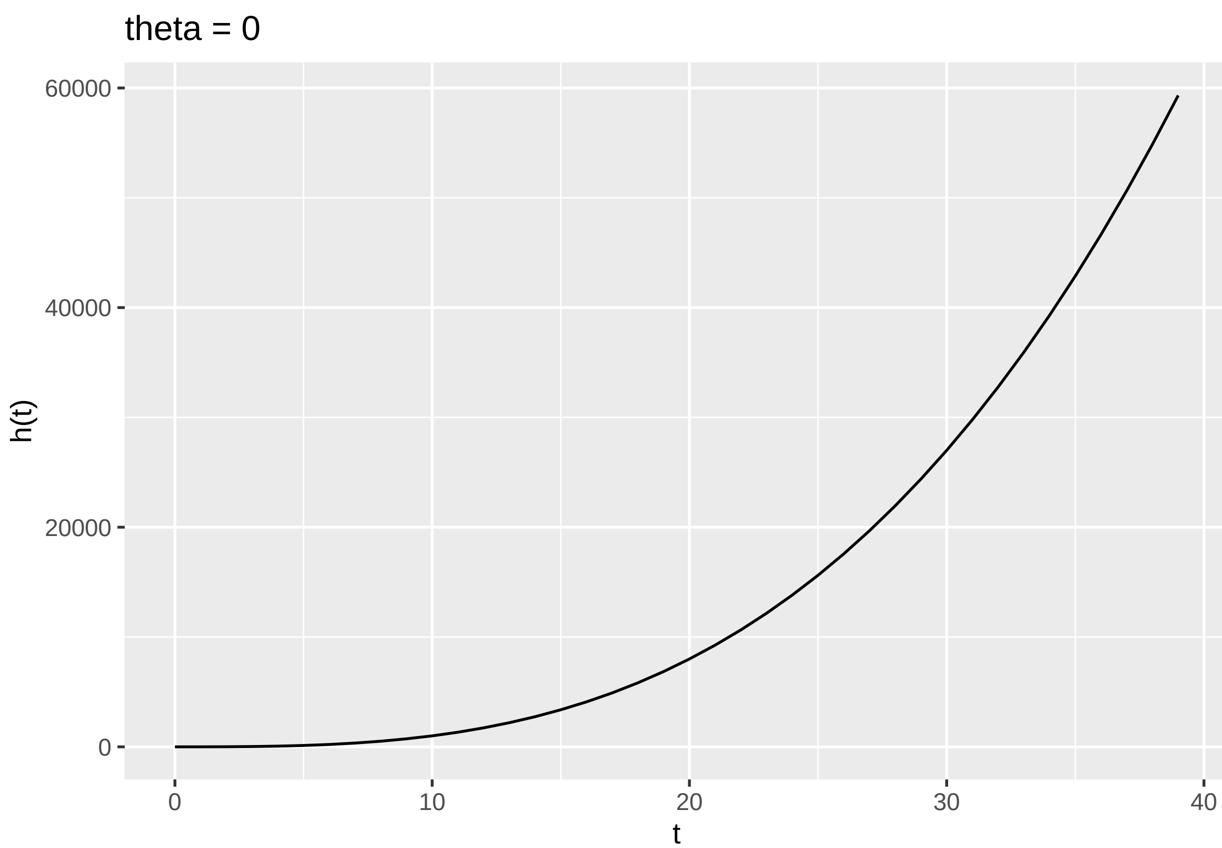

```
tibble(t, theta, beta) %>% filter(t>1) %>% mutate(h = h(t,0,theta ,beta)) %>%
  ggplot() +
  geom_line(aes(x = t, y = h )) +
  labs(title = expression("delta = 0"),
       y = "h'(t)")
```




```
tibble(t, theta) %>% mutate(h = h_prime(t,0,theta ,0)) %>%
  ggplot() +
  geom_line(aes(x = t, y = h )) +
  labs(title = expression("beta = delta = 0"),
       y = "h'(t)")
```

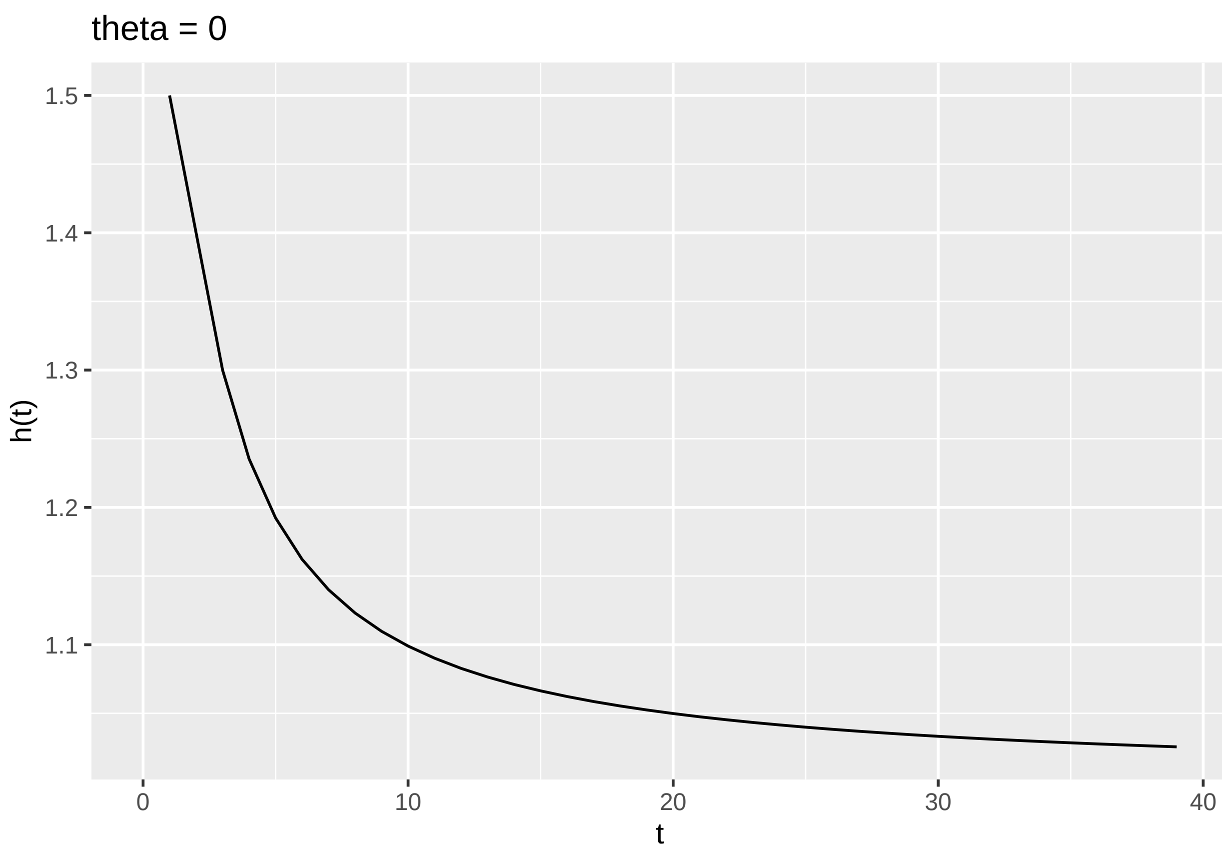



```r
# D) delta >= beta * theta
tibble(t, delta = theta*beta, theta, beta) %>%
  mutate(h = h(t, delta, theta ,beta)) %>%
  ggplot() +
  geom_line(aes(x = t, y = h )) +
  labs(title = expression("theta = 0"),
       y = "h(t)")
```

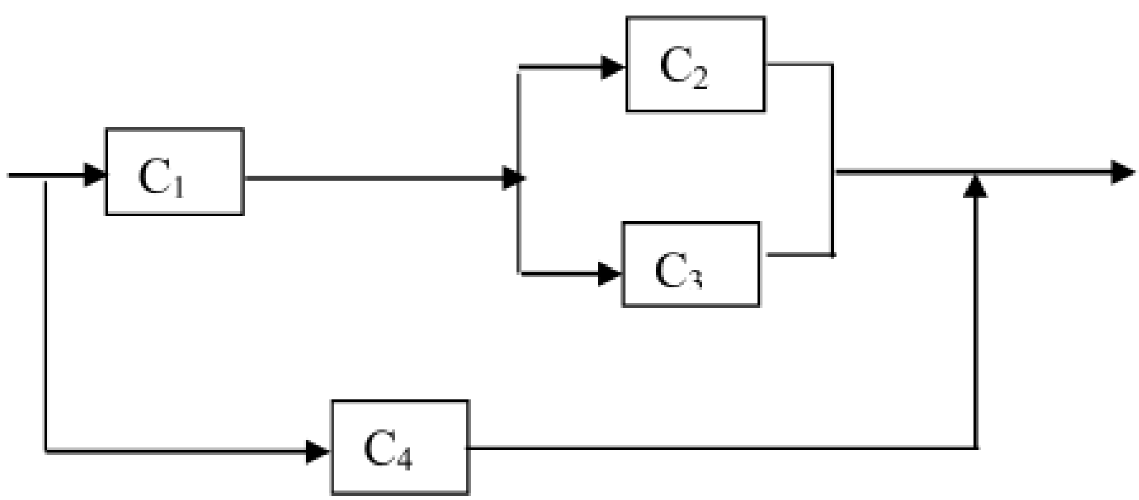



```
# E) 0 < delta <= beta * theta
tibble(t, delta = delta/(theta*beta), theta, beta) %>%
  mutate(h = h(t, delta, theta ,beta)) %>%
  ggplot() +
  geom_line(aes(x = t, y = h )) +
  labs(title = expression("theta = 0"),
       y = "h(t)")
```

```
## Warning: Removed 1 row(s) containing missing values (geom_path).
```



Exercício 1 - Lista 2



(i) confiabilidade do sistema por 8 meses

Considere T: tempo em anos até a falha do sistema.

$T_i = $ tempo até a falha do i-ésimo componente, $i = 1, ..., 4$

$$f_i(t, \beta) = \begin{cases} 2\beta t e^{-\beta t^2}, \ t > 0 \ ; \quad i = 1, ...4 \\ 0, \quad \text{c.c.} \end{cases}$$

Note que $T_i \sim$ **Weibull**$(2, \beta)$. Assim:

$$R_i(t) = e^{-\beta t^2}, \ t > 0, \ i = 1, ..., 4$$

- C_2 e C_3 estão em paralelo, logo:

$$R_{23}(t) = 1 - \prod_{i=1}^{2}(1 - R_i(t)) = 1 - (1 - e^{-\beta t^2})^2 = 1 - (1 - 2e^{-\beta t^2} + e^{-2\beta t^2}) = 2e^{-\beta t^2} - e^{-2\beta t^2}$$

- C_2 e C_3 estão em série com C_1, logo:

$$R_{123} = R_1(t)R_{23}(t) = e^{-\beta t^2}(2e^{-\beta t^2} - e^{-2\beta t^2}) = 2e^{-2\beta t^2} - e^{-3\beta t^2}$$

- C_4 está em paralelo com C_1, C_2, C_3, assim:

$$R(t) = 1 - (1 - R_4(t)(1 - R_{123}(t)) = 1 - (1 - e^{-\beta t^2})(1 - 2e^{-2\beta t^2} - e^{-3\beta t^2}) =$$

$$1 - (1 - 2e^{-2\beta t^2} - e^{-3\beta t^2} + e^{-\beta t^2} - 2e^{-3\beta t^2} - e^{-4\beta t^2}) = 2e^{-2\beta t^2} - 3e^{-3\beta t^2} + e^{-\beta t^2} + e^{-4\beta t^2}$$

Como queremos a confiabilidade por um período de 8 meses (2/3 ano):

$$R(2/3) = 2e^{-\frac{8}{9}\beta} - 3e^{-\frac{12}{9}\beta} + e^{-\frac{4}{9}\beta} + e^{-\frac{16}{9}\beta}$$

(ii) média de vida do sistema

$$E(T) = \int_0^\infty R(t)dt = \int_0^\infty 2e^{-2\beta t^2}dt - \int_0^\infty 3e^{-3\beta t^2}dt + \int_0^\infty e^{-\beta t^2}dt + \int_0^\infty e^{-4\beta t^2}dt$$

$$\int_0^\infty 2e^{-2\beta t^2}dt = 2\sqrt{\frac{2\pi}{4\beta}}\int_0^\infty \sqrt{\frac{4\beta}{2\pi}}e^{-\frac{t^2}{2(\frac{1}{4}di\beta)}}dt = \frac{3}{2}\sqrt{\frac{\pi}{2\beta}}$$

$$\int_0^\infty e^{-\beta t^2}dt = \sqrt{\frac{2\pi}{2\beta}}\int_0^\infty \sqrt{\frac{2\beta}{2\pi}}e^{-\beta t^2}dt = \frac{1}{2}\sqrt{\frac{\pi}{\beta}}$$

$$\int_0^\infty e^{-4\beta t^2}dt = \sqrt{\frac{2\pi}{8\beta}}\int_0^\infty \sqrt{\frac{8\beta}{2\pi}}e^{-\frac{t^2}{2(\frac{1}{8}\beta)}}dt = \frac{1}{2}\sqrt{\frac{\pi}{4\beta}}$$

Assim, temos:

$$E(T) = \frac{3}{2}\sqrt{\frac{\pi}{2\beta}} - \frac{3}{2}\sqrt{\frac{\pi}{2\beta}} + \frac{1}{2}\sqrt{\frac{\pi}{\beta}} + \frac{1}{2}\sqrt{\frac{\pi}{4\beta}}$$

Exercício 2 - Lista 3

(a)

Considere T : tempo em semanas até remissão da leucemia; T_i : tempo de remissão do i-ésimo paciente, com $i = 1, ..., 30$.

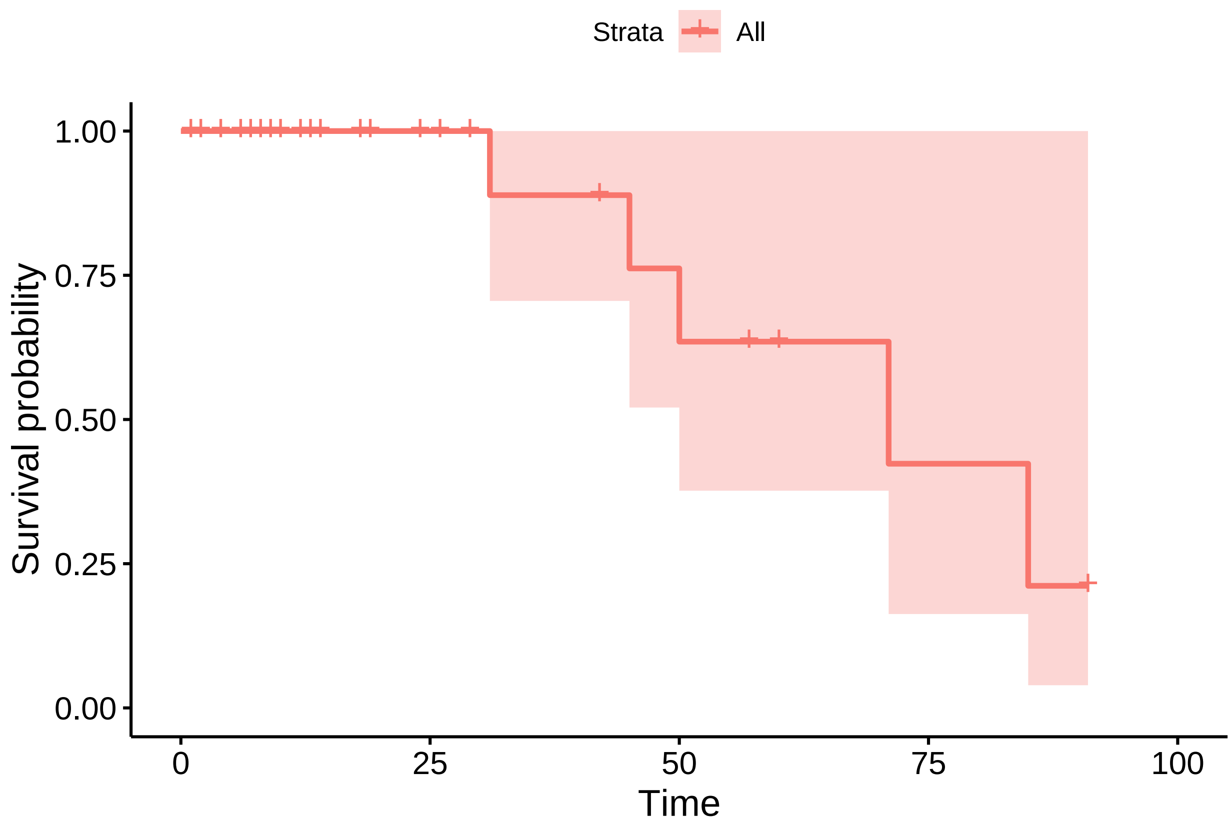
```
df <- tibble(t = c(1, 1, 2, 4, 4, 6, 6, 7, 8, 9,
                   9, 10, 12, 13, 14, 18, 19, 24, 26, 29,
                   31, 42, 45 ,50, 57, 60, 71, 85, 91),
             d = c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
                   0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
                   1, 0, 1, 1, 0, 0, 1, 1, 0))
```

```
km_fit <- survfit(Surv(t, d) ~ 0, data = df)
```

```
summary(km_fit, times = seq(0,90,10))
```

```
## Call: survfit(formula = Surv(t, d) ~ 0, data = df)
##
##  time n.risk n.event survival std.err lower 95% CI upper 95% CI
##     0     29       0    1.000   0.000       1.0000            1
##    10     18       0    1.000   0.000       1.0000            1
##    20     12       0    1.000   0.000       1.0000            1
##    30      9       0    1.000   0.000       1.0000            1
##    40      8       1    0.889   0.105       0.7056            1
##    50      6       2    0.635   0.169       0.3766            1
##    60      4       0    0.635   0.169       0.3766            1
##    70      3       0    0.635   0.169       0.3766            1
##    80      2       1    0.423   0.206       0.1628            1
##    90      1       1    0.212   0.182       0.0393            1
```

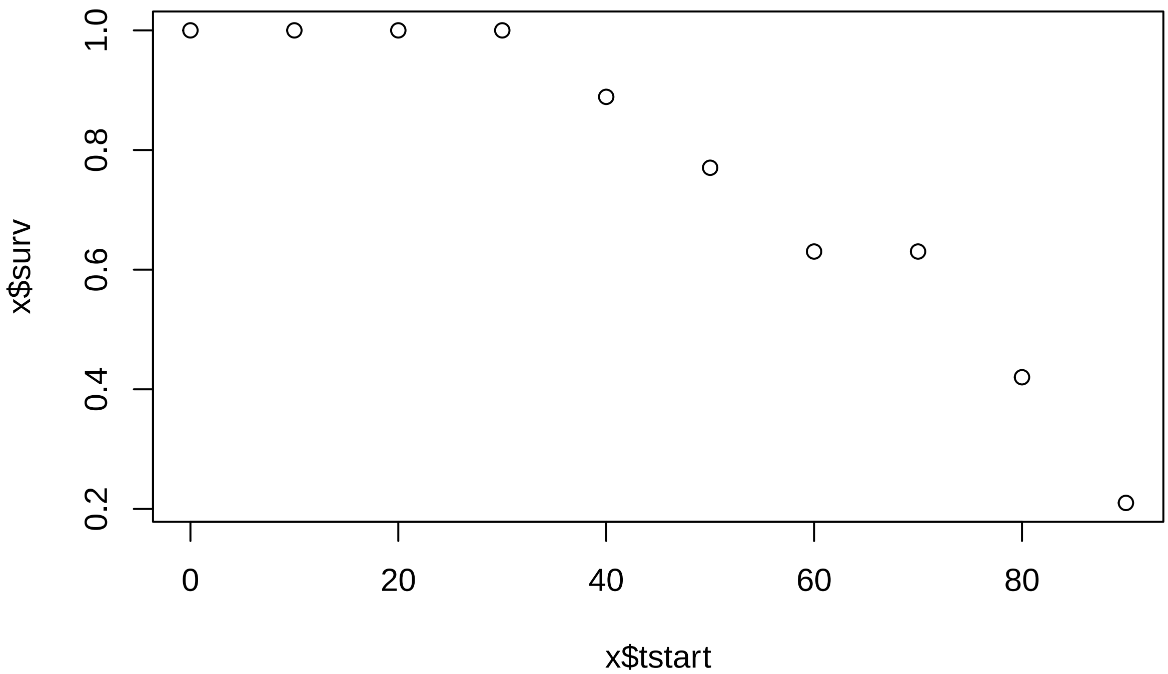
```
ggsurvplot(km_fit)
```




```
(x <- lifetab2(Surv(t,d==1)~0, df, breaks = seq(0,90,10)))
```

```
##         tstart tstop nsubs nlost nrisk nevent      surv        pdf     hazard
## 0-10         0    10    29    11  23.5      0 1.0000000 0.00000000 0.00000000
## 10-20       10    20    18     6  15.0      0 1.0000000 0.00000000 0.00000000
## 20-30       20    30    12     3  10.5      0 1.0000000 0.00000000 0.00000000
## 30-40       30    40     9     0   9.0      1 1.0000000 0.01111111 0.01176471
## 40-50       40    50     8     1   7.5      1 0.8888889 0.01185185 0.01428571
## 50-60       50    60     6     1   5.5      1 0.7703704 0.01400673 0.02000000
## 60-70       60    70     4     1   3.5      0 0.6303030 0.00000000 0.00000000
## 70-80       70    80     3     0   3.0      1 0.6303030 0.02101010 0.04000000
## 80-90       80    90     2     0   2.0      1 0.4202020 0.02101010 0.06666667
## 90-Inf      90   Inf     1     1   0.5      0 0.2101010         NA         NA
##           se.surv     se.pdf  se.hazard
## 0-10    0.0000000        NaN        NaN
## 10-20   0.0000000        NaN        NaN
## 20-30   0.0000000        NaN        NaN
## 30-40   0.0000000 0.01047566 0.01174433
## 40-50   0.1047566 0.01112153 0.01424922
## 50-60   0.1428860 0.01293318 0.01989975
## 60-70   0.1723919        NaN        NaN
## 70-80   0.1723919 0.01809154 0.03919184
## 80-90   0.2064866 0.01809154 0.06285394
## 90-Inf  0.1809154         NA         NA
```

```
biostat3::plot.lifetab2(x)
```



Ambos os estimadores obtiveram valores similares, devido ao número de intervalos ser o mesmo. Tal comportamento não seria esperado no caso de um número diferente de intervalos em cada estimador.

Exercício 6 - Lista 4

Seja $T = $ tempo(semanas) até a remissão de um grupo de pacientes com leucemia. Tem-se que $T \sim$ Weibull(α, β). Assim

$$L(\alpha, \beta, t_i) = \prod_{i=1}^{n} (h(t_i, \alpha, \beta))^{\delta_i} S(t_i; \alpha, \beta) = \prod_{i=1}^{n} (\alpha \beta t_i^{\alpha - 1})^{\delta_i} e^{-\beta t_i^{\alpha}} = \alpha^r \beta^r \exp\left(-\beta \sum_{i=1}^{n} t_i^{\alpha} + (\alpha - 1) \sum_{i=1}^{n} \delta_i \ln(t_i)\right)$$

em que $r = \sum_{i=1}^{n} \delta_i$, com δ_i um indicador que não houve censura.

$$l(\alpha, \beta) = r \ln \alpha + r \ln \beta - \beta \sum_{i=1}^{n} \delta_i \ln t_i$$

$$\frac{\partial l(\alpha, \beta)}{\partial \alpha} = \frac{r}{\alpha} - \beta \sum_{i=1}^{n} t_i^{\alpha} \ln(t_i) + \sum_{i=1}^{n} \delta_i \ln(t_i) = u_1$$

$$\frac{\partial l(\alpha, \beta)}{\partial \beta} = \frac{r}{\alpha} - \beta \sum_{i=1}^{n} t_i^{\alpha} = u_2$$

$$\frac{\partial^2 l(\alpha,\beta)}{\partial\alpha^2} = -\frac{r}{\alpha^2} - \beta\sum_{i=1}^{n} t_i^\alpha (\ln(t_i))^2$$

$$\frac{\partial^2 l(\alpha,\beta)}{\partial\beta^2} = -\frac{r}{\beta^2}$$

$$\frac{\partial^2 l(\alpha,\beta)}{\partial\alpha\partial\beta} = \sum_{i=1}^{n} t_i^\alpha \ln(t_i)$$

Agora, fazendo

$$\left[\frac{\partial l(\alpha,\beta)}{\partial\beta}\right]_{\beta=\hat\beta} = 0 \Rightarrow \hat\beta = \frac{r}{\sum_{i=1}^{n} t_i^\alpha}$$

Substituindo $\hat\beta$ em $\frac{\partial l(\alpha,\beta)}{\partial\alpha}$ segue:

$$u_1 = \frac{r}{\alpha} - \frac{r\sum_{i=1}^{n} t_i^\alpha \ln(t_i)}{\sum_{i=1}^{n} t_i^\alpha} + \sum_{i=1}^{n}\delta_i \ln(t_i)$$

Pelo método de Newton-Rapson, obtemos os EMV de α e β: $\hat\alpha = 1,407,\ \hat\beta = 0,044$

b

Sabe-se que

$$t_{0,5} = \beta^{-\frac{1}{2}}(\ln 2)^{-\frac{1}{2}} \Rightarrow t_{0,5} = 7,137$$