

# Terceira Lista - Análise de Sobrevivência

Gabriel D'assumpção de Carvalho

2025-08-05

## Introdução

Este documento apresenta uma análise exploratória de sobrevivência com dados sobre hepatite, reincidência de tumor sólido e malária. A análise utilizará os estimadores de Kaplan-Meier, Nelson-Aalen e Atuaria para comparar as curvas de sobrevivência entre os grupos de cada conjunto de dados.

## Bibliotecas

```
# Carregamento dos pacotes necessários

# Pacote para análise de sobrevivência
# install.packages("survival")
library(survival)

# Pacote para visualização dos dados
# install.packages("ggplot2")
library(ggplot2)

# Pacote para manipulação de dados
# install.packages("dplyr")
library(dplyr)
```

## Funções de análise de sobrevivência

A seguir vamos estruturar as funções de sobrevivência não paramétricas, onde vai ser utilizado o estimador Atuaria, Kaplan-Meier e Nelson-Aalen. Essas funções serão utilizadas para calcular as curvas de sobrevivência e os intervalos de confiança.

```
##### Estimador Atuaria #####
eact <- function(time, evento, classEvento = 1, alpha = 0.05, conf = "normal") {
  tempo_evento <- sort(unique(time[evento == classEvento]))

  data <- data.frame(
    time = tempo_evento,
    n_risk = NA,
    n_event = NA,
    censored = NA,
    n_risk_adj = NA,
    q_i = NA,
    p_i = NA,
    S_t_i = NA,
    std_error = NA,
    lower = NA,
    upper = NA
  )

  S_t_back <- 1

  for (i in seq_along(data$time)) {
    t <- data$time[i]

    data$n_risk[i] <- sum(time >= t)
```

```

data$n_event[i] <- sum(time == t & evento == classEvento)
data$censored[i] <- sum(time == t & evento != classEvento)
data$n_risk_adj[i] <- data$n_risk[i] - data$censored[i] / 2

if (data$n_risk_adj[i] > 0) {
  data$q_i[i] <- data$n_event[i] / data$n_risk_adj[i]
  data$p_i[i] <- 1 - data$q_i[i]
  data$S_t_i[i] <- S_t_back * data$p_i[i]
  S_t_back <- data$S_t_i[i]
} else {
  data$S_t_i[i] <- S_t_back
}

# Variância (usada nos dois métodos)
var_S_t <- data$S_t_i[i]^2 * sum(data$q_i[1:i] / (data$n_risk_adj[1:i] *
data$p_i[1:i]))
data$std_error[i] <- sqrt(var_S_t)

if (conf == "normal") {
  data$upper[i] <- data$S_t_i[i] + qnorm(1 - alpha / 2) * data$std_error[i]
  data$lower[i] <- data$S_t_i[i] - qnorm(1 - alpha / 2) * data$std_error[i]
}

if (conf == "log-log") {
  if (data$S_t_i[i] > 0 && data$S_t_i[i] < 1) {
    loglog_S <- log(-log(data$S_t_i[i]))
    se_loglog <- sqrt(sum(data$q_i[1:i] / (data$n_risk_adj[1:i] * data$p_i[1:i])))
    z <- qnorm(1 - alpha / 2)
    lower_loglog <- loglog_S - z * se_loglog
    upper_loglog <- loglog_S + z * se_loglog
    data$lower[i] <- exp(-exp(upper_loglog))
    data$upper[i] <- exp(-exp(lower_loglog))
  } else {
    data$lower[i] <- NA
    data$upper[i] <- NA
  }
}
}
cols_to_round <- c("n_risk_adj", "q_i", "p_i", "S_t_i", "std_error", "lower", "upper")
data[cols_to_round] <- round(data[cols_to_round], 3)

return(data)
}

##### Estimador de Kaplan-Meier #####
ekm <- function(time, evento, classEvento = 1, alpha = 0.05, conf = "normal") {
  tempo_evento <- sort(unique(time[evento == classEvento]))

  data <- data.frame(
    time = tempo_evento,
    n_risk = NA,
    n_event = NA,

```

```

    censored = NA,
    q_i = NA,
    p_i = NA,
    S_t_i = NA,
    std_error = NA,
    lower = NA,
    upper = NA
  )

  S_t_back <- 1

  for (i in seq_along(data$time)) {
    t <- data$time[i]

    data$n_risk[i] <- sum(time >= t)
    data$n_event[i] <- sum(time == t & evento == classEvento)
    data$censored[i] <- sum(time == t & evento != classEvento)

    if (data$n_risk[i] > 0) {
      data$q_i[i] <- data$n_event[i] / data$n_risk[i]
      data$p_i[i] <- 1 - data$q_i[i]
      data$S_t_i[i] <- S_t_back * data$p_i[i]
      S_t_back <- data$S_t_i[i]
    } else {
      data$S_t_i[i] <- S_t_back
    }

    term <- data$n_event[1:i] / (data$n_risk[1:i] * (data$n_risk[1:i] -
data$n_event[1:i]))
    term[is.infinite(term)] <- 0
    var_S_t <- data$S_t_i[i]^2 * sum(term, na.rm = TRUE)

    data$std_error[i] <- sqrt(var_S_t)

    if (conf == "log-log") {
      if (data$S_t_i[i] > 0 && data$S_t_i[i] < 1) {
        z <- qnorm(1 - alpha / 2)
        var_loglog <- sum(term, na.rm = TRUE)
        loglog_S <- log(-log(data$S_t_i[i]))

        lower_loglog <- loglog_S - z * sqrt(var_loglog)
        upper_loglog <- loglog_S + z * sqrt(var_loglog)

        data$lower[i] <- exp(-exp(upper_loglog))
        data$upper[i] <- exp(-exp(lower_loglog))
      } else {
        data$lower[i] <- NA
        data$upper[i] <- NA
      }
    }

    if (conf == "normal") {
      if (data$S_t_i[i] > 0 && data$S_t_i[i] < 1) {

```

```

    data$upper[i] <- data$S_t_i[i] + qnorm(1 - alpha / 2) * data$std_error[i]
    data$lower[i] <- data$S_t_i[i] - qnorm(1 - alpha / 2) * data$std_error[i]
  } else {
    data$upper[i] <- NA
    data$lower[i] <- NA
  }
}
}
cols_to_round <- c("q_i", "p_i", "S_t_i", "std_error", "lower", "upper")
data[cols_to_round] <- round(data[cols_to_round], 3)

return(data)
}

##### Estimador de Nelson-Aalen #####
enelson <- function(time, evento, classEvento = 1, alpha = 0.05, conf = "normal") {
  tempo_evento <- sort(unique(time[evento == classEvento]))

  data <- data.frame(
    time = tempo_evento,
    n_risk = NA,
    n_event = NA,
    q_i = NA,
    p_i = NA,
    S_t_i = NA,
    std_error = NA,
    lower = NA,
    upper = NA
  )

  for (i in seq_along(data$time)) {
    t <- data$time[i]

    data$n_risk[i] <- sum(time >= t)
    data$n_event[i] <- sum(time == t & evento == classEvento)

    if (data$n_risk[i] > 0) {
      data$q_i[i] <- data$n_event[i] / data$n_risk[i]
      data$p_i[i] <- 1 - data$q_i[i]
      estimatorNelson <- sum(data$n_event[1:i] / data$n_risk[1:i])
      data$S_t_i[i] <- exp(-estimatorNelson)
    } else {
      data$S_t_i[i] <- 1
    }
  }
  var_S_t <- data$S_t_i[i]**2 * sum(data$n_event[1:i] / data$n_risk[1:i]**2)
  data$std_error[i] <- sqrt(var_S_t)

  if (conf == "log-log") {
    z <- qnorm(1 - alpha / 2)
    if (data$S_t_i[i] > 0 && data$S_t_i[i] < 1) {
      loglog_S <- log(-log(data$S_t_i[i]))
      var_loglog <- sum(data$n_event[1:i] / (data$n_risk[1:i]^2))
    }
  }
}

```

```

    lower_loglog <- loglog_S - z * sqrt(var_loglog)
    upper_loglog <- loglog_S + z * sqrt(var_loglog)

    data$lower[i] <- exp(-exp(upper_loglog))
    data$upper[i] <- exp(-exp(lower_loglog))
  } else {
    data$lower[i] <- NA
    data$upper[i] <- NA
  }
}

if (conf == "normal") {
  if (data$S_t_i[i] > 0 && data$S_t_i[i] < 1) {
    data$upper[i] <- data$S_t_i[i] + qnorm(1 - alpha / 2) * data$std_error[i]
    data$lower[i] <- data$S_t_i[i] - qnorm(1 - alpha / 2) * data$std_error[i]
  } else {
    data$upper[i] <- NA
    data$lower[i] <- NA
  }
}
}

cols_to_round <- c("q_i", "p_i", "S_t_i", "std_error", "lower", "upper")
data[cols_to_round] <- round(data[cols_to_round], 3)

return(data)
}

```

## Hepatitis

Neste estudo, foram analisados dados de pacientes com hepatite viral aguda. O objetivo foi comparar a taxa de sobrevivência entre dois grupos, foram escolhidos 29 pacientes com hepatite viral aguda e separados aleatoriamente em ambos os conjuntos. Os pacientes foram acompanhados durante 16 semanas (4 meses) ou até a falha ou censura do paciente.

- **Tipo de estudo:** Estudo clínico randomizado;
- **Grupo de tratamento:** Pacientes que receberam terapia com esteroides;
- **Grupo controle:** Pacientes que receberam placebo.

```

##### Hepatite #####
hControle <- data.frame(
  tempo = c(
    1, 2, 3, 3, 5, 5, 16, 16, 16, 16,
    16, 16, 16, 16
  ),
  evento = c(0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)
)

hCasos <- data.frame(
  tempo = c(
    1, 1, 1, 1, 4, 5, 7, 8, 10, 10, 12, 16, 16,

```

```

16
),
evento = c(1, 1, 1, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0, 0)
)

```

Como os tempos até a falha são registrados em semanas (uma variável discreta), um gráfico de barras é a escolha ideal para visualizar a distribuição dos eventos ao longo do tempo. Isso nos permitirá comparar diretamente a frequência de falhas em cada semana entre os grupos controle e casos.

```

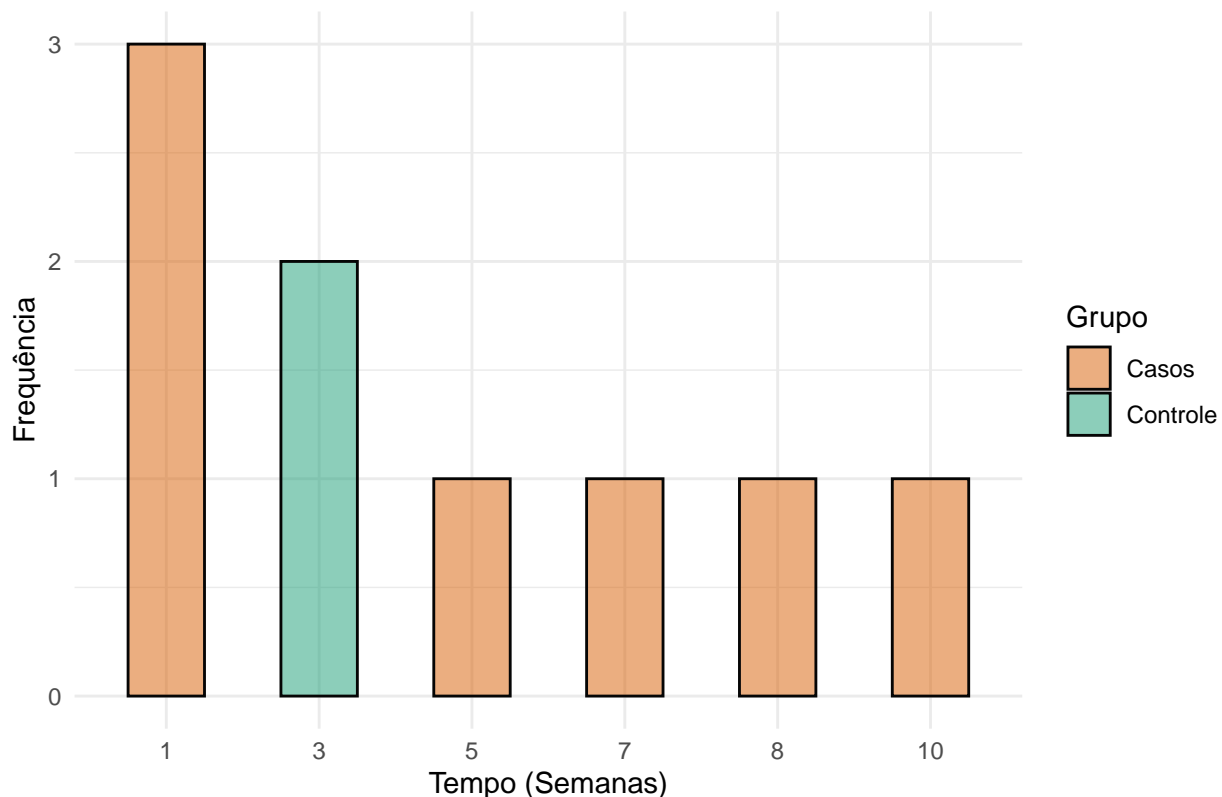
dados <- rbind(
  data.frame(tempo = hControle$tempo[hControle$evento == 1], grupo = "Controle"),
  data.frame(tempo = hCasos$tempo[hCasos$evento == 1], grupo = "Casos")
)

dados$tempo <- factor(dados$tempo, levels = sort(unique(dados$tempo)))

ggplot(dados, aes(x = tempo, fill = grupo)) +
  geom_bar(position = "identity", alpha = 0.5, color = "#050505", width = 0.5) +
  labs(
    title = "Gráfico de Barras dos Tempos até o Evento - Hepatite",
    x = "Tempo (Semanas)",
    y = "Frequência",
    fill = "Grupo"
  ) +
  scale_fill_manual(values = c("Controle" = "#1b9e77", "Casos" = "#d95f02")) +
  theme_minimal()

```

Gráfico de Barras dos Tempos até o Evento – Hepatite



```
print(length(hControle$tempo))
```

```
## [1] 15
```

```
print(length(hCasos$tempo))
```

```
## [1] 14
```

A análise do gráfico de barras revela diferenças marcantes na ocorrência de falhas entre os grupos. O grupo Controle, composto por 15 indivíduos, registrou apenas dois eventos de falha, ambos na terceira semana. Em contrapartida, o grupo Casos, com 14 indivíduos, apresentou sete eventos, distribuídos ao longo do estudo: três na primeira semana e os demais nas semanas 5, 7, 8 e 10.

A alta proporção de censura no grupo Controle, onde a maioria dos pacientes completou o estudo sem apresentar o evento, sugere uma taxa de sobrevivência consideravelmente maior em comparação ao grupo Casos. Esta disparidade será o foco da análise subsequente com as curvas de sobrevivência.

## Estimador de Kaplan-Meier

A seguir, vamos calcular as curvas de sobrevivência para os grupos controle e casos utilizando o estimador Kaplan-Meier. As curvas serão plotadas para visualização das diferenças entre os grupos.

```
ekmControle <- ekm(
  time = hControle$tempo,
  evento = hControle$evento,
  classEvento = 1,
  alpha = 0.05,
  conf = "log-log"
)
ekmCasos <- ekm(
  time = hCasos$tempo,
  evento = hCasos$evento,
  classEvento = 1,
  alpha = 0.05,
  conf = "log-log"
)
```

```
print(ekmControle)
```

```
##   time n_risk n_event censored   q_i   p_i S_t_i std_error lower upper
## 1     3     13        2         1 0.154 0.846 0.846         0.1  0.81 0.876
```

Ao analisar a tabela gerada pela função criada no início deste relatório, observa-se que o grupo controle apresenta uma taxa de sobrevivência estimada em aproximadamente 0,846 na terceira semana, com desvio padrão em torno de 0,1. O intervalo de confiança calculado pelo método log-log situa-se entre aproximadamente 0,81 e 0,876.

```
print(ekmCasos)
```

```
##   time n_risk n_event censored   q_i   p_i S_t_i std_error lower upper
## 1     1     14        3         1 0.214 0.786 0.786         0.110 0.728 0.832
## 2     5      9        1         0 0.111 0.889 0.698         0.128 0.598 0.778
## 3     7      8        1         0 0.125 0.875 0.611         0.138 0.464 0.729
## 4     8      7        1         0 0.143 0.857 0.524         0.143 0.331 0.685
## 5    10      6        1         1 0.167 0.833 0.437         0.144 0.206 0.647
```



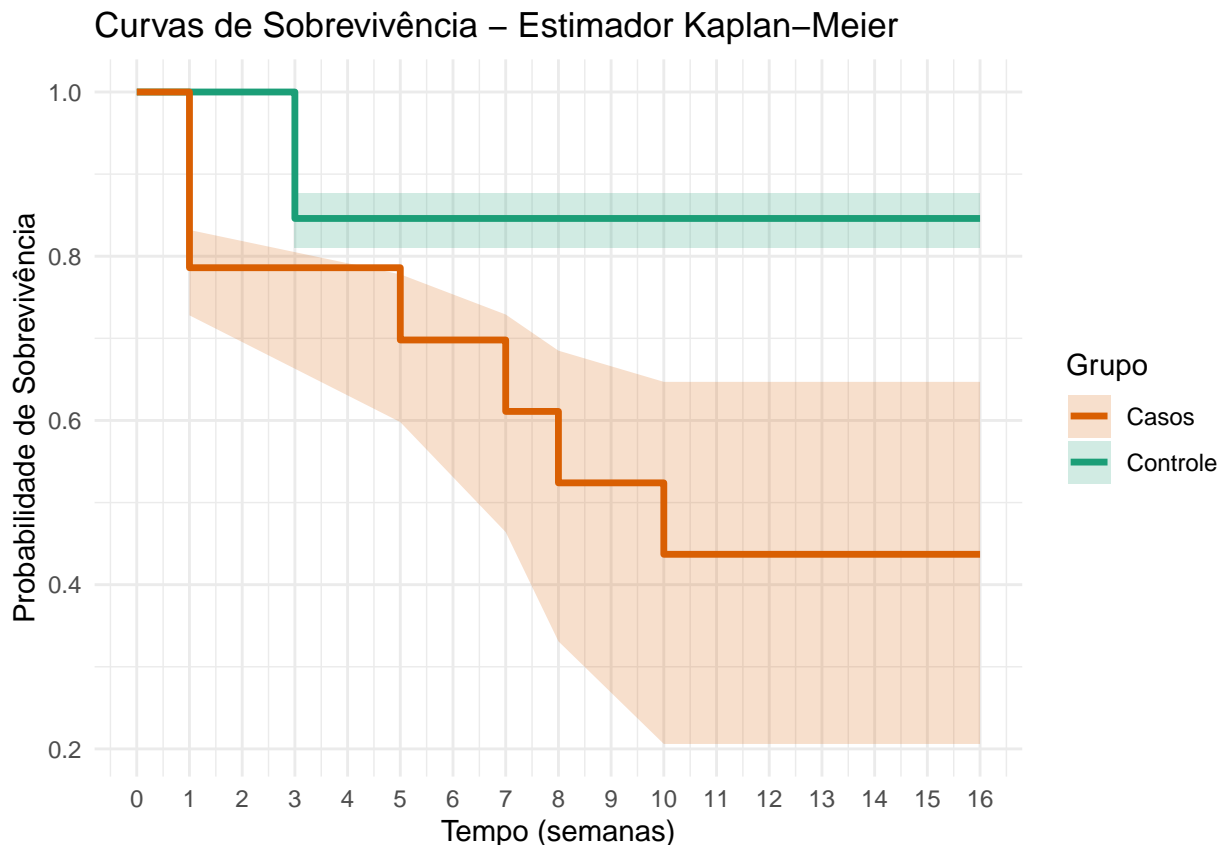
Para o grupo de casos, a taxa de sobrevivência estimada cai para aproximadamente 0,786 já na primeira semana. A sobrevida continua a diminuir, atingindo cerca de 0,436 na décima semana, indicando que menos da metade dos indivíduos deste grupo sobreviveu até esse ponto. O intervalo de confiança para esta estimativa (0,206 a 0,648) é amplo, refletindo o aumento da incerteza estatística ao longo do tempo devido ao acúmulo de variância e à redução do número de indivíduos em risco.

A seguir, vamos plotar as curvas de sobrevivência Kaplan-Meier para os grupos controle e casos. Para isso, adicionaremos manualmente os pontos iniciais e finais das curvas, garantindo que ambas comecem em 1 (100% de sobrevivência) e terminem no último tempo observado.

```
# Adiciona manualmente os pontos iniciais e finais
ekmControle_ext <- ekmControle %>%
  add_row(time = 0, S_t_i = 1, lower = NA, upper = NA, .before = 1) %>%
  add_row(
    time = 16,
    S_t_i = tail(ekmControle$S_t_i, 1),
    lower = tail(ekmControle$lower, 1),
    upper = tail(ekmControle$upper, 1)
  )

ekmCasos_ext <- ekmCasos %>%
  add_row(time = 0, S_t_i = 1, lower = NA, upper = NA, .before = 1) %>%
  add_row(
    time = 16,
    S_t_i = tail(ekmCasos$S_t_i, 1),
    lower = tail(ekmCasos$lower, 1),
    upper = tail(ekmCasos$upper, 1)
  )

# Gráfico de escada Kaplan-Meier
ggplot() +
  geom_step(data = ekmControle_ext, aes(x = time, y = S_t_i, color = "Controle"), size =
    1.2) +
  geom_step(data = ekmCasos_ext, aes(x = time, y = S_t_i, color = "Casos"), size = 1.2) +
  geom_ribbon(
    data = ekmControle_ext,
    aes(x = time, ymin = lower, ymax = upper, fill = "Controle"), alpha = 0.2
  ) +
  geom_ribbon(
    data = ekmCasos_ext,
    aes(x = time, ymin = lower, ymax = upper, fill = "Casos"), alpha = 0.2
  ) +
  scale_color_manual(values = c("Controle" = "#1b9e77", "Casos" = "#d95f02")) +
  scale_fill_manual(values = c("Controle" = "#1b9e77", "Casos" = "#d95f02")) +
  labs(
    title = "Curvas de Sobrevivência - Estimador Kaplan-Meier",
    x = "Tempo (semanas)",
    y = "Probabilidade de Sobrevivência",
    color = "Grupo",
    fill = "Grupo"
  ) +
  theme_minimal() +
  scale_x_continuous(breaks = seq(0, 16, 1))
```



O gráfico de Kaplan-Meier ilustra uma nítida separação entre as curvas de sobrevivência dos grupos controle e casos, com o grupo controle mantendo consistentemente uma taxa de sobrevivência mais elevada. A distância entre as curvas é menor no início do acompanhamento, particularmente entre a terceira e a quinta semana, mas se acentua marcadamente a partir deste ponto, indicando uma divergência crescente nos desfechos. Adicionalmente, é notável o alargamento do intervalo de confiança para o grupo de casos ao longo do tempo. Esse aumento da incerteza reflete a diminuição do número de indivíduos em risco, o que torna a estimativa da taxa de sobrevivência progressivamente menos precisa.

### Teste Logrank

### Estimador Atuarial

A diante, utilizamos o estimador atuarial para calcular as curvas de sobrevivência dos grupos controle e casos. As curvas resultantes serão plotadas para permitir uma comparação visual entre os grupos, evidenciando possíveis diferenças nos padrões de sobrevivência ao longo do tempo.

```
eactControle <- eact(
  time = hControle$tempo,
  evento = hControle$evento,
  classEvento = 1,
  conf = "log-log"
)
eactCasos <- eact(
  time = hCasos$tempo,
  evento = hCasos$evento,
  classEvento = 1,
  conf = "log-log"
)
```

```
print(eactControle)
```

```
##   time n_risk n_event censored n_risk_adj  q_i  p_i S_t_i std_error lower upper
## 1     3     13        2         1      12.5 0.16 0.84  0.84    0.104 0.801 0.872
```

O método atuarial, que ajusta o denominador de risco ao assumir que as censuras ocorrem uniformemente ao longo de cada intervalo, oferece uma perspectiva ligeiramente diferente. Para o grupo controle, a sobrevida estimada é de aproximadamente 0,84, um valor um pouco menor que o estimado por Kaplan-Meier. Essa abordagem também resulta em um intervalo de confiança mais amplo (aproximadamente 0,800 a 0,872), refletindo uma maior incerteza na estimativa.

```
print(eactCasos)
```

```
##   time n_risk n_event censored n_risk_adj  q_i  p_i S_t_i std_error lower
## 1     1     14        3         1      13.5 0.222 0.778 0.778    0.113 0.716
## 2     5      9        1         0       9.0 0.111 0.889 0.691    0.129 0.587
## 3     7      8        1         0       8.0 0.125 0.875 0.605    0.139 0.454
## 4     8      7        1         0       7.0 0.143 0.857 0.519    0.144 0.323
## 5    10      6        1         1       5.5 0.182 0.818 0.424    0.145 0.187
##   upper
## 1 0.828
## 2 0.774
## 3 0.726
## 4 0.683
## 5 0.645
```

No grupo de casos, a estimativa atuarial de sobrevida na primeira semana é de 0,778, caindo para 0,424 na décima semana, o que corrobora a observação de que menos da metade dos indivíduos sobreviveu até esse ponto. O intervalo de confiança para a décima semana é notavelmente amplo (0,187 a 0,645), destacando a crescente incerteza à medida que o número de indivíduos em risco diminui.

Para representar corretamente a curva de sobrevivência pelo método atuarial, o ideal é usar um gráfico de linha, já que esse método assume distribuição uniforme de eventos dentro dos intervalos.

```
# Adiciona manualmente os pontos iniciais e finais
eactControle_ext <- eactControle %>%
  add_row(time = 0, S_t_i = 1, lower = NA, upper = NA, .before = 1) %>%
  add_row(
    time = 16,
    S_t_i = tail(eactControle$S_t_i, 1),
    lower = tail(eactControle$lower, 1),
    upper = tail(eactControle$upper, 1)
  )

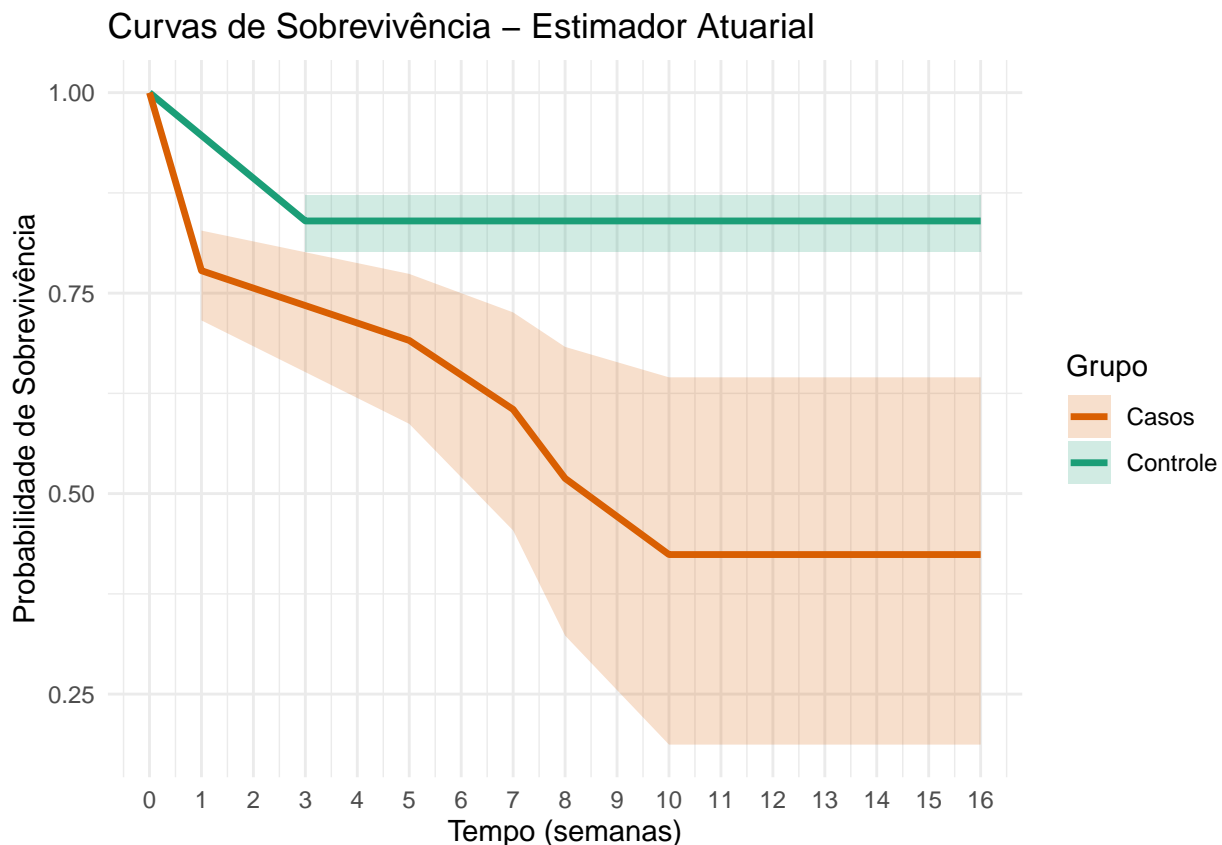
eactCasos_ext <- eactCasos %>%
  add_row(time = 0, S_t_i = 1, lower = NA, upper = NA, .before = 1) %>%
  add_row(
    time = 16,
    S_t_i = tail(eactCasos$S_t_i, 1),
    lower = tail(eactCasos$lower, 1),
    upper = tail(eactCasos$upper, 1)
  )

# Gráfico de sobrevivência atuarial com linhas suaves
ggplot() +
  geom_line(data = eactControle_ext, aes(x = time, y = S_t_i, color = "Controle"), size =
    1.2) +
```

```

geom_line(data = eactCasos_ext, aes(x = time, y = S_t_i, color = "Casos"), size = 1.2)
+
geom_ribbon(
  data = eactControle_ext,
  aes(x = time, ymin = lower, ymax = upper, fill = "Controle"), alpha = 0.2
) +
geom_ribbon(
  data = eactCasos_ext,
  aes(x = time, ymin = lower, ymax = upper, fill = "Casos"), alpha = 0.2
) +
scale_color_manual(values = c("Controle" = "#1b9e77", "Casos" = "#d95f02")) +
scale_fill_manual(values = c("Controle" = "#1b9e77", "Casos" = "#d95f02")) +
labs(
  title = "Curvas de Sobrevivência - Estimador Atuarial",
  x = "Tempo (semanas)",
  y = "Probabilidade de Sobrevivência",
  color = "Grupo",
  fill = "Grupo"
) +
theme_minimal() +
scale_x_continuous(breaks = seq(0, 16, 1))

```



O gráfico de sobrevivência atuarial revela uma tendência semelhante à observada nas curvas de Kaplan-Meier, com o grupo controle apresentando uma taxa de sobrevivência mais alta ao longo do tempo. A curva do grupo controle inicia em 1 (100% de sobrevivência) e diminui gradualmente, enquanto a curva do grupo casos mostra uma queda mais acentuada, especialmente nas primeiras semanas. O intervalo de confiança para o grupo controle é relativamente estreito, indicando maior precisão na estimativa, enquanto o intervalo para o

grupo casos é mais amplo, refletindo a incerteza crescente à medida que o número de indivíduos em risco diminui.

### Teste Logrank

### Estimador de Nelson-Aalen

A seguir, aplicaremos o estimador de Nelson-Aalen, que se baseia na função de risco acumulado para estimar a sobrevivência. Frequentemente, este método produz resultados muito próximos aos do Kaplan-Meier, embora as estimativas de sobrevivência tendam a ser ligeiramente superiores, oferecendo uma perspectiva complementar para a comparação entre os grupos.

```
enelsonControle <- enelson(
  time = hControle$tempo,
  evento = hControle$evento,
  classEvento = 1,
  alpha = 0.05,
  conf = "log-log"
)
enelsonCasos <- enelson(
  time = hCasos$tempo,
  evento = hCasos$evento,
  classEvento = 1,
  alpha = 0.05,
  conf = "log-log"
)
```

```
print(enelsonControle)
```

```
##   time n_risk n_event    q_i    p_i S_t_i std_error lower upper
## 1     3     13      2 0.154 0.846 0.857     0.093 0.827 0.883
```

A análise com o estimador de Nelson-Aalen para o grupo controle indica uma taxa de sobrevivência de aproximadamente 0,857 na terceira semana, onde ocorreu o único evento. Este valor é ligeiramente superior ao estimado por Kaplan-Meier. Adicionalmente, o estimador de Nelson-Aalen resultou em um desvio padrão menor (0,093), o que se reflete em um intervalo de confiança mais estreito, de aproximadamente 0,827 a 0,883.

```
print(enelsonCasos)
```

```
##   time n_risk n_event    q_i    p_i S_t_i std_error lower upper
## 1     1     14      3 0.214 0.786 0.807     0.100 0.761 0.845
## 2     5      9      1 0.111 0.889 0.722     0.120 0.637 0.791
## 3     7      8      1 0.125 0.875 0.637     0.133 0.508 0.741
## 4     8      7      1 0.143 0.857 0.553     0.139 0.378 0.696
## 5    10      6      1 0.167 0.833 0.468     0.141 0.253 0.657
```

Podemos notar também que a estimação da função de sobrevida para Nelson-Aalen é ligeiramente maior que a probabilidade  $p_i$ , nota-se essa diferença para o primeiro tempo, onde o estimador de Nelson-Aalen é 0,802, enquanto  $p_1$  é 0,786. Diferente do Kaplan-Meier e o estimador atuarial, onde a função de sobrevida é o produto dos  $p_i$ . Além disso, podemos ver que o estimador de Nelson-Aalen é mais suave, tendo um probabilidade de sobreviver além da décima semana de aproximadamente 0,468, mesmo assim corroborando para a observação de que menos da metade dos indivíduos sobreviveu até esse ponto. O intervalo de confiança para a décima semana é amplo (0,253 a 0,657), refletindo a crescente incerteza à medida que o número de indivíduos em risco diminui.

Abaixo conseguimos visualizar esse aumento da suavidade do estimador de Nelson-Aalen, e também o aumento do intervalo de confiança para o grupo de casos ao longo do tempo.

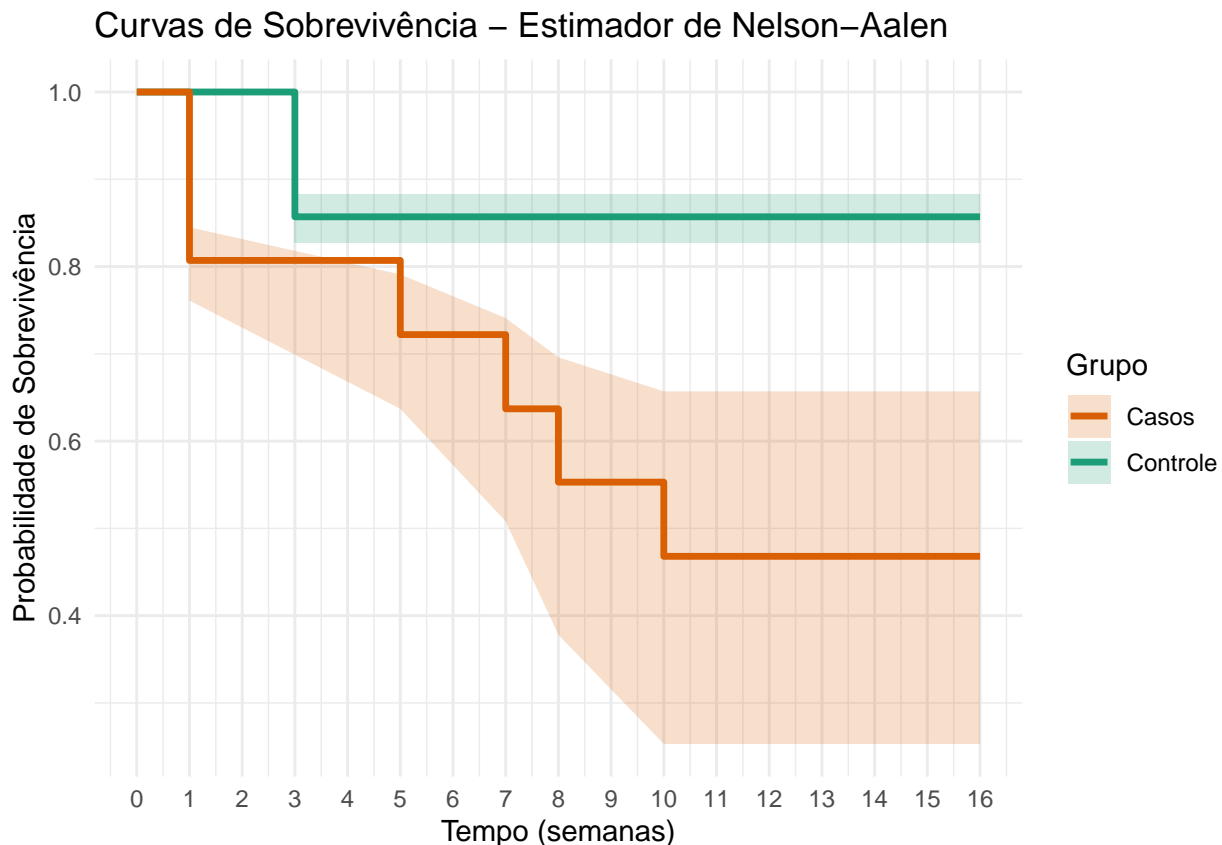
```

# Adiciona manualmente os pontos iniciais e finais
enelsonControle_ext <- enelsonControle %>%
  add_row(time = 0, S_t_i = 1, lower = NA, upper = NA, .before = 1) %>%
  add_row(
    time = 16,
    S_t_i = tail(enelsonControle$S_t_i, 1),
    lower = tail(enelsonControle$lower, 1),
    upper = tail(enelsonControle$upper, 1)
  )

enelsonCasos_ext <- enelsonCasos %>%
  add_row(time = 0, S_t_i = 1, lower = NA, upper = NA, .before = 1) %>%
  add_row(
    time = 16,
    S_t_i = tail(enelsonCasos$S_t_i, 1),
    lower = tail(enelsonCasos$lower, 1),
    upper = tail(enelsonCasos$upper, 1)
  )

# Gráfico de escada Kaplan-Meier
ggplot() +
  geom_step(data = enelsonControle_ext, aes(x = time, y = S_t_i, color = "Controle"),
    size = 1.2) +
  geom_step(data = enelsonCasos_ext, aes(x = time, y = S_t_i, color = "Casos"), size =
    1.2) +
  geom_ribbon(
    data = enelsonControle_ext,
    aes(x = time, ymin = lower, ymax = upper, fill = "Controle"), alpha = 0.2
  ) +
  geom_ribbon(
    data = enelsonCasos_ext,
    aes(x = time, ymin = lower, ymax = upper, fill = "Casos"), alpha = 0.2
  ) +
  scale_color_manual(values = c("Controle" = "#1b9e77", "Casos" = "#d95f02")) +
  scale_fill_manual(values = c("Controle" = "#1b9e77", "Casos" = "#d95f02")) +
  labs(
    title = "Curvas de Sobrevida - Estimador de Nelson-Aalen",
    x = "Tempo (semanas)",
    y = "Probabilidade de Sobrevida",
    color = "Grupo",
    fill = "Grupo"
  ) +
  theme_minimal() +
  scale_x_continuous(breaks = seq(0, 16, 1))

```



A visualização das curvas de Nelson-Aalen corrobora os achados dos métodos anteriores, confirmando o padrão de sobrevivência diferencial entre os grupos. A trajetória para o grupo controle demonstra um declínio suave, enquanto o grupo de casos experimenta uma redução mais abrupta na probabilidade de sobrevivência, principalmente no início do estudo. A precisão da estimativa, indicada pela largura da faixa de confiança, é notavelmente maior para o grupo controle. Em contrapartida, a faixa de confiança para o grupo de casos se alarga progressivamente, refletindo a maior incerteza associada à diminuição do número de indivíduos sob risco ao longo do tempo.

**Teste Logrank**

**Malária**

**Estimador de Kaplan-Meier**

**Teste Bonferroni**

**Estimador Atuarial**

**Teste Bonferroni**

**Estimador de Nelson-Aalen**

**Teste Bonferroni**

**Reincidência de tumor sólido**

```
##### Reincidência de tumor sólido #####
tumorSolido <- data.frame(
```

```
tempo = c(3, 4, 5.7, 6.5, 6.5, 8.4, 10, 10, 12, 15),  
evento = c(1, 0, 0, 1, 1, 0, 1, 0, 1, 1)  
)
```

Estimador de Kaplan-Meier

Estimador Atuarial

Estimador de Nelson-Aalen