# EPI10 - Análise de Sobrevivência

## Comparação de funções de sobrevivência

Rodrigo Citton P. dos Reis
citton.padilha@ufrgs.br

Universidade Federal do Rio Grande do Sul
Faculdade de Medicina
Programa de Pós-Graduação em Epidemiologia

Porto Alegre, 2022

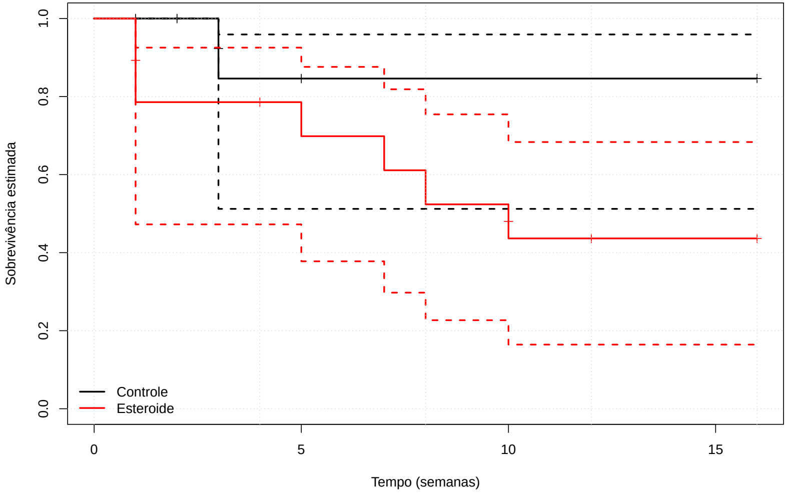
# Relembrando

# Relembrando

## Estudo de Hepatite

| Grupo | Tempo de sobrevivência em semanas |
|---|---|
| Controle | 1+, 2+, 3, 3, 3+, 5+, 5+, 16+, 16+, 16+, 16+, 16+, 16+, 16+, 16+ |
| Esteroide | 1, 1, 1, 1+, 4+, 5, 7, 8, 10, 10+, 12+, 16+, 16+, 16+ |

# Relembrando



# Relembrando

- **Pergunta:** as funções de sobrevivência do grupo Controle e Esteroide diferem?
- Em outras palavras, como podemos testar $S_1(t) = S_2(t)$.
  - Os intervalos de confiança construídos anteriormente {são pontuais}. Ou seja, para cada ponto $t$ temos um intervalo de $100(1-\alpha)\%$ confiança para $S(t)$. Mas, este coeficiente de confiança não é garantido quando olhamos para a "curva toda".
- O teste de ***log-rank***[1] pode responder esta questão adequadamente.

# Relembrando

## Teste log-rank

- Se as $k$ tabelas de contingência forem independentes, um teste aproximado para a igualdade das duas funções de sobrevivência pode ser baseado na estatística

$$\chi^2_{LR} = \frac{\left[\sum_{j=1}^{k} (d_{2j} - \overline{D}_{2j})\right]^2}{\sum_{j=1}^{k} V(D_{2j})}$$

- Sob a hipótese nula $H_0 : S_1(t) = S_2(t)$ **para todo o** $t$, em grandes amostras, tem uma **distribuição aproximada qui-quadrado com 1 grau de liberdade**.

[1]Mantel N. Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemotherapy Reports*. 1966 Mar;50(3):163-70.

# Teste log-rank: comparação de mais que dois grupos

# Teste log-rank: comparação de mais que dois grupos

- A generalização do teste **log-rank** avaliar a hipótese de igualdade entre $S_1(t), S_2(t), \ldots, S_r(t)$, $r > 2$.
  - O desenvolvimento não será demonstrado, mas salienta-se que esta estatística de teste **log-rank** generalizado para $r$ funções tem **distribuição aproximada qui-quadrado com $r-1$ graus de liberdade**.
- Neste caso, se $H_0$ é rejeitada, concluímos que pelo menos um grupo difere dos demais em relação à função de sobrevivência.
- Para identificarmos quais grupos diferem uns dos outrous, uma possibilidade é realizar comparações dos grupos, **dois a dois**, por meio do teste de **log-rank** para dois grupos.
  - O **método de Bonferroni** ($\alpha/$[número de comparações múltiplas]) pode ser utilizado para controlar as taxas de erro tipo I.

# Exemplo

# Estudo de Malária

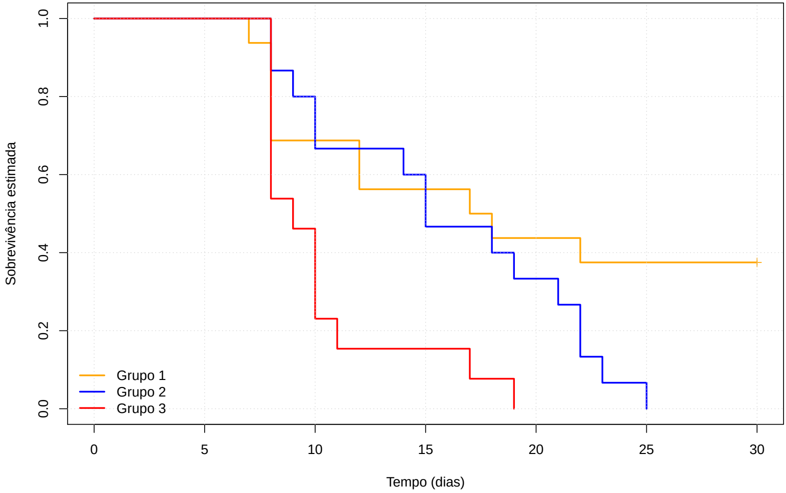
- Um estudo experimental realizado com camundongos para verificar a eficácia da imunização pela malária foi conduzido no Centro de Pesquisas Renee Rachou, Fiocruz, Minas Gerais.
- Nesse estudo, quarenta e quatro camundongos foram infectados pela malária.
  - Os camundongos do grupo 1 foram imunizados 30 dias antes da infecção.
  - Além da infecção pela malária, os camundongos dos grupos 1 e 3 foram, também, infectados pela esquistossomose.
- O desfecho de interesse nesse estudo foi o tempo (em dias) decorrido desde a infecção pela malária até a morte do camundongo.
  - O estudo teve duração de 30 dias.

# Estudo de Malária

```
df.mala <- read.table("../dados/malaria.csv",
                      sep = ";",
                      header = TRUE)
head(df.mala)
```

```
##   id grupo tempo cens
## 1  1     1     7    1
## 2  2     1     8    1
## 3  3     1     8    1
## 4  4     1     8    1
## 5  5     1     8    1
## 6  6     1    12    1
```

# Estudo de Malária



# Estudo de Malária

```r
survdiff(Surv(time = tempo, event = cens) ~ grupo,
         data = df.mala)
```

```
## Call:
## survdiff(formula = Surv(time = tempo, event = cens) ~ grupo,
##     data = df.mala)
##
##          N Observed Expected (O-E)^2/E (O-E)^2/V
## grupo=1 16       10    17.00    2.8816    6.4111
## grupo=2 15       15    14.51    0.0167    0.0317
## grupo=3 13       13     6.49    6.5190   10.4447
##
##  Chisq= 12.6  on 2 degrees of freedom, p= 0.002
```

# Estudo de Malária

- Constatada a diferença entre os grupos $(p = 0,002)$, existe a necessidade de identificar quais curvas diferem entre si.
- Se realizarmos comparações dois a dois, o método de Bonferroni ajusta o nível de significância de acordo com o número de comparações múltiplas.
  - Como temos três grupos, três comparações dois a dois são possíveis de se realizar.
  - Utilizando o nível de 5% de significância, o nível de significância ajustado por Bonferroni é $\alpha^* = \alpha/3 = 0,05/3 = 0,017$ para cada um dos testes.

# Estudo de Malária

## Grupo 1 vs. Grupo 2

```
# grupo 1 vs grupo 2
survdiff(Surv(time = tempo, event = cens) ~ grupo,
         data = df.mala,
         subset = grupo != 3)
```

```
## Call:
## survdiff(formula = Surv(time = tempo, event = cens) ~ grupo,
##     data = df.mala, subset = grupo != 3)
##
##          N Observed Expected (O-E)^2/E (O-E)^2/V
## grupo=1 16       10     13.7      1.01      2.53
## grupo=2 15       15     11.3      1.23      2.53
##
##  Chisq= 2.5  on 1 degrees of freedom, p= 0.1
```

# Estudo de Malária

## Grupo 1 vs. Grupo 3

```
# grupo 1 vs grupo 3
survdiff(Surv(time = tempo, event = cens) ~ grupo,
         data = df.mala,
         subset = grupo != 2)
```

```
## Call:
## survdiff(formula = Surv(time = tempo, event = cens) ~ grupo,
##     data = df.mala, subset = grupo != 2)
##
##          N Observed Expected (O-E)^2/E (O-E)^2/V
## grupo=1 16       10    15.34      1.86      7.86
## grupo=3 13       13     7.66      3.72      7.86
##
##  Chisq= 7.9  on 1 degrees of freedom, p= 0.005
```

# Estudo de Malária

## Grupo 2 vs. Grupo 3

```r
# grupo 2 vs grupo 3
survdiff(Surv(time = tempo, event = cens) ~ grupo,
         data = df.mala,
         subset = grupo != 1)
```

```
## Call:
## survdiff(formula = Surv(time = tempo, event = cens) ~ grupo,
##     data = df.mala, subset = grupo != 1)
##
##          N Observed Expected (O-E)^2/E (O-E)^2/V
## grupo=2 15       15    20.53      1.49      7.98
## grupo=3 13       13     7.47      4.08      7.98
##
##  Chisq= 8  on 1 degrees of freedom, p= 0.005
```

# Estudo de Malária

- **Conclusão:** ao nível de 5% de significância,
  - entre os grupos 1 e 2, não foram encontradas evidências de diferenças;
  - a diferença entre os grupos 1 e 3 atesta a eficácia da imunização pela malária na presença de infecções pela malária e pela equistossomose;
  - por outro lado, a diferença entre os grupos 2 e 3 mostra o impacto na mortalidade dos camundongos devido à infecção pela esquistossomose.

# Outros testes

# Outros testes

- Outros testes foram propostos para comparar funões de sobrevivência.
- No caso particular da comparação de duas funções de sobrevivência a seguinte forma geral inclui os testes mais importantes na literatura e generaliza a estatística $\chi^2_{LR}$

$$S = \frac{\sum_{j=1}^{k} u_j (d_{2j} - \overline{D}_{2j})^2}{\sum_{j=1}^{k} u_j V(D_{2j})},$$

com $u_j$ os **pesos** que especificam os testes.

- Sob a hipótese nula de que as funções de sobrevivência são iguais, a estatística $S$ tem distribuição qui-quadrado com 1 grau de liberdade para amostras grandes.

# Outros testes

- O teste log-rank é obtido tomando-se $u_j = 1,\ j = 1, \ldots, k$.
- Outro teste bastante utilizado na prática é o de Wilcoxon[2] obtido quando se toma $u_j = n_j$.
- O teste de Tarone e Ware[3] propõe peso $u_j = \sqrt{n_j}$, que fica entre os pesos do teste log-rank e de Wilcoxon.

| **Teste** | **Estatística** | **valor-p** |
|---|---|---|
| Log-rank | 3,67 | 0,055 |
| Wilcoxon | 3,19 | 0,074 |
| Tarone-Ware | 3,43 | 0,064 |

[2] Gehan, E. A. (1965). A Generalized Wilcoxon Test for Comparing Arbitrarily Singly-Censored Samples. *Biometrika*, 52(1/2), 203–223.

[3] Tarone, R. E., & Ware, J. (1977). On Distribution-Free Tests for Equality of Survival Distributions. *Biometrika*, 64(1), 156–160.

# Outros testes

- O teste de Wilcoxon que utiliza peso igual ao número de indivíduos sob risco, coloca mais pesos na porção inicial do eixo do tempo.
  - No início do estudo todos indivíduos estão sob risco e saindo do estado "sob risco" à medida que experimentam o evento ou são censurados.
- O teste log-rank, por outro lado, coloca mesmo peso para todo o eixo do tempo, o que reforça o enfoque nos tempos maiores quando comparado ao teste de Wilcoxon.
- O teste de Tarone-Ware se localiza em uma situação intermediária.

# Outros testes

- Peto e Peto (1972)[4] e Prentice (1978)[5] (teste Peto-Prentice) sugerem utilizar uma função do peso que depende diretamente da experiência passada de sobrevivência observada das duas amostras combinadas.
- A função do peso é uma modificação do estimador de Kaplan-Meier e é definido de tal forma que seu valor é conhecido antes do evento ocorrer.

[4] Peto, R., & Peto, J. (1972). Asymptotically Efficient Rank Invariant Test Procedures. *Journal of the Royal Statistical Society. Series A (General)*, 135(2), 185–207.

[5] Prentice, R. L. (1978). Linear Rank Tests with Right Censored Data. *Biometrika*, 65(1), 167–179.

# Outros testes

- Outra classe de pesos para $S$ foi proposta por Harrington-Fleming (1982)[6] como

$$u_j = \left[\widehat{S}(t_{j-1})\right]^{\rho}.$$

- Se $\rho = 0$, obtém-se $u_j = 1$ e tem-se então o teste log-rank.
- Se $\rho = 1$, então o peso é o Kaplan-Meier no tempo de falha anterior que é um teste similar ao de Peto-Prentice.

---

[6] Harrington, D. P. and Fleming, T. R. (1982). A class of rank test procedures for censored survival data. *Biometrika* 69, 553-566.

# Outros testes

- A função `survdiff` do pacote `survival` utiliza a classe de pesos de Harrington-Fleming, em que o argumento `rho` especifica o tipo de ponderação.

```
# Log-rank
survdiff(Surv(time = tempo, event = cens) ~ grupo,
         data = df.hep,
         rho = 0)
```

```
## Call:
## survdiff(formula = Surv(time = tempo, event = cens) ~ grupo,
##     data = df.hep, rho = 0)
##
##                  N Observed Expected (O-E)^2/E (O-E)^2/V
## grupo=Controle  15        2     4.81      1.64      3.67
## grupo=Esteroide 14        7     4.19      1.89      3.67
##
##  Chisq= 3.7  on 1 degrees of freedom, p= 0.06
```

# Outros testes

```r
# Peto-Prentice
survdiff(Surv(time = tempo, event = cens) ~ grupo,
         data = df.hep,
         rho = 1)
```

```
## Call:
## survdiff(formula = Surv(time = tempo, event = cens) ~ grupo,
##     data = df.hep, rho = 1)
##
##                   N Observed Expected (O-E)^2/E (O-E)^2/V
## grupo=Controle  15     1.79     4.16      1.35      3.43
## grupo=Esteroide 14     6.00     3.63      1.54      3.43
##
##  Chisq= 3.4  on 1 degrees of freedom, p= 0.06
```

# Considerações

# Considerações

- Com o estimador de **Kaplan-Meier** e o teste de **log-rank** é possível:
  - descrever dados de sobrevivência;
  - comparar funções de sobrevivência entre grupos.
- No entanto, estamos limitados a avaliar a influência de covariáveis (exposições ou tratamentos) discretas (categóricas ou categorizadas) na função de sobevivência em análises não ajustadas.
- Como avaliar o efeito, na função de sobrevivência, de covariáveis contínuas, e ajustando para potenciais vairáveis de confusão?
  - Uma possibilidade é proposição de modelos estatísticos com uma estrutura de regressão.
  - Um modelo muito utilizado é o **modelo de Cox**.

# Para casa

- Atividade de avaliação I.

# Próxima aula

- Modelo de Cox.

# Por hoje é só!

**Bons estudos!**