

Avaliação 5 - SME0821 - Análise de Sobrevivência e Confiabilidade

Amanda Caroline de Oliveira Pires

Junho 2024

1 Questão 1

Ajuste um modelo de regressão de Weibull utilizando as seguintes covariáveis: **yschool** - Anos de estudo; **npartner** - Número de parceiros nos últimos 30 dias. Analise os resultados obtidos e classifique as alternativas.

Obs: as covariáveis são incluídas no parâmetro μ .

- I) Após o ajuste do modelo, identificou-se que as covariáveis **yschool** e **npartner** são significativas.
- II) A estimativa obtida para a covariável **yschool** foi de 0.20882 com erro padrão de 0.04373. E a estimativa obtida para a covariável **npartner** foi de -0.07160 com erro padrão de 0.06548.
- III) Após o ajuste do modelo, identificou-se que apenas a covariável **yschool** é significativa.
- IV) Todas as estimativas obtidas neste modelo ajustado podem ser interpretadas.

Resposta: b. Apenas II e III

```

Fitting method: RS()

-----
Mu link function: log
Mu Coefficients:
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.73255    0.49018   9.655 < 2e-16 ***
yschool      0.20882    0.04373   4.776 2.1e-06 ***
npartner     -0.07160    0.06548  -1.093  0.274
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

-----
Sigma link function: log
Sigma Coefficients:
      Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.2725    0.0446  -6.109 1.5e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

-----
No. of observations in the fit: 877
Degrees of Freedom for the fit: 4
      Residual Deg. of Freedom: 873
                        at cycle: 4

Global Deviance:    5369.766
      AIC:          5377.766
      SBC:          5396.872
*****

```

Figura 1: Saída R (Questão 1).

- I) falsa, apenas a covariável yschool é significativa
- II) verdadeira, a estimativa da covariável yschool foi de 0.20882 com erro padrão de 0.04373. A estimativa para a covariável npartner foi de -0.07160 com erro padrão de 0.06548.
- III) verdadeira, yschool é significativa. Como o valor- $p < \alpha$, rejeitamos a hipótese nula de que a covariável não é significativa.
- IV) falsa, todas as estimativas obtidas neste modelo ajustado poderiam ser interpretadas se todas as covariáveis fossem significativas, como npartner não é significativa não podemos interpretar todas estimativas.

2 Questão 2

Ajuste um modelo de regressão de Weibull utilizando a covariável abdpain - Presença de dor abdominal (1=sim e 0=não). Analise os resultados obtidos e selecione a alternativa correta.

Obs: a covariável é incluída no parâmetro μ .

- a) As estimativas obtidas da modelagem são $\mu_1 = 7.063, \mu_2 = -0.393, \sigma = 0.281$.
- b) A interpretação do parâmetro associado à variável **abdpain** é: à medida que a dor abdominal aumenta em uma unidade, espera-se que o tempo de sobrevivência do paciente diminua em um fator de $e^{-0.393} \approx 0.675$.
- c) A matriz de variância-covariância é $\text{Cov} =$

$$\begin{bmatrix} 0.007310457 & 0.0063106395 & -0.0016065125 \\ 0.006310640 & 0.0373636463 & -0.0003651316 \\ -0.001606512 & -0.0003651316 & 0.0019946585 \end{bmatrix}$$

- d) O parâmetro σ representa o logaritmo da desvio padrão da distribuição Weibull.
- e) A estimativa do σ indica que a forma da função de risco associada a esta distribuição está diminuindo ao longo do tempo.

Resposta: d. VVFFV

```
Mu link function: log
Mu Coefficients:
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  7.0634      0.0855  82.612  <2e-16 ***
abdpain      -0.3929      0.1933  -2.033   0.0424 *
---
signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

-----
Sigma link function: log
Sigma Coefficients:
      Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.28067      0.04466  -6.284 5.19e-10 ***
---
signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

-----
No. of observations in the fit: 877
Degrees of Freedom for the fit: 3
Residual Deg. of Freedom: 874
                        at cycle: 4

Global Deviance:    5390.3
AIC:                5396.3
SBC:                5410.63
*****
```

Figura 2: Saída R (Questão 2).

- a) verdadeira, $\mu_1 = 7.063, \mu_2 = -0.393$ e $\sigma = -0.281$.

b) verdadeira, como $\mu_2 = -0.393$, esperamos que o tempo de sobrevivência do paciente diminua em um fator de $e^{\mu_2} = e^{-0.393} \approx 0.675$.

c) falsa, a matriz de variância-covariância é dada por:

	(Intercept)	abdpain	(Intercept)
(Intercept)	0.007310457	-0.0063106395	-0.0016065125
abdpain	-0.006310640	0.0373636463	0.0003651316
(Intercept)	-0.001606512	0.0003651316	0.0019946585

Figura 3: Saída R (Questão 2).

d) falsa, σ não representa o desvio-padrão a distribuição Weibull, mas sim o parâmetro de forma.

e) verdadeira, a função de risco é dada por:

$$\lambda(t) = \frac{\sigma}{\mu^\sigma} t^{\sigma-1}$$

Quando $\sigma > 1$ a função de taxa de falha $\lambda(t)$ é estritamente crescente, se $\sigma < 1$ a função de taxa de falha é estritamente decrescente. No caso em que $\sigma = 1$, tem-se a função de taxa de falha da distribuição exponencial, se tornando um caso especial.

Como o $\sigma = -0.28067 < 1$, concluímos que a função de taxa de falha é estritamente decrescente, ou seja, a função de risco está diminuindo ao longo do tempo.

Referência: (Colosimo and Giolo, 2006)

3 Questão 3

Ajuste o modelo de Cox e considere as cováveis: yschool - Anos de estudo; npartner - Número de parceiros nos últimos 30 dias e abdpain - Presença de dor abdominal (1=sim e 0=não). Interprete seus parâmetros acerca da razão de riscos e julgue as sentenças abaixo.

1. O coeficiente negativo indica que uma diminuição no tempo de escolaridade está associado a uma redução no risco de reinfecção. O hazard ratio de 0.84945 significa que para cada unidade adicional no tempo de escolaridade, o risco de reinfecção diminui em aproximadamente 15%. Com um p-valor menor que 0.01, esta covável é estatisticamente significativa, indicando que tempo de escolaridade é um fator

relevante para o risco de reinfecção.

2. O coeficiente positivo indica que um aumento no número de parceiros está associado a um aumento no risco de reinfecção. O hazard ratio de 1.05059 significa que para cada parceiro adicional, o risco de reinfecção aumenta em aproximadamente 5%. No entanto, esta covariável não é estatisticamente significativa no modelo, sugerindo que o número de parceiros pode não ter um impacto substancial no risco de reinfecção.

3. O coeficiente positivo indica que a presença de dor abdominal está associada a um aumento no risco de reinfecção. O hazard ratio de 1.34956 significa que a presença de dor abdominal aumenta o risco de reinfecção em aproximadamente 35%. Com um p-valor de 0.0411, esta covariável é estatisticamente significativa, indicando que a dor abdominal é um fator relevante para o risco de reinfecção.

Resposta: c. Apenas a afirmativa I é falsa

```
coxph(formula = surv(tempo, cens) ~ yschool + npartner + abdpain,
      data = dados)

n= 877, number of events= 347

              coef exp(coef) se(coef)      z Pr(>|z|)
yschool  -0.16328   0.84935  0.03324 -4.912 9.02e-07 ***
npartner   0.04937   1.05061  0.04990  0.989  0.3225
abdpain    0.30038   1.35038  0.14675  2.047  0.0407 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

              exp(coef) exp(-coef) lower .95 upper .95
yschool    0.8494    1.1774    0.7958    0.9065
npartner    1.0506    0.9518    0.9527    1.1586
abdpain     1.3504    0.7405    1.0128    1.8004

Concordance= 0.575 (se = 0.017 )
Likelihood ratio test= 29.02 on 3 df,  p=2e-06
Wald test               = 28.73 on 3 df,  p=3e-06
Score (logrank) test = 28.58 on 3 df,  p=3e-06
```

Figura 4: Saída R (Questão 3).

O `coef(yschool)` negativo sugere que quanto mais anos de estudo, o risco de reinfecção diminui. O hazard ratio (`exp(coef)`) menor do que 1 indica uma diminuição no risco. A redução pode ser calculada como

$$1 - \exp(\text{coef}) \Rightarrow 1 - 0.8494 = 0.1506$$

Aproximadamente 15%

O valor-p é 0.000000902 menor que 0.01, apontando que esta covariável é estatisticamente significativa, o tempo de escolaridade é um fator relevante para o risco de reinfecção

O **coef(npartner)** positivo mostra que a medida que o número de parceiros aumenta, o risco de reinfecção é de 1.0506, o hazard ratio é de 1.0506. O risco de reinfecção aumenta em

$$1 - \exp(\text{coef}) \Rightarrow 1 - 1.0506 = -0.0506$$

aproximadamente -5%.

Como o valor-p é 0.3225 (> 0.05), temos que essa covariável não é estatisticamente significativa, sugerindo que o número de parceiros pode não ter um impacto substancial no risco de reinfecção.

O **coef(abdpain)** positivo indica que a presença de dor abdominal está associada a um aumento no risco de reinfecção, além disso o hazard ratio maior que 1 (1.3504) o risco de reinfecção que é

$$1 - \exp(\text{coef}) \Rightarrow 1 - 1.3504 = -0.3504$$

aproximadamente -35%.

O valor-p é 0.0407, que é menor que 0.05 (nível de significância). Logo, a covariável dor abdominal é um fator estatisticamente relevante para o risco de reinfecção.

4 Questão 4

Ajuste dois modelos de cox, primeiro considere as covariáveis: npartner e abdpain. No segundo modelo adicione a interação entre as covariáveis “npartner” e “abdpain”. Acerca das afirmações abaixo classifique como verdadeira ou falsa.

1. Todas as covariáveis no primeiro modelo foram significativas a um nível de 10% de significância mas ao considerar um nível de 5% a covariável “abdpain” não é significativa.
2. O coeficiente estimado para a covariável “npartner” no primeiro modelo é positivo

mas abaixo de 0.05, indicando que conforme o valor dela aumenta, o risco de falha diminui.

3. A magnitude do coeficiente mostra a força da associação entre a covariável e o risco de falha. Coeficientes com maior valor absoluto têm um impacto mais significativo na função de risco. Com isso a covariável que indica dor abdominal é a que possui maior impacto no primeiro modelo de cox ajustado.

4. Ao ajustar o modelo que considera a interação entre as covariáveis “npartner” e “abdpain” verificamos que esta é a que possui menor p-valor, com isso é a mais significativa, portanto, sendo necessária considerá-la na modelagem final.

Resposta: d. FFVF

Modelo 1

```
coxph(formula = Surv(tempo, cens) ~ npartner + abdpain, data = dados)

n= 877, number of events= 347

              coef exp(coef) se(coef)      z Pr(>|z|)
npartner  0.04343   1.04438  0.04911  0.884  0.3766
abdpain    0.29196   1.33905  0.14671  1.990  0.0466 *
```

 signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

	exp(coef)	exp(-coef)	lower .95	upper .95
npartner	1.044	0.9575	0.9485	1.150
abdpain	1.339	0.7468	1.0044	1.785

Concordance= 0.519 (se = 0.016)
 Likelihood ratio test= 4.66 on 2 df, p=0.1
 Wald test = 5.04 on 2 df, p=0.08
 Score (logrank) test = 5.06 on 2 df, p=0.08

Figura 5: Saída R (Questão 4).

1. Falsa, no primeiro modelo ao nível de 10% de significância a única covariável significativa foi a abdpain. Ao nível 5% de significância abdpain também é estatisticamente significativa.

2. Falsa, o coeficiente estimado de npartner no primeiro modelo é de 0.04343 (positivo) sugerindo que a medida que o número de parceiros aumenta o risco de reinfeção também pode aumentar. Porém pelo valor-p observamos que essa covariável não é estatisticamente significativa.

Modelo 2

```
coxph(formula = surv(tempo, cens) ~ npartner + abdpain + npartner *  
      abdpain, data = dados)  
  
n= 877, number of events= 347  
  
              coef exp(coef) se(coef)      z Pr(>|z|)  
npartner      0.0002588 1.0002588 0.0660843 0.004 0.997  
abdpain        0.1506696 1.1626124 0.1959681 0.769 0.442  
npartner:abdpain 0.0990160 1.1040840 0.0899701 1.101 0.271  
  
              exp(coef) exp(-coef) lower .95 upper .95  
npartner              1.000      0.9997      0.8787      1.139  
abdpain                1.163      0.8601      0.7918      1.707  
npartner:abdpain      1.104      0.9057      0.9256      1.317  
  
Concordance= 0.519 (se = 0.016 )  
Likelihood ratio test= 5.79 on 3 df,  p=0.1  
Wald test              = 7.33 on 3 df,  p=0.06  
Score (logrank) test = 7.76 on 3 df,  p=0.05
```

Figura 6: Saída R (Questão 4).

3. Verdadeira, a magnitude do coeficiente mostra a força da associação entre a covariável e o risco de falha. Coeficientes com maior valor absoluto têm um impacto mais significativo na função de risco. Com isso a covariável (abdpain) que indica dor **abdominal** é a que possui maior impacto no primeiro modelo de cox ajustado, conforme se vê na figura 6.

4. Falsa, apesar da interação npartner e abdpain possuir o menor nível descritivo, se considerarmos um nível de significância de 5%, ainda assim não rejeitamos a hipótese nula de que as covariáveis não são significativas. Portanto, não deve ser considerada na modelagem final.

5 Questão 5

Ajuste dois modelos de cox, sendo o primeiro considerando apenas a covariável yschool e o segundo yschool e abdpain. Indique qual das afirmativas abaixo é verdadeira.

- a. O coeficiente de yschool no primeiro modelo indica que um aumento no tempo de escolaridade está associado a um aumento no risco de reinfecção.
- b. No segundo modelo, o coeficiente de yschool continua significativo, enquanto o

coeficiente de npartner não é estatisticamente significativo.

c. A inclusão da covariável npartner no segundo modelo torna a covariável yschool significativa.

d. A inclusão da covariável npartner no segundo modelo indica que um aumento no número de parceiros está associado a uma redução no risco de reinfeção para a covariável yschool, já que o coeficiente era -0.16047 e passou a ser -0.16161

```
coxph(formula = surv(tempo, cens) ~ yschool + abdpain, data = dados)

n= 877, number of events= 347

              coef exp(coef) se(coef)      z Pr(>|z|)
yschool -0.16229    0.85020  0.03314 -4.897 9.75e-07 ***
abdpain  0.30985    1.36323  0.14634  2.117  0.0342 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

              exp(coef) exp(-coef) lower .95 upper .95
yschool    0.8502      1.1762    0.7967    0.9073
abdpain    1.3632      0.7336    1.0233    1.8161

Concordance= 0.577 (se = 0.016 )
Likelihood ratio test= 28.13 on 2 df,  p=8e-07
Wald test               = 27.91 on 2 df,  p=9e-07
Score (logrank) test = 27.73 on 2 df,  p=1e-06
```

Figura 7: Saída R (Questão 5).

Resposta: b) No segundo modelo, o coeficiente de yschool continua significativo, enquanto o coeficiente de npartner não é estatisticamente significativo.

```
coxph(formula = surv(tempo, cens) ~ yschool, data = dados)

n= 877, number of events= 347

              coef exp(coef) se(coef)      z Pr(>|z|)
yschool -0.16058    0.85165  0.03294 -4.874 1.09e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

              exp(coef) exp(-coef) lower .95 upper .95
yschool    0.8517      1.174    0.7984    0.9085

Concordance= 0.567 (se = 0.016 )
Likelihood ratio test= 23.95 on 1 df,  p=1e-06
Wald test               = 23.76 on 1 df,  p=1e-06
Score (logrank) test = 23.5 on 1 df,  p=1e-06
```

Figura 8: Saída R (Questão 5).

a) Falsa, o coeficiente negativo indica que a covariável diminui o risco de falha. Ou seja, que um aumento no tempo de escolaridade está associado a uma diminuição no risco de reinfeção.

```
coxph(formula = Surv(tempo, cens) ~ yschool + npartner, data = dados)

n= 877, number of events= 347

              coef exp(coef) se(coef)      z Pr(>|z|)
yschool  -0.16172   0.85068  0.03305 -4.893 9.95e-07 ***
npartner   0.05650   1.05813  0.04987  1.133  0.257
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

              exp(coef) exp(-coef) lower .95 upper .95
yschool    0.8507      1.1755    0.7973    0.9076
npartner    1.0581      0.9451    0.9596    1.1668

Concordance= 0.565 (se = 0.017 )
Likelihood ratio test= 25.1 on 2 df,  p=4e-06
Wald test              = 24.83 on 2 df,  p=4e-06
Score (logrank) test = 24.6 on 2 df,  p=5e-06
```

Figura 9: Saída R (Questão 5).

b) Verdadeira, como já foi visto nos sumários, yschool é significativa e a covariável npartner não é estatisticamente significativa.

c) Falsa, a inclusão da covariável npartner no segundo modelo não torna a covariável yschool significativa.

d) Falsa, a inclusão da variável npartner não influencia na significância estatística da covariável yschool, a escolaridade já era estatisticamente significativa.

A Códigos

A.1 Questão 1

```
# Atividade 5 - Amanda Caroline de Oliveira Pires
#Bibliotecas e pacotes
library(car)
library(pwr)
library(gamlss)
library(survival)
library(gamlss.cens)
```

```

library(broom)
library(tidyverse)
library(pammtools)
library(numDeriv)
library(flexsurv)
library(survminer)
library(stats4)

# Leitura dos dados
data("std", package = "KMsurv")
head(std)

# Selecionando as colunas de interesse
dados<- subset(std, select = c("time", "rinfct", "yschool",
                              "npartner", "abdpain", "os12m"))

# Questao 1
tempo <- std$time
cens <- std$rinfct

# Ajuste do modelo de regressao Weibull
gen.cens(WEI)
mod = gamlss(formula=Surv(tempo,cens)~ yschool + npartner, family=
             cens(WEI), data=dados)
summary(mod)

```

A.2 Questão 2

```

# Questao 2
gen.cens(WEI)
mod_abdpain <- gamlss(formula = Surv(tempo, cens) ~ abdpain,
                      family = cens(WEI), data = dados)
summary(mod_abdpain)

# Verificando o item b)
# Calcular  $e^{(-\text{beta\_abdpain})}$ 

```

```
exp_coef_abdpain <- exp(-0.3929)
exp_coef_abdpain
# verificando matriz de variancia-covariancia
vcov_matrix <- vcov(mod_abdpain)
vcov_matrix
```

A.3 Questão 3

```
# Questao 3 - Ajustar o modelo de regressao de Cox
Coxfit1 = coxph(Surv(tempo,cens) ~ yschool + npartner + abdpain,
  data=dados)
summary(Coxfit1)
```

A.4 Questão 4

```
# Questao 4 - Ajustar o modelo de regressao de Cox
Coxfit2 = coxph(Surv(tempo, cens) ~ npartner + abdpain, data = dados)
summary(Coxfit2)

# Ajustar o modelo de regressao de Cox segundo modelo (interacao entre as
  covariaveis)
Coxfit3 = coxph(Surv(tempo, cens) ~ npartner + abdpain + npartner*abdpain,
  data = dados)
summary(Coxfit3)
```

A.5 Questão 5

```
# Questao 5 - Ajustar o modelo de regressao de Cox (yschool)
Coxfit4 = coxph(Surv(tempo, cens) ~ yschool, data = dados)
summary(Coxfit4)
```

```
# Ajustar o modelo de regressao de Cox ()
Coxfit5 = coxph(Surv(tempo, cens) ~ yschool + abdpain, data = dados)
summary(Coxfit5)

Coxfit6 = coxph(Surv(tempo, cens) ~ yschool + npartner, data = dados)
summary(Coxfit6)

Coxfit7 = coxph(Surv(tempo, cens) ~ npartner, data = dados)
summary(Coxfit7)
```

Referências

E. A. Colosimo and S. R. Giolo. *Análise de Sobrevivência Aplicada*. Editora Blucher, São Paulo, 2006.