

ANÁLISE DE SOBREVIVÊNCIA

Teoria e aplicações em saúde

Caderno de Respostas

Capítulo 7

Análise de Resíduos

Neste capítulo, serão feitas as análises dos resíduos dos modelos ajustados no Capítulo 6. Por isso, é necessário primeiro realizar os exercícios do capítulo anterior.

Exercício 7.1: No capítulo anterior, encontramos, ao ajustar o modelo de Cox aos dados de transplante de medula óssea (banco de dados `tmoclas.csv` – Apêndice 12.4), que o melhor modelo explicativo da sobrevivência incluía as covariáveis idade, sexo, fase, `deag` e `decr`. Adotando-o como o modelo final, refaça a análise de resíduos, refazendo o que foi apresentado ao longo do texto

- a. Obtenha os resíduos de Schoenfeld para verificar a proporcionalidade dos riscos em cada variável do modelo final

Resposta:

```
# Exercício 7.1
# item a
# lendo o arquivo de dados do transplante de medula óssea – formato clássico
tmo <- read.table("tmoclas.dat", header = T)
names(tmo)
# Acertando as variáveis
tmo$sexo <- factor(tmo$sexo)
tmo$decr <- factor(tmo$decr)
tmo$deag <- factor(tmo$deag)
tmo$fase <- factor(tmo$fase)

# Chamando a biblioteca da sobrevivência
library(survival)
# Ajustando o modelo
mod4 <- coxph(Surv(os, status) ~ idade + sexo + fase + deag + decr, data = tmo)
mod4

Call:
coxph(formula = Surv(os, status) ~ idade + sexo + fase + deag +
      decr, data = tmo)

      coef exp(coef) se(coef)      z      p
idade -0.00502    0.995  0.0149 -0.337 0.74000
sexo2  -0.27198    0.762  0.3321 -0.819 0.41000
fase2   0.59397    1.811  0.3714  1.599 0.11000
fase3   0.93841    2.556  0.5224  1.796 0.07200
deag1   1.19038    3.288  0.3275  3.635 0.00028
decr1  -1.06175    0.346  0.3385 -3.137 0.00170

# Calculando os resíduos de Schoenfeld
zph <- cox.zph(mod4)
# Teste da correlação linear Global do modelo e de cada variável
zph

      rho      chisq      p
idade -0.02226 2.92e-02 0.8644
sexo2  -0.18004 1.86e+00 0.1721
fase2  -0.00212 2.81e-04 0.9866
fase3   0.20766 2.91e+00 0.0881
deag1   0.05110 1.52e-01 0.6971
decr1   0.35133 7.22e+00 0.0072
GLOBAL      NA 1.35e+01 0.0362
```

- b. Verifique a correlação linear do tempo com cada variável e global do modelo (proporcionalidade)

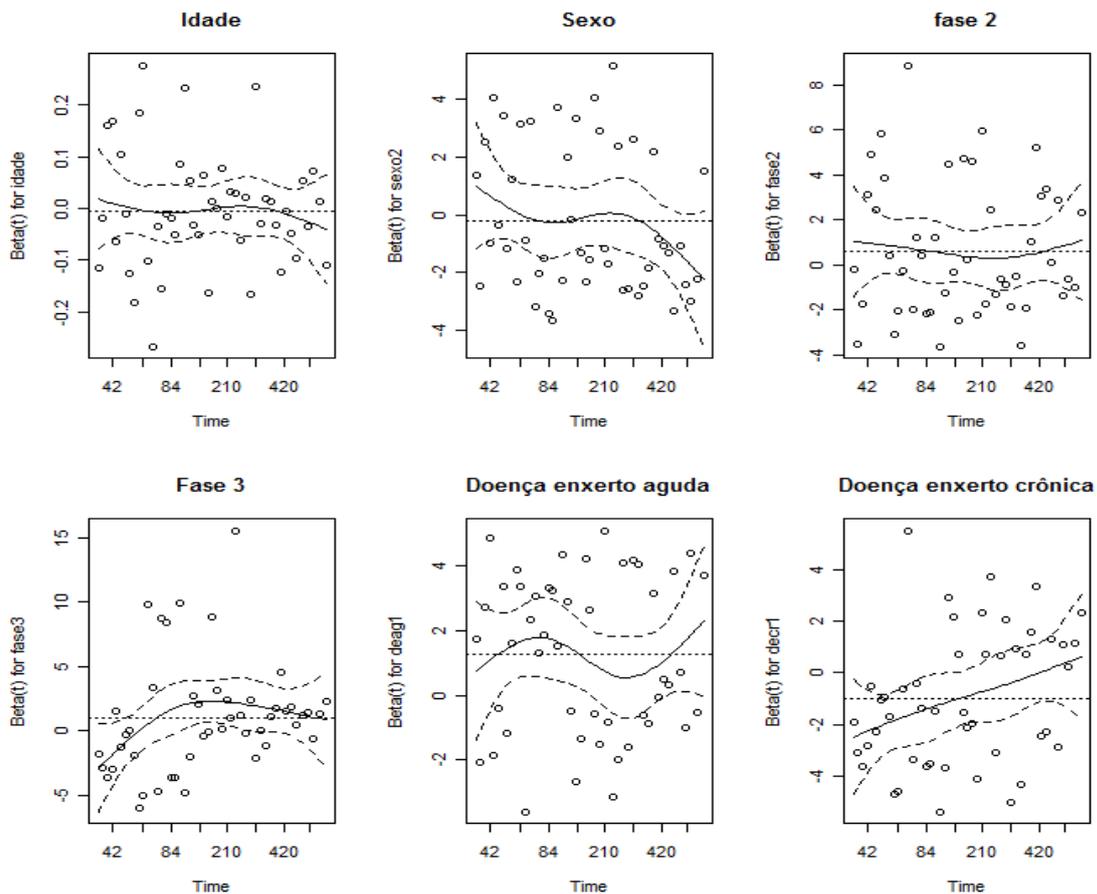
Resposta:

A variável *decr* (doença do enxerto crônica) é a única para a qual se rejeita a hipótese de proporcionalidade. O modelo global também teve a hipótese de proporcionalidade rejeitada ($p=0,036$).

- c. Faça o gráfico do resíduo de Schoenfeld de cada variável (altere a escala gráfica do tempo, se necessário)

Resposta:

```
# item c
# Gráfico dos Resíduos Schoenfeld de cada variável do modelo
par(mfrow = c(2, 3))
plot(zph[1], main = "Idade")
abline(h = mod4$coef[1], lty = 3)
plot(zph[2], main = "Sexo")
abline(h = mod4$coef[2], lty = 3)
plot(zph[3], main = "fase 2")
abline(h = mod4$coef[3], lty = 3)
plot(zph[4], , main = "Fase 3")
abline(h = mod4$coef[4], lty = 3)
plot(zph[5], main = "Doença enxerto aguda")
abline(h = mod4$coef[5], lty = 3)
plot(zph[6], main = "Doença enxerto crônica")
abline(h = mod4$coef[6], lty = 3)
```



```
# Podemos refazer os Gráficos dos Resíduos Schoenfeld de cada variável do modelo
# com objetivo de gerar imagens tipo jpg
```

```
jpeg(file="Shoenfed_tmo.jpg")
par(mfrow = c(2, 3))
plot(zph[1], main = "Idade")
abline(h = mod4$coef[1], lty = 3)
plot(zph[2], main = "Sexo")
abline(h = mod4$coef[2], lty = 3)
plot(zph[3], main = "fase 2")
abline(h = mod4$coef[3], lty = 3)
plot(zph[4], , main = "Fase 3")
abline(h = mod4$coef[4], lty = 3)
plot(zph[5], main = "Doença enxerto aguda")
abline(h = mod4$coef[5], lty = 3)
plot(zph[6], main = "Doença enxerto crônica")
abline(h = mod4$coef[6], lty = 3)

dev.off()
```

Analisando os gráficos podemos dizer que são indiscutivelmente proporcionais ao longo do tempo a idade e fase 2. Por outro lado o efeito da variável sexo e fase 3 variam um pouco ao longo do tempo, especialmente ao final, quando poucos indivíduos permanecem em acompanhamento e, portanto, esta variação pode ser facilmente desconsiderada. Também pode ser desconsiderada a variação observada na variável deag que oscila bastante no período mas que não ultrapassa os limites dos valores estimados do parâmetro (3,5). Já a variável decr não pode ser considerada proporcional ao longo do tempo por apresentar tendência linear acentuada.

- d. Trate a não-proporcionalidade da variável doença Crônica particionando o tempo em t=100 dias.

Resposta:

```
# item d
tmo$statusaguda <- ifelse(tmo$os < 100, tmo$status, 0)
tmo$statuscronica <- ifelse(tmo$os >= 100, tmo$status, 0)
y1 <- Surv(tmo$os, tmo$statuscronica)
y2 <- Surv(tmo$os, tmo$statusaguda)
```

Modelo 1 - CRONICA

```
modelocr <- coxph(y1 ~ idade + sexo+ fase + deag + decr, data = tmo, x = T)
summary(modelocr)
```

Call:

```
coxph(formula = Surv(os, statuscronica) ~ idade + sexo + fase +
      deag + decr, data = tmo, x = T)
```

```
n= 96, number of events= 30
```

	coef	exp(coef)	se(coef)	z	Pr(> z)
idade	0.0001514	1.0001514	0.0212918	0.007	0.99433
sexo2	-0.2653976	0.7669010	0.4203029	-0.631	0.52775

```

fase2  0.6364085  1.8896819  0.4983644  1.277  0.20160
fase3  2.3998012 11.0209857  0.7808279  3.073  0.00212 **
deag1  1.0974076  2.9963880  0.4000078  2.743  0.00608 **
decr1  -0.4915840  0.6116568  0.4417682  -1.113  0.26581
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

      exp(coef) exp(-coef) lower .95 upper .95
idade      1.0002   0.99985   0.9593   1.043
sexo2      0.7669   1.30395   0.3365   1.748
fase2      1.8897   0.52919   0.7115   5.019
fase3     11.0210   0.09074   2.3855  50.917
deag1      2.9964   0.33374   1.3681   6.563
decr1      0.6117   1.63490   0.2573   1.454

```

```

Rsquare= 0.196 (max possible= 0.916 )
Likelihood ratio test= 20.98 on 6 df, p=0.001851
Wald test          = 25.92 on 6 df, p=0.00023
Score (logrank) test = 38.11 on 6 df, p=1.071e-06

```

Modelo 2 -AGUDA

```

modeloag <- coxph(y2~ idade + sexo+ fase + deag + decr, data = tmo, x = T)
summary(modeloag)

```

Call:

```

coxph(formula = Surv(os, statusaguda) ~ idade + sexo + fase +
      deag + decr, data = tmo, x = T)

```

n= 96, number of events= 19

```

      coef exp(coef) se(coef)      z Pr(>|z|)
idade -0.02187   0.97837  0.02266  -0.965  0.33456
sexo2  -0.23079   0.79391  0.52299  -0.441  0.65901
fase2   0.66544   1.94535  0.57526   1.157  0.24736
fase3   0.13839   1.14842  0.70472   0.196  0.84432
deag1   1.10587   3.02186  0.55125   2.006  0.04485 *
decr1  -2.91453   0.05423  1.04105  -2.800  0.00512 **

```

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

      exp(coef) exp(-coef) lower .95 upper .95
idade      0.97837   1.0221  0.935857   1.0228
sexo2      0.79391   1.2596  0.284841   2.2128
fase2      1.94535   0.5140  0.629997   6.0070
fase3      1.14842   0.8708  0.288565   4.5704
deag1      3.02186   0.3309  1.025761   8.9023
decr1      0.05423  18.4402  0.007048   0.4172

```

```

Rsquare= 0.289 (max possible= 0.828 )
Likelihood ratio test= 32.72 on 6 df, p=1.187e-05
Wald test          = 16.6 on 6 df, p=0.01088
Score (logrank) test = 29.96 on 6 df, p=3.999e-05

```

```

zphcr <- cox.zph(modelocr)

```

zphcr

```

      rho chisq      p
idade  -0.2631  2.4438  0.1180
sexo2  -0.4107  5.4823  0.0192

```

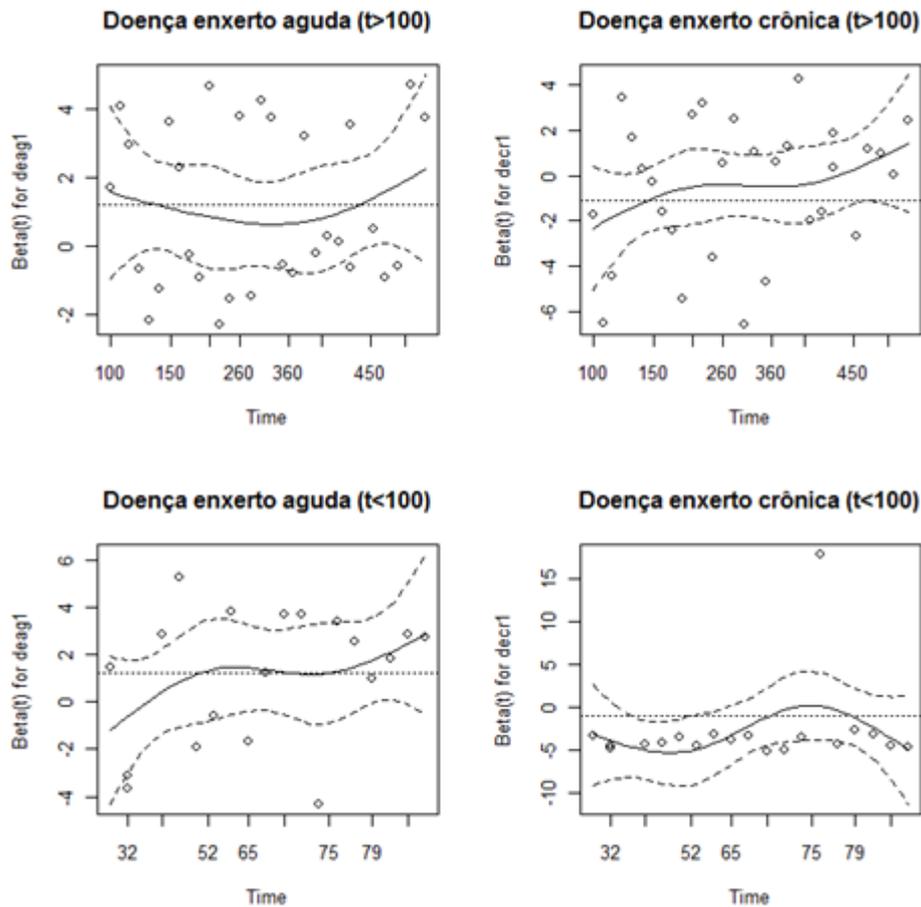
```
fase2 0.0869 0.2906 0.5899
fase3 0.1547 0.9676 0.3253
deag1 0.0428 0.0601 0.8063
decr1 0.2616 3.0390 0.0813
GLOBAL NA 8.5536 0.2003
```

```
zphag <- cox.zph(modeloag)
zphag
```

```
      rho  chisq    p
idade -0.0797 0.1909 0.662
sexo2 -0.2374 1.3514 0.245
fase2 -0.0381 0.0412 0.839
fase3 0.2451 1.3135 0.252
deag1 0.3164 2.5002 0.114
decr1 0.1785 0.7161 0.397
GLOBAL NA 8.2433 0.221
```

Observamos que, após a divisão do tempo de seguimento, nenhuma variável é considerada não proporcional.

```
# Refazendo os gráficos de Schoenfeld para os modelos com tempos particionados
# Resíduos de Schoenfeld para os modelos com o tempo particionado
par(mfrow = c(2, 2))
# "Tempo >= 100 dias"
plot(zphcr[5], main = "Doença enxerto aguda (t>100)")
abline(h = mod4$coef[5], lty = 3)
plot(zphcr[6], main = "Doença enxerto crônica (t>100)")
abline(h = mod4$coef[6], lty = 3)
# "Tempo < 100 dias"
plot(zphag[5], main = "Doença enxerto aguda (t<100)")
abline(h = mod4$coef[5], lty = 3)
plot(zphag[6], main = "Doença enxerto crônica (t<100)")
abline(h = mod4$coef[6], lty = 3)
```



Observamos agora que os gráficos não apresentam nada que indique forte desproporcionalidade como foi observado anteriormente.

e. Avalie a existência de pontos aberrantes usando os resíduos deviance

Resposta:

Avaliação da existência de Pontos aberrantes: resíduos deviance

```
res.dev <- resid(mod4,type="deviance")
```

```
res.dev
```

```

1          2          3          4          5          6          7
-1.17675131  1.89094550  0.21977956  1.25112191  0.15460311  0.56042862 -1.23732404
8          9         10         11         12         13         14
 0.87770052  0.72877771  1.56834372 -2.20466057 -1.33805988 -1.41980118  0.64552773
15         16         17         18         19         20         21
-0.23278871 -0.01884814 -0.69377498  0.79967829 -0.72401503 -0.11544114 -0.76932548
22         23         24         25         26         27         28
-1.78277076  1.09849164 -1.80981895 -0.04693387  0.57654135 -0.69551832  1.93159012
29         30         31         32         33         34         35
-0.13076360  0.63436879  1.20455938  1.17755653  1.54192431 -0.11794914 -1.41047129
36         37         38         39         40         41         42
-0.53502536 -0.70962365  1.13176790 -0.78099768 -0.69551832  1.77670031  0.54665531
43         44         45         46         47         48         49
-0.92580846  1.01281045  1.09480515 -0.76355496 -0.80285860 -0.78099768 -0.67150529
50         51         52         53         54         55         56
-0.66146948 -0.62119090  1.25644156 -1.04056616  2.45582738 -0.81503957  0.39281284

```

```

57          58          59          60          61          62          63
2.48558382 -0.73131981  1.80197059 -0.69726604 -0.47539622  0.80359099  1.31453069
64          65          66          67          68          69          70
-0.70607072 -0.79085968 -0.76932548  1.07220036  1.94413470 -0.76048464  1.31166582
71          72          73          74          75          76          77
-0.66907080  1.20508712 -0.65310289 -1.61742421  0.06471775 -1.30356703 -0.45986208
78          79          80          81          82          83          84
-0.05920295 -2.09130206  0.64970960 -1.49801185 -0.67476229 -0.88542166 -0.74711505
85          86          87          88          89          90          91
1.36372695  1.44520954 -1.13989132  1.11149931 -0.43921875 -0.43811784  0.62023579
92          93          94          95          96
-1.01343949 -0.50083098 -0.30626320  2.46566309 -0.94966792

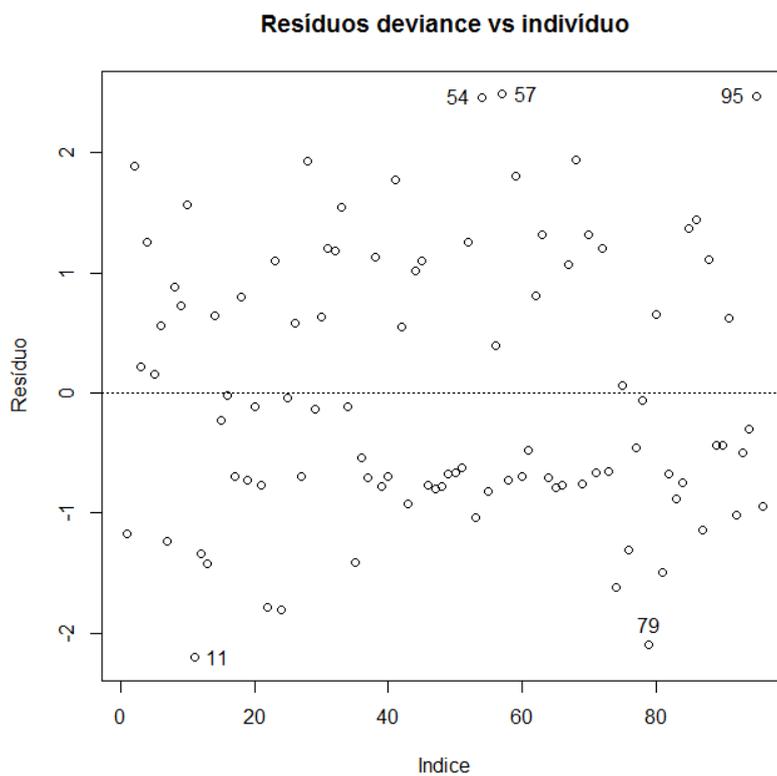
```

Gráficos dos Resíduos Deviance

```

par(mfrow = c(1,1))
plot (res.dev, xlab = "Indice", ylab = "Resíduo", main= "Resíduos deviance vs
indivíduo")
abline(h=0,lty=3)
# Apontando os indivíduos no Gráfico
identify(res.dev)

```



```

# Vendo os dados dos indivíduos mal ajustados
tmo[c(11,79, 54, 57, 95),]

```

```

id sexo idade status os plaq tempplaq deag tempdeag decr tempdecr fase
11 11 2 18 0 1000 1 30 1 12 0 3658 1
79 79 2 14 0 531 1 38 1 26 0 531 1
54 54 2 12 1 31 0 31 1 19 0 31 1
57 57 1 29 1 32 0 32 0 32 0 32 1
95 95 1 19 1 32 0 32 0 32 0 32 1

```

Observamos quatro pontos que podem ser considerados aberrantes por se encontrarem fora do intervalo [-2, 2]. Estes pontos foram identificados como sendo os pacientes de nº 11, 79, 54, 57 e 95.

Observe que os dois com resíduos negativos (11 e 79) foram censurados em tempos longos, apesar de terem doença do enxerto aguda. Os outros três, com resíduos positivos morreram, em tempo muito curtos, mesmo que em fase inicial da doença (fase=1) e, os casos 57 e 95, sem doença aguda.

- f. Verifique a existência de pontos influentes usando os resíduos escore.

Resposta:

```
# item f
```

```
# Resíduo escore
```

```
res.esco <- resid(mod4,type="dfbetas")
```

```
res.esco
```

```
      [,1]      [,2]      [,3]      [,4]      [,5]      [,6]
1 -0.0471624079 -0.117089336  0.092570216  0.0287048354  0.117791104  0.102903695
2  0.2182105337  0.162396296 -0.123722228 -0.1771975831  0.088566717 -0.155355615
3 -0.0173406000 -0.046731451  0.004258496  0.0181740668  0.051703762  0.009014743
...
94 -0.0048307905  0.004007200  0.009351228  0.0066447952  0.006617583  0.006208318
95  0.0021068551  0.002112888  0.002493096  0.0015912585  0.002957294  0.001110545
96 -0.0245822632  0.041312436  0.080839999  0.0477344404  0.052728571  0.064698117
```

```
# Gráficos dos Resíduos Escore
```

```
# Verificando a existência de pontos influentes
```

```
par(mfrow = c(3, 2))
```

```
plot(tmo$sexo, res.esco [, 2], xlab = "Sexo", ylab = "Resíduos")
```

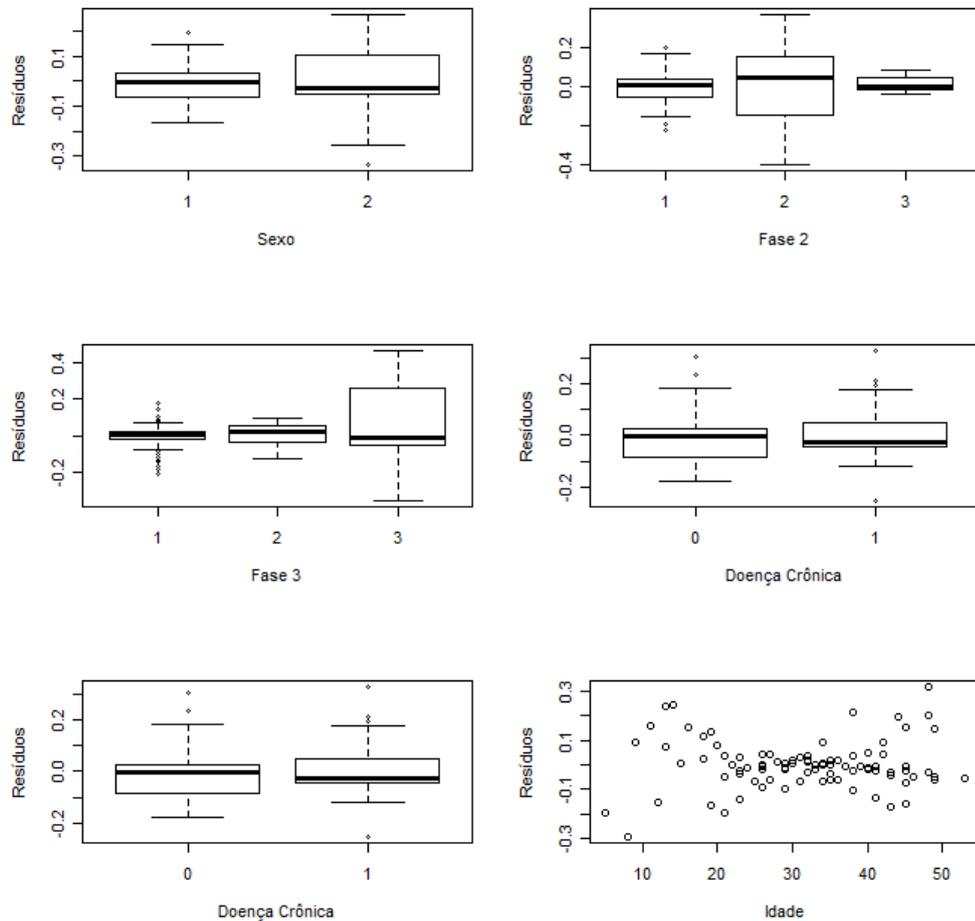
```
plot(tmo$fase, res.esco [, 3], xlab = "Fase 2", ylab = "Resíduos")
```

```
plot(tmo$fase, res.esco [, 4], xlab = "Fase 3", ylab = "Resíduos")
```

```
plot(tmo$deag, res.esco [, 5], xlab = "Doença Aguda", ylab = "Resíduos")
```

```
plot(tmo$decr, res.esco [, 6], xlab = "Doença Crônica", ylab = "Resíduos")
```

```
plot(tmo$idade, res.esco [, 1], xlab = "Idade", ylab = "Resíduos")
```



Observamos que os resíduos score ou dfbetas não identificaram pontos influentes em nenhuma das variáveis do modelo ajustado.

- g. Compare os resultados da análise desses três diferentes resíduos. Foram sempre os mesmos pontos? Se não, o que os diferencia?

Resposta:

Resíduo Schoenfeld: identificou que a variável doença crônica não deveria ser considerada diretamente no modelo por não atender o pressuposto de proporcionalidade ao longo do tempo.

Resíduo Deviance: identificou os indivíduos 11, 79, 54 e 57 como pontos aberrantes (mal ajustados).

Resíduo Score: Não identificou nenhum ponto que influenciasse fortemente as estimativas do parâmetro nas covariáveis do modelo.

- h. Analise a forma funcional das variáveis contínuas usando resíduos martingale.

Resposta:

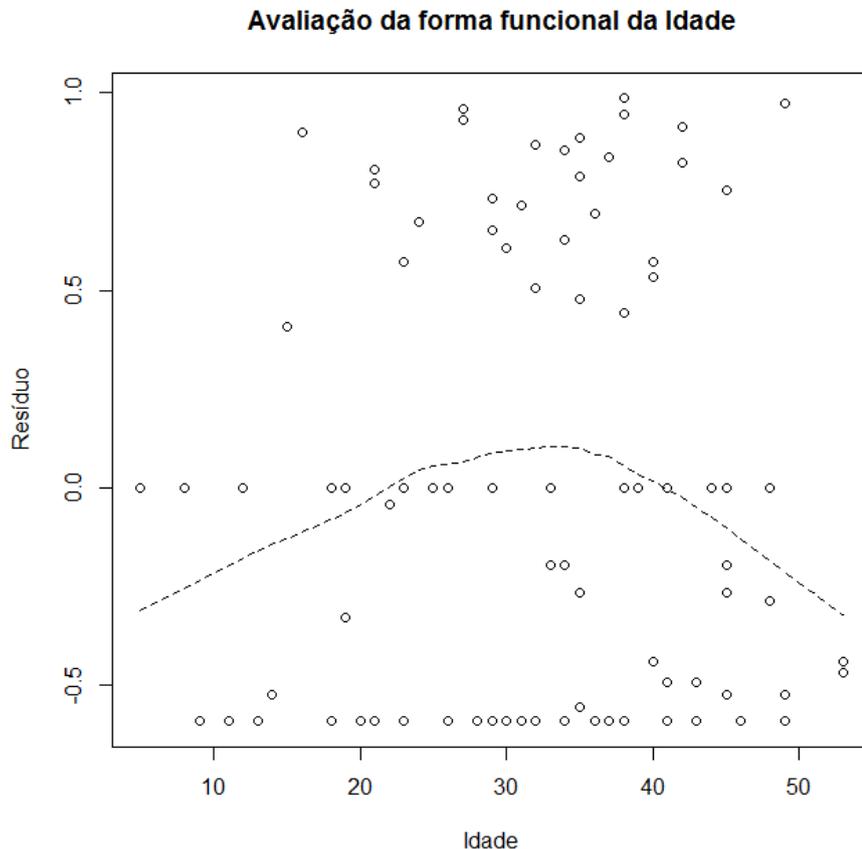
Item h

Analisando a forma funcional através do gráfico dos resíduos de Martingale x idade
`mod0 <- coxph(y1 ~ 1, data = tmo, x = T)`
`summary(mod0)`

```

mod0.mar <- resid(mod0, type= 'martingale ')
plot (tmo$idade, mod0.mar, xlab = "Idade", ylab = "Resíduo",main= "Avaliação da
forma funcional da Idade")
lines(lowess(tmo$idade, mod0.mar, iter = 0), lty = 2)

```



Não há associação entre o tempo de sobrevivência e a idade, uma vez que não há tendência linear e o lowess está em torno do zero.

Exercício 7.2: No estudo de sobrevivência de pacientes em diálise (exercício 6.2), encontramos que o modelo contendo variáveis demográficas, clínicas e ambientais foi o que melhor se ajustou aos dados (modeloIII: idade, cdiab, crim, congenita e grande). Pela análise visual do gráfico de Kaplan-Meier realizada no exercício 6.1, todas as variáveis pareciam atender ao pressuposto de Cox. A única que parecia levantar dúvidas era a variável congênita.

- a. Reavalie essas variáveis, agora utilizando os resíduos de Schoenfeld, fazendo uma análise visual dos resíduos e calculando o teste da correlação.

Resposta:

```

# Exercício 7.2
# Lendo o arquivo de dados de diálise
dialise <- read.table("dialise.csv", header = T, sep = ",")
y <- Surv(dialise$tempo, dialise$status)

```

```
# Refazendo o modelo selecionado
modeloIII <- coxph(y ~ idade + cdiab + crim + congenita + grande, data = dialise)
```

```
# item a
# calculando o Resíduo de Schoenfeld
zph <- cox.zph(modeloIII)
zph
```

	rho	chisq	p
idade	0.05367	5.0743	0.02428
cdiab	0.04419	3.1462	0.07610
crim	0.00773	0.0962	0.75641
congenita	0.05295	4.4936	0.03402
grande	-0.06718	7.2685	0.00702
GLOBAL	NA	19.7751	0.00138

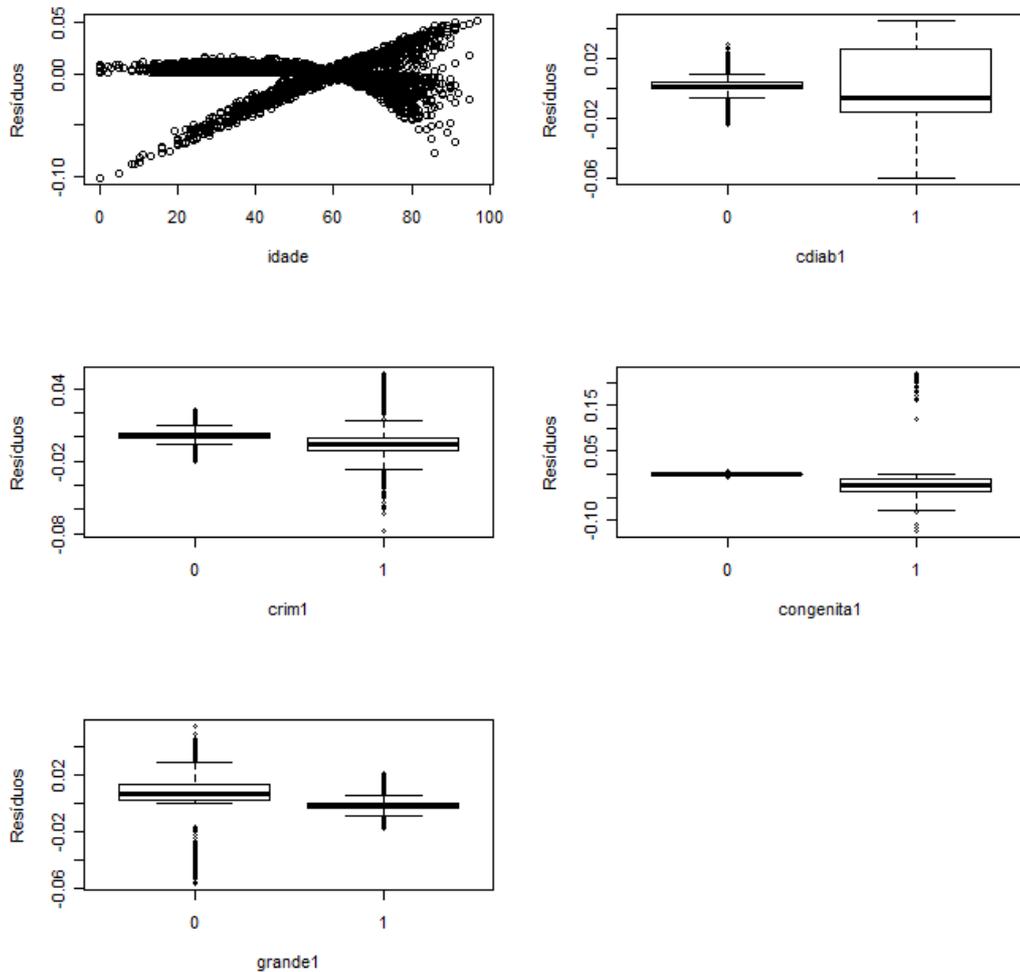
O teste da correlação linear sugere que idade, doença de base congênita e tamanho da unidade de tratamento não atendem o pressuposto de proporcionalidade. Os gráficos, por outro lado, mostram que a não proporcionalidade ocorre principalmente para os tempos muito longos (> 30 meses). É possível que, censurando estes valores, obtenha-se um modelo com proporcionalidade. A variável idade tem um padrão mais definido, no entanto, com maior variabilidade nos tempos menores.

- b. Faça também o gráfico dos resíduos escore para cada indivíduo. Há indicação de indivíduos mal ajustados pelo modelo?

Resposta:

```
# item b
# Resíduo escore
res.esc <- resid(modeloIII,type="dfbetas")

# Gráficos dos Resíduos Escore
# Verificando a existência de pontos influentes
par(mfrow = c(3, 2))
plot(dialise$idade, res.esc [, 1], xlab = "idade", ylab = "Resíduos")
plot(dialise $cdiab, res.esc [, 2], xlab = "cdiab1", ylab = "Resíduos")
plot(dialise $crim, res.esc [, 3], xlab = "crim1", ylab = "Resíduos")
plot(dialise $congenita, res.esc [, 4], xlab = "congenita1", ylab = "Resíduos")
plot(dialise $grande, res.esc [, 5], xlab = "grande1", ylab = "Resíduos")
```



O gráfico de resíduo escore sugere a ausência de indivíduos mal-ajustados pelo modelo.

Exercício 7.3: No ajuste do modelo de Cox aos dados de sobrevivência em AIDS (exercício 6.3), vimos que a variável tratamento não parece atender ao pressuposto de Cox. Faça a análise de resíduos do final (modelo III) e procure confirmar esse achado calculando os resíduos de Schoenfeld.

- a. Obtenha os resíduos de Schoenfeld para verificar a proporcionalidade dos riscos em cada variável do modelo final.

Resposta:

```
# Exercício 7.3
```

```
# Lendo os dados
```

```
ipec <- read.table("ipec.csv", header = T, sep = ";")
```

```
names(ipec)
```

```
[1] "id"      "ini"     "fim"     "tempo"  "status"  "sexo"
[7] "escola" "idade"  "risco"  "acompan" "obito"   "anotrat"
[13] "tratam" "doenca" "propcp"
```

```

# Acertando as variáveis
ipec$acompan <- factor(ipec$acompan)
ipec$tratam <- factor(ipec$tratam)
require(survival)
# Ajustando o modelo de Cox
modeloIII <- coxph(Surv(tempo, status) ~ idade + sexo + acompan + tratam, data =
ipec, x = T)
summary(modeloIII)

```

Call:

```
coxph(formula = Surv(tempo, status) ~ idade + sexo + acompan +
tratam, data = ipec, x = T)
```

n= 193, number of events= 90

	coef	exp(coef)	se(coef)	z	Pr(> z)	
idade	0.001428	1.001429	0.012120	0.118	0.90621	
sexoM	0.074236	1.077061	0.285759	0.260	0.79503	
acompan1	1.676179	5.345096	0.408363	4.105	4.05e-05	***
acompan2	2.153003	8.610675	0.467205	4.608	4.06e-06	***
tratam1	-1.241920	0.288829	0.301134	-4.124	3.72e-05	***
tratam2	-2.096740	0.122856	0.470499	-4.456	8.33e-06	***
tratam3	-2.945019	0.052601	1.018814	-2.891	0.00384	**

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

	exp(coef)	exp(-coef)	lower .95	upper .95
idade	1.0014	0.9986	0.977921	1.0255
sexoM	1.0771	0.9285	0.615179	1.8857
acompan1	5.3451	0.1871	2.400800	11.9002
acompan2	8.6107	0.1161	3.446289	21.5141
tratam1	0.2888	3.4623	0.160072	0.5212
tratam2	0.1229	8.1396	0.048855	0.3089
tratam3	0.0526	19.0110	0.007141	0.3874

Concordance= 0.789 (se = 0.033)

Rsquare= 0.372 (max possible= 0.988)

Likelihood ratio test= 89.84 on 7 df, p=1.11e-16

Wald test = 76.87 on 7 df, p=5.973e-14

Score (logrank) test = 100.1 on 7 df, p=0

item a

Calculando o resíduo de Schoenfeld e o teste da correlação linear

```
zph <- cox.zph(modeloIII)
```

	rho	chisq	p
idade	-0.123	1.24	0.2649
sexoM	0.115	1.29	0.2564
acompan1	-0.214	3.97	0.0464
acompan2	-0.153	2.31	0.1288
tratam1	0.153	2.84	0.0920
tratam2	0.177	3.08	0.0794
tratam3	-0.141	1.83	0.1756
GLOBAL	NA	12.17	0.0950

- b. Calcule a correlação linear do tempo com cada variável e o ajuste global do modelo (proporcionalidade).

Resposta:

```
# item b  
zph
```

	rho	chisq	p
idade	-0.123	1.24	0.2649
sexoM	0.115	1.29	0.2564
acompan1	-0.214	3.97	0.0464
acompan2	-0.153	2.31	0.1288
tratam1	0.153	2.84	0.0920
tratam2	0.177	3.08	0.0794
tratam3	-0.141	1.83	0.1756
GLOBAL	NA	12.17	0.0950

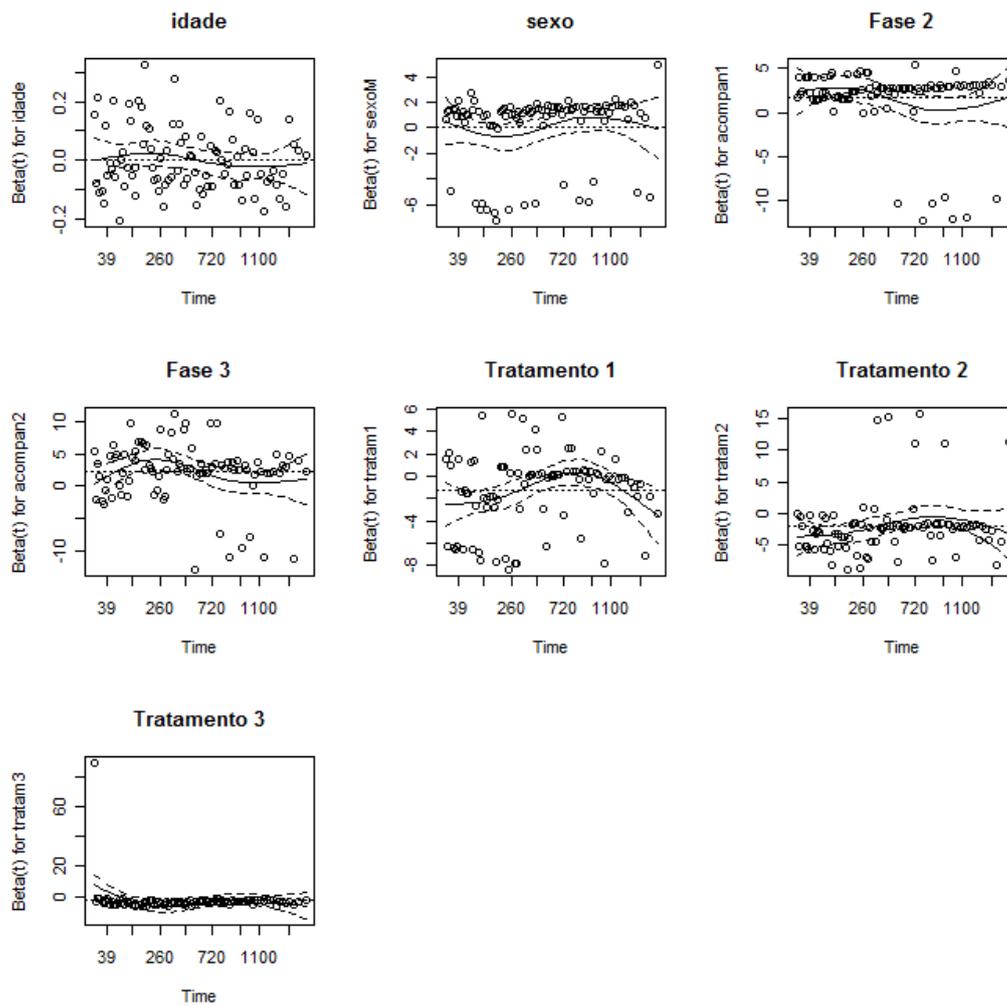
Apenas a variável `acompan1` é não proporcional, mas o ajuste global do modelo é proporcional.

- c. Faça o gráfico do resíduo de Schoenfeld de cada variável

Resposta:

```
# item c  
# Gráficos dos resíduos de Schoenfeld
```

```
par(mfrow = c(3, 3))  
plot(zph[1], main = "idade")  
abline(h = modeloIII$coef[1], lty = 3)  
plot(zph[2], main = "sexo")  
abline(h = modeloIII$coef[2], lty = 3)  
plot(zph[3], main = "Fase 2")  
abline(h = modeloIII$coef[3], lty = 3)  
plot(zph[4], main = "Fase 3")  
abline(h = modeloIII$coef[4], lty = 3)  
plot(zph[5], main = "Tratamento 1")  
abline(h = modeloIII$coef[5], lty = 3)  
plot(zph[6], main = "Tratamento 2")  
abline(h = modeloIII$coef[6], lty = 3)  
plot(zph[7], main = "Tratamento 3")  
abline(h = modeloIII$coef[7], lty = 3)
```



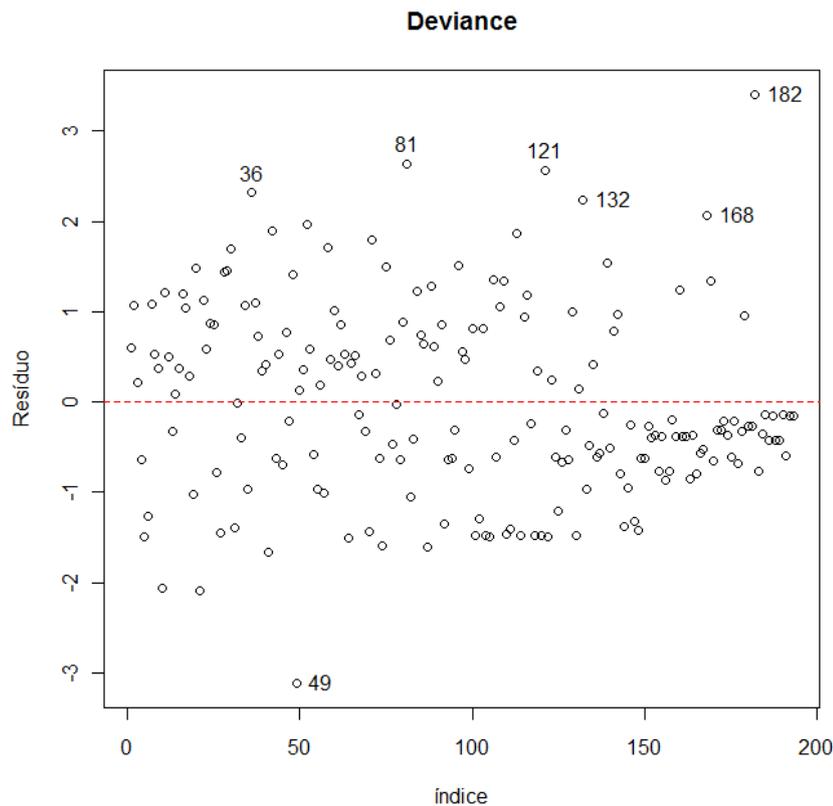
- d. Se alguma variável foi considerada não proporcional (tempo dependente), verifique se ela é realmente importante ou se não é devido a presença de pontos aberrantes (Deviance).

Resposta:

Item d

Resíduos deviance: Pontos Aberrantes

```
mod.dev <- resid(modeloIII, "deviance")
plot(mod.dev, xlab = "índice", ylab = "Resíduo", main = "Deviance")
abline(h = 0, col = 2, lty = 2)
identify(mod.dev)
```



Foram identificados muitos pacientes mal ajustados (182,81,121,36,49).
Veamos seus dados:

```
ipec[c(182,81,121,36,49),c(4,5,6,8,13)]
```

	tempo	status	sexo	idade	tratam
182	16	1	M	42	3
81	21	1	M	55	1
121	22	1	F	28	1
36	18	1	M	34	0
49	1344	0	M	30	0

Os três primeiros, apesar de terem usado terapia sobreviveram por pouco tempo. O último não usou terapia e permaneceu vivo (foi censurado) mesmo sem o uso de terapia.

- e. Faça também o gráfico dos resíduos score. Há indicação de observações mal ajustadas pelo modelo?

Resposta:

```
# item e
```

```
# Resíduos Escore
```

```
res.esc <- resid(modeloIII,type="dfbetas")
```

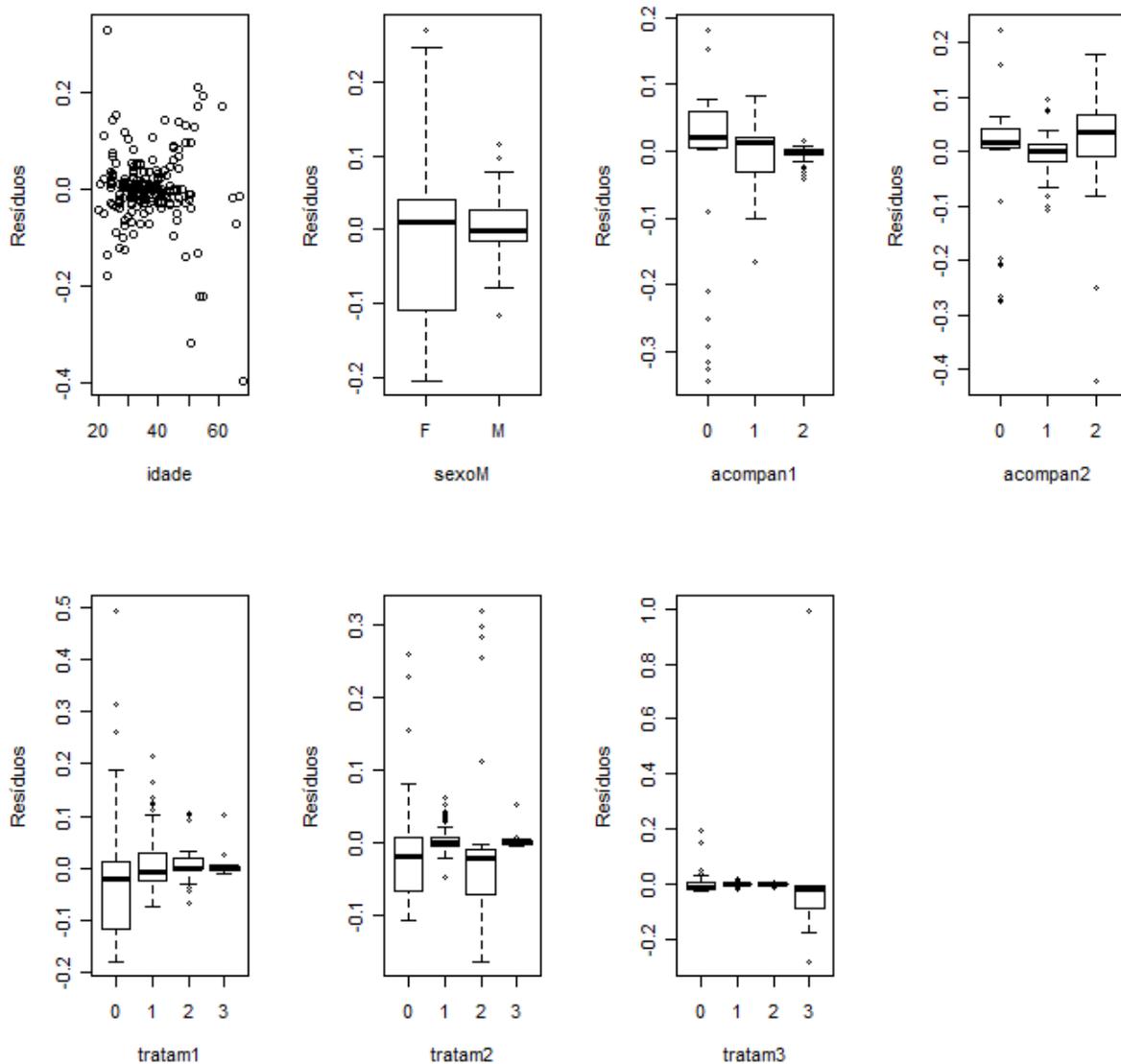
```
# Gráficos dos Resíduos Escore
```

```
# Verificando a existência de pontos influentes
```

```

par(mfrow = c(2, 4))
plot(ipec$idade, res.esc [, 1], xlab = "idade", ylab = "Resíduos")
plot(ipec$sexo, res.esc [, 2], xlab = "sexoM", ylab = "Resíduos")
plot(ipec$acompan, res.esc [, 3], xlab = "acompan1", ylab = "Resíduos")
plot(ipec$acompan, res.esc [, 4], xlab = "acompan2", ylab = "Resíduos")
plot(ipec$tratam, res.esc [, 5], xlab = "tratam1", ylab = "Resíduos")
plot(ipec$tratam, res.esc [, 6], xlab = "tratam2", ylab = "Resíduos")
plot(ipec$tratam, res.esc [, 7], xlab = "tratam3", ylab = "Resíduos")

```



Nestes últimos gráficos não observamos grandes desvios, indicando bom ajuste.