

# Lista 2

## Modelos de regressão logísticos e aditivos

Paulo Ricardo Seganfredo Campana

4 de março de 2024

**Questão 1.** O conjunto de dados descrito no arquivo brown.txt (Brown, 1980) apresenta as variáveis **cancer**, se o paciente possui ou não cancer de próstata (variável resposta), **acip**, nível de fosfatase ácida no sangue, **xray**, se há presença de manchas no raio-x, **estagio**, estágio do câncer, **grau**, grau do cancer e **idade**, idade do paciente.

- Realize o ajuste da regressão logística e selecione as variáveis. O modelo é adequado? Justifique sua escolha.
- Construa um envelope para os resíduos. Há algum ponto que não pertence ao envelope? Se sim, qual(is)?
- Construa um intervalo de confiança de 98% para os parâmetros do modelo.

```
library(tidyverse)
options(digits = 3)

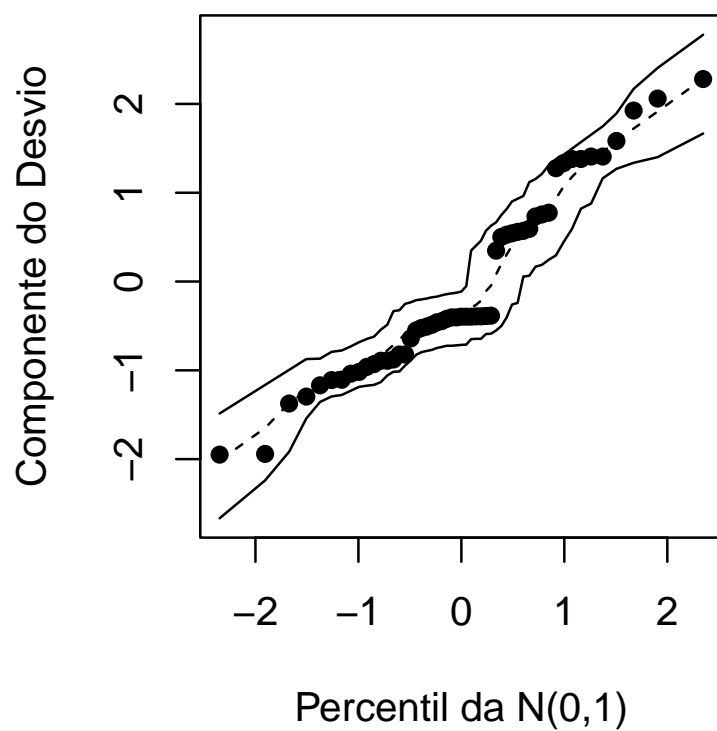
data1 <- read_csv("data/brown.csv")

menor <- glm(cancer ~ 1, family = binomial(link = "logit"), data1)
maior <- glm(cancer ~ ., family = binomial(link = "logit"), data1)
fit1 <- step(menor, scope = list(upper = maior), direction = "both", trace = 0)
# broom::tidy(step1)

# fit1 <- glm(cancer ~ xray + estagio, family = binomial(link = "logit"), data1)
# broom::tidy(fit1)
```

Por minimização do AIC, foram selecionados as variáveis **xray**, **estagio** e **acip** para explicar a variável **cancer** no modelo de regressão logística, porém **acip** não é significativo a um nível de 5%, portanto escolhi um modelo apenas com **xray** e **estagio**.

## Normal Q-Q Plot



Apenas um ponto está na borda do envelope de confiança para os resíduos do modelo

```
confint(fit1, level = 0.98)
```

Waiting for profiling to be done...

	1 %	99 %
(Intercept)	-6.72173	-1.0542
xray	0.35122	4.0831
estagio	0.13638	3.6816
acip	-0.00918	0.0552