

Trabalho de dados Binários

Acidentes de carro

Laís Hoffmann, Simone Matsubara, Yasmin Fernandes, Willian Meira

2018-11-10

1. Base de Dados

1.1 Descrição dos dados

Os dados foram retirados do pacote “DAAG”, sendo dados dos EUA, entre 1997-2002, de acidentes de carro relatados pela polícia nos quais há um evento prejudicial (pessoas ou propriedade) e do qual pelo menos um veículo foi rebocado. Os dados são restritos aos ocupantes do banco da frente, incluem apenas um subconjunto das variáveis registradas e são restritos de outras maneiras também.

A base original possui uma base de dados com 26.217 observações nas 15 variáveis a seguir.

- 1 - **veloc**: velocidades estimadas do impacto do acidente: 1-9km/h, 10-24, 25-39, 40-54, 55+
- 2 - **pesos**: Pesos de observação
- 3 - **sobrev**: Classificação se sobreviveu ao acidente: 1 = morreu ou 0 = sobreviveu
- 4 - **airbag**: Se o carro possui airbag: com ou sem airbag
- 5 - **cinto**: uso do cinto de segurança: com ou sem cinto
- 6 - **frontal**: impacto do acidente: 0 = não frontal, 1 = impacto frontal
- 7 - **sexo**: Sexo: 0 = Feminino ou 1 = Masculino
- 8 - **idade**: Idade dos ocupantes do veículo
- 9 - **anoaci**: Ano do acidente (1997-2002)
- 10 - **anovei**: Ano do veículo (1953-2003)
- 11 - **airbagcat**: Se Airbags foram acionados: deploy, nodeploy, unavail
- 12 - **ocupantes**: Posição do airbag acionado: driver, pass
- 13 - **abfunc**: Airbag acionados: 0: Se não possuía airbag ou não foi acionado, 1: Um ou mais airbags foram acionados
- 14 - **grav**: Gravidade do acidente: 0:none, 1 = Possível Lesão, 2:no incapacity, 3:incapacity, 4:killed; 5:unknown, 6:prior death
- 15 - **numcaso**: Número do caso.

No entanto, escolhemos analisar os dados do ano do acidente de 2002 e veículos de ano 2000 e retirar as variáveis weight, abcat e caseid.

2 Análise Descritiva

2.1 Medidas de Resumo

```
summary(dados[, c(1:8,10)])
```

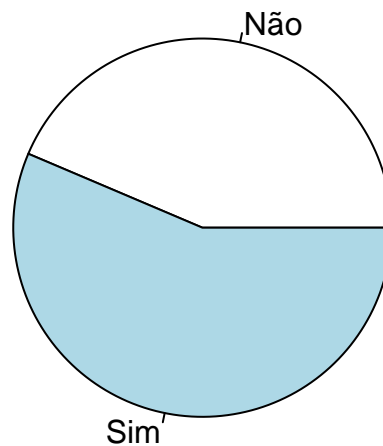
```
##          veloc      sobrev      airbag      cinto      frontal      sexo
## 01-09 mph: 12    Não: 23    Não: 1    Não:121    Não:183    Fem :254
## 10-24 mph:293    Sim:470    Sim:492    Sim:372    Sim:310    Masc:239
## 25-39 mph:121
## 40-54 mph: 46
## 55+ mph : 21
##
##
```

```
##      idade      ocupantes      grav
## Min.   :16.00  Driver:386  Min.   :0.000
## 1st Qu.:23.00  Pass  :107  1st Qu.:0.000
## Median :35.00           Median :1.000
## Mean   :37.82           Mean   :1.579
## 3rd Qu.:48.00           3rd Qu.:3.000
## Max.   :93.00           Max.   :5.000
##                                     NA's   :4
```

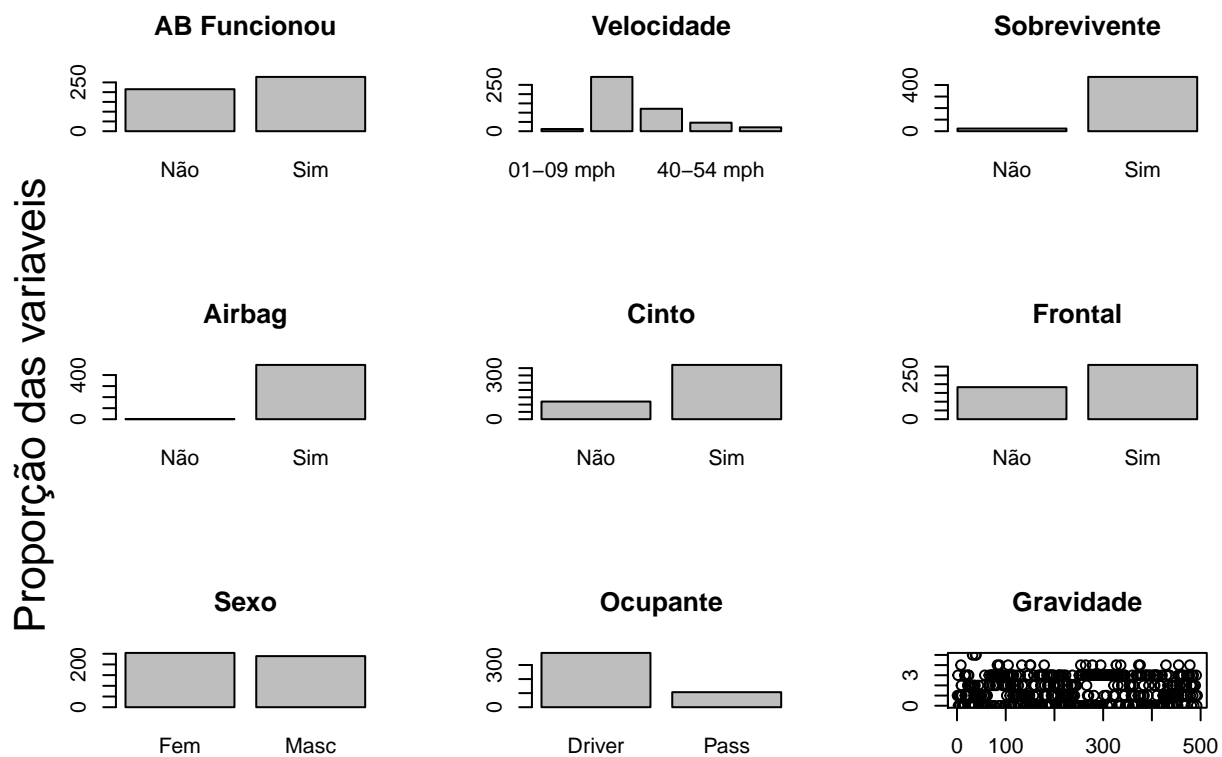
2.3 Histogramas

```
pie(table(dados$abfunc),
     main="Gráfico de setores: Grau de Instrução")
```

Gráfico de setores: Grau de Instrução



```
par(mfrow = c(3,3))
plot(dados$abfunc, xlab = '', ylab = '', main = 'AB Funcionou')
plot(dados$veloc, xlab = '', ylab = '', main = 'Velocidade')
plot(dados$sobrev, xlab = '', ylab = '', main = 'Sobrevivente')
plot(dados$airbag, xlab = '', ylab = '', main = 'Airbag')
plot(dados$cinto, xlab = '', ylab = '', main = 'Cinto')
plot(dados$frontal, xlab = '', ylab = '', main = 'Frontal')
plot(dados$sexo, xlab = '', ylab = '', main = 'Sexo')
#plot(dados$idade, xlab = '', ylab = '', main = 'Idade') ### ***** inverteu eixo
plot(dados$ocupantes, xlab = '', ylab = '', main = 'Ocupante')
plot(dados$grav, xlab = '', ylab = '', main = 'Gravidade')
mtext(side=2,cex=1.3,line=-1.5,text="Proporção das variaveis",outer=TRUE)
```



2.4 Distribuição

2.5 Análise de correlações entre covariáveis

2.6 Gráficos de Dispersão

a

NULL

3. AJUSTE DO MODELO DE REGRESSÃO

3.1 Ligação Logito

3.2 Ligação Probit

3.3 Ligação Complemento log-log

3.4 Ligação Cauchy

4. ESCOLHA DO MODELO

5. ANÁLISE DO MODELO AJUSTADO SELECIONADO

5.1 Resumo do Modelo

5.2 Reajuste do Modelo

5.3 Análise de Resíduos

5.4 Medidas de Influencia

5.5 Resíduos Quantílicos Aleatorizados

5.6 Gráfico Normal de Probabilidade com Envelope Simulado

5.7 Gráficos de Efeitos

6. PREDIÇÃO

7. AVALIAÇÃO DO PODER PREDITIVO DO MODELO

7.1 Divisão da Base de dados

7.2 Ponto de Corte

7.3 Sensibilidade e Especificidade

7.4 Curva ROC

7.5 Outra Alternativa de validação

8. REFERÊNCIAS