Análise de Dados Categorizados

Prática 4: Regressão Logística Multinomial (Politômica)

Prof. Dr. Márcio Augusto

Contents

Modelo Logístico para Resposta Nominal

Modelo Logístico categória de Base (ou de referência)

$$\log\left(\frac{\pi_j}{\pi_c}\right),\ j = 1, \ldots,\ c-1.$$

$$\log(\frac{\pi_j}{\pi_c}) = \alpha_j + \beta_j x,\ j = 1, \ldots, c-1 \tag{1}$$

$$\log(\frac{\pi_1}{\pi_2}) = \log\left(\frac{\pi_1/\pi_3}{\pi_2/\pi_3}\right) = \log\left(\pi_1/\pi_3\right) - \log\left(\pi_2/\pi_3\right)\ = (\alpha_1+\beta_1 x) - (\alpha_2+\beta_2 x) = (\alpha_1-\alpha_2) + (\beta_1-\beta_2)x \tag{2}$$

$$\log(\frac{\pi_j}{\pi_c}) = \alpha_j + \beta_{j1}x_1 + \beta_{j2}x_2 + \cdots + \beta_{jp}x_p\ j = 1, \ldots, c-1 \tag{3}$$

```r
library(VGAM)
library(tidyverse)
library(RColorBrewer)
```

Exercício

Os dados a seguir são de um estudo realizado com 338 crianças para investigar se o programa de aprendizado preferido estaria associada com a escola e o período escolar.

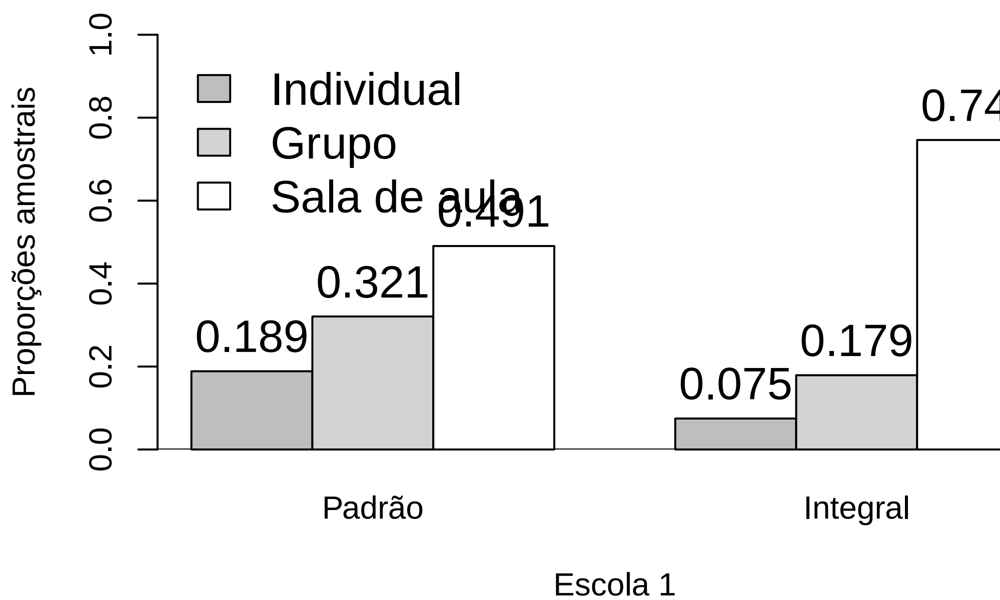
```r
#Lendo os dados
dados<-read.table("aprendizagem.txt", h=T)
dados
```

```
##   ind grupo sala escola periodo
## 1  10    17   26     E1       P
## 2   5    12   50     E1       I
## 3  21    17   26     E2       P
## 4  16    12   36     E2       I
## 5  15    15   16     E3       P
## 6  12    12   20     E3       I
```

```r
summary(dados)
```

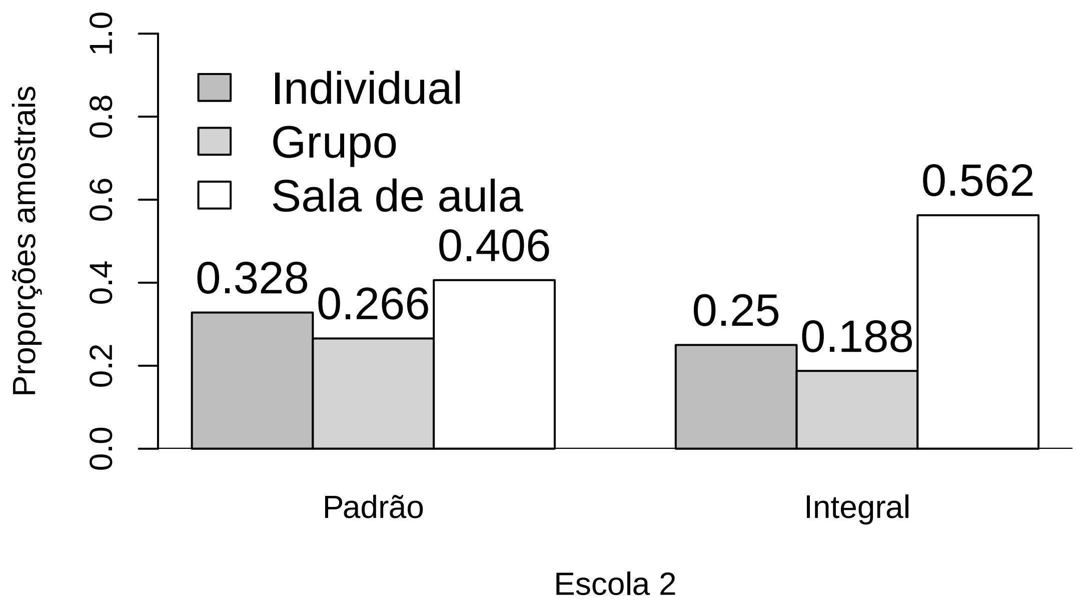
```
##        ind             grupo            sala          escola
##  Min.   : 5.00   Min.   :12.00   Min.   :16.0   Length:6
##  1st Qu.:10.50   1st Qu.:12.00   1st Qu.:21.5   Class :character
##  Median :13.50   Median :13.50   Median :26.0   Mode  :character
##  Mean   :13.17   Mean   :14.17   Mean   :29.0
##  3rd Qu.:15.75   3rd Qu.:16.50   3rd Qu.:33.5
##  Max.   :21.00   Max.   :17.00   Max.   :50.0
##    periodo
##  Length:6
##  Class :character
##  Mode  :character
##
##
##
```

```r
# par(mfrow=c(1,3))
   data<-cbind(P = c(10,17,26)/53, I = c(5,12,50)/67)
   bp<- barplot(height = data, beside = TRUE,
        col = c("gray","lightgray","white"), ylim=range(c(0,1)),
        names.arg = c("Padrão", "Integral"), xlab="Escola 1", ylab="Proporções amostrais",
        legend.text = c("Individual", "Grupo", "Sala de aula"),
        args.legend = list(x = "topleft", bty="n", cex=1.4))
   abline(h=0)
   text(bp, c(0.189,0.320,0.491,0.075,0.179,0.746), round(data,3), cex=1.4, pos=3)
```

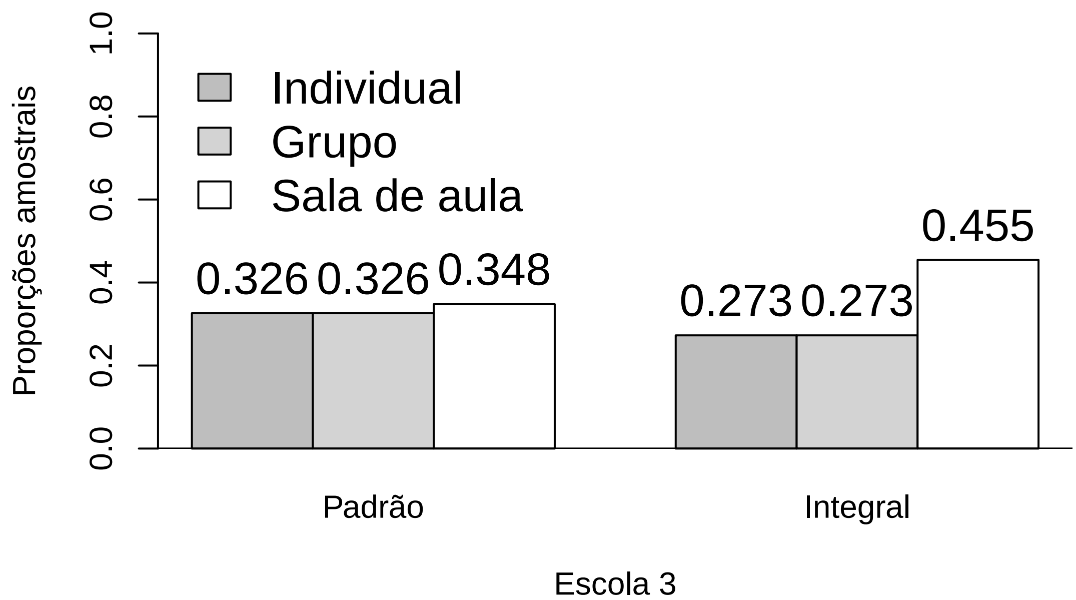


Representação gráfica dos dados

```r
data<-cbind(P = c(21,17,26)/64, I = c(16,12,36)/64)
bp<- barplot(height = data, beside = TRUE,
     col = c("gray","lightgray","white"), ylim=range(c(0,1)),
     names.arg = c("Padrão", "Integral"), xlab="Escola 2", ylab="Proporções amostrais",
     legend.text = c("Individual", "Grupo", "Sala de aula"),
     args.legend = list(x = "topleft", bty="n", cex=1.4))
abline(h=0)
text(bp, c(0.328,0.266,0.406,0.250,0.187,0.563), round(data,3), cex=1.4, pos=3)
```




```r
data<-cbind(P = c(15,15,16)/46, I = c(12,12,20)/44)
bp<- barplot(height = data, beside = TRUE,
     col = c("gray","lightgray","white"), ylim=range(c(0,1)),
     names.arg = c("Padrão", "Integral"), xlab="Escola 3", ylab="Proporções amostrais",
     legend.text = c("Individual", "Grupo", "Sala de aula"),
     args.legend = list(x = "topleft", bty="n", cex=1.4))
abline(h=0)
text(bp, c(0.326,0.326,0.348,0.273,0.273,0.454), round(data,3), cex=1.4, pos=3)
```



Ajuste do Modelo Para ajustar o modelo vamos utilizar a função 'vglm' do pacote 'VGAM' ()

```r
# Modelo sem covariaveis
mlcr0<-vglm(cbind(ind,grupo,sala)~1, multinomial, dados)
summary(mlcr0)
```

```
##
## Call:
## vglm(formula = cbind(ind, grupo, sala) ~ 1, family = multinomial,
##     data = dados)
##
## Coefficients:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept):1  -0.7896     0.1357  -5.820 5.88e-09 ***
## (Intercept):2  -0.7164     0.1323  -5.414 6.17e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Names of linear predictors: log(mu[,1]/mu[,3]), log(mu[,2]/mu[,3])
##
## Residual deviance: 30.248 on 10 degrees of freedom
##
## Log-likelihood: -39.4576 on 10 degrees of freedom
##
## Number of Fisher scoring iterations: 4
##
## No Hauck-Donner effect found in any of the estimates
##
```

```
##
## Reference group is level  3  of the response
```

```r
# Modelo com apenas a covariável escola
mlcr1<-vglm(cbind(ind,grupo,sala)~factor(escola), multinomial, dados)
summary(mlcr1)
```

```
##
## Call:
## vglm(formula = cbind(ind, grupo, sala) ~ factor(escola), family = multinomial,
##     data = dados)
##
## Coefficients:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept):1        -1.6227     0.2825  -5.743 9.28e-09 ***
## (Intercept):2        -0.9634     0.2183  -4.414 1.01e-05 ***
## factor(escola)E2:1    1.1065     0.3507   3.155 0.001604 **
## factor(escola)E2:2    0.2036     0.3135   0.650 0.515991
## factor(escola)E3:1    1.3350     0.3803   3.510 0.000448 ***
## factor(escola)E3:2    0.6758     0.3353   2.015 0.043893 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Names of linear predictors: log(mu[,1]/mu[,3]), log(mu[,2]/mu[,3])
##
## Residual deviance: 12.8716 on 6 degrees of freedom
##
## Log-likelihood: -30.7694 on 6 degrees of freedom
##
## Number of Fisher scoring iterations: 4
##
## No Hauck-Donner effect found in any of the estimates
##
##
## Reference group is level  3  of the response
```

```r
# Modelo com as covariáveis escola e periodo
mlcr2<-vglm(cbind(ind,grupo,sala)~factor(escola)+factor(periodo), multinomial, dados)
```

```
## Warning in vglm.fitter(x = x, y = y, w = w, offset = offset, Xm2 = Xm2, : some
## quantities such as z, residuals, SEs may be inaccurate due to convergence at a
## half-step
```

```r
summary(mlcr2)
```

```
##
## Call:
## vglm(formula = cbind(ind, grupo, sala) ~ factor(escola) + factor(periodo),
##     family = multinomial, data = dados)
##
## Coefficients:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept):1        -1.9708     0.3204  -6.152 7.67e-10 ***
## (Intercept):2        -1.3088     0.2596  -5.042 4.62e-07 ***
## factor(escola)E2:1    1.0828     0.3539   3.059 0.002217 **
## factor(escola)E2:2    0.1801     0.3172   0.568 0.570165
```

```
## factor(escola)E3:1   1.3147     0.3839   3.424 0.000616 ***
## factor(escola)E3:2   0.6556     0.3395   1.931 0.053456 .
## factor(periodo)P:1   0.7474     0.2820   2.651 0.008028 **
## factor(periodo)P:2   0.7426     0.2706   2.745 0.006057 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Names of linear predictors: log(mu[,1]/mu[,3]), log(mu[,2]/mu[,3])
##
## Residual deviance: 1.7776 on 4 degrees of freedom
##
## Log-likelihood: -25.2224 on 4 degrees of freedom
##
## Number of Fisher scoring iterations: 4
##
## No Hauck-Donner effect found in any of the estimates
##
##
## Reference group is level  3  of the response
```

```r
#Modelo com interação
mlcr3<-vglm(cbind(ind,grupo,sala)~factor(escola)+factor(periodo)+factor(escola)*factor(periodo), multin
```

```
## Warning in vglm.fitter(x = x, y = y, w = w, offset = offset, Xm2 = Xm2, :
## iterations terminated because half-step sizes are very small

## Warning in vglm.fitter(x = x, y = y, w = w, offset = offset, Xm2 = Xm2, : some
## quantities such as z, residuals, SEs may be inaccurate due to convergence at a
## half-step
```

```r
summary(mlcr3)
```

```
##
## Call:
## vglm(formula = cbind(ind, grupo, sala) ~ factor(escola) + factor(periodo) +
##     factor(escola) * factor(periodo), family = multinomial, data = dados)
##
## Coefficients:
##                                    Estimate Std. Error z value Pr(>|z|)
## (Intercept):1                       -2.3026     0.4690  -4.909 9.15e-07 ***
## (Intercept):2                       -1.4271     0.3215  -4.440 9.01e-06 ***
## factor(escola)E2:1                   1.4917     0.5570   2.678  0.00741 **
## factor(escola)E2:2                   0.3285     0.4631   0.709  0.47808
## factor(escola)E3:1                   1.7918     0.5944   3.014  0.00258 **
## factor(escola)E3:2                   0.9163     0.4865   1.883  0.05963 .
## factor(periodo)P:1                   1.3471     0.5987   2.250  0.02445 *
## factor(periodo)P:2                   1.0022     0.4479   2.238  0.02525 *
## factor(escola)E2:factor(periodo)P:1  -0.7497     0.7313  -1.025  0.30529
## factor(escola)E2:factor(periodo)P:2  -0.3285     0.6395  -0.514  0.60749
## factor(escola)E3:factor(periodo)P:1  -0.9008     0.7880  -1.143  0.25299
## factor(escola)E3:factor(periodo)P:2  -0.5559     0.6805  -0.817  0.41397
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Names of linear predictors: log(mu[,1]/mu[,3]), log(mu[,2]/mu[,3])
##
```

```
## Residual deviance: 1.06e-14 on 0 degrees of freedom
##
## Log-likelihood: -24.3336 on 0 degrees of freedom
##
## Number of Fisher scoring iterations: 6
##
## No Hauck-Donner effect found in any of the estimates
##
##
## Reference group is level  3  of the response
```

```r
AIC(mlcr0)
```

```
## [1] 82.91528
```

```r
  AIC(mlcr1)
```

```
## [1] 73.53885
```

```r
  AIC(mlcr2)
```

```
## [1] 66.44485
```

```r
  AIC(mlcr3)
```

```
## [1] 72.66724
```

```r
#função anova não está disponível em vglm( ); construir a partir das saídas.

deviance(mlcr2)
```

```
## [1] 1.777612
```

```r
  df.residual(mlcr2)
```

```
## [1] 4
```

```r
summary(mlcr2)
```

```
##
## Call:
## vglm(formula = cbind(ind, grupo, sala) ~ factor(escola) + factor(periodo),
##     family = multinomial, data = dados)
##
## Coefficients:
##                    Estimate Std. Error z value Pr(>|z|)
## (Intercept):1       -1.9708     0.3204  -6.152 7.67e-10 ***
## (Intercept):2       -1.3088     0.2596  -5.042 4.62e-07 ***
## factor(escola)E2:1   1.0828     0.3539   3.059 0.002217 **
## factor(escola)E2:2   0.1801     0.3172   0.568 0.570165
## factor(escola)E3:1   1.3147     0.3839   3.424 0.000616 ***
## factor(escola)E3:2   0.6556     0.3395   1.931 0.053456 .
## factor(periodo)P:1   0.7474     0.2820   2.651 0.008028 **
## factor(periodo)P:2   0.7426     0.2706   2.745 0.006057 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Names of linear predictors: log(mu[,1]/mu[,3]), log(mu[,2]/mu[,3])
##
## Residual deviance: 1.7776 on 4 degrees of freedom
```

```
##
## Log-likelihood: -25.2224 on 4 degrees of freedom
##
## Number of Fisher scoring iterations: 4
##
## No Hauck-Donner effect found in any of the estimates
##
##
## Reference group is level  3  of the response
```

```r
coef(mlcr2, matrix=TRUE)
```

```
##                   log(mu[,1]/mu[,3]) log(mu[,2]/mu[,3])
## (Intercept)                -1.970755         -1.3088271
## factor(escola)E2            1.082824          0.1801073
## factor(escola)E3            1.314712          0.6555962
## factor(periodo)P            0.747440          0.7426318
```

```r
mlcr2@y
```

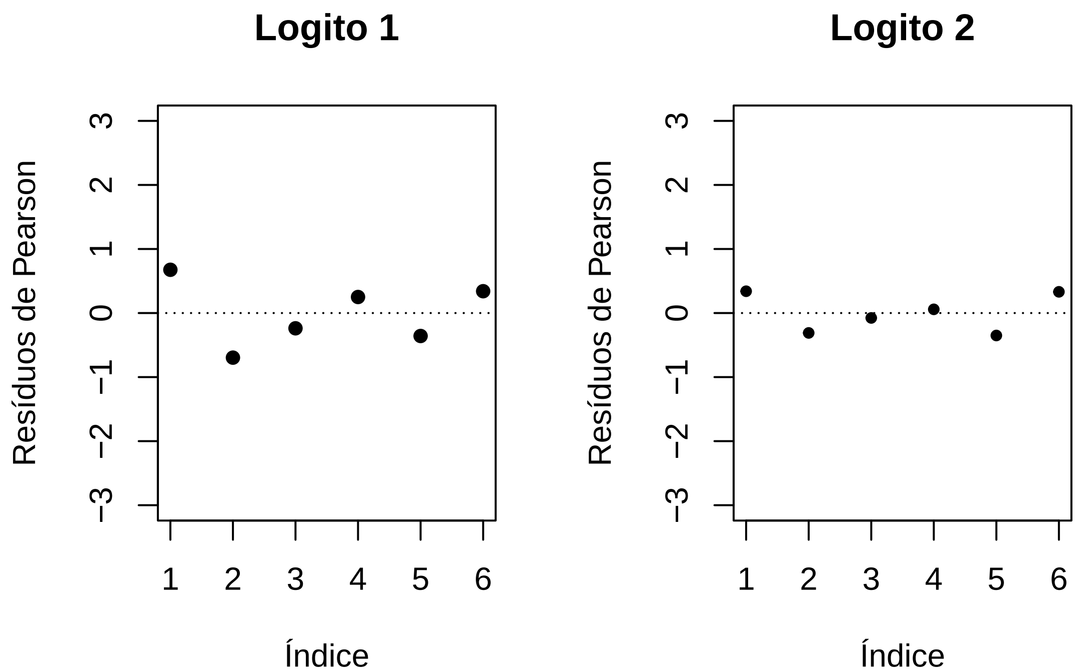
```
##          ind     grupo      sala
## 1 0.18867925 0.3207547 0.4905660
## 2 0.07462687 0.1791045 0.7462687
## 3 0.32812500 0.2656250 0.4062500
## 4 0.25000000 0.1875000 0.5625000
## 5 0.32608696 0.3260870 0.3478261
## 6 0.27272727 0.2727273 0.4545455
```

```r
fitted(mlcr2)
```

```
##          ind     grupo      sala
## 1 0.15803628 0.3048879 0.5370759
## 2 0.09886683 0.1916559 0.7094773
## 3 0.34093911 0.2666952 0.3923657
## 4 0.23718589 0.1864298 0.5763843
## 5 0.34356466 0.3428794 0.3135560
## 6 0.25445513 0.2551716 0.4903733
```

```r
#   mlcr2@y - fitted(mlcr2)
```

```r
par(mfrow=c(1,2))
   rp<-resid(mlcr2, type = "pearson")
   plot(rp[,1], pch=16, ylim=c(-3,3), xlab="Índice", ylab="Resíduos de Pearson")
   title("Logito 1")
   abline(h=0, lty=3)
   plot(rp[,2], pch=20, ylim=c(-3,3), xlab="Índice", ylab="Resíduos de Pearson")
   abline(h=0, lty=3)
   title("Logito 2")
```



O modelo ajustado tem a seguinte expressão

$$\log(\frac{\pi_1}{\pi_3}) = \alpha_1 + \beta_{11}x_1 + \beta_{12}x_2 + \beta_{13}x_{13} \tag{4}$$

e

$$\log(\frac{\pi_2}{\pi_3}) = \alpha_2 + \beta_{21}x_1 + \beta_{22}x_2 + \beta_{23}x_3 \tag{5}$$

Em que, as covariáveis escola e período escolar foram consideradas no modelo por meio de variáveis *dummy* tal que:

- $x_1 = 1$ se escola 2 e 0 caso contrário;
- $x_2 = 1$ se escola 3 e 0 caso contrário;
- $x_3 = 1$ se padrão e 0 se integral

Assim, o modelo ajustado fica

$$\log(\frac{\pi_1}{\pi_3}) = -1,9708 + 1,0828x_1 + 1,3147x_2 + 0,7474x_{13} \tag{6}$$

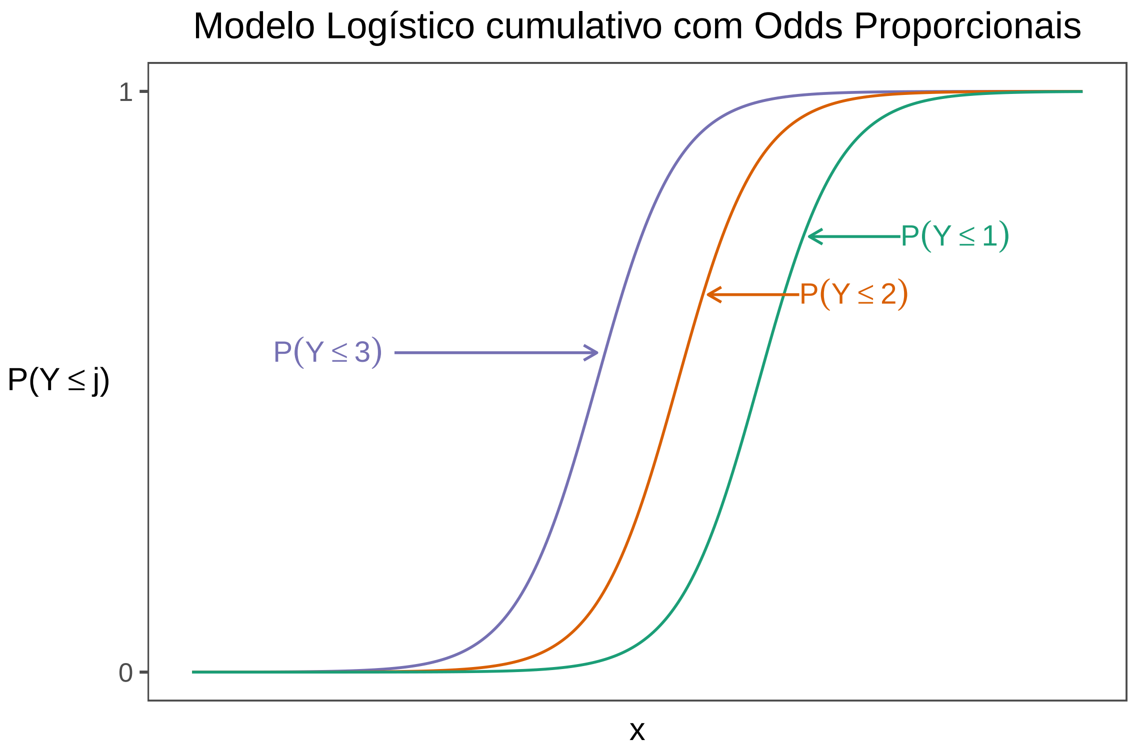
$$\log(\frac{\pi_2}{\pi_3}) = -1,3088 + 0,1801x_1 + 0,3556x_2 + 0,7426x_{13} \tag{7}$$

Modelo Logístico para Resposta Ordinal

Modelo Logístico cumulativo com Odds Proporcionais

$$\text{logit}[P(Y \leq j)] = \log\left(\frac{P(Y \leq j)}{P(Y > j)}\right) = \alpha_j + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p,\ j = 1, \ldots,\ c - 1 \tag{8}$$

```r
cdf=function(x,mu,s){
  k=(x-mu)/s
  return(1/(1+exp(-k)))
}

colors <- brewer.pal(n = 4, name = "Dark2")

theData <- tibble(x=seq(-10,12,0.01)) |>
  mutate(curve0 = cdf(x, 0, 1)) |>
  mutate(curve2 = cdf(x, 2, 1)) |>
  mutate(curve4 = cdf(x, 4, 1))

library(ggthemes)  # theme_few
theData %>%
  ggplot(aes(x = x)) +
  geom_line(aes(y = curve0), color = colors[3]) +
  geom_line(aes(y = curve2), color = colors[2]) +
  geom_line(aes(y = curve4), color = colors[1]) +
  ggtitle("Modelo Logístico cumulativo com Odds Proporcionais") +
  ylab(expression("P(Y" <= "j)")) +
  scale_y_continuous(breaks=c(0,1),
        labels=c("0", "1")) +
  theme_few() +
  theme(axis.text.x = element_blank(),
        axis.ticks.x = element_blank(),
        axis.title.y = element_text(angle = 0, vjust = 0.5),
        plot.title = element_text(hjust = 0.5)) +
    annotate(geom = "segment", x =-5, y = 0.55, xend = 0.0, yend = .55,
          arrow = arrow(length = unit(2, "mm")), color = colors[3]) +
    annotate(geom = "text", x = -8, y = 0.55, label = paste("P(Y <= 3)"),
            parse =  TRUE,hjust = "left", color = colors[3])  +

    annotate(geom = "segment", x =5, y = 0.65, xend = 2.75, yend = .65,
          arrow = arrow(length = unit(2, "mm")) , color = colors[2]) +
    annotate(geom = "text", x = 5, y = 0.65, label = paste("P(Y <= 2)"),
            parse = TRUE, hjust = "left", color = colors[2]) +

    annotate(geom = "segment", x =7.5, y = 0.75, xend = 5.25, yend = .75,
          arrow = arrow(length = unit(2, "mm")) , color = colors[1]) +
    annotate(geom = "text", x = 7.5, y = 0.75,  label = paste("P(Y <= 1)"),
            parse = TRUE, hjust = "left", color = colors[1])
```



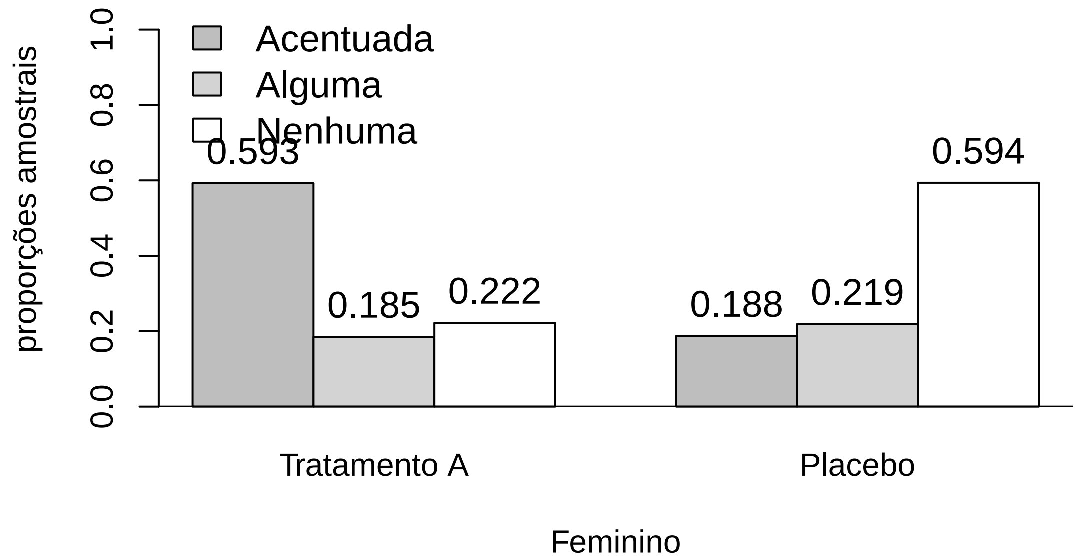
Exercício 2

O arquivo 'artrite1.txt' apresenta os dados de um ensaio clínico, sobre tratamento para dores de artrite, com 84 pacientes do sexo masculino e feminino que receberam o tratamento A ou placebo. A variável de interesse foi o grau de melhora de suas dores de artrite em uma de três categorias: melhora acentuada, alguma melhora ou nenhuma melhora.

```
#Lendo os dados
dados<-read.table("artrite1.txt", h=T)
#dados
summary(dados)
```

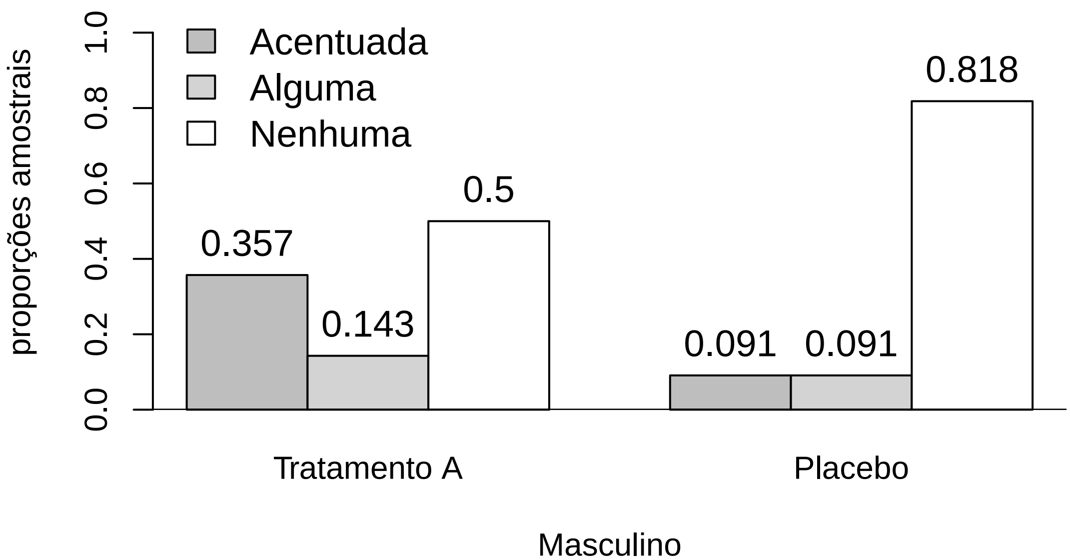
```
##     melhora           sexo          tratamento
##  Min.   :1.000   Min.   :1.000   Min.   :1.000
##  1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1.000
##  Median :1.500   Median :2.000   Median :1.000
##  Mean   :1.833   Mean   :1.702   Mean   :1.488
##  3rd Qu.:3.000   3rd Qu.:2.000   3rd Qu.:2.000
##  Max.   :3.000   Max.   :2.000   Max.   :2.000
```

```
#par(mfrow=c(1,2))
   data<-cbind(P = c(16,5,6)/27, I = c(6,7,19)/32)
   bp<- barplot(height = data, beside = TRUE,
        col = c("gray","lightgray", "white"), ylim=range(c(0,1.1)),
        names.arg = c("Tratamento A", "Placebo"), xlab="Feminino", ylab="proporções amostrais",
        legend.text = c("Acentuada", "Alguma", "Nenhuma"),
        args.legend = list(x = "topleft", bty="n", cex=1.2))
   abline(h=0)
   text(bp,c(16/27, 5/27, 6/27, 6/32, 7/32, 19/32), round(data,3), cex=1.2, pos=3)
```



Representação Gráfica

```
data<-cbind(P = c(5,2,7)/14, I = c(1,1,9)/11)
bp<- barplot(height = data, beside = TRUE,
     col = c("gray","lightgray", "white"), ylim=range(c(0,1.1)),
     names.arg = c("Tratamento A", "Placebo"), xlab="Masculino", ylab="proporções amostrais",
     legend.text = c("Acentuada", "Alguma", "Nenhuma"),
     args.legend = list(x = "topleft", bty="n", cex=1.2))
abline(h=0)
text(bp,c(5/14, 2/14, 7/14, 1/11, 1/11, 9/11), round(data,3), cex=1.2, pos=3)
```




```r
## Odds não proporcionais
mlc<-vglm(melhora ~factor(sexo)+factor(tratamento), cumulative(parallel=FALSE,reverse=FALSE), dados)
summary(mlc)
```

```
##
## Call:
## vglm(formula = melhora ~ factor(sexo) + factor(tratamento), family = cumulative(parallel = FALSE,
##     reverse = FALSE), data = dados)
##
## Coefficients:
##                       Estimate Std. Error z value Pr(>|z|)
## (Intercept):1           1.9956     0.5982   3.336 0.000850 ***
## (Intercept):2           2.3607     0.6482   3.642 0.000270 ***
## factor(sexo)2:1        -1.5505     0.5757  -2.693 0.007078 **
## factor(sexo)2:2        -0.9566     0.5872  -1.629 0.103281
## factor(tratamento)2:1  -1.8694     0.5154  -3.627 0.000287 ***
## factor(tratamento)2:2  -1.7733     0.5300  -3.346 0.000821 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Names of linear predictors: logitlink(P[Y<=1]), logitlink(P[Y<=2])
##
## Residual deviance: 148.5601 on 162 degrees of freedom
##
## Log-likelihood: -74.28 on 162 degrees of freedom
##
## Number of Fisher scoring iterations: 10
##
## No Hauck-Donner effect found in any of the estimates
```

```
##
##
## Exponentiated coefficients:
##       factor(sexo)2:1       factor(sexo)2:2 factor(tratamento)2:1
##             0.2121370             0.3841832             0.1542131
## factor(tratamento)2:2
##             0.1697681
```

```r
coef(mlc, matrix = TRUE)
```

```
##                     logitlink(P[Y<=1]) logitlink(P[Y<=2])
## (Intercept)                   1.995580          2.3607401
## factor(sexo)2                -1.550523         -0.9566359
## factor(tratamento)2          -1.869420         -1.7733221
```

```r
## Odds proporcionais
mop<-vglm(melhora ~factor(sexo)+factor(tratamento), cumulative(parallel=TRUE), dados)
summary(mop)
```

```
##
## Call:
## vglm(formula = melhora ~ factor(sexo) + factor(tratamento), family = cumulative(parallel = TRUE),
##     data = dados)
##
## Coefficients:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept):1         1.8128     0.5566   3.257 0.001127 **
## (Intercept):2         2.6672     0.5997   4.448 8.68e-06 ***
## factor(sexo)2        -1.3187     0.5292  -2.492 0.012702 *
## factor(tratamento)2  -1.7973     0.4728  -3.801 0.000144 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Names of linear predictors: logitlink(P[Y<=1]), logitlink(P[Y<=2])
##
## Residual deviance: 150.0294 on 164 degrees of freedom
##
## Log-likelihood: -75.0147 on 164 degrees of freedom
##
## Number of Fisher scoring iterations: 5
##
## No Hauck-Donner effect found in any of the estimates
##
##
## Exponentiated coefficients:
##       factor(sexo)2 factor(tratamento)2
##           0.2674697           0.1657452
```

```r
coef(mop, matrix = TRUE)
```

```
##                     logitlink(P[Y<=1]) logitlink(P[Y<=2])
## (Intercept)                   1.812799           2.667192
## factor(sexo)2                -1.318749          -1.318749
## factor(tratamento)2          -1.797303          -1.797303
```

Verificando a suposiçao de odds proporcionais (Teste da razão de verossimilhanças)

```
TRV<- 2*(logLik(mlc)-logLik(mop))
   gl <- length(coef(mlc))-length(coef(mop)); p<-1-pchisq(TRV,gl)
   cbind(TRV, gl, p)
```

```
##           TRV gl         p
## [1,] 1.469327  2 0.4796668
```

ou de forma equivalente,

```
TRV<- deviance(mop)-deviance(mlc)
   gl <- df.residual(mop)-df.residual(mlc)
   p  <- 1-pchisq(TRV,gl)
   cbind(TRV, gl, p)
```

```
##           TRV gl         p
## [1,] 1.469327  2 0.4796668
```

Portanto ficamos com o modelo de Odds proporcionais

```

mop0<-vglm(melhora ~1, cumulative(parallel=T,reverse=F), dados)
 #  mop0
   mop1<-vglm(melhora ~factor(sexo), cumulative(parallel=T,reverse=F), dados)
#   mop1
   mop2<-vglm(melhora ~factor(sexo)+factor(tratamento), cumulative(parallel=T,reverse=F), dados)
 #  mop2
   mop3<-vglm(melhora ~factor(sexo)+factor(tratamento)+factor(sexo)*factor(tratamento),
              cumulative(parallel=T,reverse=F), dados)
 #  mop3

   AIC(mop0)
```

Escolha: MOP

```
## [1] 173.9159
```

```
   AIC(mop1)
```

```
## [1] 172.1106
```

```
   AIC(mop2)
```

```
## [1] 158.0294
```

```
   AIC(mop3)
```

```
## [1] 159.721
```

```
TRV<- deviance(mop2)-deviance(mop3)
   gl <- df.residual(mop2)-df.residual(mop3)
   p  <- 1-pchisq(TRV,gl)
   cbind(TRV, gl, p)
```

```
##            TRV gl         p
## [1,] 0.3084549  1 0.5786298
```

```
   TRV<- deviance(mop0)-deviance(mop1)
   gl <- df.residual(mop0)-df.residual(mop1)
```

```r
p  <- 1-pchisq(TRV,gl)
cbind(TRV, gl, p)
```

```
##           TRV gl          p
## [1,] 3.805276  1 0.05109137
```

```r
TRV<- deviance(mop1)-deviance(mop2)
gl <- df.residual(mop1)-df.residual(mop2)
p  <- 1-pchisq(TRV,gl)
cbind(TRV, gl, p)
```

```
##           TRV gl           p
## [1,] 16.08123  1 6.06826e-05
```

Portanto, o modelo selecionado é o modelo com apenas os feitos principais de sexo e tratamento. Logo,

```r
summary(mop2)
```

```
##
## Call:
## vglm(formula = melhora ~ factor(sexo) + factor(tratamento), family = cumulative(parallel = T,
##     reverse = F), data = dados)
##
## Coefficients:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept):1         1.8128     0.5566   3.257 0.001127 **
## (Intercept):2         2.6672     0.5997   4.448 8.68e-06 ***
## factor(sexo)2        -1.3187     0.5292  -2.492 0.012702 *
## factor(tratamento)2  -1.7973     0.4728  -3.801 0.000144 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Names of linear predictors: logitlink(P[Y<=1]), logitlink(P[Y<=2])
##
## Residual deviance: 150.0294 on 164 degrees of freedom
##
## Log-likelihood: -75.0147 on 164 degrees of freedom
##
## Number of Fisher scoring iterations: 5
##
## No Hauck-Donner effect found in any of the estimates
##
##
## Exponentiated coefficients:
##       factor(sexo)2 factor(tratamento)2
##           0.2674697           0.1657452
```

```r
coef(mop2,matrix = TRUE)
```

```
##                     logitlink(P[Y<=1]) logitlink(P[Y<=2])
## (Intercept)                   1.812799           2.667192
## factor(sexo)2                -1.318749          -1.318749
## factor(tratamento)2          -1.797303          -1.797303
```

```r
QL<-deviance(mop2);
```

```r
 p <-  1-pchisq(QL,4)
cbind(QL,p)
```

Análise dos resíduos

```
##            QL p
## [1,] 150.0294 0
```

```r
   rp<-residuals(mop2, type="pearson")
   Qp<-sum(rp[,1]^2) + sum(rp[,2]^2)
   p <-  1-pchisq(Qp,4)
cbind(Qp,p)
```

```
##            Qp p
## [1,] 166.7277 0
```