

Aula 08

Matheus Cougias e Klysman Rezende

18/09/2020

Leitura de dados

Neste banco de dados temos a quantidade de animais mortos, a quantidade que foi submetida a teste, o tipo de veneno e o logaritmo da quantidade dosada.

Podemos perceber uma média de animais testados bem próxima de 50 animais sendo que a maioria dos animais envenenados morreram (em boa parte dos testes).

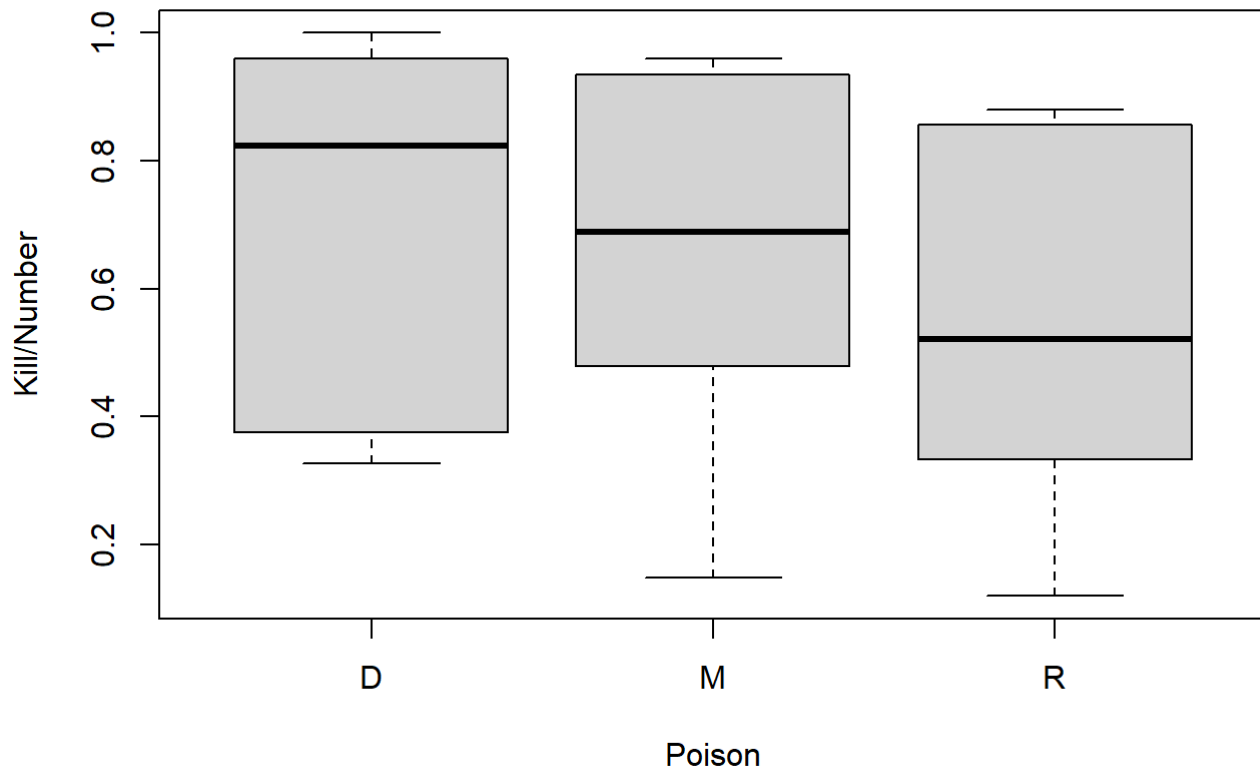
```
dados <- read.csv("potency.dat", sep="\t")
dados$Poison <- as.factor(dados$Poison)
```

Análise Exploratória

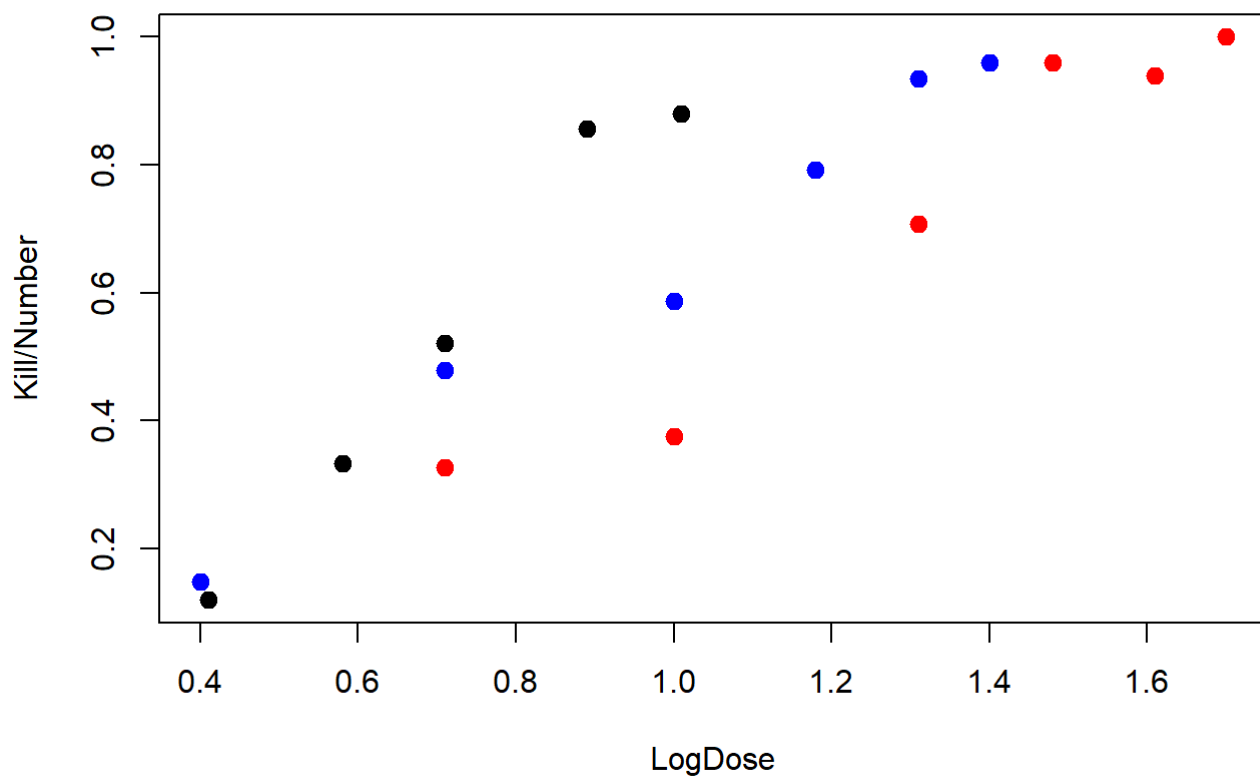
Pela análise de qual dos venenos seria o melhor aplicar, o veneno D apresenta uma taxa de mortes entre 20% e 100% (com média próxima a 80%), o veneno M encontra-se entre 10% e 90% (com média em 70%) e o veneno R encontra-se entre 5% e 90% (com média em 60%).

No segundo gráfico, vermelho=D, azul=M e preto=R. percebemos que o veneno R foi utilizado nas doses maiores (talvez explicando o motivo da maior taxa de mortes). A melhor maneira para tirar conclusões é utilizando a regressão e analisando seus p-valores, entendendo melhor qual veneno escolher.

```
plot(Kill/Number ~ Poison, data=dados)
```



```
plot(Kill/Number ~ LogDose, data=dados, pch=19, cex=1.2, col=c("red", "blue", "black")[dados
$Poison])
```



Regressão SEM interação

Nesse caso, avaliaremos qual o comportamento separadamente do tipo do veneno e a dose. Para facilitar a análise, consideramos que as doses foram multiplicadas por 10.

Como esperado, o logaritmo da dose possui valor positivo, mostrando que quanto maior a dose aplicada no animal, maior as chances dele morrer. Assim, a cada aumento de 0.1 na dose, temos um acréscimo de chances do animal morrer em $\exp(0.48277) = 62.05\%$.

Além disso, através dos resultados existe uma tendência em escolher o veneno de tipo R para comprar, já que apresenta a maior probabilidade de matar o animal. Esse fato rebate a ideia da análise exploratória, podendo ser explicada pelo fato das diferenças entre os tamanhos das doses, já que o veneno D foi utilizado em maior quantidade.

```
modelo <- glm(cbind(Kill, Number-Kill) ~ Poison + I(LogDose*10), data=dados, family=binomial)
summary(modelo)
```

```
##
## Call:
## glm(formula = cbind(Kill, Number - Kill) ~ Poison + I(LogDose *
##      10), family = binomial, data = dados)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7725  -1.0948   0.5153   1.4039   2.2419
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -4.86876    0.42009  -11.590   < 2e-16 ***
## PoisonM         0.91230    0.24494   3.725 0.000196 ***
## PoisonR         1.60339    0.26562   6.036 1.58e-09 ***
## I(LogDose * 10) 0.48277    0.03395  14.222   < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 350.069  on 16  degrees of freedom
## Residual deviance:  31.971  on 13  degrees of freedom
## AIC: 98.955
##
## Number of Fisher Scoring iterations: 4
```

Regressão COM interação

Através da nova análise, encontramos os efeitos marginais do M e do R, além do efeito da dose, todos encontrados na regressão anterior. Será mais interessante analisar nesse modelo os valores encontrados para interação entre o Veneno e o LogDose. Ainda assim encontramos que o veneno do tipo R continua sendo o mais aconselhável a comprar.

```
modelo <- glm(cbind(Kill, Number-Kill) ~ Poison + I(LogDose*10) + Poison:I(LogDose*10), data=
dados, family=binomial)
summary(modelo)
```

```
##
## Call:
## glm(formula = cbind(Kill, Number - Kill) ~ Poison + I(LogDose *
##      10) + Poison:I(LogDose * 10), family = binomial, data = dados)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.77331  -0.67995   0.03622   1.00671   2.02531
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -4.45094     0.61044  -7.291 3.07e-13 ***
## PoisonM         1.04397     0.77946   1.339  0.1805
## PoisonR        -0.38778     0.88275  -0.439  0.6605
## I(LogDose * 10)  0.44602     0.05155   8.652 < 2e-16 ***
## PoisonM:I(LogDose * 10) -0.02244     0.07140  -0.314  0.7533
## PoisonR:I(LogDose * 10)  0.26079     0.10221   2.551  0.0107 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 350.069  on 16  degrees of freedom
## Residual deviance:  22.723  on 11  degrees of freedom
## AIC: 93.707
##
## Number of Fisher Scoring iterations: 4
```