

Bioestatística em R - Exercícios

Nome: _____

Data: _____

Matricula: _____

Inferência em Dados Categóricos

Em um estudo sobre hipertensão, os pesquisadores revisaram a literatura e identificaram quatro polimorfismos no gene eNOS (enosa, enosb, enosc, enosd) que foram previamente associados a suscetibilidade ao desenvolvimento de hipertensão (hbp). Os pesquisadores coletaram 50 amostras de pacientes durante um mês de atividades para os quais genotiparam os quatro polimorfismos e diagnosticaram hipertensão.

Sobre este estudo, responda:

```
##dados disponíveis no arquivo cf.tsv
cf <- read.table('cf.tsv', header = T)
## resumo dos dados disponíveis
summary(cf)
```

```
##  enosa  enosb  enosc  enosd  hbp
## aa:39  bb:22  cc: 2  dd:16  não:39
## Aa:11  Bb:19  Cc:17  Dd: 2  sim:11
##          BB: 9  CC:31  DD:32
```

1. Sabendo que na mesma clínica os médicos haviam constatado que pacientes hipertensos se consultam com uma frequência de 16% a cada mês. O número de hipertensos observado difere da proporção normalmente observada?

```
has_hbp <- cf$hbp == "sim"
## frequência de hipertensão
mean(has_hbp)
```

```
## [1] 0.22
```

```
## Visto que estamos comparando a frequência de uma variável
## binomial, usamos o teste binomial para comparar com a
## frequência observada.
binom.test(table(has_hbp), p=(1-0.16))
```

```
##
## Exact binomial test
##
```

```
## data: table(has_hbp)
## number of successes = 39, number of trials = 50, p-value = 0.2469
## alternative hypothesis: true probability of success is not equal to 0.84
## 95 percent confidence interval:
## 0.6403881 0.8847342
## sample estimates:
## probability of success
## 0.78
```

2. Em estudos de genética humana, geralmente espera-se que todos os polimorfismos estudados adiram ao equilíbrio de Hardy-Weinberg, visto que o desvio indica uma forte pressão evolutiva, ou cruzamento seletivo está ocorrendo e ambos são incomuns em populações urbanas. Portanto, uma boa medida de diagnostico para averiguar se nenhum erro de metodologia ocorreu é testar o equilíbrio e excluir os genes que apresentam desvio. Avalie o desequilíbrio de Hardy-Weinberg para as mutações do eNOS e indique se algum polimorfismo precisa ser excluído ilustrando a diferença observada.

```
## definir função para calcular hardy-weinberg
hw <- function(obs) {
  n <- sum(obs)
  p <- (2 * obs[1] + obs[2]) / (2*n)
  exp <- c(p^2, 2 * p * (1-p), (1-p)^2) * n
  chi <- sum((obs-exp)^2/exp)
  c("chi.squared" = chi, "p.value" = pchisq(chi, df=1, lower.tail = FALSE))
}

## calcula-se o desequilíbrio para os polimorfismos:
## enosa, mas primeiramente precisamos corrigir a variavel
## para incluir AA
cf$enosa <- factor(cf$enosa, levels = c("aa", "Aa", "AA"))
hw(table(cf$enosa))

## chi.squared      p.value
## 0.7637925      0.3821443

## enosb
hw(table(cf$enosb))

## chi.squared      p.value
## 1.7093854      0.1910651

## enosc
hw(table(cf$enosc))

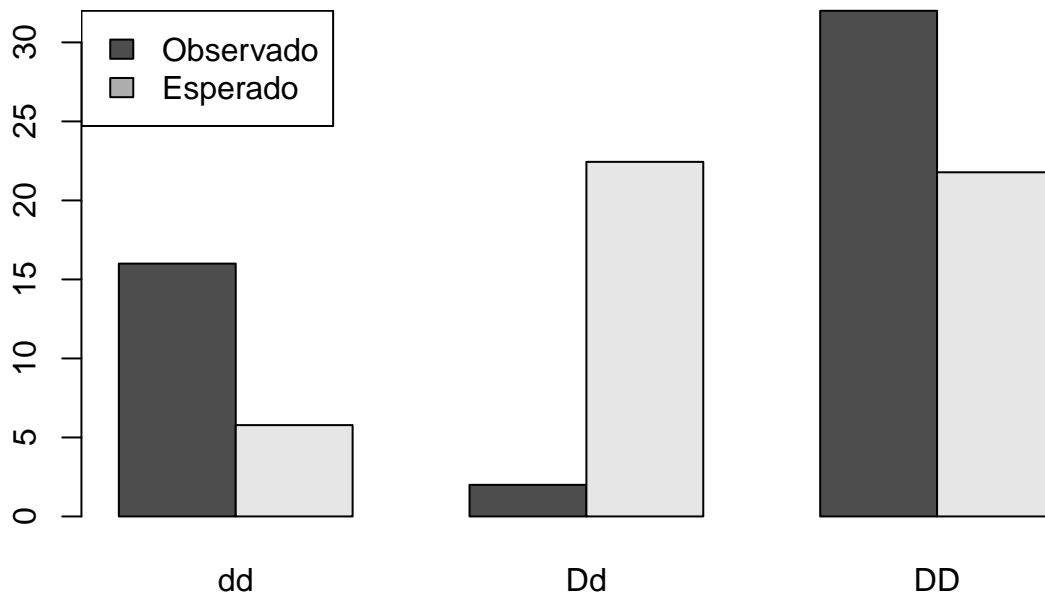
## chi.squared      p.value
## 0.03053831      0.86127441
```

```
## enosd
hw(table(cf$enosd))

##  chi.squared      p.value
## 4.148452e+01 1.188079e-10

## como enosd apresentou um desvio significativo, vamos ilustrar
## a frequência observada ao lado da esperada para visualizar a
## diferença.
obs <- table(cf$enosd)
p <- (2 * obs[1] + obs[2]) / (2*sum(obs))
exp <- c(p^2, 2 * p * (1-p), (1-p)^2) * sum(obs)

barplot(rbind(obs, exp), beside = T)
legend("topleft", c("Observado", "Esperado"), fill = gray.colors(3))
```



3. Para as mutações que se encontram em equilíbrio de Hardy-Weinberg, avalie se elas estão associadas com a ocorrência de hipertensão. Justifique a abordagem utilizada e ilustre a relação.

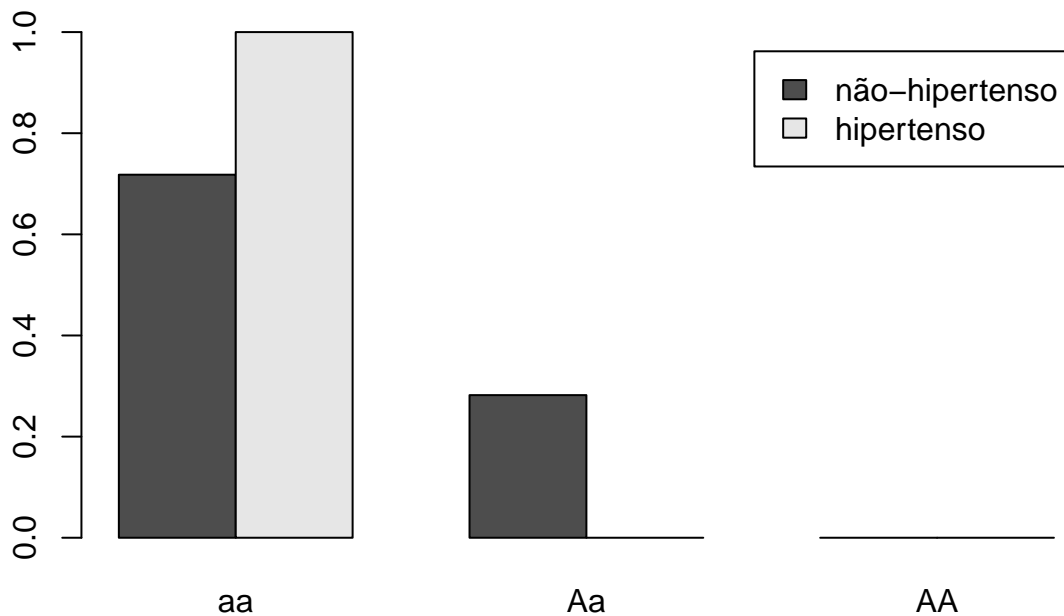
```
## enosa
## gera e imprime a tabela de confusão.
addmargins(enosa <- table(cf$hbp, cf$enosa))
```

```
##
##      aa Aa AA Sum
## não 28 11  0 39
## sim 11  0  0 11
## Sum 39 11  0 50
```

```
## avalia se existe associação.
chisq.test(enosa)
```

```
##
## Pearson's Chi-squared test
##
## data:  enosa
## X-squared = NaN, df = 2, p-value = NA
```

```
## ilustra a relação.
barplot(prop.table(enosa, 1), beside = TRUE,
        legend = c("não-hipertenso", "hipertenso"))
```



```
## enosb
## gera e imprime a tabela de confusão.
addmargins(enosb <- table(cf$hbp, cf$enosb))
```

```
##
##      bb Bb BB Sum
## não 21 14  4 39
## sim  1  5  5 11
## Sum 22 19  9 50
```

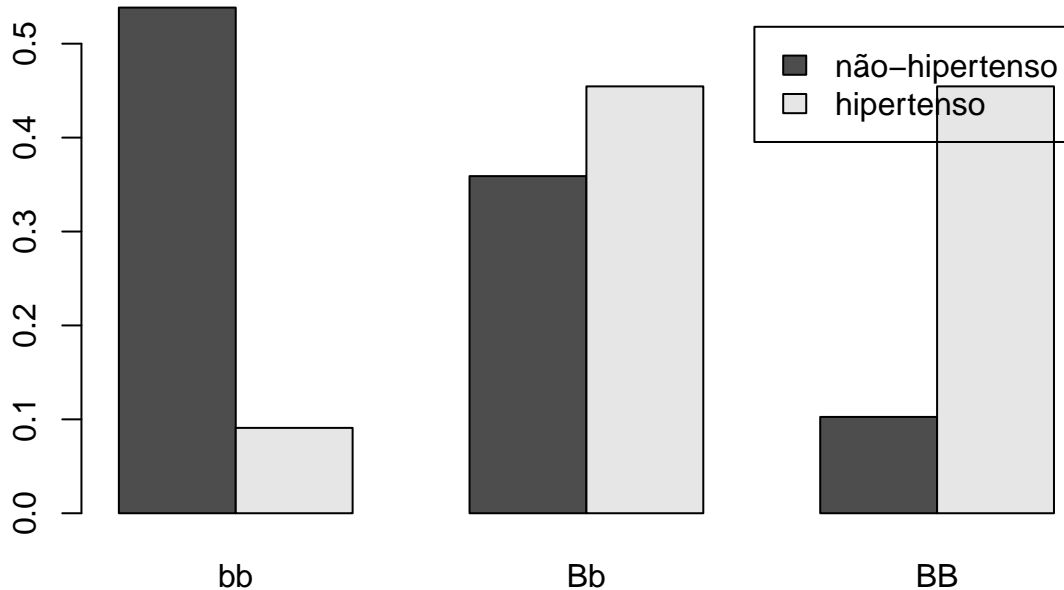
```
## avalia se existe associação.
chisq.test(enosb)
```

```
##
## Pearson's Chi-squared test
##
## data:  enosb
```

```
## X-squared = 10.018, df = 2, p-value = 0.006679
```

```
## ilustra a relação.
```

```
barplot(prop.table(enosb, 1), beside = TRUE,  
        legend = c("não-hipertenso", "hipertenso"))
```



```
## enosc  
## gera e imprime a tabela de confusão.  
addmargins(enosc <- table(cf$hbp, cf$enosc))
```

```
##  
##      cc Cc CC Sum  
## não  2 14 23  39  
## sim  0  3  8  11  
## Sum  2 17 31  50
```

```
## avalia se existe associação.  
chisq.test(enosc)
```

```
##  
## Pearson's Chi-squared test  
##  
## data:  enosc  
## X-squared = 1.0136, df = 2, p-value = 0.6024
```

```
## ilustra a relação.
```

```
barplot(prop.table(enosc, 1), beside = TRUE)  
legend("topleft", c("não-hipertenso", "hipertenso"), fill = gray.colors(2))
```

