Laboratório 8: Regressão Linear Múltipla

Dando Nota ao Professor

Vários cursos universitários dão aos alunos a oportunidade de avaliar o curso e o professor de maneira anônima ao final do semestre. Contudo, o uso das avaliações dos alunos como um indicador da qualidade do curso e a eficácia do ensino é frequentemente criticado porque essas medidas podem refletir a influência de características não relacionadas à docência, tal como a aparência física do professor. O artigo intitulado "Beauty in the classroom: instructors' pulchritude and putative pedagogical productivity" (Hamermesh & Parker, 2005)* descreve como professores que são vistos como tendo melhor aparência recebem avaliações mais altas.†

Neste laboratório analisaremos os dados deste estudo para aprender o que influencia uma avaliação positiva de um professor.

Os Dados

Os dados foram coletados a partir das avaliações discentes de final de semestre de uma grande amostra de professores da Universidade do Texas em Austin. Além disso, seis estudantes avaliaram a aparência física dos professores.‡ O resultado é um banco de dados no qual cada linha contém diferentes disciplinas e cada coluna representa as variáveis sobre as disciplinas e os professores.

```
download.file("http://www.openintro.org/stat/data/evals.RData", destfile = "evals.RData")

load("evals.RData")
```

Explorando os Dados

Exercício 1 Esse estudo é observacional ou experimental? O pergunta de pesquisa original proposta no artigo é se a beleza influencia diretamente as avaliações das disciplinas. Levando em consideração o desenho da pesquisa, é possível responder a essa pergunta tal como ela está formulada? Se não, reformule a pergunta.

Exercício 2 Descreva a distribuição da variável `score`. A distribuição é assimétrica? O que sua forma permite dizer sobre a maneira como os alunos avaliam as disciplinas? A forma corresponde ao que você esperava ver? Por quê, ou por que não?

Exercício 3 Com exceção da variável `score`, escolha duas outras variáveis e descreva sua relação utilizando as técnicas apropriadas (gráfico de dispersão, gráfico de caixas lado-a-lado, ou gráfico de mosaico).

 Este laboratório foi adaptado para a OpenIntro por Andrew Bray e Mine Çetinkaya-Rundel de um laboratório escrito por Mark Hansen do departamento de Estatística da UCLA. Tradução para o português por Erikson Kaszubowski.

*"Beleza na sala de aula: a pulcritude do professor e produtividade pedagógica putativa"

†Daniel S. Hamermesh, Amy Parker, Beauty in the classroom: instructors' pulchritude and putative pedagogical productivity, *Economics of Education Review*, Volume 24, Issue 4, August 2005, Pages 369-376, ISSN 0272-7757, 10.1016/j.econedurev.2004.07.013. (http://www.sciencedirect.com/science/article/pii/S0272775704001165).

‡Esta é uma versão levemente modificada do conjunto de dados original que foi publicado como parte dos dados de reprodução para o livro *Data Analysis Using Regression and Multilevel/Hierarchical Models* (Gelman and Hill, 2007).

score	pontuação média da avaliação do docente: (1) muito insatisfatório - (5) excelente.
rank	nível do professor: horista (teaching), assistente (tenure track), titular (tenured).*
ethnicity	etnia do professor: não-minoria, minoria.
gender	sexo do professor: feminino, masculino.
language	língua da universidade frequentada pelo professor: inglês ou não-inglês.
age	idade do professor.
cls_perc_eval	percentual de alunos na turma que completaram a avaliação.
cls_did_eval	número de alunos na turma que completaram a avaliação.
cls_students	número total de alunos na turma.
cls_level	nível da disciplina: introdutória, avançada.
cls_profs	número de professores ministrando módulos na disciplina dentro da amostra: único, múltiplos.
cls_credits	número de créditos da disciplina: um crédito, múltiplos créditos.
bty_f1lower	avaliação da beleza do professor por aluna de nível inicial: (1) mais baixo - (10) mais alto.
bty_f1upper	avaliação da beleza do professor por aluna de nível avançado: (1) mais baixo - (10) mais alto.
bty_f2upper	avaliação da beleza do professor por segunda aluna de nível avançado: (1) mais baixo - (10) mais alto.
bty_m1lower	avaliação da beleza do professor por aluno de nível inicial: (1) mais baixo - (10) mais alto.
bty_m1upper	avaliação da beleza do professor por aluno de nível avançado: (1) mais baixo - (10) mais alto.
bty_m2upper	avaliação da beleza do professor por segundo aluno de nível avançado: (1) mais baixo - (10) mais alto.
bty_avg	média da avaliação da beleza do professor.
pic_outfit	roupa do professor na foto avaliada: informal, formal.
pic_color	cor da foto avaliada: colorida, preto e branco.

Regressão Linear Simples

O fenômeno proposto pelo estudo é que professores com melhor aparência são avaliados de maneira mais favorável. Vamos criar um gráfico de dispersão para verificar se isso é verdade:

```
plot(evals$score ~ evals$bty_avg)
```

Antes de tirar conclusões sobre a tendência, compare o número de observações no banco de dados com o número de pontos no gráfico de dispersão. Há algo de errado?

Exercício 4 Refaça o gráfico de dispersão, mas agora utilize a função `jitter()` no eixo y ou x. (Utilize o comando `?jitter` para aprender mais a respeito.) O que estava errado no gráfico de dispersão inicial?

Exercício 5 Vamos verificar se a tendência aparente no gráfico é algo além de variação natural. Ajuste um modelo linear denominado `m_bty` para predizer a avaliação média de um professor a partir da média da avaliação da beleza e adicione a linha ao gráfico utilizando o comando `abline(m_bty)`. Escreva a equação do modelo linear e interprete a inclinação da reta. A média da avaliação da beleza é um preditor estatisticamente significante? Essa variável parecer ser um preditor com significância prática?

Exercício 6 Utilize gráficos de resíduos para avaliar se as condições para uma regressão utilizando mínimos quadrados são plausíveis. Utilize gráficos e comente cada uma deles (retome o Laboratório 7 para relembrar como criá-los).

Regressão Linear Múltipla

O conjunto de dados contém diversas variáveis sobre a avaliação de beleza do professor: avaliações individuais de cada um dos seis estudantes que foram convidados a avaliar a aparência física dos professores e a média dessas seis avaliações. Vamos dar uma olhada na relação entre uma dessas avaliações e a média da avaliação da beleza.

```
plot(evals$bty_avg ~ evals$bty_f1lower)
cor(evals$bty_avg, evals$bty_f1lower)
```

Como esperado, a relação é bem forte – afinal, a média das avaliações é calculada utilizando a avaliações individuais. Podemos dar uma olhada nas relações entre todas as variáveis relativas à beleza (colunas 13 a 19) utilizando o seguinte comando:

```
plot(evals[,13:19])
```

Essas variáveis são colineares (correlacionadas), e adicionar mais do que uma delas ao modelo não agregaria muito valor. Neste caso, com esses preditores com altos índices de correlação, é melhor utilizar a média das avaliações da beleza como o único representante dessas variáveis.

Para verificar se a beleza ainda é um preditor significante da avaliação docente depois que consideramos o sexo do professor, podemos adicionar um termo para o sexo no modelo.

```
m_bty_gen <- lm(score ~ bty_avg + gender, data = evals)
summary(m_bty_gen)
```

> **Exercício 7** Valores p e estimativas dos parâmetros só são confiáveis se as condições para a regressão são plausíveis. Verifique se as condições para esse modelo são plausíveis utilizando gráficos de diagnóstico.

> **Exercício 8** A variável `bty_avg` continua sendo um preditor significante de `score`? A adição da variável `gender` ao modelo alterou a estimativa do parâmetro de `bty_avg`?

Perceba que a estimativa para `gender` é agora denominada de `gendermale`. Você verá essa mudança de nome sempre que adicionar uma variável categorial ao modelo. O motivo é que o R recodifica `gender`, alterando seus valores iniciais `female` (feminino) e `male` (masculino) para uma variável indicativa denominada `gendermale` que tem o valor 0 para mulheres e o valor 1 para homens (tais variáveis são frequentemente chamada de variável "dummy" (falsa ou postiça)).

O resultado, para mulheres, é que o parâmetro estimado é multiplicado por zero, deixando a forma do intercepto e da inclinação similar à regressão simples.

$$\begin{aligned}\widehat{score} &= \hat{\beta}_0 + \hat{\beta}_1 \times bty\_avg + \hat{\beta}_2 \times (0) \\ &= \hat{\beta}_0 + \hat{\beta}_1 \times bty\_avg\end{aligned}$$

Podemos traçar essa linha e a linha correspondente aos homens com a seguinte função personalizada:

```
multiLines(m_bty_gen)
```

> **Exercício 9** Qual é a equação da linha correspondente aos homens? (*Dica:* Para os homens, a estimativa do parâmetro é multiplicada por 1.) Para dois professores que receberam a mesma avaliação de beleza, qual gênero tende a ter as avaliações mais altas?

A decisão de chamar a variável indicativa de `gendermale` ao invés de `genderfemale` não tem nenhum significado profundo. O R simplesmente codifica a categoria que vem em primeiro lugar na ordem alfabética

como um 0.[§]

> **Exercício 10** Crie um novo modelo denominado `m_bty_rank` removendo a variável `gender` e adicionando a variável `rank`. Como o R maneja variáveis categoriais que tem mais de dois níveis? Perceba que a variável `rank` tem três níveis: horista (teaching), assistente (tenure track) e titular (tenured).

A interpretação dos coeficientes na regressão múltipla é um pouco diferente da regressão simples. A estimativa do coeficiente da variável `bty_avg` reflete quanto mais um grupo de professores deve receber na avaliação da disciplina se sua avaliação de beleza é um ponto maior *mantendo todas as outras variáveis constantes*. Neste caso, isso significa considerar somente professores do mesmo nível com avaliações de `bty_avg` que estão separadas por um ponto.

A Busca pelo Melhor Modelo

Vamos começar com um modelo completo que prediz a avaliação docente com base no nível, etnia, sexo, língua da universidade onde obteve seu diploma, idade, proporção de alunos que completaram as avaliações, tamanho da turma, nível da disciplina, número de professores, número de créditos, média da avaliação da beleza, roupa e cor da foto avaliada.

> **Exercício 11** Qual variável você acha que teria o maior valor p neste modelo? Por quê? *Dica:* Pense em qual variável você esperaria não estar associada à avaliação docente.

Vamos rodar o modelo...

```
m_full <- lm(score ~ rank + ethnicity + gender + language + age + cls_perc_eval
             + cls_students + cls_level + cls_profs + cls_credits + bty_avg
             + pic_outfit + pic_color, data = evals)
summary(m_full)
```

> **Exercício 12** Verifique suas suspeitas do exercício anterior. Inclua os resultados do modelo em sua resposta.

> **Exercício 13** Interprete o coeficiente associado à variável etnia.

> **Exercício 14** Retire a variável com o maior valor p e reajuste o modelo. Os coeficientes e suas significâncias para as outras variáveis explicativas se alteraram? (Uma das coisas que torna a regressão múltipla interessante é que a estimativa dos coeficientes dependem das outras variáveis que são incluídas no modelo.) Se não, o que isso implica para questão de se a variável retirada era ou não colinear com outras variáveis explicativas?

> **Exercício 15** Utilizando seleção inversa e o valor p como critério de seleção, determine qual é o melhor modelo. Você não precisa mostrar todos os passos na sua resposta, apenas o resultado do modelo final. Também escreva a equação do modelo linear para predizer a avaliação docente com base no modelo final que você estabeleceu.

> **Exercício 16** Verifique se as condições para esse modelo são plausíveis utilizando gráficos de diagnóstico.

[§]Você pode mudar o nível de referência de uma variável categorial, que é o nível codificado como um 0, utilizando a função `relevel`. Utilize o comando `?relevel` para aprender mais a respeito.

Exercício 17 O artigo original descreve como os dados foram obtidos a partir de amostras de professores da Universidade do Texas em Austin e incluindo todas as disciplinas que eles ministraram. Considerando que cada linha representa uma disciplina, essa nova informação poderia ter algum impacto em alguma das condições para a regressão linear?

Exercício 18 Com base no seu modelo final, descreva as características de um professor e de uma disciplina da Universidade do Texas em Austin que estariam associadas com uma avaliação alta.

Exercício 19 Você se sentiria confiante em generalizar suas conclusões para todos os professores, de modo geral (e em qualquer universidade)? Por quê ou por que não?