

# Atividade 8

## Variáveis dummy e interações

Paulo Ricardo Seganfredo Campana

10 de outubro de 2023

Os dados da **Tabela B.3** do livro do *Montgomery* podem ser encontrados no pacote *MPV* do **R** como o nome *table.b3*. Estes dados consistem de 32 observações sobre o consumo de combustível de diferentes automóveis, com variáveis sobre o desempenho da quilometragem da gasolina e características físicas e/ou de performance.

Considerando estes dados e fixando o nível de significância em 5%, responda as questões abaixo:

- a) Estime um modelo de regressão linear que relaciona o consumo de combustível em milhas/galão,  $y$ , com o volume de deslocamento do motor (cilindrada),  $x_1$ , e o tipo de transmissão,  $x_{11}$ . O tipo de transmissão tem um efeito significativo no consumo de combustível?

```
data <- MPV::table.b3
fit1 <- lm(y ~ x1 + x11, data)
```

Tabela 1: Teste  $t$  para os coeficientes do primeiro modelo:  $y \sim x_1 + x_{11}$

Variável	Estimativa	Erro padrão	Estatística	p-valor
(Intercepto)	33.618	1.5396	21.837	$1.49 \times 10^{-19}$
$x_1$	-0.046	0.0087	-5.268	0.000012
$x_{11}$	-0.499	2.2282	-0.224	0.824

Não, o teste  $t$  para o coeficiente angular de  $x_{11}$  trás a informação de que o tipo da transmissão não é significativo para um modelo que mede o consumo de combustível

- b) Modifique o modelo ajustado na letra a) incluindo uma interação entre a cilindrada e o tipo de transmissão. Quais conclusões você pode fazer sobre o efeito do tipo de transmissão no consumo de combustível? Interprete os parâmetros destes modelos.

Tabela 2: Teste  $t$  para os coeficientes do segundo modelo:  $y \sim x_1 + x_{11} + x_1x_{11}$

Variável	Estimativa	Erro padrão	Estatística	p-valor
(Intercepto)	42.920	2.735	15.69	$2.10 \times 10^{-15}$
$x_1$	-0.117	0.020	-5.89	$2.49 \times 10^{-6}$
$x_{11}$	-13.464	3.844	-3.50	$1.57 \times 10^{-3}$
$x_1x_{11}$	0.082	0.021	3.84	$6.47 \times 10^{-4}$

Adicionando o termo de interação entre a cilindrada e o tipo de transmissão do veículo, temos um modelo onde todos os coeficientes são significativos, inclusive o efeito da transmissão por si só que não era significativo no modelo anterior. O novo modelo é expresso por  $\hat{y} = 42.92 - 0.117x_1$  para veículos de transmissão manual e por  $\hat{y} = 29.456 - 0.0351x_1$  para os de transmissão automática.

Para transmissão manual, a eficiência do combustível cai em média 0.117 milhas por galão para cada polegada cúbica no volume da cilindrada enquanto que os veículos automáticos essa diminuição é menos acentuada em 0.0351 milhas por galão, porém os veículos manuais possuem maior intercepto, o que tornam eles em maior parte mais eficientes:

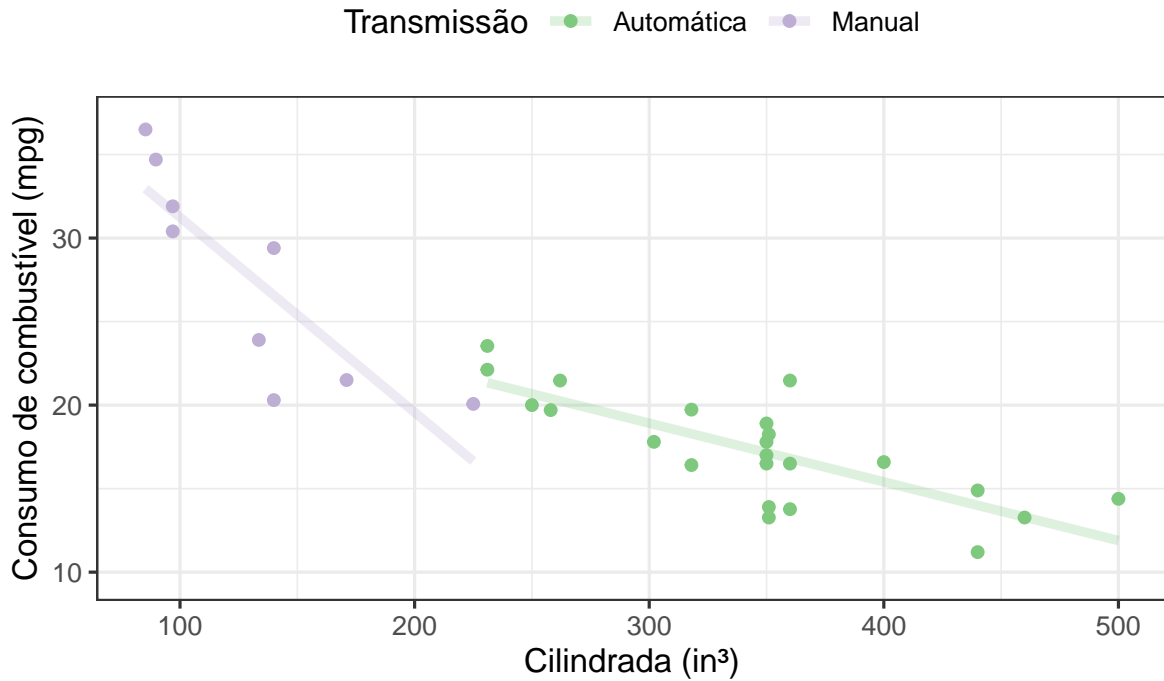


Figura 1: Relação entre o consumo de combustível e a cilindrada do motor para veículos de transmissão manual e automática

- c) Considere agora o modelo de regressão linear que relaciona o consumo de combustível,  $y$ , com o peso do veículo,  $x_{10}$  e o tipo de transmissão,  $x_{11}$ . O tipo de transmissão tem um efeito significativo no consumo de combustível neste modelo?

```
fit3 <- lm(y ~ x10 + x11, data)
```

Tabela 3: Teste  $t$  para os coeficientes do terceiro modelo:  $y \sim x_{10} + x_{11}$

Variável	Estimativa	Erro padrão	Estatística	p-valor
(Intercepto)	39.1919	2.5570	15.33	$1.95 \times 10^{-15}$
$x_{10}$	-0.0047	0.0009	-4.98	$2.72 \times 10^{-5}$
$x_{11}$	-2.6958	1.9805	-1.36	0.184

Como visto anteriormente, o tipo de transmissão do veículo por si só não é significante pelo teste  $t$  mesmo no modelo utilizando o peso do veículo ao invés da cilindrada.

- d) Modifique o modelo ajustado na letra c) incluindo uma interação entre o peso do veículo e o tipo de transmissão. Quais conclusões você pode fazer sobre o efeito do tipo de transmissão no consumo de combustível? Interprete os parâmetros destes modelos.

```
fit4 <- lm(y ~ x10 * x11, data)
```

Tabela 4: Teste  $t$  para os coeficientes do quarto modelo:  $y \sim x_{10} + x_{11} + x_{10}x_{11}$

Variável	Estimativa	Erro padrão	Estatística	p-valor
(Intercepto)	58.108	5.078	11.44	$4.53 \times 10^{-12}$
$x_{10}$	-0.013	0.002	-6.09	$1.44 \times 10^{-6}$
$x_{11}$	-26.725	6.107	-4.38	0.000152
$x_{10}x_{11}$	0.009	0.002	4.08	0.000342

Com a interação entre o peso e o tipo de transmissão do veículo sobre o consumo de combustível, o novo modelo torna mais significativo, com todas as variáveis contribuindo para medir o consumo de combustível. O novo ajuste é dado por  $\hat{y} = 58.108 - 0.0125x_1$  para veículos de transmissão manual e  $31.384 - 0.00348x_1$  para transmissão automática.

De maneira similar ao modelo utilizando a cilindrada, os veículos manuais tem maior intercepto porém o peso exerce influência maior na diminuição da eficiência do combustível, para cada libra adicional no peso do veículo ele perde em média 0.0125 milhas por galão, para veículos automáticos, essa perda é de 0.00348.

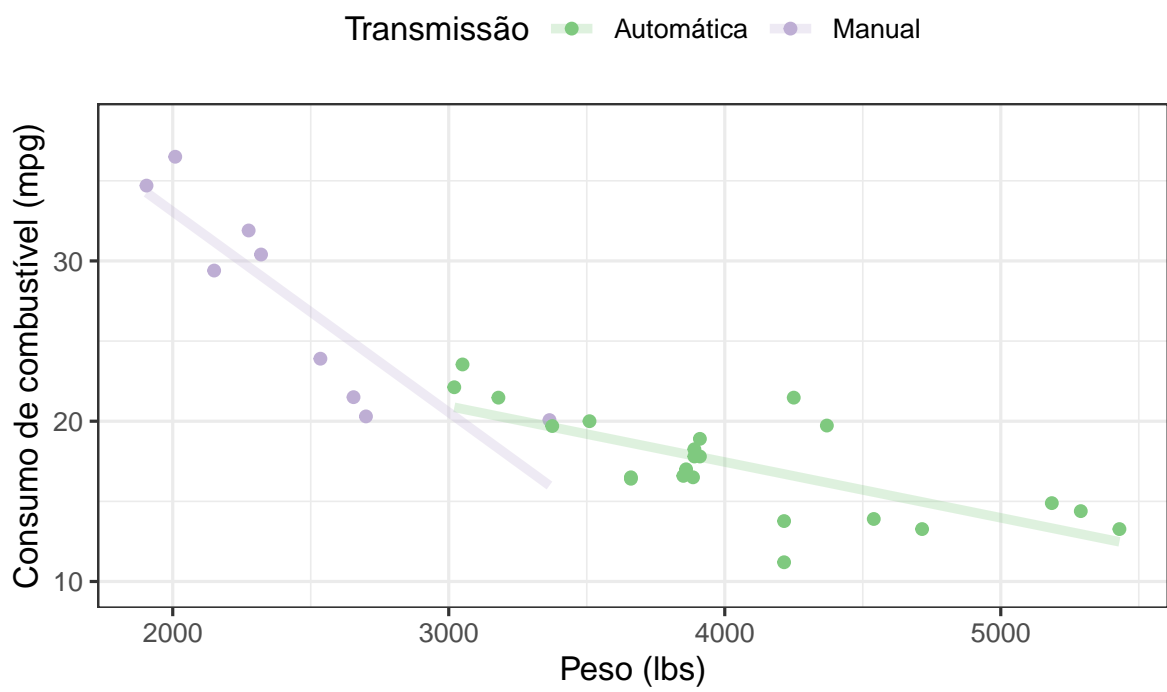


Figura 2: Relação entre o consumo de combustível e o peso do automóvel para veículos de transmissão manual e automática