



Regressão Logística

Felipe Figueiredo

Discussão da aula passada

Regressão Linear Múltipla

Regressão Logística

Aprofundamento

## Regressão Logística

Modelos com desfecho categórico binário

Felipe Figueiredo

Instituto Nacional de Traumatologia e Ortopedia

## Sumário



Regressão Logística

Felipe Figueiredo

Discussão da aula passada

Regressão Linear Múltipla

Regressão Logística

Aprofundamento

- 1 Discussão da aula passada
  - Discussão da aula passada
- 2 Regressão Linear Múltipla
  - Regressão Linear Múltipla
- 3 Regressão Logística
  - Regressão Logística
- 4 Aprofundamento
  - Aprofundamento



Regressão Logística

Felipe Figueiredo

Discussão da aula passada

Discussão da aula passada

Regressão Linear Múltipla

Regressão Logística

Aprofundamento

## Discussão da aula passada

Discussão da leitura obrigatória da aula passada



Regressão Logística

Felipe Figueiredo

Discussão da aula passada

Regressão Linear Múltipla

Regressão Linear Múltipla

Regressão Logística

Aprofundamento

JOURNAL OF WOMEN'S HEALTH  
Volume 15, Number 9, 2006  
© Mary Ann Liebert, Inc.

### The Association between Body Mass Index and Osteoporosis in Patients Referred for a Bone Mineral Density Examination

KOFI ASOMANING, M.B.Ch.B., M.S.,<sup>1</sup> ELIZABETH R. BERTONE-JOHNSON, Sc.D.,<sup>2</sup>  
PHILIP C. NASCA, Ph.D.,<sup>2</sup> FREDERICK HOOVEN, Ph.D.,<sup>3</sup>  
and PENELOPE S. PEKOW, Ph.D.<sup>2</sup>

Hoje vamos interpretar os resultados do abstract

## ABSTRACT

**Purpose:** Osteoporosis affects 4–6 million (13%–18%) postmenopausal white women in the United States. Most studies to date on risk factors for osteoporosis have considered body mass index (BMI) only as a possible confounder. In this study, we assess the direct relationship between BMI and osteoporosis.

**Methods:** We conducted a cross-sectional study among women aged 50–84 years referred by their physicians for a bone mineral density (BMD) examination at Baystate Medical Center between October 1998 and September 2000. BMI was determined prior to the BMD examination in the clinic. Information on other risk factors was obtained through a mailed questionnaire. Ordinal logistic regression was used to model the association between BMI and osteoporosis, controlling for confounding factors.

**Results:** BMI was inversely associated with BMD status. After adjustment for age, prior hormone replacement therapy (HRT) use, and other factors, odds ratios (OR) for low, high, and obese compared with moderate BMI women were 1.8 (95% CI 1.2–2.7), 0.46 (95% CI 0.29–0.71), and 0.22 (95% CI 0.14–0.36), respectively, with a significant linear trend ( $p < 0.0001$ ) across BMI categories. Evaluating BMI as a continuous variable, the odds of bone loss decreased 12% for each unit increase in BMI (OR = 0.88, 95% CI 0.85–0.91).

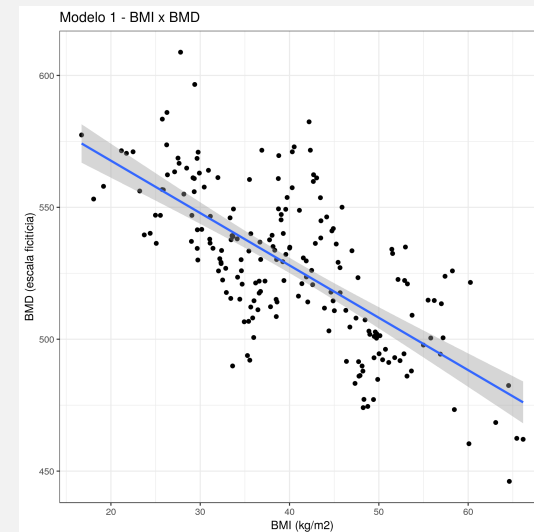
**Conclusions:** Women with low BMI are at increased risk of osteoporosis. The change in risk associated with a 1 unit change in BMI (~5–8 lb) is of greater magnitude than most other modifiable risk factors. To help reduce the risk of osteoporosis, patients should be advised to maintain a normal weight.

## Enunciado 1

Os pesquisadores querem investigar se o grupo etnicidade das participantes tem algum efeito detectável na associação entre a densidade mineral óssea (BMD) e o índice de massa corpórea (BMI).

Para isto selecionaram 100 mulheres brancas e 100 mulheres pardas, de meia idade. Mensuraram a BMD e calcularam o BMI delas.

## Modelo 1



## Quais são as variáveis?



Regressão Logística

Felipe Figueiredo

Discussão da aula passada

Regressão Linear Múltipla

Regressão Linear Múltipla

Regressão Logística

Aprofundamento

- Dependente: BMD (contínua)
- Independente: BMI (contínua)

Esta relação pode ser expressa como

$$\text{BMD} \sim \text{BMI}$$

## Componentes do modelo 1



Regressão Logística

Felipe Figueiredo

Discussão da aula passada

Regressão Linear Múltipla

Regressão Linear Múltipla

Regressão Logística

Aprofundamento

Versão simplificada (apenas variáveis)

$$\text{BMD} \sim \text{BMI}$$

Modelo completo

$$\text{BMD} = \beta_0 + \beta_1(\text{BMI}) + \varepsilon$$

Hipótese:  $\varepsilon$  é um erro aleatório <sup>1</sup> normalmente distribuído e centrado em zero – a incerteza que não pode ser controlada.

<sup>1</sup>residual – não é explicado pela relação entre as variáveis do modelo



Regressão Logística

Felipe Figueiredo

Discussão da aula passada

Regressão Linear Múltipla

Regressão Linear Múltipla

Regressão Logística

Aprofundamento

### Modelo 1

```
Residuals:
    Min       1Q   Median       3Q      Max
-50.770 -14.570  -2.449   14.357   58.778

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  607.3708     5.9857   101.47  <2e-16 ***
BMI          -1.9848     0.1444   -13.75  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 20.71 on 198 degrees of freedom
Multiple R-squared:  0.4884, Adjusted R-squared:  0.4858
F-statistic: 189 on 1 and 198 DF, p-value: < 2.2e-16
```

## Modelo 1



Regressão Logística

Felipe Figueiredo

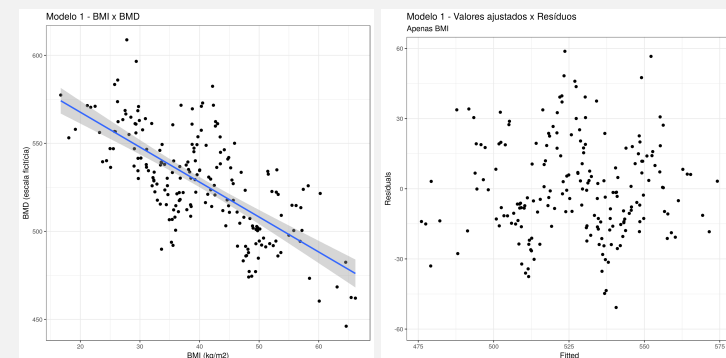
Discussão da aula passada

Regressão Linear Múltipla

Regressão Linear Múltipla

Regressão Logística

Aprofundamento

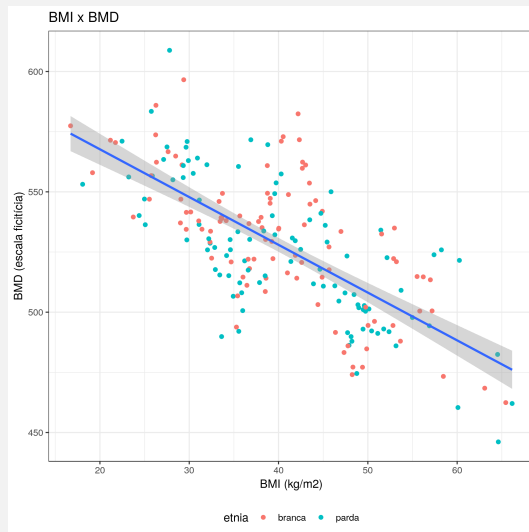


As participantes perdem, na média, 1.98 unidades de BMD para cada incremento unitário do BMI.

Este é o chamado resultado bruto. Agora vamos ajustá-lo com outros preditores.

Agora vamos ver se a etnia tem algum efeito

## Modelo 2



## Quais são as variáveis?

- Dependente: BMD (contínua)
- Independente: BMI (contínua)
- Independente: etnia (categórica – binária)

Esta relação pode ser expressa como

$$\text{BMD} \sim \text{BMI} + \text{etnia}$$



Versão simplificada (apenas variáveis)

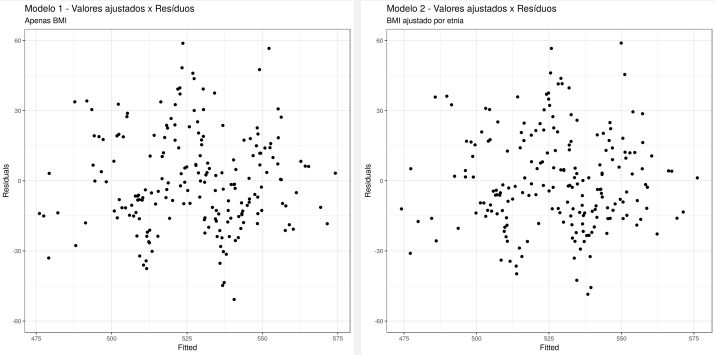
$BMD \sim BMI + etnia$

Modelo completo

$BMD = \beta_0 + \beta_1(BMI) + \beta_2(etnia) + \varepsilon$

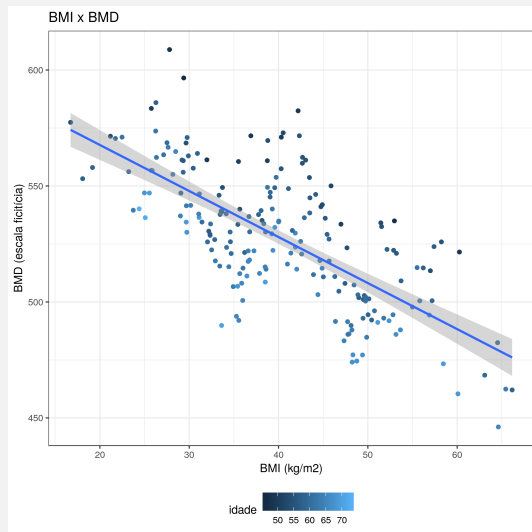
Modelo 2

Residuals:				
Min	1Q	Median	3Q	Max
-48.543	-14.299	-2.799	14.044	58.875
Coefficients:				
	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	609.236	6.097	99.917	<2e-16 ***
BMI	-1.977	0.144	-13.729	<2e-16 ***
etniaparda	-4.351	2.922	-1.489	0.138
---				
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
Residual standard error: 20.65 on 197 degrees of freedom				
Multiple R-squared: 0.4941, Adjusted R-squared: 0.489				
F-statistic: 96.21 on 2 and 197 DF, p-value: < 2.2e-16				



Que outra variável os pesquisadores deveriam ter investigado?

## Modelo 2.1 – idade



Regressão Logística

Felipe Figueiredo

Discussão da aula passada

Regressão Linear Múltipla

Regressão Linear Múltipla

Regressão Logística

Aprofundamento

## Quais são as variáveis?



Regressão Logística

Felipe Figueiredo

Discussão da aula passada

Regressão Linear Múltipla

Regressão Linear Múltipla

Regressão Logística

Aprofundamento

- Dependente: BMD (contínua)
- Independente: BMI (contínua)
- Independente: idade (contínua)

Esta relação pode ser expressa como

$$\text{BMD} \sim \text{BMI} + \text{idade}$$

## Componentes do modelo 2.1



Regressão Logística

Felipe Figueiredo

Discussão da aula passada

Regressão Linear Múltipla

Regressão Linear Múltipla

Regressão Logística

Aprofundamento

Versão simplificada (apenas variáveis)

$$\text{BMD} \sim \text{BMI} + \text{idade}$$

Modelo completo

$$\text{BMD} = \beta_0 + \beta_1(\text{BMI}) + \beta_2(\text{idade}) + \varepsilon$$



Regressão Logística

Felipe Figueiredo

Discussão da aula passada

Regressão Linear Múltipla

Regressão Linear Múltipla

Regressão Logística

Aprofundamento

### Modelo 2.1

Residuals:

	Min	1Q	Median	3Q	Max
	-30.484	-12.318	0.865	11.618	32.679

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	790.51012	12.57103	62.88	<2e-16 ***
BMI	-2.02336	0.09757	-20.74	<2e-16 ***
idade	-3.01619	0.19603	-15.39	<2e-16 ***

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 14 on 197 degrees of freedom  
Multiple R-squared: 0.7676, Adjusted R-squared: 0.7653  
F-statistic: 325.4 on 2 and 197 DF, p-value: < 2.2e-16

## Modelo 2.1



Regressão  
Logística

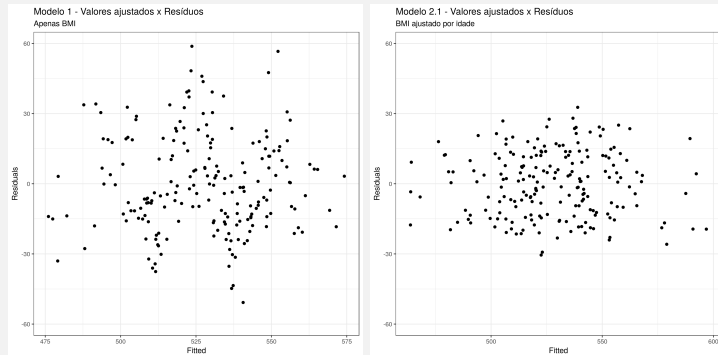
Felipe  
Figueiredo

Discussão da  
aula passada

Regressão  
Linear  
Múltipla  
Regressão Linear  
Múltipla

Regressão  
Logística

Aprofundamento



Que outra variável os pesquisadores deveriam ter investigado?

Regressão  
Logística

Felipe  
Figueiredo

Discussão da  
aula passada

Regressão  
Linear  
Múltipla  
Regressão Linear  
Múltipla

Regressão  
Logística

Aprofundamento



Regressão  
Logística

Felipe  
Figueiredo

Discussão da  
aula passada

Regressão  
Linear  
Múltipla  
Regressão Linear  
Múltipla

Regressão  
Logística

Aprofundamento

As participantes perdem, na média, 1.98 unidades de BMD para cada incremento unitário do BMI (resultado bruto).

Após ajustar pela idade, o resultado é 2.02.

## Modelo 2.2 – quantidade de hormônio



Regressão  
Logística

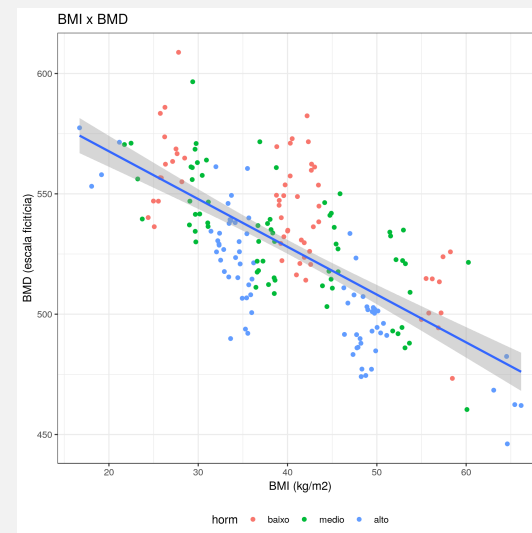
Felipe  
Figueiredo

Discussão da  
aula passada

Regressão  
Linear  
Múltipla  
Regressão Linear  
Múltipla

Regressão  
Logística

Aprofundamento



Regressão  
Logística

Felipe  
Figueiredo

Discussão da  
aula passada

Regressão  
Linear  
Múltipla  
Regressão Linear  
Múltipla

Regressão  
Logística

Aprofundamento

## Quais são as variáveis?

- Dependente: BMD (contínua)
- Independente: BMI (contínua)
- Independente: hormônio (categórica – 3 níveis)

Esta relação pode ser expressa como

$$\text{BMD} \sim \text{BMI} + \text{hormônio}$$



Regressão Logística

Felipe Figueiredo

Discussão da aula passada

Regressão Linear Múltipla

Regressão Linear Múltipla

Regressão Logística

Aprofundamento

## Componentes do modelo 2.2

Versão simplificada (apenas variáveis)

$$\text{BMD} \sim \text{BMI} + \text{hormônio}$$

Modelo completo

$$\text{BMD} = \beta_0 + \beta_1(\text{BMI}) + \beta_2(\text{hormônio}) + \varepsilon$$



Regressão Logística

Felipe Figueiredo

Discussão da aula passada

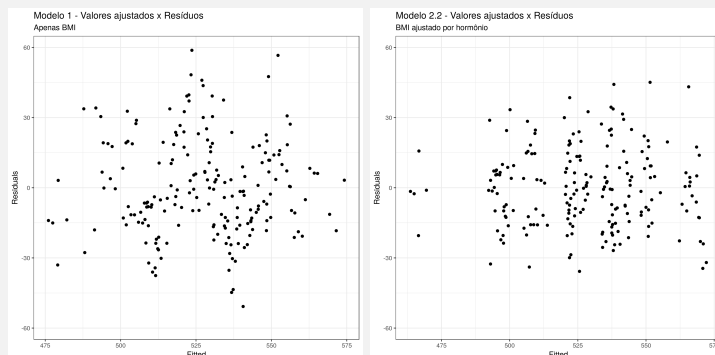
Regressão Linear Múltipla

Regressão Linear Múltipla

Regressão Logística

Aprofundamento

## Modelo 2.2



Regressão Logística

Felipe Figueiredo

Discussão da aula passada

Regressão Linear Múltipla

Regressão Linear Múltipla

Regressão Logística

Aprofundamento

## Modelo 2.2

```
Residuals:
    Min       1Q   Median       3Q      Max
-35.757 -12.239  -0.626   10.176   45.081

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  589.6897    5.2782  111.722  < 2e-16 ***
BMI          -1.9055     0.1187  -16.049  < 2e-16 ***
hormmedio    28.9182     2.9694   9.739  < 2e-16 ***
hormalto    17.7913     2.8720   6.195  3.37e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 16.98 on 196 degrees of freedom
Multiple R-squared:  0.6595, Adjusted R-squared:  0.6543
F-statistic: 126.6 on 3 and 196 DF, p-value: < 2.2e-16
```



Regressão Logística

Felipe Figueiredo

Discussão da aula passada

Regressão Linear Múltipla

Regressão Linear Múltipla

Regressão Logística

Aprofundamento



As participantes perdem, na média, 1.98 unidades de BMD para cada incremento unitário do BMI (resultado bruto).

Após ajustar pelo hormônio, o resultado é 1.90.

Agora um modelo maior (ajustando para todas as variáveis relevantes)

## Componentes do modelo 3

### Versão simplificada (apenas variáveis)

$$\text{BMD} \sim \text{BMI} + \text{idade} + \text{hormônio}$$

### Modelo completo

$$\text{BMD} = \beta_0 + \beta_1(\text{BMI}) + \beta_2(\text{idade}) + \beta_3(\text{hormônio}) + \varepsilon$$

### Modelo 3

Residuals:

	Min	1Q	Median	3Q	Max
	-18.3964	-4.1337	-0.5398	4.4640	25.6025

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	774.02125	6.48702	119.32	<2e-16 ***
BMI	-1.94720	0.04992	-39.00	<2e-16 ***
idade	-3.02783	0.10013	-30.24	<2e-16 ***
hormmedio	29.35318	1.24821	23.52	<2e-16 ***
hormalto	16.32422	1.20813	13.51	<2e-16 ***

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7.139 on 195 degrees of freedom  
Multiple R-squared: 0.9402, Adjusted R-squared: 0.9389  
F-statistic: 765.9 on 4 and 195 DF, p-value: < 2.2e-16

## Modelo 3



Regressão Logística

Felipe Figueiredo

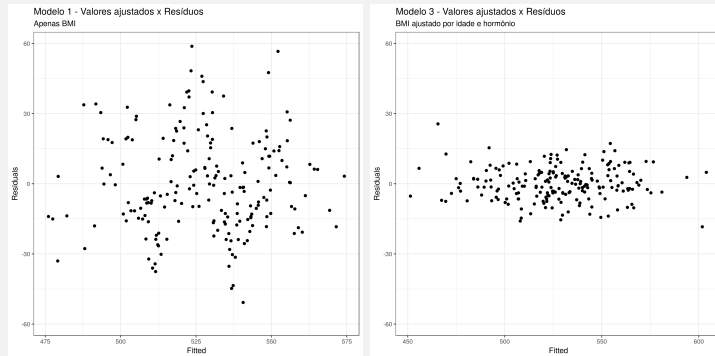
Discussão da aula passada

Regressão Linear Múltipla

Regressão Linear Múltipla

Regressão Logística

Aprofundamento



Regressão Logística

Felipe Figueiredo

Discussão da aula passada

Regressão Linear Múltipla

Regressão Linear Múltipla

Regressão Logística

Aprofundamento

As participantes perdem, na média, 1.98 unidades de BMD para cada incremento unitário do BMI (resultado bruto).

Após ajustar pela idade e pelo hormônio, o resultado é 1.95.

## Modelo 4



Regressão Logística

Felipe Figueiredo

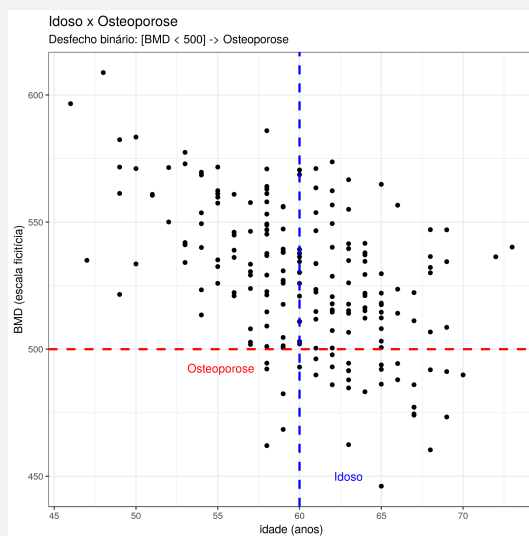
Discussão da aula passada

Regressão Linear Múltipla

Regressão Logística

Regressão Logística

Aprofundamento



## Quais são as variáveis?



Regressão Logística

Felipe Figueiredo

Discussão da aula passada

Regressão Linear Múltipla

Regressão Logística

Regressão Logística

Aprofundamento

- Dependente: Osteoporose (categórica – binária)
- Independente: Idoso (categórica – binária)

Esta relação pode ser expressa como

Osteoporose ~ Idoso

## Componentes do modelo 4



Regressão Logística

Felipe Figueiredo

Discussão da aula passada

Regressão Linear Múltipla

Regressão Logística

Regressão Logística

Aprofundamento

### Versão simplificada (apenas variáveis)

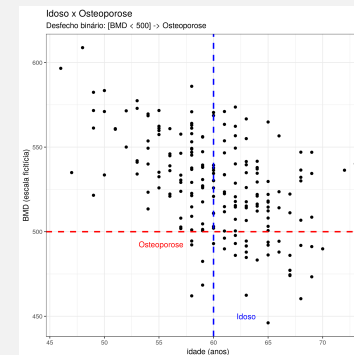
Osteoporose ~ Idoso

### Modelo completo

$$\text{Osteoporose} = \beta_0 + \beta_1(\text{Idoso}) + \varepsilon$$

### Tabela de contingência Idoso x Osteoporose

idoso	osteo	Sadio	Osteoporose
Nao Idoso	98		6
Idoso	68		28



Regressão Logística

Felipe Figueiredo

Discussão da aula passada

Regressão Linear Múltipla

Regressão Logística

Regressão Logística

Aprofundamento

### Modelo 4

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.8305 -0.8305 -0.3447 -0.3447  2.3886
```

#### Coefficients:

```
      Estimate Std. Error z value Pr(>|z|)
(Intercept)  -2.7932     0.4206  -6.642 3.10e-11 ***
idosoIdoso    1.9059     0.4767   3.998 6.39e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(Dispersion parameter for binomial family taken to be 1)

```
Null deviance: 182.35  on 199  degrees of freedom
Residual deviance: 161.78  on 198  degrees of freedom
AIC: 165.78
```

Number of Fisher Scoring iterations: 5

### log da OR de um idoso x osteoporose

$$\log(\text{OR}) = 1.9059$$



Regressão Logística

Felipe Figueiredo

Discussão da aula passada

Regressão Linear Múltipla

Regressão Logística

Regressão Logística

Aprofundamento

### Transformando o log da OR na OR

$$\log(\text{OR}) \approx 1.90...$$

... portanto...

$$e^{1.90} \approx 6.7$$

### Resultado

- (Idoso) OR: 6.73, IC: [2.64, 17.12]



Regressão Logística

Felipe Figueiredo

Discussão da aula passada

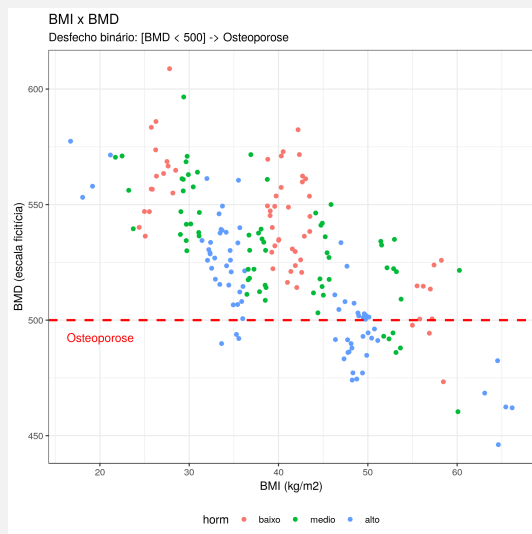
Regressão Linear Múltipla

Regressão Logística

Regressão Logística

Aprofundamento

## Modelo 5



Regressão Logística

Felipe Figueiredo

Discussão da aula passada

Regressão Linear Múltipla

Regressão Logística

Regressão Logística

Aprofundamento

## Quais são as variáveis?



Regressão Logística

Felipe Figueiredo

Discussão da aula passada

Regressão Linear Múltipla

Regressão Logística

Regressão Logística

Aprofundamento

- Dependente: Osteoporose (categórica – binária)
- Independente: BMI (contínua)
- Independente: idade (contínua)
- Independente: hormônio (categórica – 3 níveis)

Esta relação pode ser expressa como

Osteoporose ~ BMI + idade + hormônio

## Componentes do modelo 5



Regressão Logística

Felipe Figueiredo

Discussão da aula passada

Regressão Linear Múltipla

Regressão Logística

Regressão Logística

Aprofundamento

Versão simplificada (apenas variáveis)

Osteoporose ~ BMI + idade + hormônio

Modelo completo

Osteoporose =  $\beta_0 + \beta_1(\text{BMI}) + \beta_2(\text{idade}) + \beta_3(\text{hormônio}) + \varepsilon$

## Modelo 5

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.85523  -0.08866  -0.01263  -0.00068   2.05724

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  -77.4340    18.0040  -4.301 1.70e-05 ***
BMI             0.4554     0.1089   4.183 2.88e-05 ***
idade          0.8285     0.2086   3.972 7.13e-05 ***
hormmedio      1.9614     1.4716   1.333 0.182590
hormalto       5.5205     1.5101   3.656 0.000257 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 182.354  on 199  degrees of freedom
Residual deviance:  43.436  on 195  degrees of freedom
AIC: 53.436

Number of Fisher Scoring iterations: 9
```



Regressão Logística

Felipe Figueiredo

Discussão da aula passada

Regressão Linear Múltipla

Regressão Logística

Regressão Logística

Aprofundamento

## Resultado

- (BMI) **OR: 1.58, IC: [1.27, 1.95]**
- (idade) **OR: 2.29, IC: [1.52, 3.45]**
- (hormônio médio x baixo) **OR: 7.11, IC: [0.40, 127.19]**
- (hormônio alto x baixo) **OR: 249.76, IC: [12.94, 4818.87]**

## Interpretação

Após ajustar pela idade e pelo hormônio, as participantes tem chance aumentada de desenvolver osteoporose para cada incremento unitário do BMI.

## Aprofundamento

## Leitura obrigatória

- Capítulo 31
- Capítulo 32

## Exercícios selecionados

Não há.

## Leitura recomendada

Capítulo 25: seção teste t de uma razão (sobre o uso do logaritmo)