

# Modelos Lineares

## Variáveis Explicativas Qualitativas

Susana Faria

20 de Janeiro de 2017

# Notas Iniciais

- O uso destas notas como único material de estudo é fortemente desaconselhado.

# Variáveis Explicativas Qualitativas

- Vai apresentar-se uma extensão do modelo de regressão linear que permite introduzir variáveis qualitativas para explicar o comportamento da variável dependente.
- O tratamento desta questão exige a definição de um novo tipo de variáveis, designadas por **variáveis artificiais (variável dummy)**

## Exemplos de Variáveis Explicativas Qualitativas:

- compras (sim; não)
- género (masculino e feminino)
- tipo de firma (valores, ações, capital e comercial)
- regiões (nordeste, centro e sul)
- estação do ano (verão, outono, inverno e primavera).

# Variáveis Explicativas Qualitativas

**Exemplo:** Um economista deseja relacionar a velocidade com que um novo seguro é adoptado ( $Y$ ) com a dimensão da firma ( $X_1$ ) e o tipo de firma ( $X_2$ ). A variável resposta é medida em número de meses entre o tempo que a primeira empresa adoptou o seguro e o tempo que uma dada empresa adoptou. A variável  $X_1$  é dada em milhões de dólares. A segunda variável explicativa é qualitativa: firmas de capital e firmas comerciais.

- Para o exemplo da inovação de um seguro, onde a variável qualitativa tem duas classes, podemos definir duas variáveis indicadoras,  $X_2$  e  $X_3$  do seguinte modo:

$$X_2 = \begin{cases} 1 & \text{firma de capital;} \\ 0 & \text{outros casos.} \end{cases}$$

$$X_3 = \begin{cases} 1 & \text{firma de comércio;} \\ 0 & \text{outros casos.} \end{cases}$$

# Variáveis Explicativas Qualitativas

Para o exemplo, pensaríamos em usar um modelo de primeira ordem, dado por:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \epsilon_i$$

Esta abordagem intuitiva de designar variáveis indicadoras para cada classe da variável qualitativa, infelizmente, **origina grandes dificuldades computacionais**. Por exemplo,

- Em geral, vamos seguir o seguinte procedimento: **Uma variável qualitativa com c classes será representada por c-1 variáveis indicadoras, cada uma delas recebendo os valores 0 e 1.**

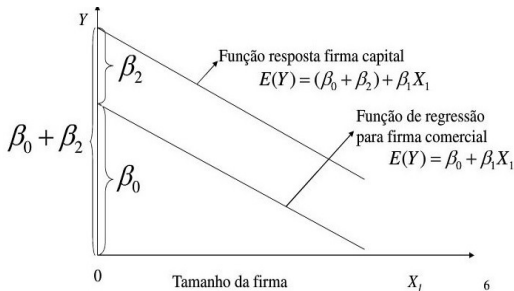
# Interpretação dos coeficientes de regressão

Considere que se retira a variável indicadora  $X_3$  do modelo de regressão, assim o modelo fica:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \epsilon_i$$

onde:  $X_{i1}$  = dimensão da firma e  $X_{i2} = 1$  se for firma de capital e  $X_{i2} = 0$  em outros casos.

Neste caso considera-se que a **variável qualitativa tem apenas impacto no termo independente e não na variável explicativa quantitativa.**



# Interpretação dos coeficientes de regressão

No exemplo tem-se :

- o valor esperado de  $Y$  é uma função linear do tamanho da firma ( $X_1$ ), com o mesmo coeficiente angular,  $\beta_1$ , para ambas as firmas;
- $\beta_0$  representa o termo independente para as firmas comerciais;
- $\beta_0 + \beta_2$  representa o termo independente para as firmas capital;
- $\beta_2$  representa a diferença entre os dois termos independentes;
- $\beta_2 > 0$  significa que, para idêntico valor  $x_i$ , o valor esperado de  $Y_i$  é superior nas firmas capital.
- $\beta_2 < 0$  significa que, para idêntico valor  $x_i$ , o valor esperado de  $Y_i$  é inferior nas firmas capital.

# Interpretação dos coeficientes de regressão - Exemplo

Dados do exemplo de seguros			
Firma	(1) Número de meses	(2) Tamanho da firma	(3) Tipo de firma
$i$	$Y_i$	$X_{ij}$	
1	17	151	Comercial
2	26	92	Comercial
3	21	175	Comercial
4	30	31	Comercial
5	22	104	Comercial
6	0	277	Comercial
7	12	210	Comercial
8	19	120	Comercial
9	4	290	Comercial
10	16	238	Comercial
11	28	164	Capitais
12	15	272	Capitais
13	11	295	Capitais
14	38	68	Capitais
15	31	85	Capitais
16	21	224	Capitais
17	20	166	Capitais
18	13	305	Capitais
19	30	124	Capitais
20	14	246	Capitais



# Modelo Efeito da Interação

Considere o seguinte modelo:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i1} X_{i2} + \epsilon_i$$

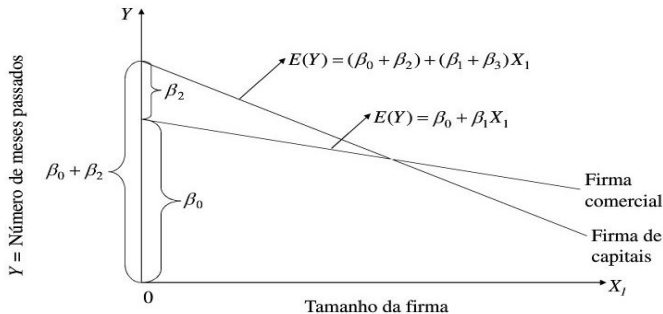
onde:  $X_{i1}$  = dimensão da firma e  $X_{i2} = 1$  se for firma de capital e  $X_{i2} = 0$  em outros casos.

Para a firma de comércio vem:  $E[Y_i] = \beta_0 + \beta_1 X_{i1}$

Para a firma de capital vem:  $E[Y_i] = \beta_0 + \beta_2 + (\beta_1 + \beta_3) X_{i1}$

# Modelo Efeito da Interação

As funções anteriores são apresentadas na seguinte figura:



Pela figura observamos que o efeito do tipo de firma depende do tamanho da mesma. Para firmas pequenas, as companhias de comércio adotam o seguro mais rapidamente, porém, para firmas maiores, as companhias de capitais adotam o seguro antes do que as de comércio. É o efeito da interação.

# Modelo Efeito da Interação

## Exercício 1:

- O modelo explicativo será o mesmo para os dois tipos de firma?
- Teste a existência de interação no modelo.

## Exercício 2:

Num modelo procurou-se verificar quais os factores que determinam os salários das pessoas com mais de um emprego. Os dados foram obtidos a partir de 318 trabalhadores e a regressão estimada foi a seguinte:

$$\hat{W}_a = 37.07 + 0.403W_p - 90.06D_1 + 75.51D_2 + 47.33D_3 + 2.26A$$

$W_a$  - salário do emprego adicional

$W_p$  - salário do emprego principal

$A$  - idade

$D_1 = 0$  se branco; 1 se não branco

$D_2 = 0$  proveniente zona urbana; 1 proveniente zona rural

$D_3 = 0$  Não tem Curso Superior; 1 Tem Curso Superior.

- Interprete os resultados dos parâmetros das variáveis binárias.
- Escreva um modelo de regressão para o salário médio adicional de um trabalhador não-branco, rural e com formação superior.

## Modelo com variável qualitativa com mais de 2 classes

Considere o modelo de regressão da durabilidade de uma ferramenta ( $Y$ ), sobre a velocidade ( $X_1$ ) e o modelo da ferramenta, onde, esta é uma variável qualitativa com 4 classes (M1, M2, M3, M4). Para trabalhar com esta variável precisamos definir as seguintes variáveis indicadoras:

$$X_2 = \begin{cases} 1 & \text{para o modelo 1} \\ 0 & \text{para outros casos.} \end{cases}$$

$$X_3 = \begin{cases} 1 & \text{para o modelo 2} \\ 0 & \text{para outros casos.} \end{cases}$$

$$X_4 = \begin{cases} 1 & \text{para o modelo 3} \\ 0 & \text{para outros casos.} \end{cases}$$

Modelo de primeira ordem é dado por:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \epsilon_i$$

Exercicio: Interpretar o modelo.

# Modelo com mais do que uma variável explicativa qualitativa

## Exemplo:

Considere os dados do ficheiro P124.txt, em que S-salário, X- experiência, E-Educação (1 - curso superior, 2 - bacharelato, 3- curso secundário), M - função de gestão (1= função de gestão, 0= não tem função de gestão). Pretende-se estimar o efeito das três variáveis explicativas(X,E,M) na variável salário.