

A) Formato de regressão linear

$$y = b_0 + b_1 x$$

$$b_0 = \frac{\sum XY - (\sum X) \cdot (\sum Y)}{n}$$
$$\frac{\sum X^2 - \frac{(\sum X)^2}{n}}$$

$$b_1 = \bar{y} - b_0 \bar{X}, \text{ logo } b_1 = \frac{\sum Y_i}{n} - b_0 \frac{\sum X_i}{n}$$

b) Existem três tipos de suposições mais necessárias para o modelo de regressão linear múltipla, que também enquadram o modelo de regressão linear simples.

I) O erro tem média = zero e variância = σ^2 desconhecida

II) Os erros não são correlacionados (são aleatórios)

III) Os erros tem distribuição normal

Sugiro essas quatro maneiras de checar as suposições. Primeiro deve-se analisar a linearidade das variáveis resposta/explicativa. Um bom jeito de se fazer isso é plotando um gráfico dos valores observados pelos valores gerados pela sua regressão. A forma desse gráfico deveria ser ~~mais~~ ou menos ~~uma~~ uma reta diagonal.

Para testar a independência dos erros é bom plotar um gráfico serial dos erros entre o valor observado e seu modelo. Uma vez

que esse gráfico apresenta uma correlação próxima ou igual a zero, assume-se a independência.

Para verificar se a variância do erro é constante, sugere-se plotar um gráfico dos erros pelos valores gerados pelo modelo. Nesse plot, espera-se que os erros não fiquem maiores à medida que o eixo x desloca-se para a direita.

Verificar a normalidade dos erros é o teste mais simples, caso o plot dos erros siga o formato de uma normal, a suposição está verificada.

c) Estimar que os erros seguem uma distribuição normal é parte do que deve ser feito. Além dessa estimativa, duas suposições adicionais devem ser consideradas:

- considerar uma amostra suficientemente grande
- supor que $e_i \sim N(0, \sigma^2)$

É o teste para a poder ser feito graças aos conceitos já estudados no TLC.

Aceitar ou rejeitar um teste de hipóteses numa regressão linear é saber se dada um certo nível de significância (α) que você escolheu, o seu modelo de ajuste corresponde em média ao comportamento da correlação entre sua variável resposta e a variável explicativa que você esperava.

d) Sim, é possível. Quando tratamos de uma regressão múltipla, o modelo passa a ser o seguinte:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$$

As suposições para o modelo são as mesmas que as da regressão linear simples, adicionando apenas mais uma suposição: as variáveis regressoras x_1, x_2, \dots, x_p assumem valores fixos.

Os testes de hipóteses que devem ser feitos são iguais as quantidades de variáveis explicativas, ou seja, deve ser feito um teste para cada um.