

7b. Exercícios de Regressão Múltipla

Exercício 7b.01

Uma estimativa da incerteza na previsão do modelo

Antes de passarmos para os exercícios de regressões múltiplas, vamos revisar a regressão simples mais uma vez. Uma forma de avaliar a incerteza na previsão de um modelo de regressão é fazer um gráfico com o intervalo de confiança de 95% da previsão do modelo. Para isso é preciso estimar esse intervalo da previsão e plotar um gráfico com os dados, o modelo e as curvas do intervalo de confiança. Para estimar o intervalo de confiança é preciso calcular o erro padrão da previsão do modelo e multiplicar pelo valor t de Student, para $n-2$ graus de liberdade. O erro padrão não é constante para todos os valores previstos, pode variar de acordo com o valor da variável preditora de acordo com a fórmula:

$$se_{\{\bar{y}\}} = \sqrt{s^2 \left[\frac{1}{n} + \frac{(x - \bar{x})^2}{SSX} \right]}$$

- onde:

- SSX : soma dos desvios quadráticos dos valores de x observados;
- s^2 : variância dos valores de y observados;
- n : número de observações;
- x : valor da variável preditora para o qual se quer estimar o erro padrão;
- \bar{x} : média dos valores de x observados.

Construa esse gráfico para avaliar a incerteza do modelo para a relação **peso do bebê (kg) ~ tempo de gestação (dias)** dos dados [Massa ao Nascer e Dados da mãe](#), usando a fórmula acima para estimar o intervalo de confiança de 95%.

Dicas:

- Antes de começar os cálculos, pense no gráfico que quer produzir e planeje as etapas necessárias para plotar o gráfico;
- Estime o intervalo de confiança da previsão do modelo para diversos valores da variável preditora, regularmente espalhados ao longo do eixo x do gráfico. Isso pode ser feito criando um vetor para representar a variável preditora como uma sequência de 100 valores compreendidos entre os valores máximo e mínimo da variável preditora observada.

Galileu estava Certo?

Partindo do tutorial [Ajuste de Polinômios](#), avalie se um polinômio de terceiro grau é um melhor modelo para descrever os dados do experimento de Galileu. A equação para este modelo é:

$$y = a + b \cdot x + c \cdot x^2 + d \cdot x^3$$

Massa de Recém-Nascidos

Utilize todas as variáveis e interações que fazem sentido para a construção do modelo cheio ¹⁾ no conjunto de dados [Massa ao Nascer e Dados da mãe](#), em seguida parta desse modelo para chegar ao melhor modelo de previsão da massa de bebês ao nascer.

Parta do arquivo acima e não do que foi usado no tutorial. Antes de começar, elimine as observações com dados faltantes, veja descrição do arquivo.

¹⁾

o modelo que contempla todas as variáveis e interações que permitem o entendimento do processo

From:

<http://ecor.ib.usp.br/> - **ecoR**

Permanent link:

http://ecor.ib.usp.br/doku.php?id=01_curso_atual:exercicios8&rev=1601754541



Last update: **2020/10/03 16:49**