

Universidade Federal da Fronteira Sul – UFFS

PROJETO DE EXTENSÃO – Curso avançado 2018.2

Software R: Capacitação em análise estatística de dados utilizando um software livre.

**Nome:**

**ATIVIDADE 3**

Encontro 3 – Modelos de Regressão Múltipla

**Orientações:** Nomeie o arquivo com seu nome completo e envie no respectivo link até o dia indicado.

**QUESTÃO 1)**

A Tabela 1 apresenta os dados para um estudo da relação da quantidade de gordura corporal (*Y*) a várias variáveis preditoras possíveis, baseados em uma amostra de 20 mulheres saudáveis de idade entre 25 a 34 anos. As variáveis preditoras possíveis são espessura da epiderme no tríceps (*X*1), circunferência da coxa (*X*2) e circunferência do braço (*X*3). A quantidade de gordura corporal na Tabela 1 para cada uma das 20 pessoas foi obtida por um procedimento incômodo e caro que necessitava da imersão da pessoa na água. Portanto seria muito útil que um modelo de regressão com algumas ou todas as variáveis preditoras pudessem prover estimativas confiáveis da quantidade de gordura corporal já que as medidas necessárias para as variáveis preditoras são fáceis de serem obtidas.

# Tabela 1: Dados

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sujeito  *i* | Espessura da epiderme no tríceps  (*Xi*1) | Circunferência da coxa  (*Xi*2) | Circunferência do braço  (*Xi*3) | Gordura Corporal  (*Yi*) |
| 1 | 19,5 | 43,1 | 29,1 | 11,9 |
| 2 | 24,7 | 49,8 | 28,2 | 22,8 |
| 3 | 30,7 | 51,9 | 37,0 | 18,7 |
| 4 | 29,8 | 54,3 | 31,1 | 20,1 |
| 5 | 19,1 | 42,2 | 30,9 | 12,9 |
| 6 | 25,6 | 53,9 | 23,7 | 21,7 |
| 7 | 31,4 | 58,5 | 27,6 | 27,1 |
| 8 | 27,9 | 52,1 | 30,6 | 25,4 |
| 9 | 22,1 | 49,9 | 23,2 | 21,3 |
| 10 | 25,5 | 53,5 | 24,8 | 19,3 |
| 11 | 31,1 | 56,6 | 30,0 | 25,4 |
| 12 | 30,4 | 56,7 | 28,3 | 27,2 |
| 13 | 18,7 | 46,5 | 23,0 | 11,7 |
| 14 | 19,7 | 44,2 | 28,6 | 17,8 |
| 15 | 14,6 | 42,7 | 21,3 | 12,8 |
| 16 | 29,5 | 54,4 | 30,1 | 23,9 |
| 17 | 27,7 | 55,3 | 25,7 | 22,6 |
| 18 | 30,2 | 58,6 | 24,6 | 25,4 |
| 19 | 22,7 | 48,2 | 27,1 | 14,8 |
| 20 | 25,2 | 51,0 | 27,5 | 21,1 |

1. Encontre o coeficiente de correlação entre a gordura corporal e cada variável preditora. Interprete.
2. Ajuste um modelo de regressão múltipla com os preditores disponíveis no banco. Interprete.
3. Teste a significância dos coeficientes. Quais coeficientes são significativos? Qual(is) coeficientes permanecem para o modelo final?
4. Teste a significância da equação ajustada final. Interprete.
5. Encontre o coeficiente de determinação para a equação ajustada final. Interprete.
6. Faça análise completa dos resíduos. Interprete.
7. Existe resíduos outliers ou influentes?
8. Encontre um valor predito para determinados (você define) valores de cada variável preditora.
9. Ajuste um novo modelo de regressão múltipla usando uma variável dummy para a variável circunferência da coxa, sendo 0 para “<=50” e 1 para “>50”. Interprete todos os resultados, incluindo significância dos coeficientes, da equação, coeficiente de determinação, análise de resíduos.
10. Apresente os resultados na forma de um resumo de evento científico (título, autores, introdução, objetivo, material e métodos, resultados e conclusão), com no máximo 300 palavras (sem incluir o título e autores).