**Lista 09 – Análise de Regressão – Estatística 2023**

1. Para cada item abaixo, responda verdadeiro ou falso e justifique sua escolha.
2. O coeficiente de correlação tem o mesmo sinal do coeficiente angular de uma regressão linear simples;
3. Um alto coeficiente de determinação indica que as variáveis dependente e independente tem necessariamente uma relação linear forte;
4. Numa regressão linear simples, o teste F e o teste t para o coeficiente angular são correspondentes;
5. As melhores estimativas de *Y* numa regressão linear simples são obtidas próximas de *X* = 0.
6. A variância do erro ε*i* deve ser a mesma para todos os valores de *Xi*.
7. Com base nos dados abaixo, estime a equação que representa a relação entre a variável dependente *Y* e as variáveis independentes *X*1, *X*2. e *X*3. Verifique se as relações entre a variável dependente e as independentes são lineares e proponha transformações caso necessário. Através dos testes F e t, verifique se a regressão é realmente válida. Caso seja necessário, simplifique o modelo descartando as variáveis independentes que não contribuem significativamente (adote o nível de significância de 5%). De quanto a variação total de *Y* é explicada pelo modelo final ajustado? Aplique os testes de Shapiro-Wilk, Brush-Pagan e apresente os gráficos de dispersão dos erros padronizados para verificar a adequabilidade do modelo proposto. Comente os resultados.

|  |  |  |  |
| --- | --- | --- | --- |
| *Y* | *X*1 | *X*2 | *X*3 |
| 73,27 | 9,25 | 21,24 | 56,56 |
| 73,43 | 8,40 | 21,13 | 53,92 |
| 78,3 | 9,57 | 22,42 | 59,31 |
| 78,98 | 9,17 | 23,33 | 58,67 |
| 83,41 | 11,13 | 21,15 | 54,32 |
| 76,73 | 9,67 | 22,18 | 55,06 |
| 80,40 | 10,54 | 20,24 | 54,49 |
| 74,31 | 9,18 | 22,19 | 57,43 |
| 75,30 | 8,52 | 22,99 | 56,87 |
| 76,58 | 9,35 | 21,56 | 53,83 |
| 84,14 | 11,30 | 23,79 | 61,38 |
| 72,19 | 9,11 | 20,73 | 52,13 |
| 78,04 | 10,66 | 21,48 | 56,71 |
| 80,90 | 10,74 | 22,77 | 60,42 |
| 75,80 | 9,06 | 23,38 | 60,72 |

1. Para avaliar a relação linear entre duas variáveis *X* e *Y*, 10 pontos foram coletados sobre dois alvos distintos *A* e *B*. Verifique se os alvos possuem a mesma relação e, caso afirmativo, qual o modelo que representa a relação entre estas variáveis? Dica: utilize uma variável indicadora para gerar um único modelo para os alvos a fim de testar a hipótese de que ambos possuem a mesma relação.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Alvo *A* | |  | Alvo *B* | |
| *Y* | *X* |  | *Y* | *X* |
| 43,4 | 7,9 |  | 45,1 | 8,7 |
| 56,7 | 11,7 |  | 55,5 | 12,0 |
| 52,6 | 12,7 |  | 42,5 | 8,8 |
| 59,6 | 14,3 |  | 61,2 | 15,2 |
| 40,1 | 6,8 |  | 47,8 | 11,7 |
| 63,0 | 14,9 |  | 57,8 | 14,8 |
| 49,7 | 10,8 |  | 41,8 | 7,8 |
| 43,8 | 7,8 |  | 36,1 | 6,2 |
| 46,9 | 9,8 |  | 54,9 | 13,5 |
| 50,8 | 11,5 |  | 33,2 | 5,8 |

1. De acordo com os dados abaixo, determine o melhor modelo que relaciona a variável dependente *Y* com as variáveis independentes *X*1, *X*2 e *X*3. Observe que nem todas as variáveis independentes necessitam ser utilizadas no modelo. Faça a análise dos resíduos a fim de constatar se o modelo encontrado parece adequado e proponha transformações nas variáveis caso seja necessário.

|  |  |  |  |
| --- | --- | --- | --- |
| *Y* | *X*1 | *X*2 | *X*3 |
| 104,43 | 0,15 | 37,22 | 0,66 |
| 105,35 | 11,71 | 38,45 | 0,55 |
| 103,28 | 13,95 | 38,52 | 0,44 |
| 106,67 | 2,82 | 34,27 | 0,77 |
| 97,43 | 0,21 | 35,87 | -0,82 |
| 104,33 | 4,07 | 32,62 | -0,22 |
| 107,88 | 5,98 | 30,86 | 0,60 |
| 99,07 | 14,49 | 38,43 | -0,18 |
| 97,21 | 14,69 | 31,06 | -0,95 |
| 95,39 | 16,05 | 33,51 | -0,20 |
| 100,78 | 1,92 | 34,69 | -0,21 |
| 106,40 | 6,89 | 30,48 | -0,01 |
| 90,65 | 19,37 | 38,73 | -0,75 |
| 106,03 | 9,22 | 35,61 | -0,38 |
| 101,71 | 4,52 | 31,66 | -0,51 |
| 96,05 | 17,04 | 31,28 | 0,22 |
| 101,39 | 10,07 | 33,99 | -0,93 |
| 93,27 | 16,93 | 32,55 | -0,46 |
| 107,64 | 3,76 | 36,60 | 0,24 |
| 99,33 | 11,52 | 36,80 | -0,92 |