Exercício 3 – Lohrs – p.165

Pesquisadores querem estimar a idade de árvores de uma certa espécie, mas isso exige contar os anéis da árvore, o que exige seu corte. Logo, o pesquisador deseja estimar a idade média de árvores a partir das alturas. Em uma floresta com 1132 árvores, sabe-se que a altura média populacional das árvores é igual a 10,3. Uma amostra de 20 árvores (amostra sem reposição) foi estudada e os dados estão na tabela a seguir.

|  |  |  |  |
| --- | --- | --- | --- |
|   |   |   | Resíduos ao quadrado |
|  | Diameter (x) | Idade (y) | Razão | Regressão |
| Tree |   |   | (y-rx)^2 | (y-yh)^2 |
| 1 | 12 | 125 | 144.8049 | 201.2958585 |
| 2 | 11.4 | 119 | 125.0331 | 164.8162718 |
| 3 | 7.9 | 83 | 52.0377 | 35.57235214 |
| 4 | 9 | 85 | 315.9549 | 304.1147527 |
| 5 | 10.5 | 99 | 436.99 | 475.823621 |
| 6 | 7.9 | 117 | 717.505 | 786.0029299 |
| 7 | 7.3 | 69 | 206.2682 | 159.1245512 |
| 8 | 10.2 | 133 | 272.961 | 251.5877342 |
| 9 | 11.7 | 154 | 415.8477 | 341.7699798 |
| 10 | 11.3 | 168 | 1517.892 | 1397.779348 |
| 11 | 5.7 | 61 | 16.73554 | 1.030217254 |
| 12 | 8 | 80 | 128.9511 | 103.8202567 |
| 13 | 10.3 | 114 | 13.1074 | 19.03968687 |
| 14 | 12 | 147 | 99.33127 | 61.02926316 |
| 15 | 9.2 | 122 | 286.9971 | 292.7925169 |
| 16 | 8.5 | 106 | 79.82724 | 93.81754219 |
| 17 | 7 | 82 | 4.259253 | 16.48718171 |
| 18 | 10.7 | 88 | 1168.833 | 1243.501344 |
| 19 | 9.3 | 97 | 84.65761 | 83.0610828 |
| 20 | 8.2 | 99 | 28.73439 | 40.46034524 |
| média | 9.405 | 107.4 |  |  |
| soma |  |  | 6116.728 | 6072.926837 |
| dp\_amostral |   | 28.6621 |   |   |

1. Faça um diagrama de dispersão com os dados da tabela.

Há relação linear entre as variáveis. Quanto maior a altura, em média maior é a árvore.

1. Estime a idade média das árvores usando o estimador razão e calcule seu erro padrão.

N= 1132; n=20

A razão estimada é 11,42=(107,4/9,405).

A idade média estimada será 117,62 (=10,3\*razão estimada).

A variância estimada é 15,81.

O erro padrão é 3,98.

1. Avalie se o estimador razão pode ser utilizado.

O estimador razão pode ser usado se há relação linear entre y e x e ao ajustarmos um modelo de regressão linear concluímos que o intercepto não é significativo.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | *Estimativa* | *Erro padrão* | *t* | *P-value* | *Inf. 95%* | *Sup. 95%* |
| Intercept0 | -7.808 | 22.056 | -0.354 | 0.727 | -54.146 | 38.530 |
| Diameter | 12.250 | 2.304 | 5.316 | 0.000 | 7.409 | 17.090 |

O intercepto não é significativo (p=0,727).

Verifiquem se há dúvidas sobre como fazer esse teste. Eu passarei para vocês só a estimativa e o erro padrão (ou variância) e vocês quem farão o teste em prova. Tanto usando região crítica quanto valor p.

Saibam também como estimar a variância (sigma2) dos erros do modelo, pois isso é importante para predição e previsão.

1. Estimem a altura média usando o estimador regressão e calculem o erro padrão.

A altura média estimada é 118,36 (= intercepto + inclinação \* altura média pop = -7,8+12,25\*10,3).

A variância usando o resultado em Lohrs é 20,75 com erro padrão 4,56.

1. Estimem a altura média só usando os dados da altura com AASs e calculem o erro padrão.

A estimativa é 107,4 (=média amostral de y das 20 árvores).

A variância é 40,35 com erro padrão 6,35.

1. Comparem os estimadores.

Os estimadores razão e regressão são os melhores por terem variâncias bem menores.

O estimador regressão tem variância um pouquinho menor, mas é não viesado, assumindo-se que o modelo de regressão linear está bem ajustado (erros com média zero, como parece ser o caso).