

Condições:Correlação

MODERADA OU FORTE

A solid blue horizontal bar spanning the width of the slide, located at the bottom.

Coeficiente de Determinação (R^2)



> 0,7 : ótimo



Entre eles: ?

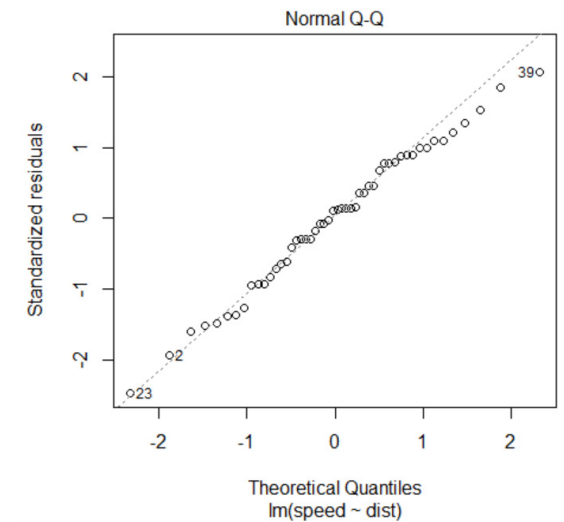
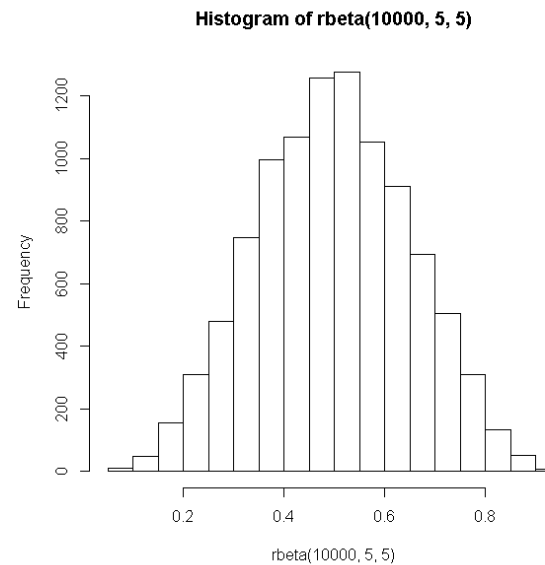


0 e 0,3 : ruins

Residuais padronizados

Residuais próximos de uma distribuição normal:

- Histograma
- Diagrama de normalidade
- Teste de Shapiro-Wilk



Simple e Múltipla

Simple

- Uma variável explanatória para prever uma variável dependente
- $Y \sim X$

Múltipla

- Duas ou mais variáveis explanatórias para prever uma variável dependente
- $Y \sim X_1 + X_2 + X_n$

Analisar Cada x Com y

- Analisar cada variável independente com y individualmente
- Gerar gráficos de dispersão individuais
- Buscar redundâncias (mesmos efeitos de x sobre y)

Coeficiente de Determinação (R^2)

- Lembrando que R^2 é o percentual de variação da variável de resposta que é explicada pelo modelo
- Quando se colocam mais variáveis no modelo, a tendência é que R^2 aumente, mesmo que a adição da variável não aumente a precisão do modelo
- Para isso, utiliza-se R^2 ajustado, que ajusta a variação do modelo de acordo com o número de variáveis independentes que é incluída no modelo
- R^2 ajustado vai ser sempre menor que R^2

Colinearidade e Parcimônia

- Colinearidade: duas variáveis independentes que são correlacionadas
- Incluir variáveis independentes colineares pode prejudicar o modelo, criando previsões não confiáveis
- Parcimônia: não colocar variáveis que não melhorem o modelo em nada: criar modelos parcimoniosos

Requisitos Básicos

1. Linearidade entre a variável dependente e as variáveis independentes
2. Pouca ou nenhuma Colinearidade

Resíduos

- Próximos a distribuição normal
- Variância constante em relação a linha de melhor ajuste
- Independentes (sem padrão)

Correlograma

