Probabilidade e estatística - Aula 22

Regressão linear simples: Aula 2 - Testes de hipóteses e intervalos de confiança em regressão linear simples

Dr. Giannini Italino Alves Vieira

Universidade Federal do Ceará - Campus de Crateús

2024

1. Testes de hipóteses na regressão linear simples

2. IC para o intercepto e a inclinação

Regressão linear simples

- Recorde que na aula passada vimos o modelo de regressão linear simples.
- Ou seja, vimos que o modelo considera que cada observação Y possa ser descrita pelo modelo

$$Y = \beta_0 + \beta_1 x + \epsilon$$

  em que β_0 é chamado de o intercepto, β_1 é chamado de inclinação e ϵ é um erro aleatório com média zero e variância σ^2 (desconhecida).
- Vimos que a estimação de β_0 e β_1 pode ser feita por meio do método de mínimos quadrados, ou seja, minimizando a soma dos quadrados dos desvios.

Regressão linear simples

- Recorde que na aula passada vimos o modelo de regressão linear simples.
- Ou seja, vimos que o modelo considera que cada observação Y possa ser descrita pelo modelo

$$Y = \beta_0 + \beta_1 x + \epsilon$$

em que β_0 é chamado de o intercepto, β_1 é chamado de inclinação e ϵ é um erro aleatório com média zero e variância σ^2 (desconhecida).

- Vimos que a estimação de β_0 e β_1 pode ser feita por meio do método de mínimos quadrados, ou seja, minimizando a soma dos quadrados dos desvios.

Regressão linear simples

- Recorde que na aula passada vimos o modelo de regressão linear simples.
- Ou seja, vimos que o modelo considera que cada observação Y possa ser descrita pelo modelo

$$Y = \beta_0 + \beta_1 x + \epsilon$$

em que β_0 é chamado de o intercepto, β_1 é chamado de inclinação e ϵ é um erro aleatório com média zero e variância σ^2 (desconhecida).

- Vimos que a estimação de β_0 e β_1 pode ser feita por meio do método de mínimos quadrados, ou seja, minimizando a soma dos quadrados dos desvios.

Regressão linear simples

- Recorde que na aula passada vimos o modelo de regressão linear simples.
- Ou seja, vimos que o modelo considera que cada observação Y possa ser descrita pelo modelo

$$Y = \beta_0 + \beta_1 x + \epsilon$$

em que β_0 é chamado de o intercepto, β_1 é chamado de inclinação e ϵ é um erro aleatório com média zero e variância σ^2 (desconhecida).

- Vimos que a estimação de β_0 e β_1 pode ser feita por meio do método de mínimos quadrados, ou seja, minimizando a soma dos quadrados dos desvios.

Regressão linear simples

- Recorde que na aula passada vimos o modelo de regressão linear simples.
- Ou seja, vimos que o modelo considera que cada observação Y possa ser descrita pelo modelo

$$Y = \beta_0 + \beta_1 x + \epsilon$$

  em que β_0 é chamado de o intercepto, β_1 é chamado de inclinação e ϵ é um erro aleatório com média zero e variância σ^2 (desconhecida).
- Vimos que a estimação de β_0 e β_1 pode ser feita por meio do método de mínimos quadrados, ou seja, minimizando a soma dos quadrados dos desvios.

Estimadores de β_0 e β_1 - Método dos mínimos Quadrados

- Vimos que as estimativas de mínimos quadrados do intercepto e da inclinação no modelo de regressão linear simples são

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

e

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$$

em que $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ e $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

- Dessa forma, a linha de regressão ajustada (modelo ajustado ou estimado) é consequentemente

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

Estimadores de β_0 e β_1 - Método dos mínimos Quadrados

- Vimos que as estimativas de mínimos quadrados do intercepto e da inclinação no modelo de regressão linear simples são

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

e

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$$

em que $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ e $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

- Dessa forma, a linha de regressão ajustada (modelo ajustado ou estimado) é consequentemente

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

Estimadores de β_0 e β_1 - Método dos mínimos Quadrados

- Vimos que as estimativas de mínimos quadrados do intercepto e da inclinação no modelo de regressão linear simples são

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

e

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$$

em que $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ e $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

- Dessa forma, a linha de regressão ajustada (modelo ajustado ou estimado) é consequentemente

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

Estimadores de β_0 e β_1 - Método dos mínimos Quadrados

- Vimos que as estimativas de mínimos quadrados do intercepto e da inclinação no modelo de regressão linear simples são

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

e

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$$

em que $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ e $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

- Dessa forma, a linha de regressão ajustada (modelo ajustado ou estimado) é consequentemente

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

Estimadores de β_0 e β_1 - Método dos mínimos Quadrados

- Vimos que as estimativas de mínimos quadrados do intercepto e da inclinação no modelo de regressão linear simples são

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}$$

e

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$$

em que $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ e $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

- Dessa forma, a linha de regressão ajustada (modelo ajustado ou estimado) é consequentemente

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

Estimadores de β_0 e β_1 - Método dos mínimos Quadrados

- Vimos que as estimativas de mínimos quadrados do intercepto e da inclinação no modelo de regressão linear simples são

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

e

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$$

em que $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ e $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

- Dessa forma, a linha de regressão ajustada (modelo ajustado ou estimado) é consequentemente

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

Um outro parâmetro

- Note que em regressão linear simples há ainda um outro parâmetro desconhecido, que é a variância do termo de erro ϵ, ou seja, σ^2.
- Para obter uma estimativa para σ^2 usamos os resíduos $e_i = y_i - \hat{y}_i$ (note que o resíduo descreve o erro no ajuste do modelo para a i-ésima observação y_i).
- Temos que a soma dos quadrados dos resíduos é

$$SQ_E = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

- Pode-se mostrar que $E(SQ_E) = (n-2)\sigma^2$. Logo, um estimador não viesado para σ^2 é

$$\hat{\sigma}^2 = \frac{SQ_E}{n-2}$$

Um outro parâmetro

- Note que em regressão linear simples há ainda um outro parâmetro desconhecido, que é a variância do termo de erro ϵ, ou seja, σ^2.
- Para obter uma estimativa para σ^2 usamos os resíduos $e_i = y_i - \hat{y}_i$ (note que o resíduo descreve o erro no ajuste do modelo para a i-ésima observação y_i).
- Temos que a soma dos quadrados dos resíduos é

$$SQ_E = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

- Pode-se mostrar que $E(SQ_E) = (n-2)\sigma^2$. Logo, um estimador não viesado para σ^2 é

$$\hat{\sigma}^2 = \frac{SQ_E}{n-2}$$

Um outro parâmetro

- Note que em regressão linear simples há ainda um outro parâmetro desconhecido, que é a variância do termo de erro ϵ, ou seja, σ^2.
- Para obter uma estimativa para σ^2 usamos os resíduos $e_i = y_i - \hat{y}_i$ (note que o resíduo descreve o erro no ajuste do modelo para a i-ésima observação y_i).
- Temos que a soma dos quadrados dos resíduos é

$$SQ_E = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

- Pode-se mostrar que $E(SQ_E) = (n-2)\sigma^2$. Logo, um estimador não viesado para σ^2 é

$$\hat{\sigma}^2 = \frac{SQ_E}{n-2}$$

Um outro parâmetro

- Note que em regressão linear simples há ainda um outro parâmetro desconhecido, que é a variância do termo de erro ϵ, ou seja, σ^2.
- Para obter uma estimativa para σ^2 usamos os resíduos $e_i = y_i - \hat{y}_i$ (note que o resíduo descreve o erro no ajuste do modelo para a i-ésima observação y_i).
- Temos que a soma dos quadrados dos resíduos é

$$SQ_E = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

- Pode-se mostrar que $E(SQ_E) = (n-2)\sigma^2$. Logo, um estimador não viesado para σ^2 é

$$\hat{\sigma}^2 = \frac{SQ_E}{n-2}$$

Um outro parâmetro

- Note que em regressão linear simples há ainda um outro parâmetro desconhecido, que é a variância do termo de erro ϵ, ou seja, σ^2.
- Para obter uma estimativa para σ^2 usamos os resíduos $e_i = y_i - \hat{y}_i$ (note que o resíduo descreve o erro no ajuste do modelo para a i-ésima observação y_i).
- Temos que a soma dos quadrados dos resíduos é

$$SQ_E = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

- Pode-se mostrar que $E(SQ_E) = (n-2)\sigma^2$. Logo, um estimador não viesado para σ^2 é

$$\hat{\sigma}^2 = \frac{SQ_E}{n-2}$$

Um outro parâmetro

- Note que em regressão linear simples há ainda um outro parâmetro desconhecido, que é a variância do termo de erro ϵ, ou seja, σ^2.
- Para obter uma estimativa para σ^2 usamos os resíduos $e_i = y_i - \hat{y}_i$ (note que o resíduo descreve o erro no ajuste do modelo para a i-ésima observação y_i).
- Temos que a soma dos quadrados dos resíduos é

$$SQ_E = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

- Pode-se mostrar que $E(SQ_E) = (n-2)\sigma^2$. Logo, um estimador não viesado para σ^2 é

$$\hat{\sigma}^2 = \frac{SQ_E}{n-2}$$

Um outro parâmetro

- Só que calcular $SQ_E = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ pode ser bastante trabalhoso.
- Logo, pode-se mostrar que uma maneira alternativa, mais conveniente, de se obter SQ_E é

$$SQ_E = SQ_T - \hat{\beta}_1 S_{xy}$$

em que

$$SQ_T = \sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 \qquad \text{(Soma total dos quadrados da variável } y\text{)}$$

e

$$S_{xy} = \sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y} \qquad \text{(Numerador de } \hat{\beta}_1\text{)}$$

Um outro parâmetro

- Só que calcular $SQ_E = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ pode ser bastante trabalhoso.
- Logo, pode-se mostrar que uma maneira alternativa, mais conveniente, de se obter SQ_E é

$$SQ_E = SQ_T - \hat{\beta}_1 S_{xy}$$

em que

$$SQ_T = \sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 \qquad \text{(Soma total dos quadrados da variável } y\text{)}$$

e

$$S_{xy} = \sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y} \qquad \text{(Numerador de } \hat{\beta}_1\text{)}$$

Um outro parâmetro

- Só que calcular $SQ_E = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ pode ser bastante trabalhoso.
- Logo, pode-se mostrar que uma maneira alternativa, mais conveniente, de se obter SQ_E é

$$SQ_E = SQ_T - \hat{\beta}_1 S_{xy}$$

em que

$$SQ_T = \sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 \qquad \text{(Soma total dos quadrados da variável } y\text{)}$$

e

$$S_{xy} = \sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y} \qquad \text{(Numerador de } \hat{\beta}_1\text{)}$$

Um outro parâmetro

- Só que calcular $SQ_E = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ pode ser bastante trabalhoso.
- Logo, pode-se mostrar que uma maneira alternativa, mais conveniente, de se obter SQ_E é

$$SQ_E = SQ_T - \hat{\beta}_1 S_{xy}$$

em que

$$SQ_T = \sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 \qquad \text{(Soma total dos quadrados da variável } y\text{)}$$

e

$$S_{xy} = \sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y} \qquad \text{(Numerador de } \hat{\beta}_1\text{)}$$

Um outro parâmetro

- Só que calcular $SQ_E = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ pode ser bastante trabalhoso.
- Logo, pode-se mostrar que uma maneira alternativa, mais conveniente, de se obter SQ_E é

$$SQ_E = SQ_T - \hat{\beta}_1 S_{xy}$$

em que

$$SQ_T = \sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 \qquad \text{(Soma total dos quadrados da variável } y\text{)}$$

e

$$S_{xy} = \sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y} \qquad \text{(Numerador de } \hat{\beta}_1\text{)}$$

Regressão linear simples

Um outro parâmetro

- Só que calcular $SQ_E = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ pode ser bastante trabalhoso.
- Logo, pode-se mostrar que uma maneira alternativa, mais conveniente, de se obter SQ_E é

$$SQ_E = SQ_T - \hat{\beta}_1 S_{xy}$$

em que

$$SQ_T = \sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 \qquad \text{(Soma total dos quadrados da variável } y\text{)}$$

e

$$S_{xy} = \sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y} \qquad \text{(Numerador de } \hat{\beta}_1\text{)}$$

Propriedades dos estimadores de mínimos quadrados de β_0 e β_1

- Pode-se mostrar que os estimadores de mínimos quadrados de β_0 e β_1, ou seja, $\hat{\beta}_0$ e $\hat{\beta}_1$ possuem as seguintes propriedades:
- Com relação a $\hat{\beta}_1$:
  - $E(\hat{\beta}_1) = \beta_1$ (ou seja, $\hat{\beta}_1$ é não viesado para β_1)
  - $Var(\hat{\beta}_1) = \frac{\sigma^2}{S_{xx}}$, em que $S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2$
- Com relação a $\hat{\beta}_0$:
  - $E(\hat{\beta}_0) = \beta_0$ (ou seja, $\hat{\beta}_0$ é não viesado para β_0)
  - $Var(\hat{\beta}_0) = \sigma^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]$.
- Note que nas variâncias de $\hat{\beta}_0$ e $\hat{\beta}_1$ aparece a variância dos erros (σ^2). Logo, podemos usar $\hat{\sigma}^2$ para fornecer estimativas das variâncias desses estimadores.

Propriedades dos estimadores de mínimos quadrados de β_0 e β_1

- Pode-se mostrar que os estimadores de mínimos quadrados de β_0 e β_1, ou seja, $\hat{\beta}_0$ e $\hat{\beta}_1$ possuem as seguintes propriedades:
- Com relação a $\hat{\beta}_1$:
  - $E(\hat{\beta}_1) = \beta_1$ (ou seja, $\hat{\beta}_1$ é não viesado para β_1)
  - $Var(\hat{\beta}_1) = \frac{\sigma^2}{S_{xx}}$, em que $S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2$
- Com relação a $\hat{\beta}_0$:
  - $E(\hat{\beta}_0) = \beta_0$ (ou seja, $\hat{\beta}_0$ é não viesado para β_0)
  - $Var(\hat{\beta}_0) = \sigma^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]$.
- Note que nas variâncias de $\hat{\beta}_0$ e $\hat{\beta}_1$ aparece a variância dos erros (σ^2). Logo, podemos usar $\hat{\sigma}^2$ para fornecer estimativas das variâncias desses estimadores.

Propriedades dos estimadores de mínimos quadrados de β_0 e β_1

- Pode-se mostrar que os estimadores de mínimos quadrados de β_0 e β_1, ou seja, $\hat{\beta}_0$ e $\hat{\beta}_1$ possuem as seguintes propriedades:
- Com relação a $\hat{\beta}_1$:
  - $E(\hat{\beta}_1) = \beta_1$ (ou seja, $\hat{\beta}_1$ é não viesado para β_1)
  - $Var(\hat{\beta}_1) = \frac{\sigma^2}{S_{xx}}$, em que $S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2$
- Com relação a $\hat{\beta}_0$:
  - $E(\hat{\beta}_0) = \beta_0$ (ou seja, $\hat{\beta}_0$ é não viesado para β_0)
  - $Var(\hat{\beta}_0) = \sigma^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]$.
- Note que nas variâncias de $\hat{\beta}_0$ e $\hat{\beta}_1$ aparece a variância dos erros (σ^2). Logo, podemos usar $\hat{\sigma}^2$ para fornecer estimativas das variâncias desses estimadores.

Propriedades dos estimadores de mínimos quadrados de β_0 e β_1

- Pode-se mostrar que os estimadores de mínimos quadrados de β_0 e β_1, ou seja, $\hat{\beta}_0$ e $\hat{\beta}_1$ possuem as seguintes propriedades:
- Com relação a $\hat{\beta}_1$:
  - $E(\hat{\beta}_1) = \beta_1$ (ou seja, $\hat{\beta}_1$ é não viesado para β_1)
  - $Var(\hat{\beta}_1) = \frac{\sigma^2}{S_{xx}}$, em que $S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2$
- Com relação a $\hat{\beta}_0$:
  - $E(\hat{\beta}_0) = \beta_0$ (ou seja, $\hat{\beta}_0$ é não viesado para β_0)
  - $Var(\hat{\beta}_0) = \sigma^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]$.
- Note que nas variâncias de $\hat{\beta}_0$ e $\hat{\beta}_1$ aparece a variância dos erros (σ^2). Logo, podemos usar $\hat{\sigma}^2$ para fornecer estimativas das variâncias desses estimadores.

Propriedades dos estimadores de mínimos quadrados de β_0 e β_1

- Pode-se mostrar que os estimadores de mínimos quadrados de β_0 e β_1, ou seja, $\hat{\beta}_0$ e $\hat{\beta}_1$ possuem as seguintes propriedades:
- Com relação a $\hat{\beta}_1$:
  - $E(\hat{\beta}_1) = \beta_1$ (ou seja, $\hat{\beta}_1$ é não viesado para β_1)
  - $Var(\hat{\beta}_1) = \frac{\sigma^2}{S_{xx}}$, em que $S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2$
- Com relação a $\hat{\beta}_0$:
  - $E(\hat{\beta}_0) = \beta_0$ (ou seja, $\hat{\beta}_0$ é não viesado para β_0)
  - $Var(\hat{\beta}_0) = \sigma^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]$.
- Note que nas variâncias de $\hat{\beta}_0$ e $\hat{\beta}_1$ aparece a variância dos erros (σ^2). Logo, podemos usar $\hat{\sigma}^2$ para fornecer estimativas das variâncias desses estimadores.

Propriedades dos estimadores de mínimos quadrados de β_0 e β_1

- Pode-se mostrar que os estimadores de mínimos quadrados de β_0 e β_1, ou seja, $\hat{\beta}_0$ e $\hat{\beta}_1$ possuem as seguintes propriedades:
- Com relação a $\hat{\beta}_1$:
  - $E(\hat{\beta}_1) = \beta_1$ (ou seja, $\hat{\beta}_1$ é não viesado para β_1)
  - $Var(\hat{\beta}_1) = \frac{\sigma^2}{S_{xx}}$, em que $S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2$
- Com relação a $\hat{\beta}_0$:
  - $E(\hat{\beta}_0) = \beta_0$ (ou seja, $\hat{\beta}_0$ é não viesado para β_0)
  - $Var(\hat{\beta}_0) = \sigma^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]$.
- Note que nas variâncias de $\hat{\beta}_0$ e $\hat{\beta}_1$ aparece a variância dos erros (σ^2). Logo, podemos usar $\hat{\sigma}^2$ para fornecer estimativas das variâncias desses estimadores.

Propriedades dos estimadores de mínimos quadrados de β_0 e β_1

- Pode-se mostrar que os estimadores de mínimos quadrados de β_0 e β_1, ou seja, $\hat{\beta}_0$ e $\hat{\beta}_1$ possuem as seguintes propriedades:
- Com relação a $\hat{\beta}_1$:
  - $E(\hat{\beta}_1) = \beta_1$ (ou seja, $\hat{\beta}_1$ é não viesado para β_1)
  - $Var(\hat{\beta}_1) = \frac{\sigma^2}{S_{xx}}$, em que $S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2$
- Com relação a $\hat{\beta}_0$:
  - $E(\hat{\beta}_0) = \beta_0$ (ou seja, $\hat{\beta}_0$ é não viesado para β_0)
  - $Var(\hat{\beta}_0) = \sigma^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]$.
- Note que nas variâncias de $\hat{\beta}_0$ e $\hat{\beta}_1$ aparece a variância dos erros (σ^2). Logo, podemos usar $\hat{\sigma}^2$ para fornecer estimativas das variâncias desses estimadores.

Propriedades dos estimadores de mínimos quadrados de β_0 e β_1

- Pode-se mostrar que os estimadores de mínimos quadrados de β_0 e β_1, ou seja, $\hat{\beta}_0$ e $\hat{\beta}_1$ possuem as seguintes propriedades:
- Com relação a $\hat{\beta}_1$:
  - $E(\hat{\beta}_1) = \beta_1$ (ou seja, $\hat{\beta}_1$ é não viesado para β_1)
  - $Var(\hat{\beta}_1) = \frac{\sigma^2}{S_{xx}}$, em que $S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2$
- Com relação a $\hat{\beta}_0$:
  - $E(\hat{\beta}_0) = \beta_0$ (ou seja, $\hat{\beta}_0$ é não viesado para β_0)
  - $Var(\hat{\beta}_0) = \sigma^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]$.
- Note que nas variâncias de $\hat{\beta}_0$ e $\hat{\beta}_1$ aparece a variância dos erros (σ^2). Logo, podemos usar $\hat{\sigma}^2$ para fornecer estimativas das variâncias desses estimadores.

Testes de hipóteses na regressão linear simples

- Uma importante parte na verificação da adequação do modelo de regressão linear simples consiste na realização de testes de hipóteses em relação aos parâmetros do modelo.
- Para testar hipóteses sobre o intercepto e a inclinação do modelo temos de fazer a suposição adicional de que a componente do erro do modelo, ϵ, seja distribuída normalmente.
- Ou seja, supomos que os erros são normal e independentemente distribuídos com média zero e variância σ^2.

Testes de hipóteses na regressão linear simples

- Uma importante parte na verificação da adequação do modelo de regressão linear simples consiste na realização de testes de hipóteses em relação aos parâmetros do modelo.
- Para testar hipóteses sobre o intercepto e a inclinação do modelo temos de fazer a suposição adicional de que a componente do erro do modelo, ϵ, seja distribuída normalmente.
- Ou seja, supomos que os erros são normal e independentemente distribuídos com média zero e variância σ^2.

Testes de hipóteses na regressão linear simples

- Uma importante parte na verificação da adequação do modelo de regressão linear simples consiste na realização de testes de hipóteses em relação aos parâmetros do modelo.
- Para testar hipóteses sobre o intercepto e a inclinação do modelo temos de fazer a suposição adicional de que a componente do erro do modelo, ϵ, seja distribuída normalmente.
- Ou seja, supomos que os erros são normal e independentemente distribuídos com média zero e variância σ^2.

Testes de hipóteses para β_1

- Suponha que queremos testar a hipótese de que a inclinação seja igual a uma constante, digamos $\beta_{1,0}$. Ou seja, queremos testar, ao nível de significância α as hipóteses:

$$H_0 : \beta_1 = \beta_{1,0} \quad \text{contra} \quad H_1 : \beta_1 \neq \beta_{1,0}$$

- Note que uma vez que estamos supondo que $\epsilon_i \sim N(0, \sigma^2)$, então as observações Y_i também são $N(\beta_0 + \beta_1 x_i, \sigma^2)$. Como $\hat{\beta}_1$ é uma combinação linear das variáveis normais, então $\hat{\beta}_1$ é também $N(\beta_1, \frac{\sigma^2}{S_{xx}})$.
- Logo, para realizar um teste da forma acima, usamos a seguinte estatística de teste

$$T_0 = \frac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}}$$

que tem distribuição t com $n - 2$ graus de liberdade, sujeito a H_0.

- Logo, com um argumento análogo ao que usamos em testes de hipóteses, temos que a região de rejeição desse teste é

$$A_1 = \{t_0 > t_{\frac{\alpha}{2}, n-2} \ \text{ ou } \ t_0 < -t_{\frac{\alpha}{2}, n-2}\}$$

em que t_0 é o valor observado da estatística de teste T_0.

Testes de hipóteses para β_1

- Suponha que queremos testar a hipótese de que a inclinação seja igual a uma constante, digamos $\beta_{1,0}$. Ou seja, queremos testar, ao nível de significância α as hipóteses:
$$H_0 : \beta_1 = \beta_{1,0} \quad \text{contra} \quad H_1 : \beta_1 \neq \beta_{1,0}$$
- Note que uma vez que estamos supondo que $\epsilon_i \sim N(0, \sigma^2)$, então as observações Y_i também são $N(\beta_0 + \beta_1 x_i, \sigma^2)$. Como $\hat{\beta}_1$ é uma combinação linear das variáveis normais, então $\hat{\beta}_1$ é também $N(\beta_1, \frac{\sigma^2}{S_{xx}})$.
- Logo, para realizar um teste da forma acima, usamos a seguinte estatística de teste
$$T_0 = \frac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}}$$
que tem distribuição t com $n-2$ graus de liberdade, sujeito a H_0.
- Logo, com um argumento análogo ao que usamos em testes de hipóteses, temos que a região de rejeição desse teste é
$$A_1 = \{t_0 > t_{\frac{\alpha}{2}, n-2} \ \text{ ou } \ t_0 < -t_{\frac{\alpha}{2}, n-2}\}$$
em que t_0 é o valor observado da estatística de teste T_0.

Testes de hipóteses para β_1

- Suponha que queremos testar a hipótese de que a inclinação seja igual a uma constante, digamos $\beta_{1,0}$. Ou seja, queremos testar, ao nível de significância α as hipóteses:

$$H_0 : \beta_1 = \beta_{1,0} \quad \text{contra} \quad H_1 : \beta_1 \neq \beta_{1,0}$$

- Note que uma vez que estamos supondo que $\epsilon_i \sim N(0, \sigma^2)$, então as observações Y_i também são $N(\beta_0 + \beta_1 x_i, \sigma^2)$. Como $\hat{\beta}_1$ é uma combinação linear das variáveis normais, então $\hat{\beta}_1$ é também $N(\beta_1, \frac{\sigma^2}{S_{xx}})$.
- Logo, para realizar um teste da forma acima, usamos a seguinte estatística de teste

$$T_0 = \frac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}}$$

que tem distribuição t com $n-2$ graus de liberdade, sujeito a H_0.

- Logo, com um argumento análogo ao que usamos em testes de hipóteses, temos que a região de rejeição desse teste é

$$A_1 = \{t_0 > t_{\frac{\alpha}{2}, n-2} \;\; \text{ou} \;\; t_0 < -t_{\frac{\alpha}{2}, n-2}\}$$

em que t_0 é o valor observado da estatística de teste T_0.

Testes de hipóteses para β_1

- Suponha que queremos testar a hipótese de que a inclinação seja igual a uma constante, digamos $\beta_{1,0}$. Ou seja, queremos testar, ao nível de significância α as hipóteses:

$$H_0 : \beta_1 = \beta_{1,0} \quad \text{contra} \quad H_1 : \beta_1 \neq \beta_{1,0}$$

- Note que uma vez que estamos supondo que $\epsilon_i \sim N(0, \sigma^2)$, então as observações Y_i também são $N(\beta_0 + \beta_1 x_i, \sigma^2)$. Como $\hat{\beta}_1$ é uma combinação linear das variáveis normais, então $\hat{\beta}_1$ é também $N(\beta_1, \frac{\sigma^2}{S_{xx}})$.
- Logo, para realizar um teste da forma acima, usamos a seguinte estatística de teste

$$T_0 = \frac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}}$$

que tem distribuição t com $n-2$ graus de liberdade, sujeito a H_0.

- Logo, com um argumento análogo ao que usamos em testes de hipóteses, temos que a região de rejeição desse teste é

$$A_1 = \{t_0 > t_{\frac{\alpha}{2}, n-2} \ \text{ ou } \ t_0 < -t_{\frac{\alpha}{2}, n-2}\}$$

em que t_0 é o valor observado da estatística de teste T_0.

Testes de hipóteses para β_1

- Suponha que queremos testar a hipótese de que a inclinação seja igual a uma constante, digamos $\beta_{1,0}$. Ou seja, queremos testar, ao nível de significância α as hipóteses:

$$H_0 : \beta_1 = \beta_{1,0} \quad \text{contra} \quad H_1 : \beta_1 \neq \beta_{1,0}$$

- Note que uma vez que estamos supondo que $\epsilon_i \sim N(0, \sigma^2)$, então as observações Y_i também são $N(\beta_0 + \beta_1 x_i, \sigma^2)$. Como $\hat{\beta}_1$ é uma combinação linear das variáveis normais, então $\hat{\beta}_1$ é também $N(\beta_1, \frac{\sigma^2}{S_{xx}})$.
- Logo, para realizar um teste da forma acima, usamos a seguinte estatística de teste

$$T_0 = \frac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}}$$

que tem distribuição t com $n-2$ graus de liberdade, sujeito a H_0.

- Logo, com um argumento análogo ao que usamos em testes de hipóteses, temos que a região de rejeição desse teste é

$$A_1 = \{t_0 > t_{\frac{\alpha}{2}, n-2} \quad \text{ou} \quad t_0 < -t_{\frac{\alpha}{2}, n-2}\}$$

em que t_0 é o valor observado da estatística de teste T_0.

Testes de hipóteses para β_1

- Suponha que queremos testar a hipótese de que a inclinação seja igual a uma constante, digamos $\beta_{1,0}$. Ou seja, queremos testar, ao nível de significância α as hipóteses:

$$H_0 : \beta_1 = \beta_{1,0} \quad \text{contra} \quad H_1 : \beta_1 \neq \beta_{1,0}$$

- Note que uma vez que estamos supondo que $\epsilon_i \sim N(0, \sigma^2)$, então as observações Y_i também são $N(\beta_0 + \beta_1 x_i, \sigma^2)$. Como $\hat{\beta}_1$ é uma combinação linear das variáveis normais, então $\hat{\beta}_1$ é também $N(\beta_1, \frac{\sigma^2}{S_{xx}})$.
- Logo, para realizar um teste da forma acima, usamos a seguinte estatística de teste

$$T_0 = \frac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}}$$

que tem distribuição t com $n-2$ graus de liberdade, sujeito a H_0.

- Logo, com um argumento análogo ao que usamos em testes de hipóteses, temos que a região de rejeição desse teste é

$$A_1 = \{t_0 > t_{\frac{\alpha}{2}, n-2} \;\; \text{ou} \;\; t_0 < -t_{\frac{\alpha}{2}, n-2}\}$$

em que t_0 é o valor observado da estatística de teste T_0.

Testes de hipóteses para β_1

- Suponha que queremos testar a hipótese de que a inclinação seja igual a uma constante, digamos $\beta_{1,0}$. Ou seja, queremos testar, ao nível de significância α as hipóteses:
$$H_0 : \beta_1 = \beta_{1,0} \quad \text{contra} \quad H_1 : \beta_1 \neq \beta_{1,0}$$
- Note que uma vez que estamos supondo que $\epsilon_i \sim N(0, \sigma^2)$, então as observações Y_i também são $N(\beta_0 + \beta_1 x_i, \sigma^2)$. Como $\hat{\beta}_1$ é uma combinação linear das variáveis normais, então $\hat{\beta}_1$ é também $N(\beta_1, \frac{\sigma^2}{S_{xx}})$.
- Logo, para realizar um teste da forma acima, usamos a seguinte estatística de teste
$$T_0 = \frac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}}$$
que tem distribuição t com $n-2$ graus de liberdade, sujeito a H_0.
- Logo, com um argumento análogo ao que usamos em testes de hipóteses, temos que a região de rejeição desse teste é
$$A_1 = \{t_0 > t_{\frac{\alpha}{2}, n-2} \quad \text{ou} \quad t_0 < -t_{\frac{\alpha}{2}, n-2}\}$$
em que t_0 é o valor observado da estatística de teste T_0.

Testes de hipóteses para β_1

- Suponha que queremos testar a hipótese de que a inclinação seja igual a uma constante, digamos $\beta_{1,0}$. Ou seja, queremos testar, ao nível de significância α as hipóteses:

$$H_0 : \beta_1 = \beta_{1,0} \quad \text{contra} \quad H_1 : \beta_1 \neq \beta_{1,0}$$

- Note que uma vez que estamos supondo que $\epsilon_i \sim N(0, \sigma^2)$, então as observações Y_i também são $N(\beta_0 + \beta_1 x_i, \sigma^2)$. Como $\hat{\beta}_1$ é uma combinação linear das variáveis normais, então $\hat{\beta}_1$ é também $N(\beta_1, \frac{\sigma^2}{S_{xx}})$.
- Logo, para realizar um teste da forma acima, usamos a seguinte estatística de teste

$$T_0 = \frac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}}$$

que tem distribuição t com $n-2$ graus de liberdade, sujeito a H_0.

- Logo, com um argumento análogo ao que usamos em testes de hipóteses, temos que a região de rejeição desse teste é

$$A_1 = \{t_0 > t_{\frac{\alpha}{2}, n-2} \quad \text{ou} \quad t_0 < -t_{\frac{\alpha}{2}, n-2}\}$$

em que t_0 é o valor observado da estatística de teste T_0.

Testes de hipóteses para β_1

- Suponha que queremos testar a hipótese de que a inclinação seja igual a uma constante, digamos $\beta_{1,0}$. Ou seja, queremos testar, ao nível de significância α as hipóteses:

$$H_0 : \beta_1 = \beta_{1,0} \quad \text{contra} \quad H_1 : \beta_1 \neq \beta_{1,0}$$

- Note que uma vez que estamos supondo que $\epsilon_i \sim N(0, \sigma^2)$, então as observações Y_i também são $N(\beta_0 + \beta_1 x_i, \sigma^2)$. Como $\hat{\beta}_1$ é uma combinação linear das variáveis normais, então $\hat{\beta}_1$ é também $N(\beta_1, \frac{\sigma^2}{S_{xx}})$.
- Logo, para realizar um teste da forma acima, usamos a seguinte estatística de teste

$$T_0 = \frac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}}$$

que tem distribuição t com $n-2$ graus de liberdade, sujeito a H_0.

- Logo, com um argumento análogo ao que usamos em testes de hipóteses, temos que a região de rejeição desse teste é

$$A_1 = \{t_0 > t_{\frac{\alpha}{2}, n-2} \;\text{ ou }\; t_0 < -t_{\frac{\alpha}{2}, n-2}\}$$

em que t_0 é o valor observado da estatística de teste T_0.

Testes de hipóteses para β_1

- Suponha que queremos testar a hipótese de que a inclinação seja igual a uma constante, digamos $\beta_{1,0}$. Ou seja, queremos testar, ao nível de significância α as hipóteses:

$$H_0 : \beta_1 = \beta_{1,0} \quad \text{contra} \quad H_1 : \beta_1 \neq \beta_{1,0}$$

- Note que uma vez que estamos supondo que $\epsilon_i \sim N(0, \sigma^2)$, então as observações Y_i também são $N(\beta_0 + \beta_1 x_i, \sigma^2)$. Como $\hat{\beta}_1$ é uma combinação linear das variáveis normais, então $\hat{\beta}_1$ é também $N(\beta_1, \frac{\sigma^2}{S_{xx}})$.
- Logo, para realizar um teste da forma acima, usamos a seguinte estatística de teste

$$T_0 = \frac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}}$$

que tem distribuição t com $n-2$ graus de liberdade, sujeito a H_0.

- Logo, com um argumento análogo ao que usamos em testes de hipóteses, temos que a região de rejeição desse teste é

$$A_1 = \{t_0 > t_{\frac{\alpha}{2}, n-2} \quad \text{ou} \quad t_0 < -t_{\frac{\alpha}{2}, n-2}\}$$

em que t_0 é o valor observado da estatística de teste T_0.

Testes de hipóteses para β_0

- Um procedimento similar pode ser usado para testar hipóteses sobre o intercepto, ou seja, se queremos testar as hipóteses

$$H_0 : \beta_0 = \beta_{0,0} \quad \text{contra} \quad H_1 : \beta_0 \neq \beta_{0,0}$$

- Então usamos a estatística de teste

$$T_0 = \frac{\hat{\beta}_0 - \beta_{0,0}}{\sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]}}$$

- E a região de rejeição desse teste é

$$A_1 = \{t_0 > t_{\frac{\alpha}{2},n-2} \ \text{ ou } \ t_0 < -t_{\frac{\alpha}{2},n-2}\}$$

em que t_0 é o valor observado da estatística de teste T_0.

Testes de hipóteses para β_0

- Um procedimento similar pode ser usado para testar hipóteses sobre o intercepto, ou seja, se queremos testar as hipóteses

$$H_0 : \beta_0 = \beta_{0,0} \quad \text{contra} \quad H_1 : \beta_0 \neq \beta_{0,0}$$

- Então usamos a estatística de teste

$$T_0 = \frac{\hat{\beta}_0 - \beta_{0,0}}{\sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]}}$$

- E a região de rejeição desse teste é

$$A_1 = \{t_0 > t_{\frac{\alpha}{2},n-2} \ \text{ ou } \ t_0 < -t_{\frac{\alpha}{2},n-2}\}$$

em que t_0 é o valor observado da estatística de teste T_0.

Testes de hipóteses para β_0

- Um procedimento similar pode ser usado para testar hipóteses sobre o intercepto, ou seja, se queremos testar as hipóteses

$$H_0 : \beta_0 = \beta_{0,0} \quad \text{contra} \quad H_1 : \beta_0 \neq \beta_{0,0}$$

- Então usamos a estatística de teste

$$T_0 = \frac{\hat{\beta}_0 - \beta_{0,0}}{\sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]}}$$

- E a região de rejeição desse teste é

$$A_1 = \{t_0 > t_{\frac{\alpha}{2},n-2} \ \text{ ou } \ t_0 < -t_{\frac{\alpha}{2},n-2}\}$$

em que t_0 é o valor observado da estatística de teste T_0.

Testes de hipóteses para β_0

- Um procedimento similar pode ser usado para testar hipóteses sobre o intercepto, ou seja, se queremos testar as hipóteses

$$H_0 : \beta_0 = \beta_{0,0} \quad \text{contra} \quad H_1 : \beta_0 \neq \beta_{0,0}$$

- Então usamos a estatística de teste

$$T_0 = \frac{\hat{\beta}_0 - \beta_{0,0}}{\sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]}}$$

- E a região de rejeição desse teste é

$$A_1 = \{t_0 > t_{\frac{\alpha}{2},n-2} \;\; \text{ou} \;\; t_0 < -t_{\frac{\alpha}{2},n-2}\}$$

em que t_0 é o valor observado da estatística de teste T_0.

Testes de hipóteses para β_0

- Um procedimento similar pode ser usado para testar hipóteses sobre o intercepto, ou seja, se queremos testar as hipóteses

$$H_0 : \beta_0 = \beta_{0,0} \quad \text{contra} \quad H_1 : \beta_0 \neq \beta_{0,0}$$

- Então usamos a estatística de teste

$$T_0 = \frac{\hat{\beta}_0 - \beta_{0,0}}{\sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]}}$$

- E a região de rejeição desse teste é

$$A_1 = \{t_0 > t_{\frac{\alpha}{2},n-2} \quad \text{ou} \quad t_0 < -t_{\frac{\alpha}{2},n-2}\}$$

em que t_0 é o valor observado da estatística de teste T_0.

Testes de hipóteses para β_0

- Um procedimento similar pode ser usado para testar hipóteses sobre o intercepto, ou seja, se queremos testar as hipóteses

$$H_0 : \beta_0 = \beta_{0,0} \quad \text{contra} \quad H_1 : \beta_0 \neq \beta_{0,0}$$

- Então usamos a estatística de teste

$$T_0 = \frac{\hat{\beta}_0 - \beta_{0,0}}{\sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]}}$$

- E a região de rejeição desse teste é

$$A_1 = \{t_0 > t_{\frac{\alpha}{2},n-2} \;\text{ ou }\; t_0 < -t_{\frac{\alpha}{2},n-2}\}$$

em que t_0 é o valor observado da estatística de teste T_0.

Testes de hipóteses para β_0

- Um procedimento similar pode ser usado para testar hipóteses sobre o intercepto, ou seja, se queremos testar as hipóteses

$$H_0 : \beta_0 = \beta_{0,0} \quad \text{contra} \quad H_1 : \beta_0 \neq \beta_{0,0}$$

- Então usamos a estatística de teste

$$T_0 = \frac{\hat{\beta}_0 - \beta_{0,0}}{\sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]}}$$

- E a região de rejeição desse teste é

$$A_1 = \{t_0 > t_{\frac{\alpha}{2},n-2} \;\; \text{ou} \;\; t_0 < -t_{\frac{\alpha}{2},n-2}\}$$

em que t_0 é o valor observado da estatística de teste T_0.

Um caso especial do testes de hipóteses para β_1

- Um caso especial e importante das hipóteses sobre β_1 é testar as hipóteses

$$H_0 : \beta_1 = 0 \quad \text{contra} \quad H_1 : \beta_1 \neq 0$$

- Essas hipóteses se relacionam com à significância da regressão.
- Aceitar $H_0 : \beta_1 = 0$ é equivalente a concluir que não há relação linear entre x e Y.
  - Isso pode implicar que x seja de pouco valor para explicar a variação em Y.
- Rejeitar $H_0 : \beta_1 = 0$ implica que x é importante para explicar a variabilidade de Y.

Um caso especial do testes de hipóteses para β_1

- Um caso especial e importante das hipóteses sobre β_1 é testar as hipóteses

$$H_0 : \beta_1 = 0 \quad \text{contra} \quad H_1 : \beta_1 \neq 0$$

- Essas hipóteses se relacionam com à significância da regressão.
- Aceitar $H_0 : \beta_1 = 0$ é equivalente a concluir que não há relação linear entre x e Y.
  - Isso pode implicar que x seja de pouco valor para explicar a variação em Y.
- Rejeitar $H_0 : \beta_1 = 0$ implica que x é importante para explicar a variabilidade de Y.

Um caso especial do testes de hipóteses para β_1

- Um caso especial e importante das hipóteses sobre β_1 é testar as hipóteses

$$H_0 : \beta_1 = 0 \quad \text{contra} \quad H_1 : \beta_1 \neq 0$$

- Essas hipóteses se relacionam com à significância da regressão.
- Aceitar $H_0 : \beta_1 = 0$ é equivalente a concluir que não há relação linear entre x e Y.
  - Isso pode implicar que x seja de pouco valor para explicar a variação em Y.
- Rejeitar $H_0 : \beta_1 = 0$ implica que x é importante para explicar a variabilidade de Y.

Um caso especial do testes de hipóteses para β_1

- Um caso especial e importante das hipóteses sobre β_1 é testar as hipóteses

$$H_0 : \beta_1 = 0 \quad \text{contra} \quad H_1 : \beta_1 \neq 0$$

- Essas hipóteses se relacionam com à significância da regressão.
- Aceitar $H_0 : \beta_1 = 0$ é equivalente a concluir que não há relação linear entre x e Y.
  - Isso pode implicar que x seja de pouco valor para explicar a variação em Y.
- Rejeitar $H_0 : \beta_1 = 0$ implica que x é importante para explicar a variabilidade de Y.

Um caso especial do testes de hipóteses para β_1

- Um caso especial e importante das hipóteses sobre β_1 é testar as hipóteses

$$H_0 : \beta_1 = 0 \quad \text{contra} \quad H_1 : \beta_1 \neq 0$$

- Essas hipóteses se relacionam com à significância da regressão.
- Aceitar $H_0 : \beta_1 = 0$ é equivalente a concluir que não há relação linear entre x e Y.
  - Isso pode implicar que x seja de pouco valor para explicar a variação em Y.
- Rejeitar $H_0 : \beta_1 = 0$ implica que x é importante para explicar a variabilidade de Y.

Um caso especial do testes de hipóteses para β_1

- Um caso especial e importante das hipóteses sobre β_1 é testar as hipóteses

$$H_0 : \beta_1 = 0 \quad \text{contra} \quad H_1 : \beta_1 \neq 0$$

- Essas hipóteses se relacionam com à significância da regressão.
- Aceitar $H_0 : \beta_1 = 0$ é equivalente a concluir que não há relação linear entre x e Y.
  - Isso pode implicar que x seja de pouco valor para explicar a variação em Y.
- Rejeitar $H_0 : \beta_1 = 0$ implica que x é importante para explicar a variabilidade de Y.

Exemplo

Exemplo: Um artigo na revista *IEEE transactions on instrumentation and measurement* *["Direct, fast, and accurate measurement of V_T and K of an MOS transistor using a V_T-sift circuit" (Vol. 40, 1991, pp. 951-955)]* descreveu o uso de um modelo de regressão linear simples para expressar a corrente y (em miliamperes), como função da diferença de voltagem x (em volts). Os dados são fornecidos a seguir:

x	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2
y	0.734	0.886	1.04	1.19	1.35	1.5	1.66	1.81	1.97	2.12

(a) Desenhe um diagrama de dispersão desses dados. Um modelo de regressão linear simples parece plausível?

(b) Ajuste um modelo de regressão linear simples a esses dados.

(c) Teste a hipótese de que $H_0 : \beta_0 = 0$ contra $H_1 : \beta_0 \neq 0$, ao nível de significância $\alpha = 5\%$.

Exemplo

Exemplo: Um artigo na revista *IEEE transactions on instrumentation and measurement ["Direct, fast, and accurate measurement of V_T and K of an MOS transistor using a V_T-sift circuit" (Vol. 40, 1991, pp. 951-955)]* descreveu o uso de um modelo de regressão linear simples para expressar a corrente y (em miliamperes), como função da diferença de voltagem x (em volts). Os dados são fornecidos a seguir:

x	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2
y	0.734	0.886	1.04	1.19	1.35	1.5	1.66	1.81	1.97	2.12

(a) Desenhe um diagrama de dispersão desses dados. Um modelo de regressão linear simples parece plausível?

(b) Ajuste um modelo de regressão linear simples a esses dados.

(c) Teste a hipótese de que $H_0 : \beta_0 = 0$ contra $H_1 : \beta_0 \neq 0$, ao nível de significância $\alpha = 5\%$.

Exemplo

Exemplo: Um artigo na revista *IEEE transactions on instrumentation and measurement ["Direct, fast, and accurate measurement of V_T and K of an MOS transistor using a V_T-sift circuit" (Vol. 40, 1991, pp. 951-955)]* descreveu o uso de um modelo de regressão linear simples para expressar a corrente y (em miliamperes), como função da diferença de voltagem x (em volts). Os dados são fornecidos a seguir:

x	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2
y	0.734	0.886	1.04	1.19	1.35	1.5	1.66	1.81	1.97	2.12

(a) Desenhe um diagrama de dispersão desses dados. Um modelo de regressão linear simples parece plausível?

(b) Ajuste um modelo de regressão linear simples a esses dados.

(c) Teste a hipótese de que $H_0 : \beta_0 = 0$ contra $H_1 : \beta_0 \neq 0$, ao nível de significância $\alpha = 5\%$.

Exemplo

Exemplo: Um artigo na revista *IEEE transactions on instrumentation and measurement ["Direct, fast, and accurate measurement of V_T and K of an MOS transistor using a V_T-sift circuit" (Vol. 40, 1991, pp. 951-955)]* descreveu o uso de um modelo de regressão linear simples para expressar a corrente y (em miliamperes), como função da diferença de voltagem x (em volts). Os dados são fornecidos a seguir:

x	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2
y	0.734	0.886	1.04	1.19	1.35	1.5	1.66	1.81	1.97	2.12

(a) Desenhe um diagrama de dispersão desses dados. Um modelo de regressão linear simples parece plausível?

(b) Ajuste um modelo de regressão linear simples a esses dados.

(c) Teste a hipótese de que $H_0 : \beta_0 = 0$ contra $H_1 : \beta_0 \neq 0$, ao nível de significância $\alpha = 5\%$.

Exemplo

Exemplo: Um artigo na revista *IEEE transactions on instrumentation and measurement ["Direct, fast, and accurate measurement of V_T and K of an MOS transistor using a V_T-sift circuit" (Vol. 40, 1991, pp. 951-955)]* descreveu o uso de um modelo de regressão linear simples para expressar a corrente y (em miliamperes), como função da diferença de voltagem x (em volts). Os dados são fornecidos a seguir:

x	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2
y	0.734	0.886	1.04	1.19	1.35	1.5	1.66	1.81	1.97	2.12

(a) Desenhe um diagrama de dispersão desses dados. Um modelo de regressão linear simples parece plausível?

(b) Ajuste um modelo de regressão linear simples a esses dados.

(c) Teste a hipótese de que $H_0 : \beta_0 = 0$ contra $H_1 : \beta_0 \neq 0$, ao nível de significância $\alpha = 5\%$.

Sol. (a)

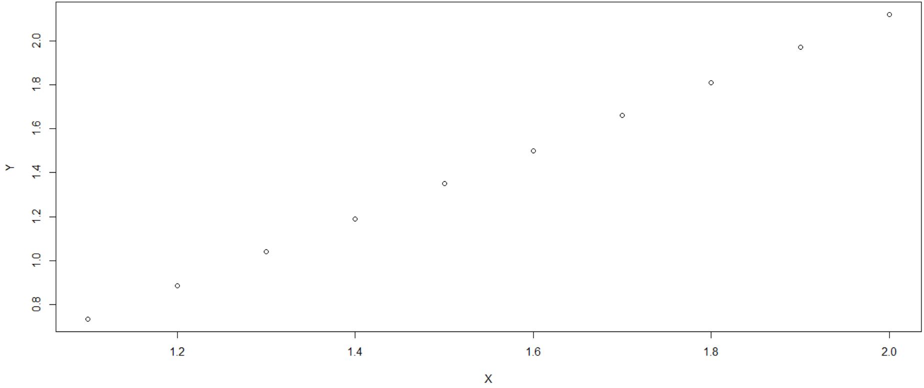

Figure 1: Dados do problema. X: Diferença de voltagem, Y: Corrente

Sol. (b)

Recorde que para ajustar o modelo precisamos obter $\hat{\beta}_0$ e $\hat{\beta}_1$.

- Temos que $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$ e $\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$.
- Temos ainda que
  - $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1.1+1.2+1.3+1.4+1.5+1.6+1.7+1.8+1.9+2}{10} = 1.55$;
  - $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{0.734+0.886+1.04+1.19+1.35+1.5+1.66+1.81+1.97+2.12}{10} = 1.426$;
  - $\sum_{i=1}^{n} y_i x_i = (1.1)(0.734) + (1.2)(0.886) + (1.3)(1.04) + (1.4)(1.19) + (1.5)(1.35) + (1.6)(1.5) + (1.7)(1.66) + (1.8)(1.81) + (1.9)(1.97) + (2)(2.12) = 23.3766$;
  - $\sum_{i=1}^{n} x_i^2 = 1.1^2 + 1.2^2 + 1.3^2 + 1.4^2 + 1.5^2 + 1.6^2 + 1.7^2 + 1.8^2 + 1.9^2 + 2^2 = 24.85$.
- Logo, temos que

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2} = \frac{23.3766 - 10(1.55)(1.426)}{24.85 - 10(1.55)^2} \approx 1.54376$$

Sol. (b)

Recorde que para ajustar o modelo precisamos obter $\hat{\beta}_0$ e $\hat{\beta}_1$.

- Temos que $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}$ e $\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$.
- Temos ainda que
  - $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1.1+1.2+1.3+1.4+1.5+1.6+1.7+1.8+1.9+2}{10} = 1.55$;
  - $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{0.734+0.886+1.04+1.19+1.35+1.5+1.66+1.81+1.97+2.12}{10} = 1.426$;
  - $\sum_{i=1}^{n} y_i x_i = (1.1)(0.734)+(1.2)(0.886)+(1.3)(1.04)+(1.4)(1.19)+(1.5)(1.35)+(1.6)(1.5)+(1.7)(1.66)+(1.8)(1.81)+(1.9)(1.97)+(2)(2.12) = 23.3766$;
  - $\sum_{i=1}^{n} x_i^2 = 1.1^2+1.2^2+1.3^2+1.4^2+1.5^2+1.6^2+1.7^2+1.8^2+1.9^2+2^2 = 24.85$.
- Logo, temos que

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2} = \frac{23.3766 - 10(1.55)(1.426)}{24.85 - 10(1.55)^2} \approx 1.54376$$

Sol. (b)

Recorde que para ajustar o modelo precisamos obter $\hat{\beta}_0$ e $\hat{\beta}_1$.

- Temos que $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}$ e $\hat{\beta}_1 = \frac{\sum_{i=1}^n y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}$.
- Temos ainda que
  - $\bar{x} = \frac{1}{n}\sum_{i=1}^n x_i = \frac{1.1+1.2+1.3+1.4+1.5+1.6+1.7+1.8+1.9+2}{10} = 1.55;$
  - $\bar{y} = \frac{1}{n}\sum_{i=1}^n y_i = \frac{0.734+0.886+1.04+1.19+1.35+1.5+1.66+1.81+1.97+2.12}{10} = 1.426;$
  - $\sum_{i=1}^n y_i x_i = (1.1)(0.734) + (1.2)(0.886) + (1.3)(1.04) + (1.4)(1.19) + (1.5)(1.35) + (1.6)(1.5)+(1.7)(1.66)+(1.8)(1.81)+(1.9)(1.97)+(2)(2.12) = 23.3766;$
  - $\sum_{i=1}^n x_i^2 = 1.1^2 + 1.2^2 + 1.3^2 + 1.4^2 + 1.5^2 + 1.6^2 + 1.7^2 + 1.8^2 + 1.9^2 + 2^2 = 24.85.$
- Logo, temos que

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^n x_i^2 - n\bar{x}^2} = \frac{23.3766 - 10(1.55)(1.426)}{24.85 - 10(1.55)^2} \approx 1.54376$$

Sol. (b)

Recorde que para ajustar o modelo precisamos obter $\hat{\beta}_0$ e $\hat{\beta}_1$.

- Temos que $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$ e $\hat{\beta}_1 = \frac{\sum_{i=1}^n y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}$.
- Temos ainda que
  - $\bar{x} = \frac{1}{n}\sum_{i=1}^n x_i = \frac{1.1+1.2+1.3+1.4+1.5+1.6+1.7+1.8+1.9+2}{10} = 1.55$;
  - $\bar{y} = \frac{1}{n}\sum_{i=1}^n y_i = \frac{0.734+0.886+1.04+1.19+1.35+1.5+1.66+1.81+1.97+2.12}{10} = 1.426$;
  - $\sum_{i=1}^n y_i x_i = (1.1)(0.734) + (1.2)(0.886) + (1.3)(1.04) + (1.4)(1.19) + (1.5)(1.35) + (1.6)(1.5)+(1.7)(1.66)+(1.8)(1.81)+(1.9)(1.97)+(2)(2.12) = 23.3766$;
  - $\sum_{i=1}^n x_i^2 = 1.1^2 + 1.2^2 + 1.3^2 + 1.4^2 + 1.5^2 + 1.6^2 + 1.7^2 + 1.8^2 + 1.9^2 + 2^2 = 24.85$.
- Logo, temos que

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^n x_i^2 - n\bar{x}^2} = \frac{23.3766 - 10(1.55)(1.426)}{24.85 - 10(1.55)^2} \approx 1.54376$$

Sol. (b)

Recorde que para ajustar o modelo precisamos obter $\hat{\beta}_0$ e $\hat{\beta}_1$.

- Temos que $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$ e $\hat{\beta}_1 = \frac{\sum_{i=1}^n y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}$.
- Temos ainda que
  - $\bar{x} = \frac{1}{n}\sum_{i=1}^n x_i = \frac{1.1+1.2+1.3+1.4+1.5+1.6+1.7+1.8+1.9+2}{10} = 1.55$;
  - $\bar{y} = \frac{1}{n}\sum_{i=1}^n y_i = \frac{0.734+0.886+1.04+1.19+1.35+1.5+1.66+1.81+1.97+2.12}{10} = 1.426$;
  - $\sum_{i=1}^n y_i x_i = (1.1)(0.734) + (1.2)(0.886) + (1.3)(1.04) + (1.4)(1.19) + (1.5)(1.35) + (1.6)(1.5) + (1.7)(1.66) + (1.8)(1.81) + (1.9)(1.97) + (2)(2.12) = 23.3766$;
  - $\sum_{i=1}^n x_i^2 = 1.1^2 + 1.2^2 + 1.3^2 + 1.4^2 + 1.5^2 + 1.6^2 + 1.7^2 + 1.8^2 + 1.9^2 + 2^2 = 24.85$.
- Logo, temos que

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^n x_i^2 - n\bar{x}^2} = \frac{23.3766 - 10(1.55)(1.426)}{24.85 - 10(1.55)^2} \approx 1.54376$$

Sol. (b)

Recorde que para ajustar o modelo precisamos obter $\hat{\beta}_0$ e $\hat{\beta}_1$.

- Temos que $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$ e $\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$.
- Temos ainda que
  - $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1.1+1.2+1.3+1.4+1.5+1.6+1.7+1.8+1.9+2}{10} = 1.55$;
  - $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{0.734+0.886+1.04+1.19+1.35+1.5+1.66+1.81+1.97+2.12}{10} = 1.426$;
  - $\sum_{i=1}^{n} y_i x_i = (1.1)(0.734) + (1.2)(0.886) + (1.3)(1.04) + (1.4)(1.19) + (1.5)(1.35) + (1.6)(1.5) + (1.7)(1.66) + (1.8)(1.81) + (1.9)(1.97) + (2)(2.12) = 23.3766$;
  - $\sum_{i=1}^{n} x_i^2 = 1.1^2 + 1.2^2 + 1.3^2 + 1.4^2 + 1.5^2 + 1.6^2 + 1.7^2 + 1.8^2 + 1.9^2 + 2^2 = 24.85$.
- Logo, temos que

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2} = \frac{23.3766 - 10(1.55)(1.426)}{24.85 - 10(1.55)^2} \approx 1.54376$$

Sol. (b)

Recorde que para ajustar o modelo precisamos obter $\hat{\beta}_0$ e $\hat{\beta}_1$.

- Temos que $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}$ e $\hat{\beta}_1 = \frac{\sum_{i=1}^n y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}$.
- Temos ainda que
  - $\bar{x} = \frac{1}{n}\sum_{i=1}^n x_i = \frac{1.1+1.2+1.3+1.4+1.5+1.6+1.7+1.8+1.9+2}{10} = 1.55$;
  - $\bar{y} = \frac{1}{n}\sum_{i=1}^n y_i = \frac{0.734+0.886+1.04+1.19+1.35+1.5+1.66+1.81+1.97+2.12}{10} = 1.426$;
  - $\sum_{i=1}^n y_i x_i = (1.1)(0.734) + (1.2)(0.886) + (1.3)(1.04) + (1.4)(1.19) + (1.5)(1.35) + (1.6)(1.5) + (1.7)(1.66) + (1.8)(1.81) + (1.9)(1.97) + (2)(2.12) = 23.3766$;
  - $\sum_{i=1}^n x_i^2 = 1.1^2 + 1.2^2 + 1.3^2 + 1.4^2 + 1.5^2 + 1.6^2 + 1.7^2 + 1.8^2 + 1.9^2 + 2^2 = 24.85$.
- Logo, temos que

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^n x_i^2 - n\bar{x}^2} = \frac{23.3766 - 10(1.55)(1.426)}{24.85 - 10(1.55)^2} \approx 1.54376$$

Cont. Sol. (b)

- E ainda
$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x} = (1.426) - (1.54376)(1.55) \approx -0.966824.$$

- Portanto, o modelo ajustado é $\hat{y}_i = -0.966824 + 1.54376x_i$

Cont. Sol. (b)

- E ainda
$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = (1.426) - (1.54376)(1.55) \approx -0.966824.$$

- Portanto, o modelo ajustado é $\hat{y}_i = -0.966824 + 1.54376x_i$

Cont. Sol. (b)

- E ainda
$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = (1.426) - (1.54376)(1.55) \approx -0.966824.$$

- Portanto, o modelo ajustado é $\hat{y}_i = -0.966824 + 1.54376 x_i$

Modelo ajustado

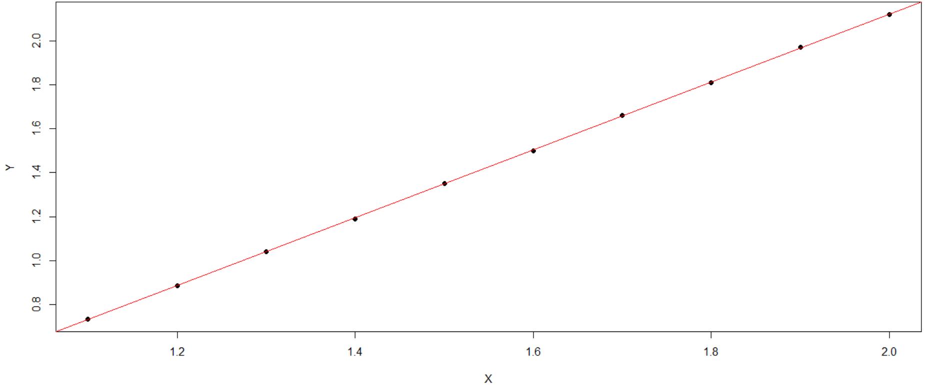

Figure 2: Modelo ajustado

Sol. (c)

- No item (c) queremos testar hipóteses com respeito ao intercepto do modelo, ou seja, queremos testar, ao nível de significância $\alpha = 5\%$, as hipóteses
  $H_0 : \beta_0 = 0$ contra $H_1 : \beta_0 \neq 0$.
- Vimos que um teste de nível α para β_0 tem região de rejeição

$$A_1 = \{t_0 > t_{\frac{\alpha}{2},n-2} \quad \text{ou} \quad t_0 < -t_{\frac{\alpha}{2},n-2}\}$$

em que t_0 é o valor observado da estatística de teste T_0, dada por

$$T_0 = \frac{\hat{\beta}_0 - \beta_{0,0}}{\sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]}}$$

- Dessa forma, precisamos obter a estimativa $\hat{\sigma}^2$.

Sol. (c)

- No item (c) queremos testar hipóteses com respeito ao intercepto do modelo, ou seja, queremos testar, ao nível de significância $\alpha = 5\%$, as hipóteses $H_0 : \beta_0 = 0$ contra $H_1 : \beta_0 \neq 0$.
- Vimos que um teste de nível α para β_0 tem região de rejeição

$$A_1 = \{t_0 > t_{\frac{\alpha}{2},n-2} \ \text{ ou } \ t_0 < -t_{\frac{\alpha}{2},n-2}\}$$

em que t_0 é o valor observado da estatística de teste T_0, dada por

$$T_0 = \frac{\hat{\beta}_0 - \beta_{0,0}}{\sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]}}$$

- Dessa forma, precisamos obter a estimativa $\hat{\sigma}^2$.

Sol. (c)

- No item (c) queremos testar hipóteses com respeito ao intercepto do modelo, ou seja, queremos testar, ao nível de significância $\alpha = 5\%$, as hipóteses $H_0 : \beta_0 = 0$ contra $H_1 : \beta_0 \neq 0$.
- Vimos que um teste de nível α para β_0 tem região de rejeição

$$A_1 = \{t_0 > t_{\frac{\alpha}{2},n-2} \text{ ou } t_0 < -t_{\frac{\alpha}{2},n-2}\}$$

em que t_0 é o valor observado da estatística de teste T_0, dada por

$$T_0 = \frac{\hat{\beta}_0 - \beta_{0,0}}{\sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]}}$$

- Dessa forma, precisamos obter a estimativa $\hat{\sigma}^2$.

Sol. (c)

- No item (c) queremos testar hipóteses com respeito ao intercepto do modelo, ou seja, queremos testar, ao nível de significância $\alpha = 5\%$, as hipóteses $H_0 : \beta_0 = 0$ contra $H_1 : \beta_0 \neq 0$.
- Vimos que um teste de nível α para β_0 tem região de rejeição

$$A_1 = \{t_0 > t_{\frac{\alpha}{2},n-2} \text{ ou } t_0 < -t_{\frac{\alpha}{2},n-2}\}$$

em que t_0 é o valor observado da estatística de teste T_0, dada por

$$T_0 = \frac{\hat{\beta}_0 - \beta_{0,0}}{\sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]}}$$

- Dessa forma, precisamos obter a estimativa $\hat{\sigma}^2$.

Sol. (c)

- No item (c) queremos testar hipóteses com respeito ao intercepto do modelo, ou seja, queremos testar, ao nível de significância $\alpha = 5\%$, as hipóteses
  $H_0 : \beta_0 = 0$ contra $H_1 : \beta_0 \neq 0$.
- Vimos que um teste de nível α para β_0 tem região de rejeição

$$A_1 = \{t_0 > t_{\frac{\alpha}{2},n-2} \quad \text{ou} \quad t_0 < -t_{\frac{\alpha}{2},n-2}\}$$

  em que t_0 é o valor observado da estatística de teste T_0, dada por

$$T_0 = \frac{\hat{\beta}_0 - \beta_{0,0}}{\sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]}}$$

- Dessa forma, precisamos obter a estimativa $\hat{\sigma}^2$.

Cont. (c)

- Vimos que $\hat{\sigma}^2 = \frac{SQ_E}{n-2}$, em que $SQ_E = SQ_T - \hat{\beta}_1 S_{xy}$
- Sendo

$$SQ_T = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 \quad \text{e} \quad S_{xy} = \sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y}$$

- Como temos que $\sum_{i=1}^{n} y_i^2 = (0.734)^2 + (0.886)^2 + (1.04)^2 + (1.19)^2 + (1.35)^2 + (1.5)^2 + (1.66)^2 + (1.81)^2 + (1.97)^2 + (2.12)^2 \approx 22.301$, e ja vimos que $\bar{y} = 1.426$, temos que

$$SQ_T = 22.301 - 10(1.426)^2 = 1.96624$$

- Já calculamos S_{xy}, ou seja, obtivemos que $S_{xy} = 23.3766 - 10(1.55)(1.426) = 1.2736.$
- Logo, substituindo em $\hat{\sigma}^2 = \frac{SQ_E}{n-2}$ temos que

$$\hat{\sigma}^2 = \frac{1.96624 - 1.54376(1.2736)}{10-2} \approx 0.000013$$

Cont. (c)

- Vimos que $\hat{\sigma}^2 = \frac{SQ_E}{n-2}$, em que $SQ_E = SQ_T - \hat{\beta}_1 S_{xy}$
- Sendo

$$SQ_T = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 \quad \text{e} \quad S_{xy} = \sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y}$$

- Como temos que $\sum_{i=1}^{n} y_i^2 = (0.734)^2 + (0.886)^2 + (1.04)^2 + (1.19)^2 + (1.35)^2 + (1.5)^2 + (1.66)^2 + (1.81)^2 + (1.97)^2 + (2.12)^2 \approx 22.301$, e ja vimos que $\bar{y} = 1.426$, temos que
$$SQ_T = 22.301 - 10(1.426)^2 = 1.96624$$
- Já calculamos S_{xy}, ou seja, obtivemos que $S_{xy} = 23.3766 - 10(1.55)(1.426) = 1.2736.$
- Logo, substituindo em $\hat{\sigma}^2 = \frac{SQ_E}{n-2}$ temos que

$$\hat{\sigma}^2 = \frac{1.96624 - 1.54376(1.2736)}{10-2} \approx 0.000013$$

Cont. (c)

- Vimos que $\hat{\sigma}^2 = \frac{SQ_E}{n-2}$, em que $SQ_E = SQ_T - \hat{\beta}_1 S_{xy}$
- Sendo

$$SQ_T = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 \quad \text{e} \quad S_{xy} = \sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y}$$

- Como temos que $\sum_{i=1}^{n} y_i^2 = (0.734)^2 + (0.886)^2 + (1.04)^2 + (1.19)^2 + (1.35)^2 + (1.5)^2 + (1.66)^2 + (1.81)^2 + (1.97)^2 + (2.12)^2 \approx 22.301$, e ja vimos que $\bar{y} = 1.426$, temos que

$$SQ_T = 22.301 - 10(1.426)^2 = 1.96624$$

- Já calculamos S_{xy}, ou seja, obtivemos que $S_{xy} = 23.3766 - 10(1.55)(1.426) = 1.2736$.
- Logo, substituindo em $\hat{\sigma}^2 = \frac{SQ_E}{n-2}$ temos que

$$\hat{\sigma}^2 = \frac{1.96624 - 1.54376(1.2736)}{10-2} \approx 0.000013$$

Cont. (c)

- Vimos que $\hat{\sigma}^2 = \frac{SQ_E}{n-2}$, em que $SQ_E = SQ_T - \hat{\beta}_1 S_{xy}$
- Sendo

$$SQ_T = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 \quad \text{e} \quad S_{xy} = \sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y}$$

- Como temos que $\sum_{i=1}^{n} y_i^2 = (0.734)^2 + (0.886)^2 + (1.04)^2 + (1.19)^2 + (1.35)^2 + (1.5)^2 + (1.66)^2 + (1.81)^2 + (1.97)^2 + (2.12)^2 \approx 22.301$, e ja vimos que $\bar{y} = 1.426$, temos que

$$SQ_T = 22.301 - 10(1.426)^2 = 1.96624$$

- Já calculamos S_{xy}, ou seja, obtivemos que $S_{xy} = 23.3766 - 10(1.55)(1.426) = 1.2736$.
- Logo, substituindo em $\hat{\sigma}^2 = \frac{SQ_E}{n-2}$ temos que

$$\hat{\sigma}^2 = \frac{1.96624 - 1.54376(1.2736)}{10 - 2} \approx 0.000013$$

- Vimos que $\hat{\sigma}^2 = \frac{SQ_E}{n-2}$, em que $SQ_E = SQ_T - \hat{\beta}_1 S_{xy}$
- Sendo

$$SQ_T = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 \quad \text{e} \quad S_{xy} = \sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y}$$

- Como temos que $\sum_{i=1}^{n} y_i^2 = (0.734)^2 + (0.886)^2 + (1.04)^2 + (1.19)^2 + (1.35)^2 + (1.5)^2 + (1.66)^2 + (1.81)^2 + (1.97)^2 + (2.12)^2 \approx 22.301$, e ja vimos que $\bar{y} = 1.426$, temos que

$$SQ_T = 22.301 - 10(1.426)^2 = 1.96624$$

- Já calculamos S_{xy}, ou seja, obtivemos que $S_{xy} = 23.3766 - 10(1.55)(1.426) = 1.2736.$
- Logo, substituindo em $\hat{\sigma}^2 = \frac{SQ_E}{n-2}$ temos que

$$\hat{\sigma}^2 = \frac{1.96624 - 1.54376(1.2736)}{10-2} \approx 0.000013$$

Cont. (c)

- Vimos que $\hat{\sigma}^2 = \frac{SQ_E}{n-2}$, em que $SQ_E = SQ_T - \hat{\beta}_1 S_{xy}$
- Sendo

$$SQ_T = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 \quad \text{e} \quad S_{xy} = \sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y}$$

- Como temos que $\sum_{i=1}^{n} y_i^2 = (0.734)^2 + (0.886)^2 + (1.04)^2 + (1.19)^2 + (1.35)^2 + (1.5)^2 + (1.66)^2 + (1.81)^2 + (1.97)^2 + (2.12)^2 \approx 22.301$, e ja vimos que $\bar{y} = 1.426$, temos que

$$SQ_T = 22.301 - 10(1.426)^2 = 1.96624$$

- Já calculamos S_{xy}, ou seja, obtivemos que $S_{xy} = 23.3766 - 10(1.55)(1.426) = 1.2736.$
- Logo, substituindo em $\hat{\sigma}^2 = \frac{SQ_E}{n-2}$ temos que

$$\hat{\sigma}^2 = \frac{1.96624 - 1.54376(1.2736)}{10-2} \approx 0.000013$$

Cont. (c)

- Logo, como a estatística de teste é

$$T_0 = \frac{\hat{\beta}_0 - \beta_{0,0}}{\sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]}}$$

- Então o valor observado da estatística de teste é

$$t_0 = \frac{-0.966824}{\sqrt{(0.000013)\left[\frac{1}{10} + \frac{(1.55)^2}{0.825}\right]}} \approx -154.504$$

- Note que nós já calculamos S_{xx} no denominador de $\hat{\beta}_1$.
- Logo, como $\alpha = 5\%$, temos que $t_{\frac{\alpha}{2},n-2} = t_{0.025,8} = 2.306$.
- Dessa forma, temos que $t_0 < -t_{\frac{\alpha}{2},n-2}$, uma vez que $-154.504 < -2.306$. Portanto, rejeitamos H_0.

Cont. (c)

- Logo, como a estatística de teste é

$$T_0 = \frac{\hat{\beta}_0 - \beta_{0,0}}{\sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]}}$$

- Então o valor observado da estatística de teste é

$$t_0 = \frac{-0.966824}{\sqrt{(0.000013)\left[\frac{1}{10} + \frac{(1.55)^2}{0.825}\right]}} \approx -154.504$$

- Note que nós já calculamos S_{xx} no denominador de $\hat{\beta}_1$.
- Logo, como $\alpha = 5\%$, temos que $t_{\frac{\alpha}{2},n-2} = t_{0.025,8} = 2.306$.
- Dessa forma, temos que $t_0 < -t_{\frac{\alpha}{2},n-2}$, uma vez que $-154.504 < -2.306$. Portanto, rejeitamos H_0.

Cont. (c)

- Logo, como a estatística de teste é

$$T_0 = \frac{\hat{\beta}_0 - \beta_{0,0}}{\sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]}}$$

- Então o valor observado da estatística de teste é

$$t_0 = \frac{-0.966824}{\sqrt{(0.000013) \left[\frac{1}{10} + \frac{(1.55)^2}{0.825}\right]}} \approx -154.504$$

- Note que nós já calculamos S_{xx} no denominador de $\hat{\beta}_1$.
- Logo, como $\alpha = 5\%$, temos que $t_{\frac{\alpha}{2},n-2} = t_{0.025,8} = 2.306$.
- Dessa forma, temos que $t_0 < -t_{\frac{\alpha}{2},n-2}$, uma vez que $-154.504 < -2.306$. Portanto, rejeitamos H_0.

Cont. (c)

- Logo, como a estatística de teste é

$$T_0 = \frac{\hat{\beta}_0 - \beta_{0,0}}{\sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]}}$$

- Então o valor observado da estatística de teste é

$$t_0 = \frac{-0.966824}{\sqrt{(0.000013)\left[\frac{1}{10} + \frac{(1.55)^2}{0.825}\right]}} \approx -154.504$$

- Note que nós já calculamos S_{xx} no denominador de $\hat{\beta}_1$.
- Logo, como $\alpha = 5\%$, temos que $t_{\frac{\alpha}{2},n-2} = t_{0.025,8} = 2.306$.
- Dessa forma, temos que $t_0 < -t_{\frac{\alpha}{2},n-2}$, uma vez que $-154.504 < -2.306$. Portanto, rejeitamos H_0.

Cont. (c)

- Logo, como a estatística de teste é

$$T_0 = \frac{\hat{\beta}_0 - \beta_{0,0}}{\sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]}}$$

- Então o valor observado da estatística de teste é

$$t_0 = \frac{-0.966824}{\sqrt{(0.000013)\left[\frac{1}{10} + \frac{(1.55)^2}{0.825}\right]}} \approx -154.504$$

- Note que nós já calculamos S_{xx} no denominador de $\hat{\beta}_1$.
- Logo, como $\alpha = 5\%$, temos que $t_{\frac{\alpha}{2},n-2} = t_{0.025,8} = 2.306$.
- Dessa forma, temos que $t_0 < -t_{\frac{\alpha}{2},n-2}$, uma vez que $-154.504 < -2.306$. Portanto, rejeitamos H_0.

IC para o intercepto e a inclinação

- Podemos ainda obter estimativas de intervalo de confiança para os parâmetros β_0 e β_1 do modelo de regressão linear simples.
- Assumindo mais uma vez que os erros no modelo de regressão são normal e independentemente distribuídos, então as variáveis aleatórias

$$\frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}} \quad \text{e} \quad \frac{\hat{\beta}_0 - \beta_0}{\sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]}}$$

  são ambas distribuídas como uma t com n-2 graus de liberdade.
- Logo, podemos usar esses fatos para obter IC para β_0 e β_1, procedendo de forma análoga ao que estudamos em IC.

IC para o intercepto e a inclinação

- Podemos ainda obter estimativas de intervalo de confiança para os parâmetros β_0 e β_1 do modelo de regressão linear simples.
- Assumindo mais uma vez que os erros no modelo de regressão são normal e independentemente distribuídos, então as variáveis aleatórias

$$\frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}} \quad \text{e} \quad \frac{\hat{\beta}_0 - \beta_0}{\sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]}}$$

são ambas distribuídas como uma t com n-2 graus de liberdade.

- Logo, podemos usar esses fatos para obter IC para β_0 e β_1, procedendo de forma análoga ao que estudamos em IC.

IC para o intercepto e a inclinação

- Podemos ainda obter estimativas de intervalo de confiança para os parâmetros β_0 e β_1 do modelo de regressão linear simples.
- Assumindo mais uma vez que os erros no modelo de regressão são normal e independentemente distribuídos, então as variáveis aleatórias

$$\frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}} \qquad \text{e} \qquad \frac{\hat{\beta}_0 - \beta_0}{\sqrt{\hat{\sigma}^2\left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]}}$$

  são ambas distribuídas como uma t com n-2 graus de liberdade.
- Logo, podemos usar esses fatos para obter IC para β_0 e β_1, procedendo de forma análoga ao que estudamos em IC.

IC para o intercepto e a inclinação

- Podemos ainda obter estimativas de intervalo de confiança para os parâmetros β_0 e β_1 do modelo de regressão linear simples.
- Assumindo mais uma vez que os erros no modelo de regressão são normal e independentemente distribuídos, então as variáveis aleatórias

$$\frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}} \quad \text{e} \quad \frac{\hat{\beta}_0 - \beta_0}{\sqrt{\hat{\sigma}^2\left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]}}$$

  são ambas distribuídas como uma t com n-2 graus de liberdade.
- Logo, podemos usar esses fatos para obter IC para β_0 e β_1, procedendo de forma análoga ao que estudamos em IC.

IC para o intercepto e a inclinação

- Logo, temos que um IC com $100(1-\alpha)\%$ para β_1 é

$$\hat{\beta}_1 - t_{\alpha/2,n-2}\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} \leq \beta_1 \leq \hat{\beta}_1 + t_{\alpha/2,n-2}\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}$$

- De maneira similar, temos que um IC com $100(1-\alpha)\%$ para β_0 é

$$\hat{\beta}_0 - t_{\alpha/2,n-2}\sqrt{\hat{\sigma}^2\left[\frac{1}{n}+\frac{\bar{x}^2}{S_{xx}}\right]} \leq \beta_0 \leq \hat{\beta}_0 + t_{\alpha/2,n-2}\sqrt{\hat{\sigma}^2\left[\frac{1}{n}+\frac{\bar{x}^2}{S_{xx}}\right]}$$

IC para o intercepto e a inclinação

- Logo, temos que um IC com $100(1-\alpha)\%$ para β_1 é

$$\hat{\beta}_1 - t_{\alpha/2,n-2}\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} \leq \beta_1 \leq \hat{\beta}_1 + t_{\alpha/2,n-2}\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}$$

- De maneira similar, temos que um IC com $100(1-\alpha)\%$ para β_0 é

$$\hat{\beta}_0 - t_{\alpha/2,n-2}\sqrt{\hat{\sigma}^2\left[\frac{1}{n}+\frac{\bar{x}^2}{S_{xx}}\right]} \leq \beta_0 \leq \hat{\beta}_0 + t_{\alpha/2,n-2}\sqrt{\hat{\sigma}^2\left[\frac{1}{n}+\frac{\bar{x}^2}{S_{xx}}\right]}$$

Exemplo

- Por exemplo, se no exemplo anterior quisermos obter um IC com 95% de confiança para a inclinação (β_1).

  Então sabemos que o intervalo é da forma

$$\hat{\beta}_1 - t_{\alpha/2,n-2}\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} \leq \beta_1 \leq \hat{\beta}_1 + t_{\alpha/2,n-2}\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}$$

- Usando as quantidades já calculadas no exercício, temos que um IC com 95% de confiança para β_1 é

$$1.54376 - 2.306\sqrt{\frac{0.000013}{0.825}} \leq \beta_1 \leq 1.54376 + 2.306\sqrt{\frac{0.000013}{0.825}}$$

  ou seja

$$[1.53461, 1.55291].$$

Exemplo

- Por exemplo, se no exemplo anterior quisermos obter um IC com 95% de confiança para a inclinação (β_1).

  Então sabemos que o intervalo é da forma

$$\hat{\beta}_1 - t_{\alpha/2,n-2}\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} \leq \beta_1 \leq \hat{\beta}_1 + t_{\alpha/2,n-2}\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}$$

- Usando as quantidades já calculadas no exercício, temos que um IC com 95% de confiança para β_1 é

$$1.54376 - 2.306\sqrt{\frac{0.000013}{0.825}} \leq \beta_1 \leq 1.54376 + 2.306\sqrt{\frac{0.000013}{0.825}}$$

  ou seja

$$[1.53461, 1.55291].$$

Exemplo

- Por exemplo, se no exemplo anterior quisermos obter um IC com 95% de confiança para a inclinação (β_1).

  Então sabemos que o intervalo é da forma

$$\hat{\beta}_1 - t_{\alpha/2,n-2}\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} \leq \beta_1 \leq \hat{\beta}_1 + t_{\alpha/2,n-2}\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}$$

- Usando as quantidades já calculadas no exercício, temos que um IC com 95% de confiança para β_1 é

$$1.54376 - 2.306\sqrt{\frac{0.000013}{0.825}} \leq \beta_1 \leq 1.54376 + 2.306\sqrt{\frac{0.000013}{0.825}}$$

  ou seja

$$[1.53461, 1.55291].$$

Exemplo

- Por exemplo, se no exemplo anterior quisermos obter um IC com 95% de confiança para a inclinação (β_1).

  Então sabemos que o intervalo é da forma

$$\hat{\beta}_1 - t_{\alpha/2,n-2}\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} \leq \beta_1 \leq \hat{\beta}_1 + t_{\alpha/2,n-2}\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}$$

- Usando as quantidades já calculadas no exercício, temos que um IC com 95% de confiança para β_1 é

$$1.54376 - 2.306\sqrt{\frac{0.000013}{0.825}} \leq \beta_1 \leq 1.54376 + 2.306\sqrt{\frac{0.000013}{0.825}}$$

  ou seja

$$[1.53461, 1.55291].$$

Exemplo

- Por exemplo, se no exemplo anterior quisermos obter um IC com 95% de confiança para a inclinação (β_1).

  Então sabemos que o intervalo é da forma

$$\hat{\beta}_1 - t_{\alpha/2,n-2}\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} \leq \beta_1 \leq \hat{\beta}_1 + t_{\alpha/2,n-2}\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}$$

- Usando as quantidades já calculadas no exercício, temos que um IC com 95% de confiança para β_1 é

$$1.54376 - 2.306\sqrt{\frac{0.000013}{0.825}} \leq \beta_1 \leq 1.54376 + 2.306\sqrt{\frac{0.000013}{0.825}}$$

  ou seja

$$[1.53461, 1.55291].$$

Exemplo

- Por exemplo, se no exemplo anterior quisermos obter um IC com 95% de confiança para a inclinação (β_1).

  Então sabemos que o intervalo é da forma

$$\hat{\beta}_1 - t_{\alpha/2,n-2}\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} \leq \beta_1 \leq \hat{\beta}_1 + t_{\alpha/2,n-2}\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}$$

- Usando as quantidades já calculadas no exercício, temos que um IC com 95% de confiança para β_1 é

$$1.54376 - 2.306\sqrt{\frac{0.000013}{0.825}} \leq \beta_1 \leq 1.54376 + 2.306\sqrt{\frac{0.000013}{0.825}}$$

ou seja

$$[1.53461, 1.55291].$$