Probabilidade e estatística - Aula 21

Regressão linear simples - Aula 1

Dr. Giannini Italino Alves Vieira

Universidade Federal do Ceará - Campus de Crateús

2024

Regressão linear simples

- Muitos problemas em engenharia envolvem explorar relações entre duas ou mais variáveis.
- Por exemplo:
  - A pressão de um gás em um recipiente está relacionada à temperatura;
  - A velocidade da água em um canal aberto está relacionada à largura do canal;
  - O deslocamento de uma partícula, em certo tempo, está relacionada à sua velocidade.
- Note que no último exemplo, temos que se d_0 for o deslocamento da partícula a partir da origem no tempo $t = 0$ e v for a velocidade, então o deslocamento d_t no tempo t é $d_t = d_0 + vt$.
  - Esse é um exemplo de relação linear determinística, porque o modelo prevê perfeitamente o deslocamento (sem considerar erros).

- Muitos problemas em engenharia envolvem explorar relações entre duas ou mais variáveis.
- Por exemplo:
  - A pressão de um gás em um recipiente está relacionada à temperatura;
  - A velocidade da água em um canal aberto está relacionada à largura do canal;
  - O deslocamento de uma partícula, em certo tempo, está relacionada à sua velocidade.
- Note que no último exemplo, temos que se d_0 for o deslocamento da partícula a partir da origem no tempo $t = 0$ e v for a velocidade, então o deslocamento d_t no tempo t é $d_t = d_0 + vt$.
  - Esse é um exemplo de relação linear determinística, porque o modelo prevê perfeitamente o deslocamento (sem considerar erros).

Regressão linear simples

- Muitos problemas em engenharia envolvem explorar relações entre duas ou mais variáveis.
- Por exemplo:
  - A pressão de um gás em um recipiente está relacionada à temperatura;
  - A velocidade da água em um canal aberto está relacionada à largura do canal;
  - O deslocamento de uma partícula, em certo tempo, está relacionada à sua velocidade.
- Note que no último exemplo, temos que se d_0 for o deslocamento da partícula a partir da origem no tempo $t = 0$ e v for a velocidade, então o deslocamento d_t no tempo t é $d_t = d_0 + vt$.
  - Esse é um exemplo de relação linear determinística, porque o modelo prevê perfeitamente o deslocamento (sem considerar erros).

Regressão linear simples

- Muitos problemas em engenharia envolvem explorar relações entre duas ou mais variáveis.
- Por exemplo:
  - A pressão de um gás em um recipiente está relacionada à temperatura;
  - A velocidade da água em um canal aberto está relacionada à largura do canal;
  - O deslocamento de uma partícula, em certo tempo, está relacionada à sua velocidade.
- Note que no último exemplo, temos que se d_0 for o deslocamento da partícula a partir da origem no tempo $t = 0$ e v for a velocidade, então o deslocamento d_t no tempo t é $d_t = d_0 + vt$.
  - Esse é um exemplo de relação linear determinística, porque o modelo prevê perfeitamente o deslocamento (sem considerar erros).

Regressão linear simples

- Muitos problemas em engenharia envolvem explorar relações entre duas ou mais variáveis.
- Por exemplo:
  - A pressão de um gás em um recipiente está relacionada à temperatura;
  - A velocidade da água em um canal aberto está relacionada à largura do canal;
  - O deslocamento de uma partícula, em certo tempo, está relacionada à sua velocidade.
- Note que no último exemplo, temos que se d_0 for o deslocamento da partícula a partir da origem no tempo $t = 0$ e v for a velocidade, então o deslocamento d_t no tempo t é $d_t = d_0 + vt$.
  - Esse é um exemplo de relação linear determinística, porque o modelo prevê perfeitamente o deslocamento (sem considerar erros).

- Muitos problemas em engenharia envolvem explorar relações entre duas ou mais variáveis.
- Por exemplo:
  - A pressão de um gás em um recipiente está relacionada à temperatura;
  - A velocidade da água em um canal aberto está relacionada à largura do canal;
  - O deslocamento de uma partícula, em certo tempo, está relacionada à sua velocidade.
- Note que no último exemplo, temos que se d_0 for o deslocamento da partícula a partir da origem no tempo $t = 0$ e v for a velocidade, então o deslocamento d_t no tempo t é $d_t = d_0 + vt$.
  - Esse é um exemplo de relação linear determinística, porque o modelo prevê perfeitamente o deslocamento (sem considerar erros).

Regressão linear simples

- Muitos problemas em engenharia envolvem explorar relações entre duas ou mais variáveis.
- Por exemplo:
  - A pressão de um gás em um recipiente está relacionada à temperatura;
  - A velocidade da água em um canal aberto está relacionada à largura do canal;
  - O deslocamento de uma partícula, em certo tempo, está relacionada à sua velocidade.
- Note que no último exemplo, temos que se d_0 for o deslocamento da partícula a partir da origem no tempo $t = 0$ e v for a velocidade, então o deslocamento d_t no tempo t é $d_t = d_0 + vt$.
  - Esse é um exemplo de relação linear determinística, porque o modelo prevê perfeitamente o deslocamento (sem considerar erros).

Regressão linear simples

- Contudo, existem situações em que a relação entre as variáveis é não determinística.
- Por exemplo
  - O consumo de energia (y) em uma casa está relacionado com o tamanho (x) da casa.
- Note que no exemplo anterior, y não pode ser previsto perfeitamente a partir do conhecimento de x correspondente.
- A coleção de ferramentas estatísticas que são usadas para modelar e explorar relações entre variáveis que estão relacionadas de maneira não determinística é chamada de análise de regressão.
- Veremos que a ideia da regressão linear simples consiste em estudar a relação entre uma variável, chamada de variável dependente, e uma única variável independente.

Regressão linear simples

- Contudo, existem situações em que a relação entre as variáveis é não determinística.
- Por exemplo
  - O consumo de energia (y) em uma casa está relacionado com o tamanho (x) da casa.
- Note que no exemplo anterior, y não pode ser previsto perfeitamente a partir do conhecimento de x correspondente.
- A coleção de ferramentas estatísticas que são usadas para modelar e explorar relações entre variáveis que estão relacionadas de maneira não determinística é chamada de análise de regressão.
- Veremos que a ideia da regressão linear simples consiste em estudar a relação entre uma variável, chamada de variável dependente, e uma única variável independente.

Regressão linear simples

- Contudo, existem situações em que a relação entre as variáveis é não determinística.
- Por exemplo
  - O consumo de energia (y) em uma casa está relacionado com o tamanho (x) da casa.
- Note que no exemplo anterior, y não pode ser previsto perfeitamente a partir do conhecimento de x correspondente.
- A coleção de ferramentas estatísticas que são usadas para modelar e explorar relações entre variáveis que estão relacionadas de maneira não determinística é chamada de análise de regressão.
- Veremos que a ideia da regressão linear simples consiste em estudar a relação entre uma variável, chamada de variável dependente, e uma única variável independente.

Regressão linear simples

- Contudo, existem situações em que a relação entre as variáveis é não determinística.
- Por exemplo
  - O consumo de energia (y) em uma casa está relacionado com o tamanho (x) da casa.
- Note que no exemplo anterior, y não pode ser previsto perfeitamente a partir do conhecimento de x correspondente.
- A coleção de ferramentas estatísticas que são usadas para modelar e explorar relações entre variáveis que estão relacionadas de maneira não determinística é chamada de análise de regressão.
- Veremos que a ideia da regressão linear simples consiste em estudar a relação entre uma variável, chamada de variável dependente, e uma única variável independente.

Regressão linear simples

- Contudo, existem situações em que a relação entre as variáveis é não determinística.
- Por exemplo
  - O consumo de energia (y) em uma casa está relacionado com o tamanho (x) da casa.
- Note que no exemplo anterior, y não pode ser previsto perfeitamente a partir do conhecimento de x correspondente.
- A coleção de ferramentas estatísticas que são usadas para modelar e explorar relações entre variáveis que estão relacionadas de maneira não determinística é chamada de análise de regressão.
- Veremos que a ideia da regressão linear simples consiste em estudar a relação entre uma variável, chamada de variável dependente, e uma única variável independente.

Regressão linear simples

- Contudo, existem situações em que a relação entre as variáveis é não determinística.
- Por exemplo
  - O consumo de energia (y) em uma casa está relacionado com o tamanho (x) da casa.
- Note que no exemplo anterior, y não pode ser previsto perfeitamente a partir do conhecimento de x correspondente.
- A coleção de ferramentas estatísticas que são usadas para modelar e explorar relações entre variáveis que estão relacionadas de maneira não determinística é chamada de análise de regressão.
- Veremos que a ideia da regressão linear simples consiste em estudar a relação entre uma variável, chamada de variável dependente, e uma única variável independente.

Motivação

- A fim de motivar a ideia considere o exemplo a seguir:

Exemplo: Considere que um laboratório está interessado em medir o efeito da temperatura sobre a potência de um medicamento. Dez amostras de 50g cada foram guardadas a diferentes temperaturas, e após 15 dias mediu-se a potência. Os resultados estão no quadro abaixo:

Temperatura (x)	30°		50°			70°			90°	
Potência (y)	38	43	32	26	33	19	27	23	14	21

- A figura abaixo apresenta o diagrama de dispersão dos dados do exercício.
- Cada par (x_i, y_i) é representado como um ponto plotado em um sistema bidimensional de coordenadas.

Motivação

- A fim de motivar a ideia considere o exemplo a seguir:

Exemplo: Considere que um laboratório está interessado em medir o efeito da temperatura sobre a potência de um medicamento. Dez amostras de 50g cada foram guardadas a diferentes temperaturas, e após 15 dias mediu-se a potência. Os resultados estão no quadro abaixo:

Temperatura (x)	30°		50°			70°			90°	
Potência (y)	38	43	32	26	33	19	27	23	14	21

- A figura abaixo apresenta o diagrama de dispersão dos dados do exercício.
- Cada par (x_i, y_i) é representado como um ponto plotado em um sistema bidimensional de coordenadas.

Motivação

- A fim de motivar a ideia considere o exemplo a seguir:

Exemplo: Considere que um laboratório está interessado em medir o efeito da temperatura sobre a potência de um medicamento. Dez amostras de 50g cada foram guardadas a diferentes temperaturas, e após 15 dias mediu-se a potência. Os resultados estão no quadro abaixo:

Temperatura (x)	30°		50°			70°			90°	
Potência (y)	38	43	32	26	33	19	27	23	14	21

- A figura abaixo apresenta o diagrama de dispersão dos dados do exercício.
- Cada par (x_i, y_i) é representado como um ponto plotado em um sistema bidimensional de coordenadas.

Motivação

- A fim de motivar a ideia considere o exemplo a seguir:

Exemplo: Considere que um laboratório está interessado em medir o efeito da temperatura sobre a potência de um medicamento. Dez amostras de 50g cada foram guardadas a diferentes temperaturas, e após 15 dias mediu-se a potência. Os resultados estão no quadro abaixo:

Temperatura (x)	30°		50°			70°			90°	
Potência (y)	38	43	32	26	33	19	27	23	14	21

- A figura abaixo apresenta o diagrama de dispersão dos dados do exercício.
- Cada par (x_i, y_i) é representado como um ponto plotado em um sistema bidimensional de coordenadas.

Regressão linear simples

Motivação

- A fim de motivar a ideia considere o exemplo a seguir:

Exemplo: Considere que um laboratório está interessado em medir o efeito da temperatura sobre a potência de um medicamento. Dez amostras de 50g cada foram guardadas a diferentes temperaturas, e após 15 dias mediu-se a potência. Os resultados estão no quadro abaixo:

Temperatura (x)	30°		50°			70°			90°	
Potência (y)	38	43	32	26	33	19	27	23	14	21

- A figura abaixo apresenta o diagrama de dispersão dos dados do exercício.
- Cada par (x_i, y_i) é representado como um ponto plotado em um sistema bidimensional de coordenadas.

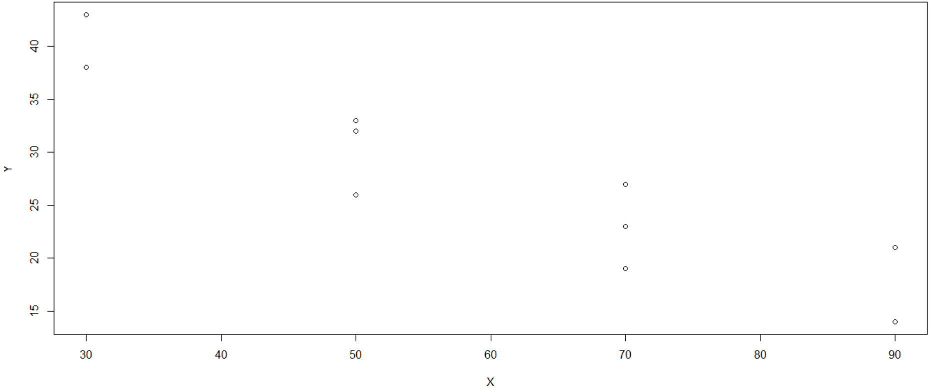

Figure 1: Dados do problema anterior. X: Temperatura, Y: Potência

Motivação

- Note que embora nenhuma curva simples passe através de todos os pontos, há uma forte indicação de que os pontos repousam, dispersos de maneira aleatória, em torno de uma reta.
- Logo, é razoável considerar que para um valor fixo de x o valor real de Y seja determinado da forma

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

- Em que ϵ é o termo de erro aleatório.
- Veremos que o modelo acima será chamado de regressão de regressão linear simples, porque ele só tem somente um regressor (variável independente).

Motivação

- Note que embora nenhuma curva simples passe através de todos os pontos, há uma forte indicação de que os pontos repousam, dispersos de maneira aleatória, em torno de uma reta.
- Logo, é razoável considerar que para um valor fixo de x o valor real de Y seja determinado da forma

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

- Em que ϵ é o termo de erro aleatório.
- Veremos que o modelo acima será chamado de regressão de regressão linear simples, porque ele só tem somente um regressor (variável independente).

Motivação

- Note que embora nenhuma curva simples passe através de todos os pontos, há uma forte indicação de que os pontos repousam, dispersos de maneira aleatória, em torno de uma reta.
- Logo, é razoável considerar que para um valor fixo de x o valor real de Y seja determinado da forma

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

- Em que ϵ é o termo de erro aleatório.
- Veremos que o modelo acima será chamado de regressão de regressão linear simples, porque ele só tem somente um regressor (variável independente).

Motivação

- Note que embora nenhuma curva simples passe através de todos os pontos, há uma forte indicação de que os pontos repousam, dispersos de maneira aleatória, em torno de uma reta.
- Logo, é razoável considerar que para um valor fixo de x o valor real de Y seja determinado da forma

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

- Em que ϵ é o termo de erro aleatório.
- Veremos que o modelo acima será chamado de regressão de regressão linear simples, porque ele só tem somente um regressor (variável independente).

Motivação

- Note que embora nenhuma curva simples passe através de todos os pontos, há uma forte indicação de que os pontos repousam, dispersos de maneira aleatória, em torno de uma reta.
- Logo, é razoável considerar que para um valor fixo de x o valor real de Y seja determinado da forma

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

- Em que ϵ é o termo de erro aleatório.
- Veremos que o modelo acima será chamado de regressão de regressão linear simples, porque ele só tem somente um regressor (variável independente).

Regressão linear simples

- O caso de regressão linear simples considera um único regressor (ou preditor) x e uma variável dependente (ou variável resposta) Y.
- Suponha que a relação verdadeira entre Y e x seja uma linha reta e que a observação Y em cada nível de x seja uma variável aleatória.
- O modelo de regressão linear simples considera que cada observação, Y, possa ser descrita pelo seguinte 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).

Regressão linear simples

- O caso de regressão linear simples considera um único regressor (ou preditor) x e uma variável dependente (ou variável resposta) Y.
- Suponha que a relação verdadeira entre Y e x seja uma linha reta e que a observação Y em cada nível de x seja uma variável aleatória.
- O modelo de regressão linear simples considera que cada observação, Y, possa ser descrita pelo seguinte 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).

Regressão linear simples

- O caso de regressão linear simples considera um único regressor (ou preditor) x e uma variável dependente (ou variável resposta) Y.
- Suponha que a relação verdadeira entre Y e x seja uma linha reta e que a observação Y em cada nível de x seja uma variável aleatória.
- O modelo de regressão linear simples considera que cada observação, Y, possa ser descrita pelo seguinte 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).

Regressão linear simples

- O caso de regressão linear simples considera um único regressor (ou preditor) x e uma variável dependente (ou variável resposta) Y.
- Suponha que a relação verdadeira entre Y e x seja uma linha reta e que a observação Y em cada nível de x seja uma variável aleatória.
- O modelo de regressão linear simples considera que cada observação, Y, possa ser descrita pelo seguinte 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).

Regressão linear simples

- O caso de regressão linear simples considera um único regressor (ou preditor) x e uma variável dependente (ou variável resposta) Y.
- Suponha que a relação verdadeira entre Y e x seja uma linha reta e que a observação Y em cada nível de x seja uma variável aleatória.
- O modelo de regressão linear simples considera que cada observação, Y, possa ser descrita pelo seguinte 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).

- Dadas n observações da variável X, digamos, $x_1, x_2, \ldots, x_n$, então teremos n variáveis $Y_1, Y_2, \ldots, Y_n$ satisfazendo a equação,

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad i = 1, 2, \ldots, n$$

- Note que uma amostra de tamanho n consiste de n pares de valores (x_i, y_i) com $i = 1, 2, \ldots, n$ que devem satisfazer o seguinte modelo, $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, $i = 1, 2, \ldots, n$.
- Suponha que tenhamos n pares de observações, digamos

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$$

- A figura abaixo ilustra um diagrama de dispersão dos dados observados e uma candidata para a linha estimada de regressão.

- Dadas n observações da variável X, digamos, $x_1, x_2, \ldots, x_n$, então teremos n variáveis $Y_1, Y_2, \ldots, Y_n$ satisfazendo a equação,

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad i = 1, 2, \ldots, n$$

- Note que uma amostra de tamanho n consiste de n pares de valores (x_i, y_i) com $i = 1, 2, \ldots, n$ que devem satisfazer o seguinte modelo, $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, $i = 1, 2, \ldots, n$.
- Suponha que tenhamos n pares de observações, digamos

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$$

- A figura abaixo ilustra um diagrama de dispersão dos dados observados e uma candidata para a linha estimada de regressão.

- Dadas n observações da variável X, digamos, $x_1, x_2, \ldots, x_n$, então teremos n variáveis $Y_1, Y_2, \ldots, Y_n$ satisfazendo a equação,

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad i = 1, 2, \ldots, n$$

- Note que uma amostra de tamanho n consiste de n pares de valores (x_i, y_i) com $i = 1, 2, \ldots, n$ que devem satisfazer o seguinte modelo, $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, $i = 1, 2, \ldots, n$.
- Suponha que tenhamos n pares de observações, digamos

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$$

- A figura abaixo ilustra um diagrama de dispersão dos dados observados e uma candidata para a linha estimada de regressão.

- Dadas n observações da variável X, digamos, $x_1, x_2, \ldots, x_n$, então teremos n variáveis $Y_1, Y_2, \ldots, Y_n$ satisfazendo a equação,

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad i = 1, 2, \ldots, n$$

- Note que uma amostra de tamanho n consiste de n pares de valores (x_i, y_i) com $i = 1, 2, \ldots, n$ que devem satisfazer o seguinte modelo, $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, $i = 1, 2, \ldots, n$.
- Suponha que tenhamos n pares de observações, digamos

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$$

- A figura abaixo ilustra um diagrama de dispersão dos dados observados e uma candidata para a linha estimada de regressão.

- Dadas n observações da variável X, digamos, $x_1, x_2, \ldots, x_n$, então teremos n variáveis $Y_1, Y_2, \ldots, Y_n$ satisfazendo a equação,

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad i = 1, 2, \ldots, n$$

- Note que uma amostra de tamanho n consiste de n pares de valores (x_i, y_i) com $i = 1, 2, \ldots, n$ que devem satisfazer o seguinte modelo, $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, $i = 1, 2, \ldots, n$.
- Suponha que tenhamos n pares de observações, digamos

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$$

- A figura abaixo ilustra um diagrama de dispersão dos dados observados e uma candidata para a linha estimada de regressão.

- Dadas n observações da variável X, digamos, $x_1, x_2, \ldots, x_n$, então teremos n variáveis $Y_1, Y_2, \ldots, Y_n$ satisfazendo a equação,

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad i = 1, 2, \ldots, n$$

- Note que uma amostra de tamanho n consiste de n pares de valores (x_i, y_i) com $i = 1, 2, \ldots, n$ que devem satisfazer o seguinte modelo, $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, $i = 1, 2, \ldots, n$.
- Suponha que tenhamos n pares de observações, digamos

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$$

- A figura abaixo ilustra um diagrama de dispersão dos dados observados e uma candidata para a linha estimada de regressão.

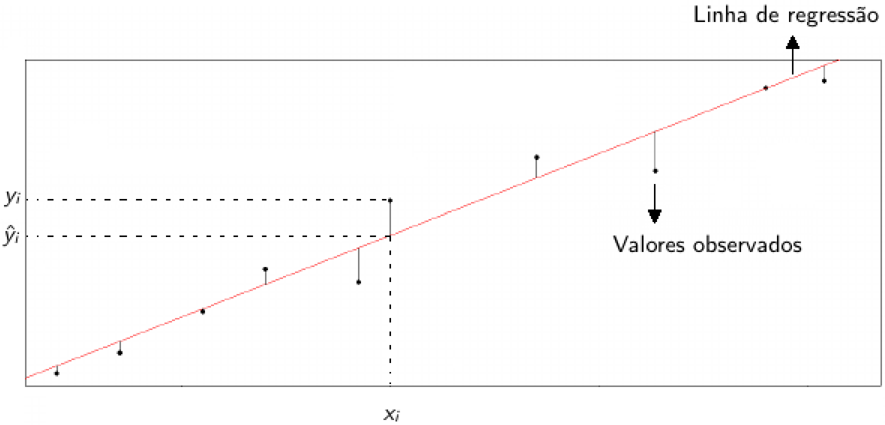

Figure 2: Desvios dos dados em relação ao modelo estimado de regressão

Estimando β_0 e β_1 - Método dos mínimos Quadrados

- Logo, o problema consiste em encontrar a linha que seja, em algum sentido, o "melhor ajuste", para os dados.
- Uma ideia, comumente usada na estimação dos parâmetros do modelo de regressão é estimar β_0 e β_1 de maneira a minimizar a soma dos quadrados dos desvios, ou seja, minimizar a soma

$$L(\beta_0, \beta_1) = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2$$

- Temos então um problema de encontrar o mínimo de uma função de duas variáveis β_0 e β_1.

Estimando β_0 e β_1 - Método dos mínimos Quadrados

- Logo, o problema consiste em encontrar a linha que seja, em algum sentido, o "melhor ajuste", para os dados.
- Uma ideia, comumente usada na estimação dos parâmetros do modelo de regressão é estimar β_0 e β_1 de maneira a minimizar a soma dos quadrados dos desvios, ou seja, minimizar a soma

$$L(\beta_0, \beta_1) = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2$$

- Temos então um problema de encontrar o mínimo de uma função de duas variáveis β_0 e β_1.

Estimando β_0 e β_1 - Método dos mínimos Quadrados

- Logo, o problema consiste em encontrar a linha que seja, em algum sentido, o "melhor ajuste", para os dados.
- Uma ideia, comumente usada na estimação dos parâmetros do modelo de regressão é estimar β_0 e β_1 de maneira a minimizar a soma dos quadrados dos desvios, ou seja, minimizar a soma

$$L(\beta_0, \beta_1) = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2$$

- Temos então um problema de encontrar o mínimo de uma função de duas variáveis β_0 e β_1.

Estimando β_0 e β_1 - Método dos mínimos Quadrados

- Logo, o problema consiste em encontrar a linha que seja, em algum sentido, o "melhor ajuste", para os dados.
- Uma ideia, comumente usada na estimação dos parâmetros do modelo de regressão é estimar β_0 e β_1 de maneira a minimizar a soma dos quadrados dos desvios, ou seja, minimizar a soma

$$L(\beta_0, \beta_1) = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2$$

- Temos então um problema de encontrar o mínimo de uma função de duas variáveis β_0 e β_1.

Estimando β_0 e β_1 - Método dos mínimos Quadrados

- Derivando $L(\beta_0, \beta_1)$ com respeito a β_0 e β_1, igualando a zero, e simplificando temos que as soluções $\hat{\beta}_0$ e $\hat{\beta}_1$ devem satisfazer as equações

$$n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$$

$$\hat{\beta}_0 \sum_{i=1}^{n} x_i + \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} y_i x_i$$

- As soluções dessas equações, para $\hat{\beta}_0$ e $\hat{\beta}_1$, resultam nos estimadores de mínimos quadrados de β_0 e β_1.

Estimando β_0 e β_1 - Método dos mínimos Quadrados

- Derivando $L(\beta_0, \beta_1)$ com respeito a β_0 e β_1, igualando a zero, e simplificando temos que as soluções $\hat{\beta}_0$ e $\hat{\beta}_1$ devem satisfazer as equações

$$n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$$

$$\hat{\beta}_0 \sum_{i=1}^{n} x_i + \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} y_i x_i$$

- As soluções dessas equações, para $\hat{\beta}_0$ e $\hat{\beta}_1$, resultam nos estimadores de mínimos quadrados de β_0 e β_1.

Estimando β_0 e β_1 - Método dos mínimos Quadrados

- Derivando $L(\beta_0, \beta_1)$ com respeito a β_0 e β_1, igualando a zero, e simplificando temos que as soluções $\hat{\beta}_0$ e $\hat{\beta}_1$ devem satisfazer as equações

$$n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$$

$$\hat{\beta}_0 \sum_{i=1}^{n} x_i + \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} y_i x_i$$

- As soluções dessas equações, para $\hat{\beta}_0$ e $\hat{\beta}_1$, resultam nos estimadores de mínimos quadrados de β_0 e β_1.

Estimando β_0 e β_1 - Método dos mínimos Quadrados

- Derivando $L(\beta_0, \beta_1)$ com respeito a β_0 e β_1, igualando a zero, e simplificando temos que as soluções $\hat{\beta}_0$ e $\hat{\beta}_1$ devem satisfazer as equações

$$n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$$

$$\hat{\beta}_0 \sum_{i=1}^{n} x_i + \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} y_i x_i$$

- As soluções dessas equações, para $\hat{\beta}_0$ e $\hat{\beta}_1$, resultam nos estimadores de mínimos quadrados de β_0 e β_1.

Regressão linear simples

Estimando β_0 e β_1 - Método dos mínimos Quadrados

- 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, temos que a linha de regressão ajustada (modelo ajustado ou estimado) é consequentemente
$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

Estimando β_0 e β_1 - Método dos mínimos Quadrados

- 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, temos que a linha de regressão ajustada (modelo ajustado ou estimado) é consequentemente

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

Regressão linear simples

Estimando β_0 e β_1 - Método dos mínimos Quadrados

- 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, temos que a linha de regressão ajustada (modelo ajustado ou estimado) é consequentemente
$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

Estimando β_0 e β_1 - Método dos mínimos Quadrados

- 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, temos que a linha de regressão ajustada (modelo ajustado ou estimado) é consequentemente

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

Estimando β_0 e β_1 - Método dos mínimos Quadrados

- 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, temos que a linha de regressão ajustada (modelo ajustado ou estimado) é consequentemente

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

Estimando β_0 e β_1 - Método dos mínimos Quadrados

- 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, temos que a linha de regressão ajustada (modelo ajustado ou estimado) é consequentemente

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

Exemplo

Exemplo: Vamos voltar ao exemplo inicial dessa aula, ou seja, ao problema de medir o efeito da temperatura sobre a potência de um medicamento. Recorde que nossos dados são:

Temperatura (x)	30°		50°			70°			90°	
Potência (y)	38	43	32	26	33	19	27	23	14	21

- Vamos ajustar um modelo de regressão para esses dados.
- Recorde que para ajustar o modelo precisamos obter $\hat{\beta}_0$ e $\hat{\beta}_1$
- Note que nossos dados são:

$$(30, 38), (30, 43), (50, 32), (50, 26), (50, 33)$$

$$(70, 19), (70, 27), (70, 23), (90, 14), (90, 21)$$

Exemplo

Exemplo: Vamos voltar ao exemplo inicial dessa aula, ou seja, ao problema de medir o efeito da temperatura sobre a potência de um medicamento. Recorde que nossos dados são:

Temperatura (x)	30°		50°			70°			90°	
Potência (y)	38	43	32	26	33	19	27	23	14	21

- Vamos ajustar um modelo de regressão para esses dados.
- Recorde que para ajustar o modelo precisamos obter $\hat{\beta}_0$ e $\hat{\beta}_1$
- Note que nossos dados são:

$$(30, 38), (30, 43), (50, 32), (50, 26), (50, 33)$$
$$(70, 19), (70, 27), (70, 23), (90, 14), (90, 21)$$

Exemplo

Exemplo: Vamos voltar ao exemplo inicial dessa aula, ou seja, ao problema de medir o efeito da temperatura sobre a potência de um medicamento. Recorde que nossos dados são:

Temperatura (x)	30°		50°			70°			90°	
Potência (y)	38	43	32	26	33	19	27	23	14	21

- Vamos ajustar um modelo de regressão para esses dados.
- Recorde que para ajustar o modelo precisamos obter $\hat{\beta}_0$ e $\hat{\beta}_1$
- Note que nossos dados são:

$$(30, 38), (30, 43), (50, 32), (50, 26), (50, 33)$$

$$(70, 19), (70, 27), (70, 23), (90, 14), (90, 21)$$

Exemplo

Exemplo: Vamos voltar ao exemplo inicial dessa aula, ou seja, ao problema de medir o efeito da temperatura sobre a potência de um medicamento. Recorde que nossos dados são:

Temperatura (x)	30°		50°			70°			90°	
Potência (y)	38	43	32	26	33	19	27	23	14	21

- Vamos ajustar um modelo de regressão para esses dados.
- Recorde que para ajustar o modelo precisamos obter $\hat{\beta}_0$ e $\hat{\beta}_1$
- Note que nossos dados são:

$$(30, 38), (30, 43), (50, 32), (50, 26), (50, 33)$$

$$(70, 19), (70, 27), (70, 23), (90, 14), (90, 21)$$

Exemplo

Exemplo: Vamos voltar ao exemplo inicial dessa aula, ou seja, ao problema de medir o efeito da temperatura sobre a potência de um medicamento. Recorde que nossos dados são:

Temperatura (x)	30°		50°			70°			90°	
Potência (y)	38	43	32	26	33	19	27	23	14	21

- Vamos ajustar um modelo de regressão para esses dados.
- Recorde que para ajustar o modelo precisamos obter $\hat{\beta}_0$ e $\hat{\beta}_1$
- Note que nossos dados são:

$$(30,38),(30,43),(50,32),(50,26),(50,33)$$

$$(70,19),(70,27),(70,23),(90,14),(90,21)$$

Exemplo

Exemplo: Vamos voltar ao exemplo inicial dessa aula, ou seja, ao problema de medir o efeito da temperatura sobre a potência de um medicamento. Recorde que nossos dados são:

Temperatura (x)	30°		50°			70°			90°	
Potência (y)	38	43	32	26	33	19	27	23	14	21

- Vamos ajustar um modelo de regressão para esses dados.
- Recorde que para ajustar o modelo precisamos obter $\hat{\beta}_0$ e $\hat{\beta}_1$
- Note que nossos dados são:

$$(30, 38), (30, 43), (50, 32), (50, 26), (50, 33)$$
$$(70, 19), (70, 27), (70, 23), (90, 14), (90, 21)$$

Cont. do exemplo

- 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
  - $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{30+30+50+50+50+70+70+70+90+90}{10} = 60;$
  - $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{38+43+32+26+33+19+27+23+14+21}{10} = 27.6;$
  - $\sum_{i=1}^{n} y_i x_i = (38\cdot 30) + (43\cdot 30) + (32\cdot 50) + (26\cdot 50) + (33\cdot 50) + (19\cdot 70) + (27\cdot 70) + (23\cdot 70) + (14\cdot 90) + (21\cdot 90) = 14960;$
  - $\sum_{i=1}^{n} x_i^2 = 30^2 + 30^2 + 50^2 + 50^2 + 50^2 + 70^2 + 70^2 + 70^2 + 90^2 + 90^2 = 40200.$
- Logo, temos que

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2} = \frac{14960 - 10(60)(27.6)}{40200 - 10(60)^2} \approx -0.38095$$

e

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = (27.6) - (-0.38095)(60) \approx 50.457.$$

- Portanto, o modelo ajustado é $\hat{y}_i = 50.457 - 0.38095 x_i$

Cont. do exemplo

- 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
  - $\bar{x} = \frac{1}{n}\sum_{i=1}^n x_i = \frac{30+30+50+50+50+70+70+70+90+90}{10} = 60;$
  - $\bar{y} = \frac{1}{n}\sum_{i=1}^n y_i = \frac{38+43+32+26+33+19+27+23+14+21}{10} = 27.6;$
  - $\sum_{i=1}^n y_i x_i = (38\cdot 30) + (43\cdot 30) + (32\cdot 50) + (26\cdot 50) + (33\cdot 50) + (19\cdot 70) + (27\cdot 70) + (23\cdot 70) + (14\cdot 90) + (21\cdot 90) = 14960;$
  - $\sum_{i=1}^n x_i^2 = 30^2 + 30^2 + 50^2 + 50^2 + 50^2 + 70^2 + 70^2 + 70^2 + 90^2 + 90^2 = 40200.$
- Logo, temos que

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^n x_i^2 - n\bar{x}^2} = \frac{14960 - 10(60)(27.6)}{40200 - 10(60)^2} \approx -0.38095$$

e

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x} = (27.6) - (-0.38095)(60) \approx 50.457.$$

- Portanto, o modelo ajustado é $\hat{y}_i = 50.457 - 0.38095x_i$

Cont. do exemplo

- 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
  - $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{30+30+50+50+50+70+70+70+90+90}{10} = 60;$
  - $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{38+43+32+26+33+19+27+23+14+21}{10} = 27.6;$
  - $\sum_{i=1}^{n} y_i x_i = (38 \cdot 30) + (43 \cdot 30) + (32 \cdot 50) + (26 \cdot 50) + (33 \cdot 50) + (19 \cdot 70) + (27 \cdot 70) + (23 \cdot 70) + (14 \cdot 90) + (21 \cdot 90) = 14960;$
  - $\sum_{i=1}^{n} x_i^2 = 30^2 + 30^2 + 50^2 + 50^2 + 50^2 + 70^2 + 70^2 + 70^2 + 90^2 + 90^2 = 40200.$
- Logo, temos que

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2} = \frac{14960 - 10(60)(27.6)}{40200 - 10(60)^2} \approx -0.38095$$

e

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = (27.6) - (-0.38095)(60) \approx 50.457.$$

- Portanto, o modelo ajustado é $\hat{y}_i = 50.457 - 0.38095 x_i$

Cont. do exemplo

- 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
  - $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{30+30+50+50+50+70+70+70+90+90}{10} = 60;$
  - $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{38+43+32+26+33+19+27+23+14+21}{10} = 27.6;$
  - $\sum_{i=1}^{n} y_i x_i = (38 \cdot 30) + (43 \cdot 30) + (32 \cdot 50) + (26 \cdot 50) + (33 \cdot 50) + (19 \cdot 70) + (27 \cdot 70) + (23 \cdot 70) + (14 \cdot 90) + (21 \cdot 90) = 14960;$
  - $\sum_{i=1}^{n} x_i^2 = 30^2 + 30^2 + 50^2 + 50^2 + 50^2 + 70^2 + 70^2 + 70^2 + 90^2 + 90^2 = 40200.$
- Logo, temos que

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2} = \frac{14960 - 10(60)(27.6)}{40200 - 10(60)^2} \approx -0.38095$$

e

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x} = (27.6) - (-0.38095)(60) \approx 50.457.$$

- Portanto, o modelo ajustado é $\hat{y}_i = 50.457 - 0.38095x_i$

Cont. do exemplo

- 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
  - $\bar{x} = \frac{1}{n}\sum_{i=1}^n x_i = \frac{30+30+50+50+50+70+70+70+90+90}{10} = 60;$
  - $\bar{y} = \frac{1}{n}\sum_{i=1}^n y_i = \frac{38+43+32+26+33+19+27+23+14+21}{10} = 27.6;$
  - $\sum_{i=1}^n y_i x_i = (38\cdot 30) + (43\cdot 30) + (32\cdot 50) + (26\cdot 50) + (33\cdot 50) + (19\cdot 70) + (27\cdot 70) + (23\cdot 70) + (14\cdot 90) + (21\cdot 90) = 14960;$
  - $\sum_{i=1}^n x_i^2 = 30^2 + 30^2 + 50^2 + 50^2 + 50^2 + 70^2 + 70^2 + 70^2 + 90^2 + 90^2 = 40200.$
- Logo, temos que

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^n x_i^2 - n\bar{x}^2} = \frac{14960 - 10(60)(27.6)}{40200 - 10(60)^2} \approx -0.38095$$

e

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x} = (27.6) - (-0.38095)(60) \approx 50.457.$$

- Portanto, o modelo ajustado é $\hat{y}_i = 50.457 - 0.38095x_i$

Cont. do exemplo

- 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
  - $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{30+30+50+50+50+70+70+70+90+90}{10} = 60$;
  - $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{38+43+32+26+33+19+27+23+14+21}{10} = 27.6$;
  - $\sum_{i=1}^{n} y_i x_i = (38 \cdot 30) + (43 \cdot 30) + (32 \cdot 50) + (26 \cdot 50) + (33 \cdot 50) + (19 \cdot 70) + (27 \cdot 70) + (23 \cdot 70) + (14 \cdot 90) + (21 \cdot 90) = 14960$;
  - $\sum_{i=1}^{n} x_i^2 = 30^2 + 30^2 + 50^2 + 50^2 + 50^2 + 70^2 + 70^2 + 70^2 + 90^2 + 90^2 = 40200$.
- Logo, temos que

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2} = \frac{14960 - 10(60)(27.6)}{40200 - 10(60)^2} \approx -0.38095$$

e

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = (27.6) - (-0.38095)(60) \approx 50.457.$$

- Portanto, o modelo ajustado é $\hat{y}_i = 50.457 - 0.38095 x_i$

Cont. do exemplo

- 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
  - $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{30+30+50+50+50+70+70+70+90+90}{10} = 60;$
  - $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{38+43+32+26+33+19+27+23+14+21}{10} = 27.6;$
  - $\sum_{i=1}^{n} y_i x_i = (38 \cdot 30) + (43 \cdot 30) + (32 \cdot 50) + (26 \cdot 50) + (33 \cdot 50) + (19 \cdot 70) + (27 \cdot 70) + (23 \cdot 70) + (14 \cdot 90) + (21 \cdot 90) = 14960;$
  - $\sum_{i=1}^{n} x_i^2 = 30^2 + 30^2 + 50^2 + 50^2 + 50^2 + 70^2 + 70^2 + 70^2 + 90^2 + 90^2 = 40200.$
- Logo, temos que

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2} = \frac{14960 - 10(60)(27.6)}{40200 - 10(60)^2} \approx -0.38095$$

e

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = (27.6) - (-0.38095)(60) \approx 50.457.$$

- Portanto, o modelo ajustado é $\hat{y}_i = 50.457 - 0.38095 x_i$

Cont. do exemplo

- 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
  - $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{30+30+50+50+50+70+70+70+90+90}{10} = 60;$
  - $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{38+43+32+26+33+19+27+23+14+21}{10} = 27.6;$
  - $\sum_{i=1}^{n} y_i x_i = (38\cdot 30) + (43\cdot 30) + (32\cdot 50) + (26\cdot 50) + (33\cdot 50) + (19\cdot 70) + (27\cdot 70) + (23\cdot 70) + (14\cdot 90) + (21\cdot 90) = 14960;$
  - $\sum_{i=1}^{n} x_i^2 = 30^2 + 30^2 + 50^2 + 50^2 + 50^2 + 70^2 + 70^2 + 70^2 + 90^2 + 90^2 = 40200.$
- Logo, temos que

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2} = \frac{14960 - 10(60)(27.6)}{40200 - 10(60)^2} \approx -0.38095$$

e

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x} = (27.6) - (-0.38095)(60) \approx 50.457.$$

- Portanto, o modelo ajustado é $\hat{y}_i = 50.457 - 0.38095x_i$

Cont. do exemplo

- 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
  - $\bar{x} = \frac{1}{n}\sum_{i=1}^n x_i = \frac{30+30+50+50+50+70+70+70+90+90}{10} = 60;$
  - $\bar{y} = \frac{1}{n}\sum_{i=1}^n y_i = \frac{38+43+32+26+33+19+27+23+14+21}{10} = 27.6;$
  - $\sum_{i=1}^n y_i x_i = (38 \cdot 30) + (43 \cdot 30) + (32 \cdot 50) + (26 \cdot 50) + (33 \cdot 50) + (19 \cdot 70) + (27 \cdot 70) + (23 \cdot 70) + (14 \cdot 90) + (21 \cdot 90) = 14960;$
  - $\sum_{i=1}^n x_i^2 = 30^2 + 30^2 + 50^2 + 50^2 + 50^2 + 70^2 + 70^2 + 70^2 + 90^2 + 90^2 = 40200.$
- Logo, temos que

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^n x_i^2 - n\bar{x}^2} = \frac{14960 - 10(60)(27.6)}{40200 - 10(60)^2} \approx -0.38095$$

e

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x} = (27.6) - (-0.38095)(60) \approx 50.457.$$

- Portanto, o modelo ajustado é $\hat{y}_i = 50.457 - 0.38095x_i$

Cont. do exemplo

- 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
  - $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{30+30+50+50+50+70+70+70+90+90}{10} = 60$;
  - $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{38+43+32+26+33+19+27+23+14+21}{10} = 27.6$;
  - $\sum_{i=1}^{n} y_i x_i = (38 \cdot 30) + (43 \cdot 30) + (32 \cdot 50) + (26 \cdot 50) + (33 \cdot 50) + (19 \cdot 70) + (27 \cdot 70) + (23 \cdot 70) + (14 \cdot 90) + (21 \cdot 90) = 14960$;
  - $\sum_{i=1}^{n} x_i^2 = 30^2 + 30^2 + 50^2 + 50^2 + 50^2 + 70^2 + 70^2 + 70^2 + 90^2 + 90^2 = 40200$.
- Logo, temos que

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2} = \frac{14960 - 10(60)(27.6)}{40200 - 10(60)^2} \approx -0.38095$$

e

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x} = (27.6) - (-0.38095)(60) \approx 50.457.$$

- Portanto, o modelo ajustado é $\hat{y}_i = 50.457 - 0.38095x_i$

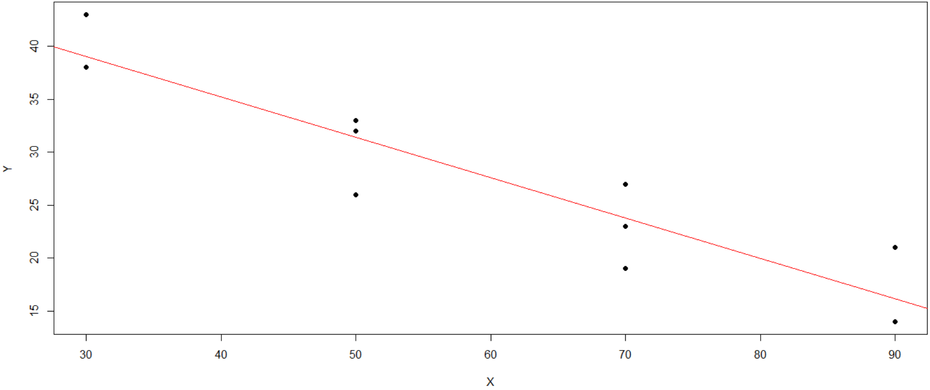

Figure 3: Dados do problema anterior. X: Temperatura, Y: Potência