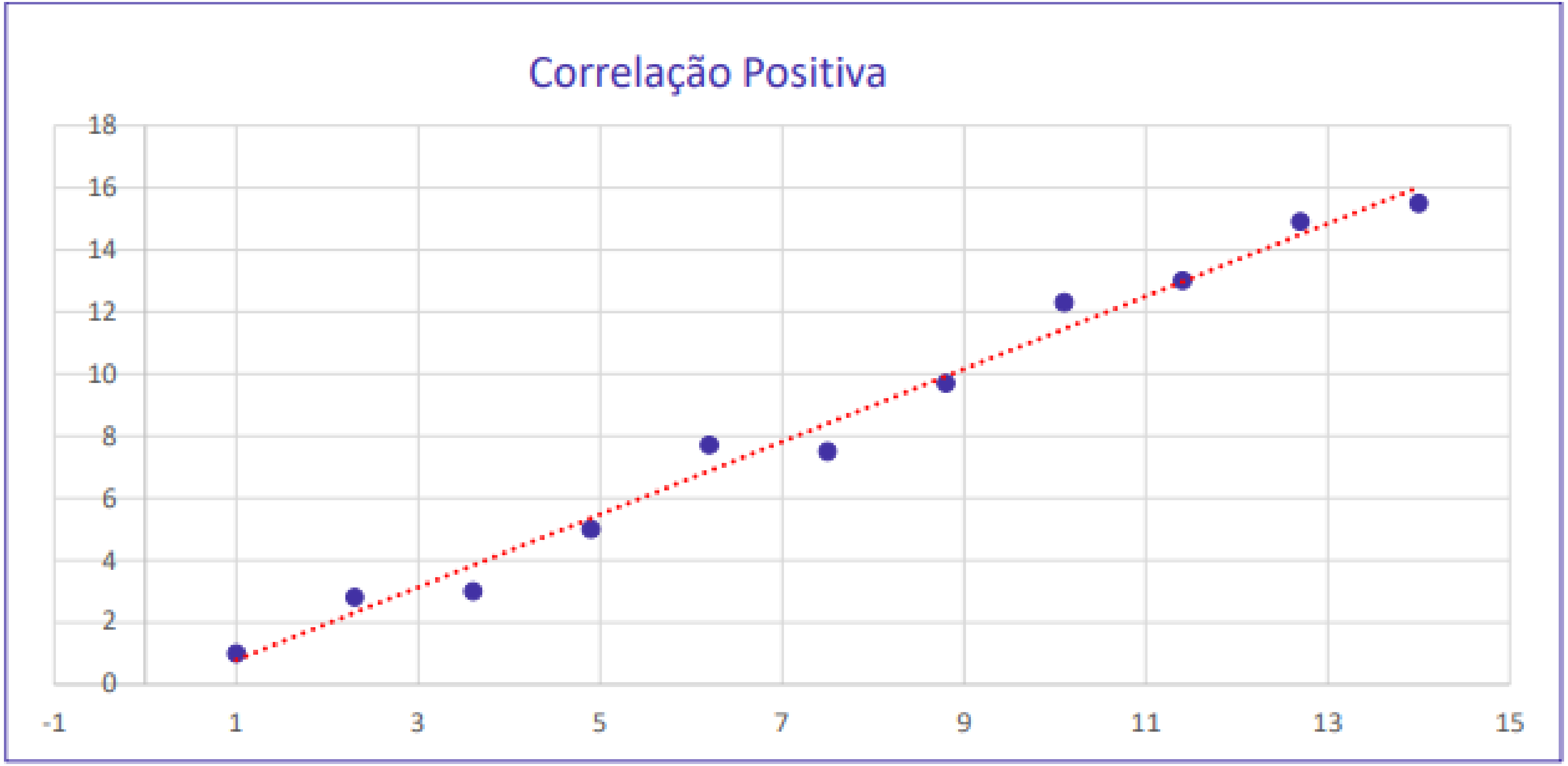
By @kakashi_copiador

Estratégia
Concursos

CORRELAÇÃO

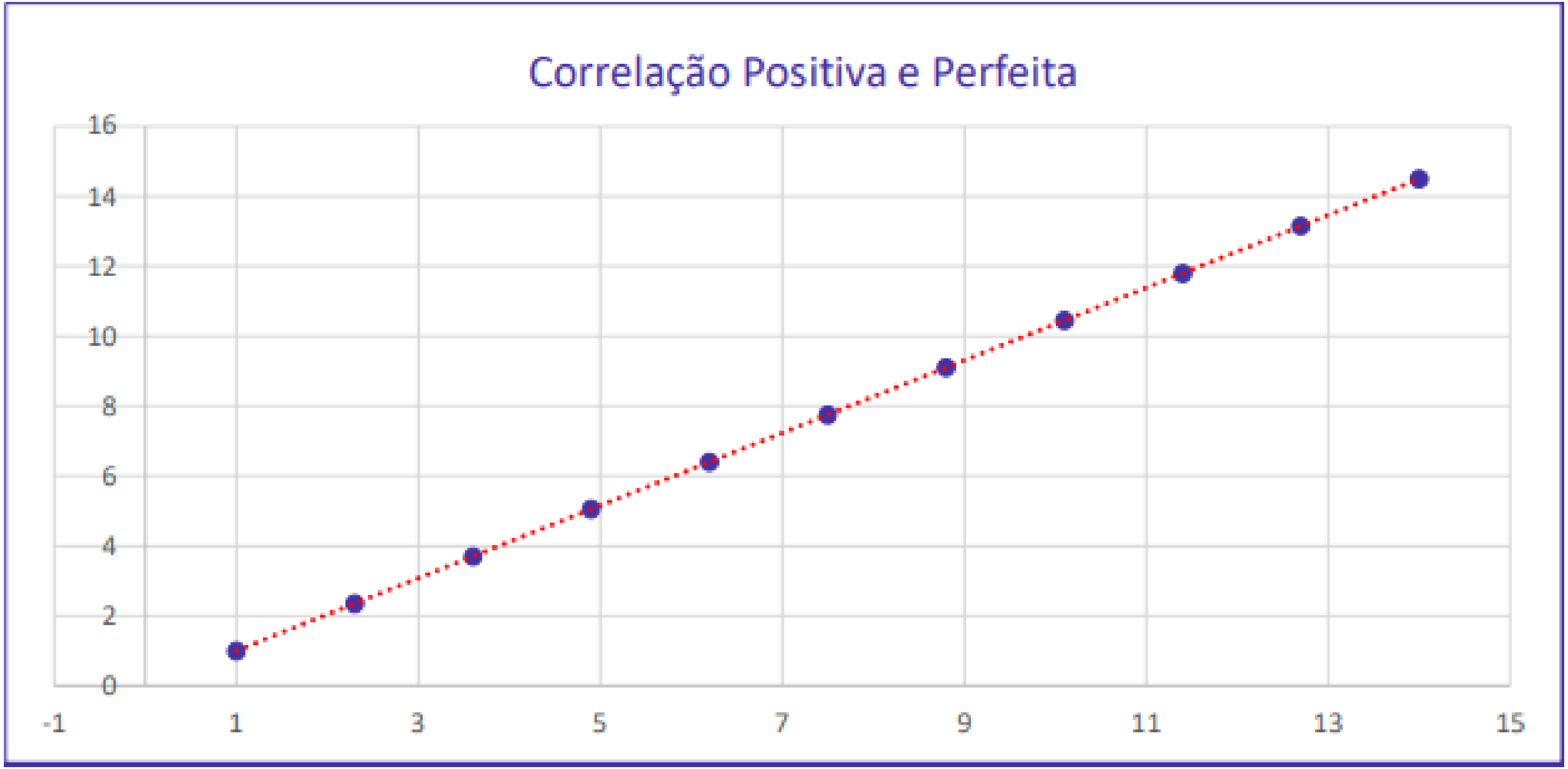
Prof. Jhoni Zini

CORRELAÇÃO POSITIVA

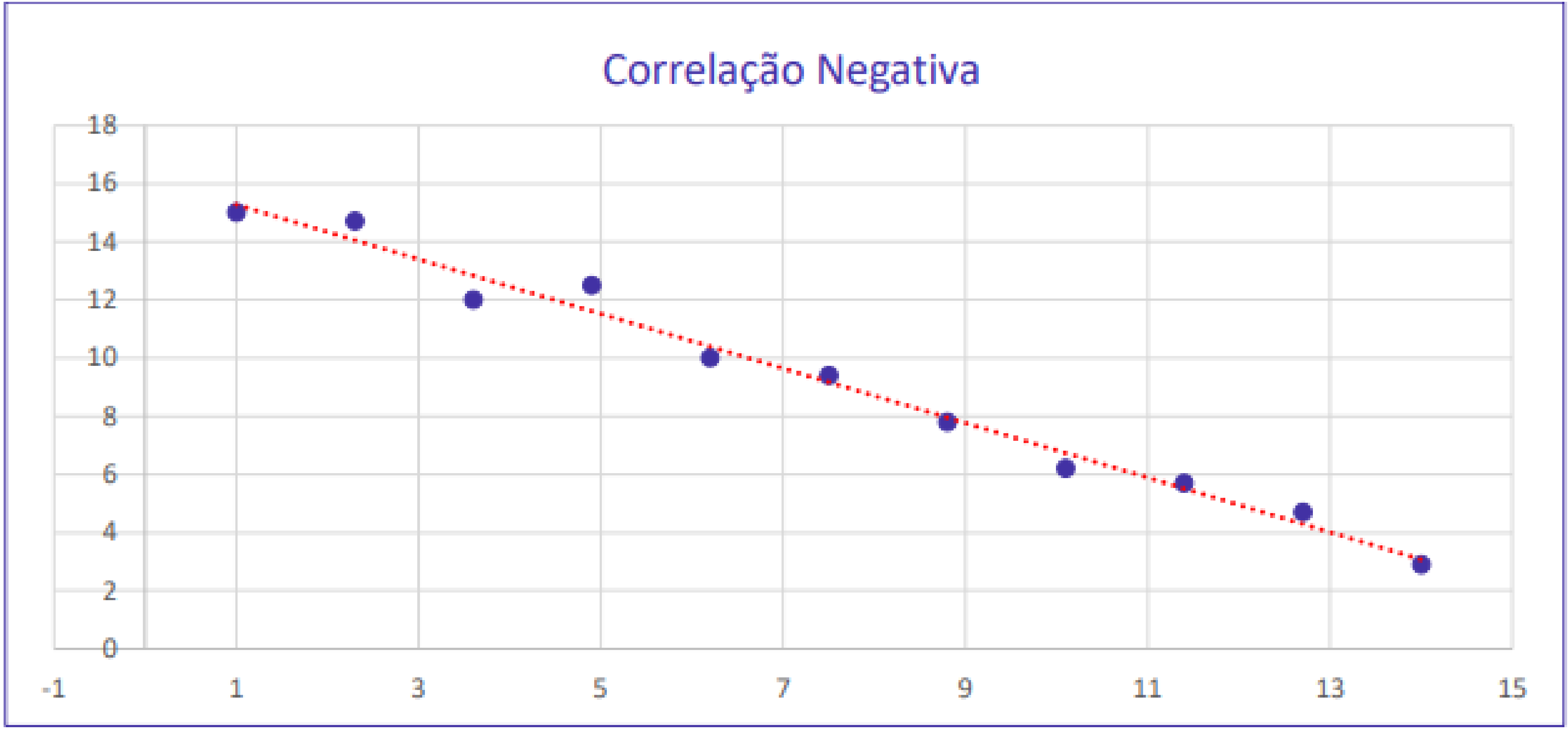


CORRELAÇÃO POSITIVA PERFEITA

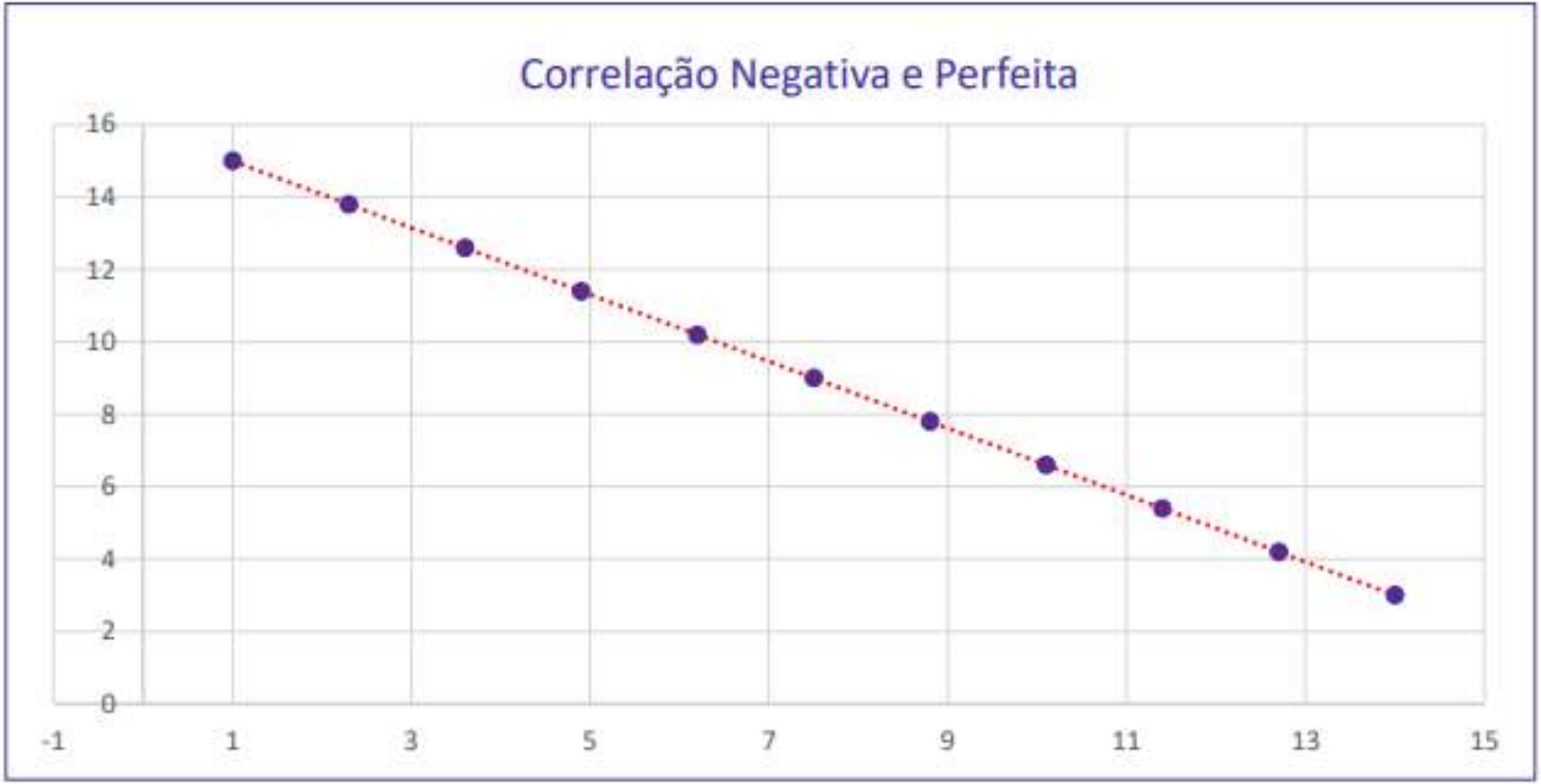
Estratégia Concursos



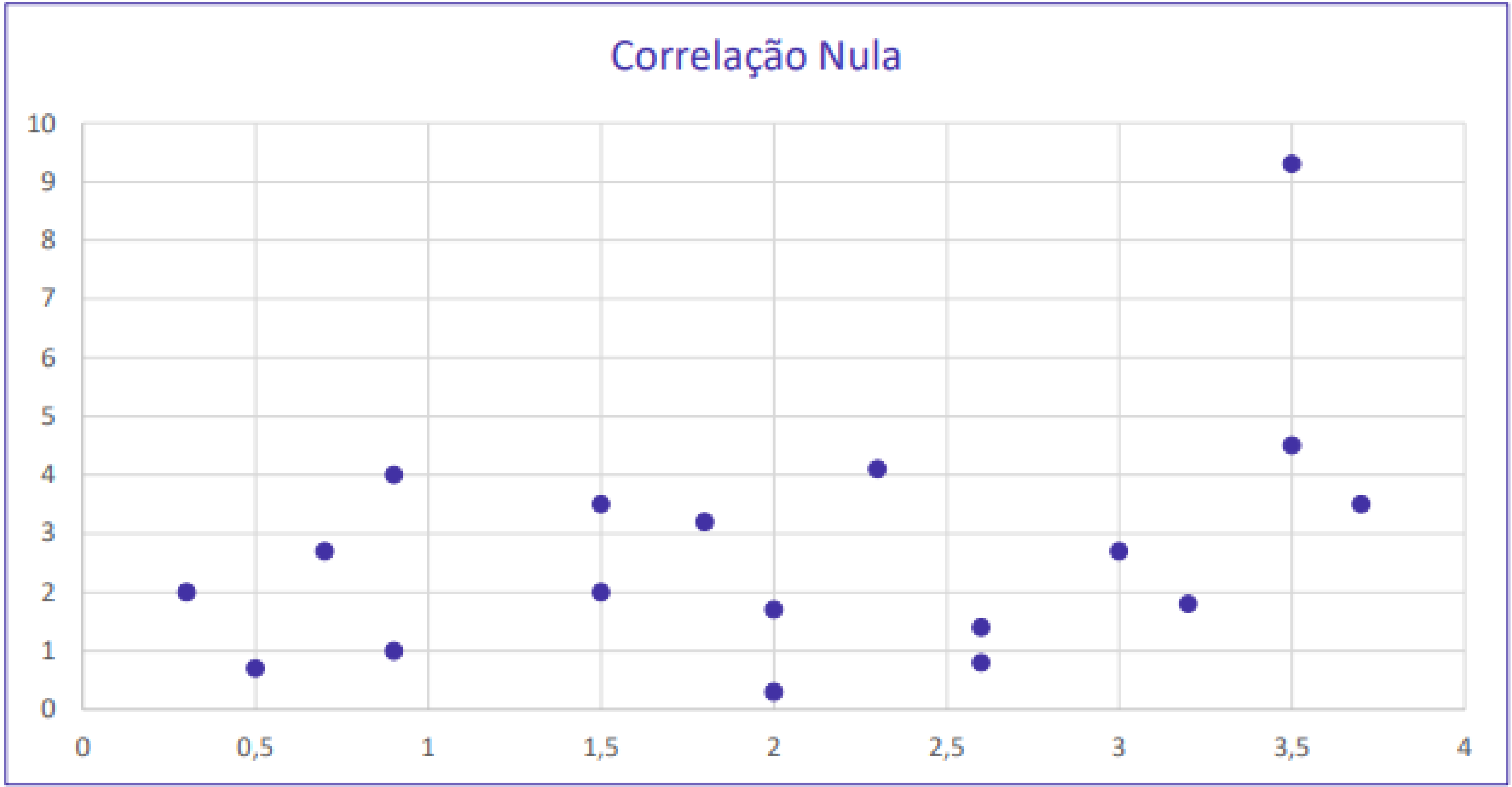
CORRELAÇÃO NEGATIVA



CORRELAÇÃO NEGATIVA PERFEITA



CORRELAÇÃO NULA



Obrigado

Prof. Jhoni Zini

REGRESSÃO LINEAR SIMPLES

Prof. Jhoni Zini

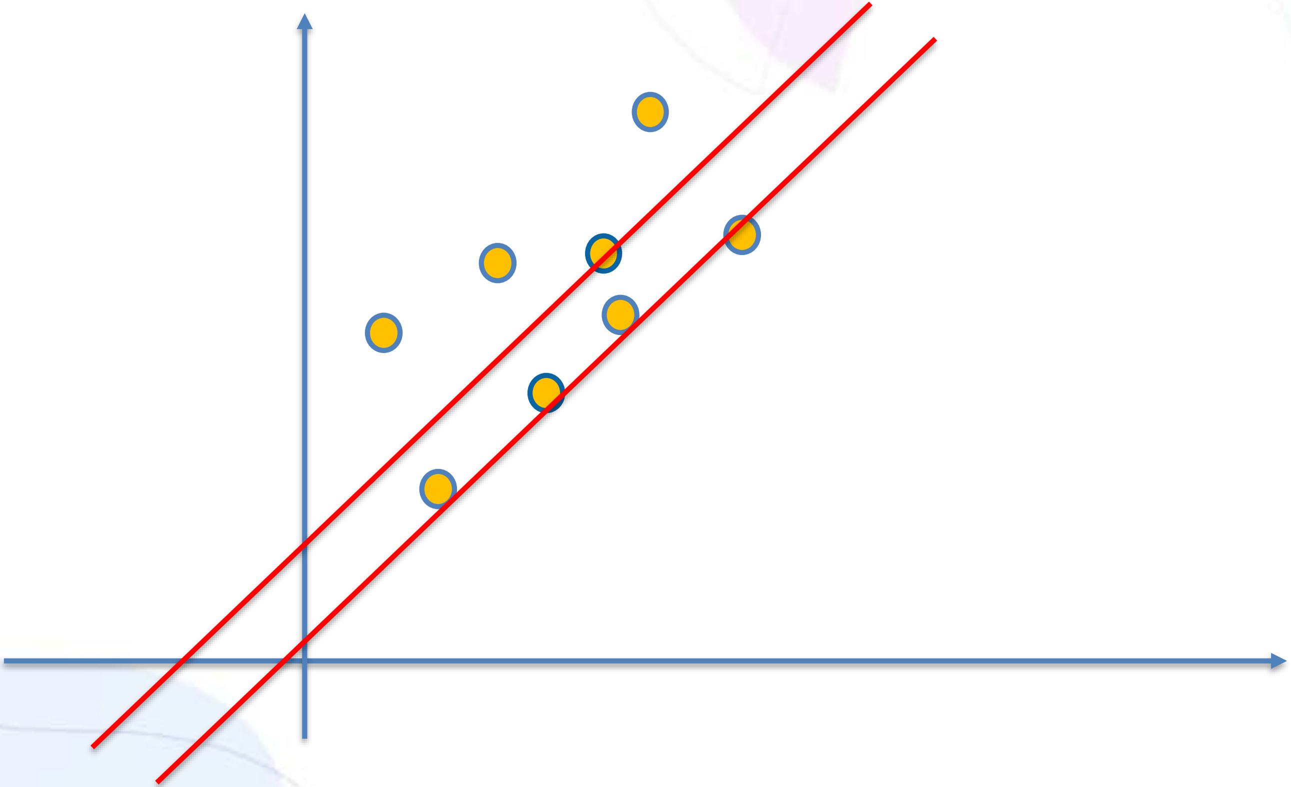
REGRESSÃO LINEAR SIMPLES

- MÉTODO DOS MÍNIMOS QUADRADOS ORDINÁRIOS

$$Y = \alpha + \beta x + \varepsilon$$

REGRESSÃO LINEAR SIMPLES

$$Y = \alpha + \beta x + \varepsilon$$



CÁLCULO DO COEFICIENTE ANGULAR

Estratégia Concursos

$$\beta = \frac{\sum xy - n\bar{x}\bar{y}}{\sum x^2 - n\bar{x}^2}$$

$$\beta = \frac{COV(X;Y)}{VAR(X)}$$

CÁLCULO DO COEFICIENTE ANGULAR

$$\beta = \frac{\sum xy - n\bar{x}\bar{y}}{\sum x^2 - n\bar{x}^2}$$

$$\beta = \frac{COV(X;Y)}{VAR(X)}$$

CÁLCULO DO COEFICIENTE ANGULAR

amostra	x	y	x.y	x^2
1	100	60		
2	80	40		
3	90	40		
4	120	50		
5	110	60		
TOTAL				

CÁLCULO DO COEFICIENTE ANGULAR

Para fazer uma regressão linear da forma $Y = \alpha + \beta X$, um analista, usando o método dos mínimos quadrados, encontrou, a partir de 20 amostras, os seguintes somatórios

$\sum X=300$; $\sum Y=400$; $\sum X^2=6.000$; $\sum$ e $\sum(XY)=8.400$

CÁLCULO DO COEFICIENTE ANGULAR

	x	y
MÉDIA	8	10
DESVIO PADRÃO	2	3
COV (X;Y)	3	

OBRIGADO

Prof. Jhoni Zini

Regressão Linear simples

Prof. Jhoni Zini

CÁLCULO DO INTERCEPTO

$$\alpha = \bar{y} - \beta \bar{x}$$

CÁLCULO DO INTERCEPTO

Para fazer uma regressão linear da forma $Y = \alpha + \beta X$, um analista, usando o método dos mínimos quadrados, encontrou, a partir de 20 amostras, os seguintes somatórios

$\sum X=300$; $\sum Y=400$; $\sum X^2=6.000$; $\sum$ e $\sum(XY)=8.400$

CÁLCULO DO INTERCEPTO

	x	y
MÉDIA	8	10
DESVIO PADRÃO	2	3
COV (X;Y)	3	

Obrigado

Prof. Jhoni Zini

Regressão Linear simples

Prof. Jhoni Zini

CÁLCULO DA EQUAÇÃO

Um grupo de 5 pessoas ingressou em um plano de dieta com o objetivo de reduzir peso. Obtenha a equação de regressão estimada que relacione a quantidade de peso perdida, Y em kg, e o número de semanas de cada um dos participantes no plano, X, sabendo que os valores registrados foram:

$\sum X=15$, $\sum Y=35$, $\sum XY=123$, $\sum X^2=55$.

CÁLCULO DA EQUAÇÃO

$\sum X=15$, $\sum Y=35$, $\sum XY=123$, $\sum X^2=55$.

CÁLCULO DA EQUAÇÃO

A tabela a seguir apresenta uma amostra aleatória simples formada por 5 pares de valores (X_i, Y_i), em que i = 1,2, ... ,5, X_i é uma variável explicativa e Y_i é uma variável dependente.

i	1	2	3	4	5
X_i	0	1	2	3	4
Y_i	0,5	2,0	2,5	5,0	3,5

CÁLCULO DA EQUAÇÃO

Considere o modelo de regressão linear simples na forma $Y_i = bX_i + \epsilon_i$, no qual ϵ representa um erro aleatório normal com média zero e variância σ^2 e b é o coeficiente do modelo.
Com base nos dados da tabela e nas informações apresentadas, é correto afirmar que o valor da estimativa de mínimos quadrados ordinários do coeficiente b é igual a

A. 0,75.
B. 0,9.
C. 1,2.
D. 1,35.
E. 1,45.

Regressão Linear simples

Prof. Jhoni Zini

ERRO ALEATÓRIO

εi é o componente aleatório de Yi que descreve os erros (ou desvios) cometidos quando tentamos aproximar uma série de observações Xi por meio de uma reta Yi .

ERRO ALEATÓRIO

i) $\boldsymbol{E(\varepsilon i\ ) = 0}$.

- A média dos erros é igual a zero. Ou seja, os desvios "para cima da reta" igualam o valor dos desvios "para baixo da reta" na média.

ERRO ALEATÓRIO

ii) $\boldsymbol{Var(\varepsilon i\,)} = \boldsymbol{\sigma}^2$.

- A variância dos erros é constante. Essa propriedade é denominada de homocedasticia.

ERRO ALEATÓRIO

iii) $\boldsymbol{Cov(\varepsilon i\ ,\ \varepsilon j) = 0\ para\ i \neq j}$.

- Os erros cometidos não são correlacionados, isto é, os desvios $\boldsymbol{\varepsilon i}$ são variáveis aleatórias independentes.

Obrigado

Prof. Jhoni Zini

Regressão Linear simples

Prof. Jhoni Zini

QUESTÃO 1

Um modelo de regressão linear simples na forma $y = ax + b + \epsilon$, no qual ϵ representa o erro aleatório com média nula e variância constante, foi ajustado para um conjunto de dados no qual as médias aritméticas das variáveis y e x são, respectivamente, $\bar{y} = 10$ e $\bar{x} = 5$. Pelo método dos mínimos quadrados ordinários, se a estimativa do intercepto (coeficiente b) for igual a 20, então a estimativa do coeficiente angular a proporcionada por esse mesmo método deverá ser igual a

A. −2.

B. 2.

C. −1.

D. 0.

E. 1.

QUESTÃO 2

A variável x tem média 4 e desvio padrão 2, enquanto a variável y tem média 3 e desvio padrão 1. A covariância entre x e y é –1.

A equação estimada da regressão linear simples de y por x é:

A. $\hat{y}$ =2−0,25×.

B. $\widehat{y}$ =3−0,5×.

C. $\hat{y}$ =3−×.

D. $\widehat{y}$ = 4−×.

E. $\hat{y}$ =4−0,25×.

QUESTÃO 3

Sejam S o valor do salário, em R$ 1.000,00, e t o respectivo tempo de serviço, em anos, de 20 empregados de uma empresa. Optou-se, com o objetivo de previsão do salário de um determinado empregado em função do seu tempo de serviço, por utilizar a relação linear $S_i = \alpha + \beta t_i + \varepsilon_i$, com i representando a i-ésima observação, α e β são parâmetros desconhecidos e ε_i é o erro aleatório com as respectivas hipóteses da regressão linear simples.

QUESTÃO 3

Utilizando o método dos mínimos quadrados, com base nas 20 observações correspondentes dos 20 empregados, obtiveram-se as estimativas de α e β (a e b, respectivamente). O valor encontrado para b foi de 1,8 e as médias dos salários dos 20 empregados e dos correspondentes tempos de serviço apresentam os valores de R$ 2.800,00 e 2 anos, respectivamente.

QUESTÃO 3

A previsão de salário para um empregado que tenha 5 anos de serviço é de

A. R$ 6.800,00

B. R$ 7.500,00

C. R$ 8.200,00

D. R$ 8.400,00

E. R$ 9.000,00

QUESTÃO 4

Um grupo de 5 pessoas ingressou em um plano de dieta com o objetivo de reduzir peso. Obtenha a equação de regressão estimada que relacione a quantidade de peso perdida, Y em kg, e o número de semanas de cada um dos participantes no plano, X, sabendo que os valores registrados foram:

QUESTÃO 4

$\sum X=15, \sum Y=35, \sum Y^2=279, \sum XY=123, \sum X^2=55.$

A. $\hat{y}=1,8+1,6X$

B. $\hat{y}=9,6+1,6X$

C. $\hat{y}=1,6+1,8X$

D. $\hat{y}=-9,6+1,8X$

E. $\hat{y}=-1,6-1,8X$

QUESTÃO 5

Seja o modelo linear $Y_i = \beta X_i + \varepsilon_i$ estabelecendo uma relação linear, sem intercepto, entre duas variáveis X e Y, em que Y_i é a variável dependente na observação i, X_i é a variável explicativa na observação i e ε_i o erro aleatório com as respectivas hipóteses para a regressão linear simples. O parâmetro β do modelo é desconhecido e sua estimativa foi obtida pelo método dos mínimos quadrados com base em 10 pares de observações (X_i, Y_i).

QUESTÃO 5

$\sum X=120$, $\sum Y=180$, $\sum XY=2.400$ e $\sum X^2 =1.500$

Considerando a equação da reta obtida pelo método dos mínimos quadrados, obtém-se que Y é igual a 24 quando X for igual a

A. 15.

B. 6.

C. 16.

D. 18.

E. 20.

QUESTÃO 6

Assinale a alternativa que apresenta a premissa da homocedasticidade que é subjacente ao método dos mínimos quadrados no modelo de regressão linear clássico.

A. Dado o valor de X, o valor médio ou esperado do distúrbio aleatório u i é zero.
B. Não há autocorrelação entre os termos de erro.
C. Ausência de covariância entre u i e X i.
D. Dado o valor de X, a variância de ui é a mesma para todas as observações.
E. Os valores de X em uma dada amostra não devem ser os mesmos.

Obrigado

Prof. Jhoni Zini

Análise da variância da regressão

Prof. Jhoni Zini

TESTE DE HIPÓTESES

$$H_0\text{: } \beta = 0$$
$$H_1\text{: } \beta \neq 0$$

TABELA ANOVA

FONTE DE VARIAÇÃO	SOMA DOS QUADRADOS	GRAUS DE LIBERDADE	QUADRADOS MÉDIOS	ESTATÍSTICA F
REGRESSÃO	SQ_R	1	QM_R	$\frac{QM_R}{QM_E}$
RESÍDUOS (ERROS)	SQ_E	N-2	QM_E	
TOTAL	SQ_T	N-1	QM_T	

SOMA DOS QUADRADOS DA REGRESSÃO

$$SQ_{REGRESSÃO} = \sum (\hat{y} - \bar{y})^2$$

SOMA DOS QUADRADOS DOS RESÍDUOS

$$SQ_{RES} = \left(Y_i - \hat{Y}\right)^2$$

SOMA DOS QUADRADOS TOTAIS

$$SQ_T = (Y_i - \bar{Y})^2$$

EXEMPLO

$Y = \alpha + \beta X + \varepsilon$ N = 21 SQM = 40 SQR = 380 SQT = 420

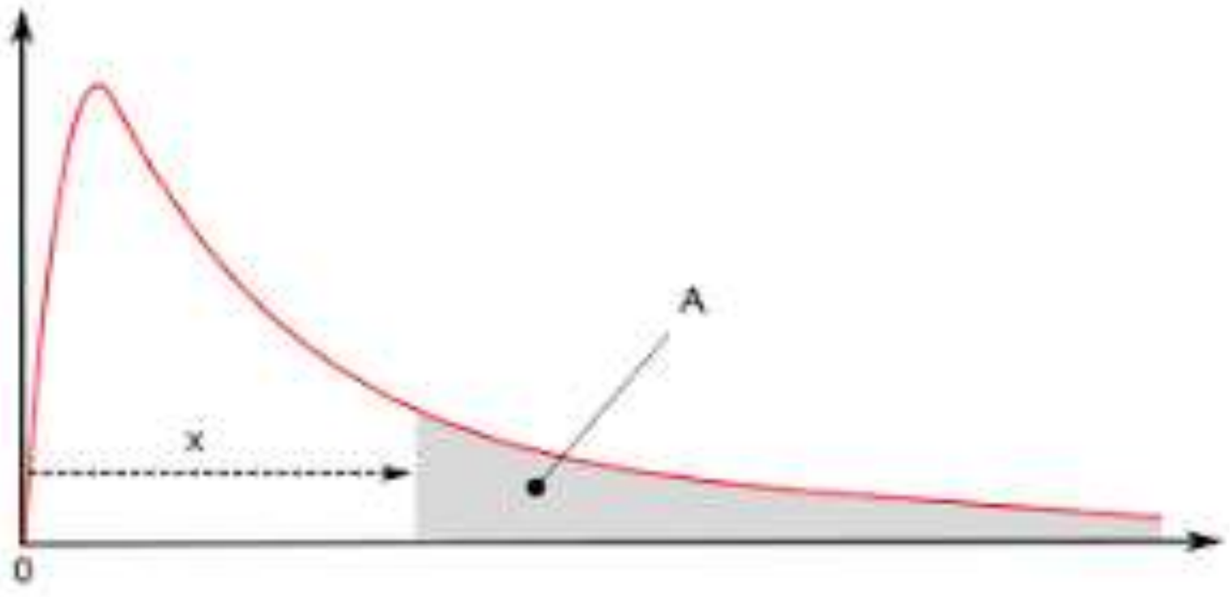
FONTE DE VARIAÇÃO	SOMA DOS QUADRADOS	GRAUS DE LIBERDADE	QUADRADOS MÉDIOS	ESTATÍSTICA F
REGRESSÃO				
RESÍDUOS				
TOTAL				

EXEMPLO

$Y = \alpha + \beta X + \varepsilon$ N = 11 SQM = 6 SQR = 18 SQT = 24

FONTE DE VARIAÇÃO	SOMA DOS QUADRADOS	GRAUS DE LIBERDADE	QUADRADOS MÉDIOS	ESTATÍSTICA F
REGRESSÃO				
RESÍDUOS				
TOTAL				

ANÁLISE DO TESTE – DISTRIBUIÇÃO F



EXEMPLO

FONTE DE VARIAÇÃO	SOMA DOS QUADRADOS	GRAUS DE LIBERDADE	QUADRADOS MÉDIOS	ESTATÍSTICA F	F TAB
REGRESSÃO	180	1			4,5
RESÍDUOS	900	15			
TOTAL					

ANÁLISE DA VARIÂNCIA DA REGRESSÃO

Prof. Jhoni Zini

COEFICIENTE DE DETERMINAÇÃO

- O coeficiente de determinação mede a qualidade do ajuste proporcionado pela reta de regressão.

$$R^2 = \frac{SQM}{SQT}$$

$$R^2 = 1 - \frac{SQE}{SQT}$$

COEFICIENTE DE DETERMINAÇÃO

FONTE DE VARIAÇÃO	SOMA DOS QUADRADOS	GRAUS DE LIBERDADE
REGRESSÃO	225	1
RESÍDUOS	175	15
TOTAL	400	16

COEFICIENTE DE DETERMINAÇÃO AJUSTADO

- Sempre um pouco menor que o padrão

$$\overline{R^2} = 1 - \frac{QME}{QMT}$$

$$\overline{R^2} = 1 - \frac{(1 - R^2).(n - 1)}{n - 2}$$

COEFICIENTE DE DETERMINAÇÃO AJUSTADO

FONTE DE VARIAÇÃO	SOMA DOS QUADRADOS	GRAUS DE LIBERDADE
REGRESSÃO	40	1
RESÍDUOS	380	20
TOTAL	420	21

QUESTÃO 1

Numa regressão linear simples em que foi utilizada uma amostra com 52 observações, a soma dos quadrados totais é de 50 e a soma dos quadrados dos resíduos é de 20. O coeficiente de determinação e a estatística F dessa regressão são, respectivamente:

A. 0,6 e 75.

B. 0,6 e 12.

C. 0,8 e 1,5.

D. 0,8 e 12.

E. 0,8 e 75.

Obrigado

Prof. Jhoni Zini

Análise da variância da regressão

Prof. Jhoni Zini

QUESTÃO 2

Considerando-se que, em uma regressão de dados estatísticos, a soma dos quadrados da regressão seja igual a 60.000 e a soma dos quadrados dos erros seja igual a 15.000, é correto afirmar que o coeficiente de determinação — R^2 — é igual a

A. 0,75.

B. 0,25.

C. 0,50.

D. 0,20.

E. 0,80.

QUESTÃO 3

Para entender a relação entre a variável independente X e a variável dependente Y, foi calculado o coeficiente de correlação linear de Pearson r=0,90. Sabe-se que existe uma relação de causa-efeito entre X e Y, então foi proposto um modelo de regressão linear simples. Acerca da explicação que este modelo será capaz de fornecer sobre a variabilidade da variável resposta, assinale a alternativa correta.

A. 90%

B. B. 84%

C. 81%

D. 100%

E. 88%

QUESTÃO 4

Para uma variável resposta Y e uma variável explicativa X, o ajuste de um modelo de regressão linear simples resultou em soma de quadrados dos desvios não explicados igual a 50 e soma de quadrados dos desvios totais igual a 200.

De acordo com os dados, é correto afirmar que o coeficiente de correlação linear entre as variáveis X e Y é igual a

A. 3/4

B. 1/4

C. $\sqrt{3/4}$

D. $\sqrt{1/4}$

QUESTÃO 5

Estratégia Concursos

Um estudo considerou um modelo de regressão linear simples na forma $y=0{,}8x+b+\epsilon$, em que y é a variável dependente, x representa a variável explicativa do modelo, o coeficiente b denomina-se intercepto e ϵ é um erro aleatório que possui média nula e desvio padrão σ. Sabe-se que a variável y segue a distribuição normal padrão e que o modelo apresenta coeficiente de determinação R^2 igual a 85%.

Com base nessas informações, julgue o item que se segue.

A correlação linear entre as variáveis x e y é superior a 0,9.

QUESTÃO 6

Numa regressão linear, as afirmativas a seguir, acerca do coeficiente de determinação, estão corretas, exceto uma. Assinale-a.

A. Mede a porcentagem da variação total da variável resposta que é explicada pela regressão.
B. É o quadrado do coeficiente de correlação estimado.
C. É um número entre 0 e 1.
D. Determina se as estimativas e predições dos coeficientes são tendenciosas.
E. Em geral, mas nem sempre, quanto maior seu valor, melhor o modelo se ajusta aos dados.

QUESTÃO 6

O coeficiente de determinação (R^2) e o desvio-padrão σ podem ser estimados através das estatísticas ρ^2 e $\sqrt{s^2}$ respectivamente.

Contudo, esses estimadores tenderão a apresentar alguns problemas de tal forma que:

A. R^2 será superestimado enquanto σ será subestimado;

B. ambos, R^2 e σ, serão estimados sem viés;

C. R^2 será subestimado enquanto σ será superestimado;

D. ambos, R^2 e σ, serão superestimados;

E. ambos, R^2 e σ, serão subestimados.

Obrigado

Prof. Jhoni Zini

Análise da variância da regressão

Prof. Jhoni Zini

QUADRADOS MÉDIOS

Estratégia Concursos

- SÃO ESTIMATIVAS DAS VARIÂNCIAS

QUADRADO MÉDIO

QMM = ESTIMATIVA DA VARIÂNCIA DA VARIÁVEL X

QME = ESTIMATIVA DA VARIÂNCIA DO ERROS

QMT = ESTIMATIVA DA VARIÂNCIA DA VARIÁVEL Y

QUADRADOS MÉDIOS

- SÃO ESTIMATIVAS DAS VARIÂNCIAS

QUADRADO MÉDIO

QMM = ESTIMATIVA DA VARIÂNCIA DA VARIÁVEL X

QME = ESTIMATIVA DA VARIÂNCIA DO ERROS

QMT = ESTIMATIVA DA VARIÂNCIA DA VARIÁVEL Y

QUESTÃO 1

Após formular e estimar um modelo de regressão simples, o estatístico responsável pela análise trabalha nos resultados, defrontando-se com a tabela a seguir:

Fonte	S. Quadrados	G.L	Q. Médio	F-Snedecor	p-valor
Equação	450	1	450	12,00	0,21%
Resíduos	300	8	37,50		
Total	750	9	83,33		

QUESTÃO 1

A partir desses números, é correto concluir que:

A. com uma amostra de tamanho 10, o modelo é capaz de explicar 60% da variação total;

B. a variância estimada do erro aleatório é inferior a 6;

C. apesar de um poder de explicação de 60%, o modelo não passa no teste de significância da estatística F;

D. a variância da variável explicativa do modelo é igual a 75;

E. a estimativa do coeficiente angular da equação é igual a 2.

QUESTÃO 2

Determinado estudo considerou um modelo de regressão linear simples na forma $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, em que y_i representa o número de leitos por habitante existente no município i; xi representa um indicador de qualidade de vida referente a esse mesmo município i, para i = 1, ..., n. A componente ε_i representa um erro aleatório com média 0 e variância σ^2. A tabela a seguir mostra a tabela ANOVA resultante do ajuste desse modelo pelo método dos mínimos quadrados ordinários.

QUESTÃO 2

fonte de variação	soma dos quadrados	graus de liberdade	média dos quadrados	razão F	P-valor
modelo	900	1	900	90	<0,001
erro	100	10	10		
total	1.000	11			

A partir das informações e da tabela apresentadas, julgue o item subsequente.

A estimativa de σ^2 foi igual a 10.

QUESTÃO 3

fonte de variação	soma dos quadrados	graus de liberdade	média dos quadrados	razão F	P-valor
modelo	900	1	900	90	<0,001
erro	100	10	10		
total	1.000	11			

A partir das informações e da tabela apresentadas, julgue o item subsequente.

O desvio padrão amostral do número de leitos por habitante foi superior a 10 leitos por habitante.

OBRIGADO

Prof. Jhoni Zini

Estratégia
Concursos