



Métodos quantitativos de apoio à decisão



Teste de hipóteses, regressão linear simples e correlação

Testes de hipóteses

Bloco 1

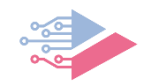
Rafaela Rodrigues Oliveira Amaro



Vamos refletir?

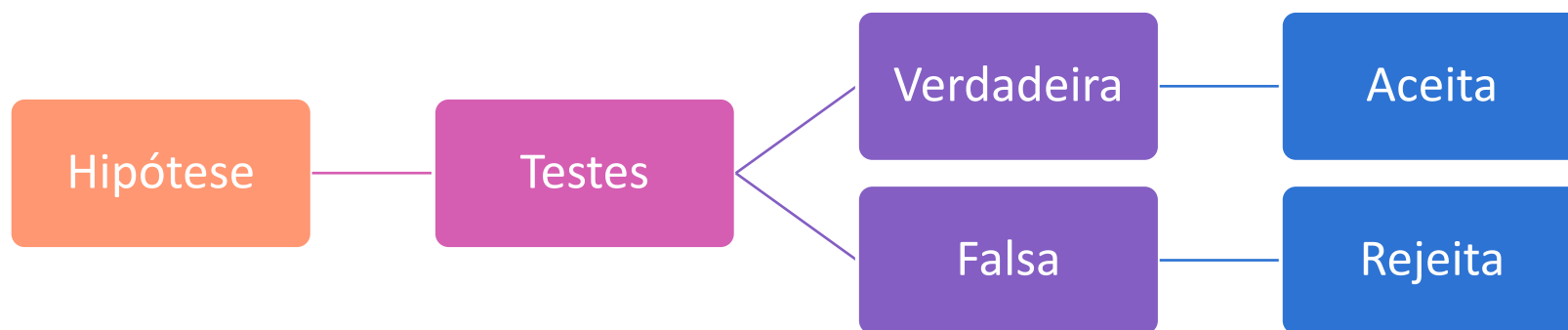
Você sabia que a hipótese é uma parte integrante do método científico? É uma afirmativa que procura responder ao problema da pesquisa e, por isso, forma a base de qualquer experimento.





Teste de hipóteses

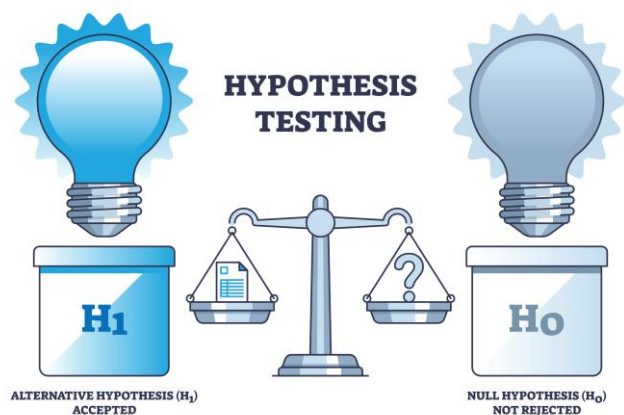
- Fornece ferramentas que permitem rejeitar ou não uma hipótese estatística, por meio da evidência fornecida pela amostra.
- Permite fazer inferências sobre uma população, com base nos dados obtidos de uma amostra desta população perante diferentes parâmetros (μ, σ^2, p) .





Teste de hipóteses: exemplo

Componentes de um teste de hipótese



Fonte: Vector Mine/ adobe.stock.com.

Verificar se uma moeda é honesta:

- Hipótese nula (H_0): a moeda é honesta.
- Hipótese alternativa (H_1): a moeda não é honesta.



$$\begin{cases} H_0: p = \frac{1}{2} = 0,5 \\ H_1: p \neq \frac{1}{2} \neq 0,5 \end{cases}$$



Teste de hipóteses: erros

Erro tipo I:

- Rejeitar H_0 quando H_0 é verdadeira.
- Exemplo: constata-se que a moeda não é honesta, quando na verdade ela é.

Erro tipo II:

- Não rejeitar H_0 quando H_0 é falsa.
- Exemplo: constata-se que a moeda é honesta, quando na verdade ela não é.

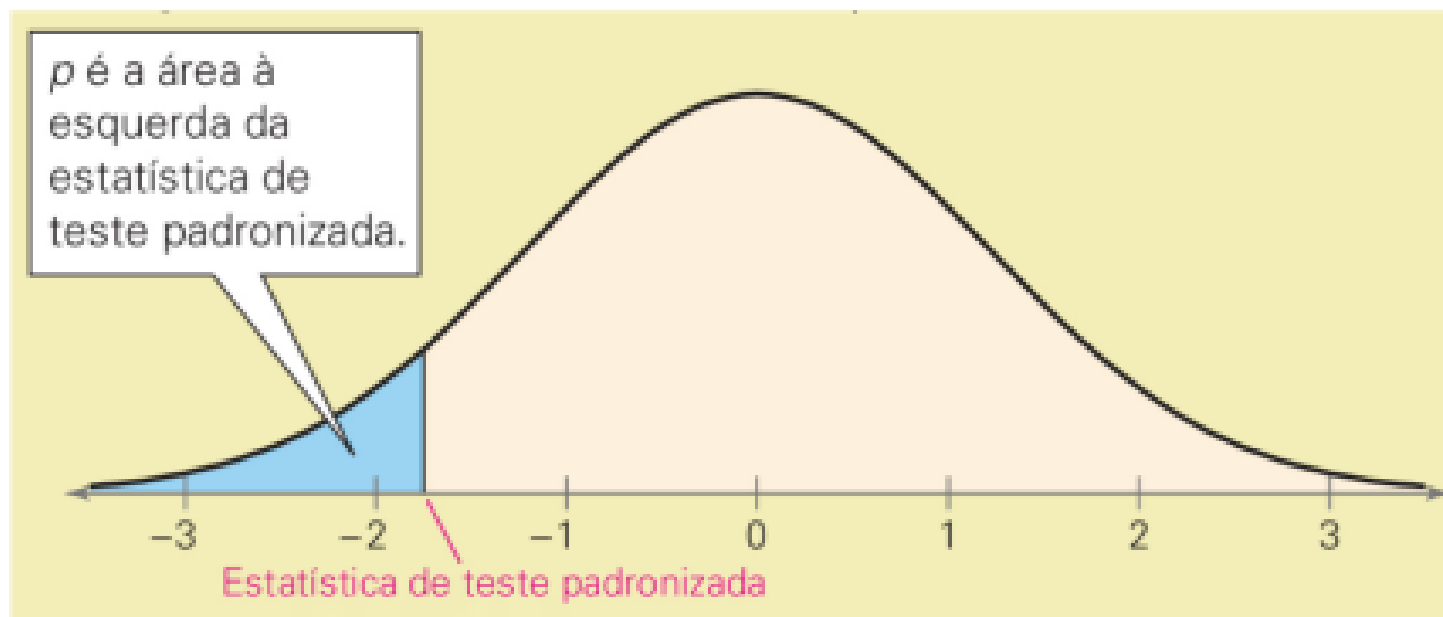




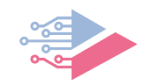
Testes estatísticos e valores p

- p (p-value) indica a probabilidade de observar uma diferença tão grande ou maior do que a que foi observada sob a hipótese nula.
- Rejeita-se H_0 quando o valor p for menor ou igual ao nível de significância estabelecido.

Teste unilateral à esquerda

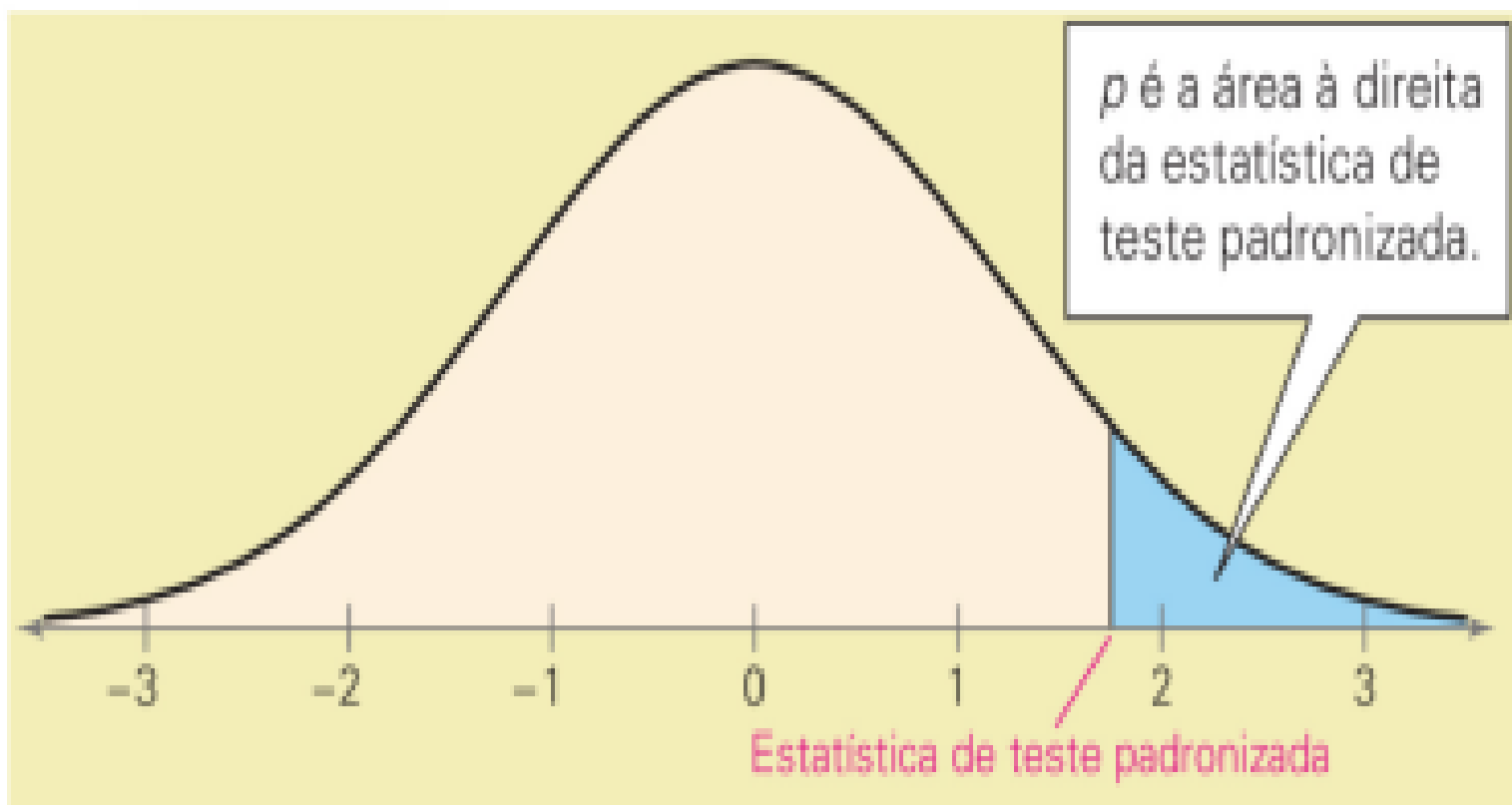


Fonte: Larson e Farber (2015, [n. p.]).



Testes estatísticos e valores p

Teste unilateral à direita

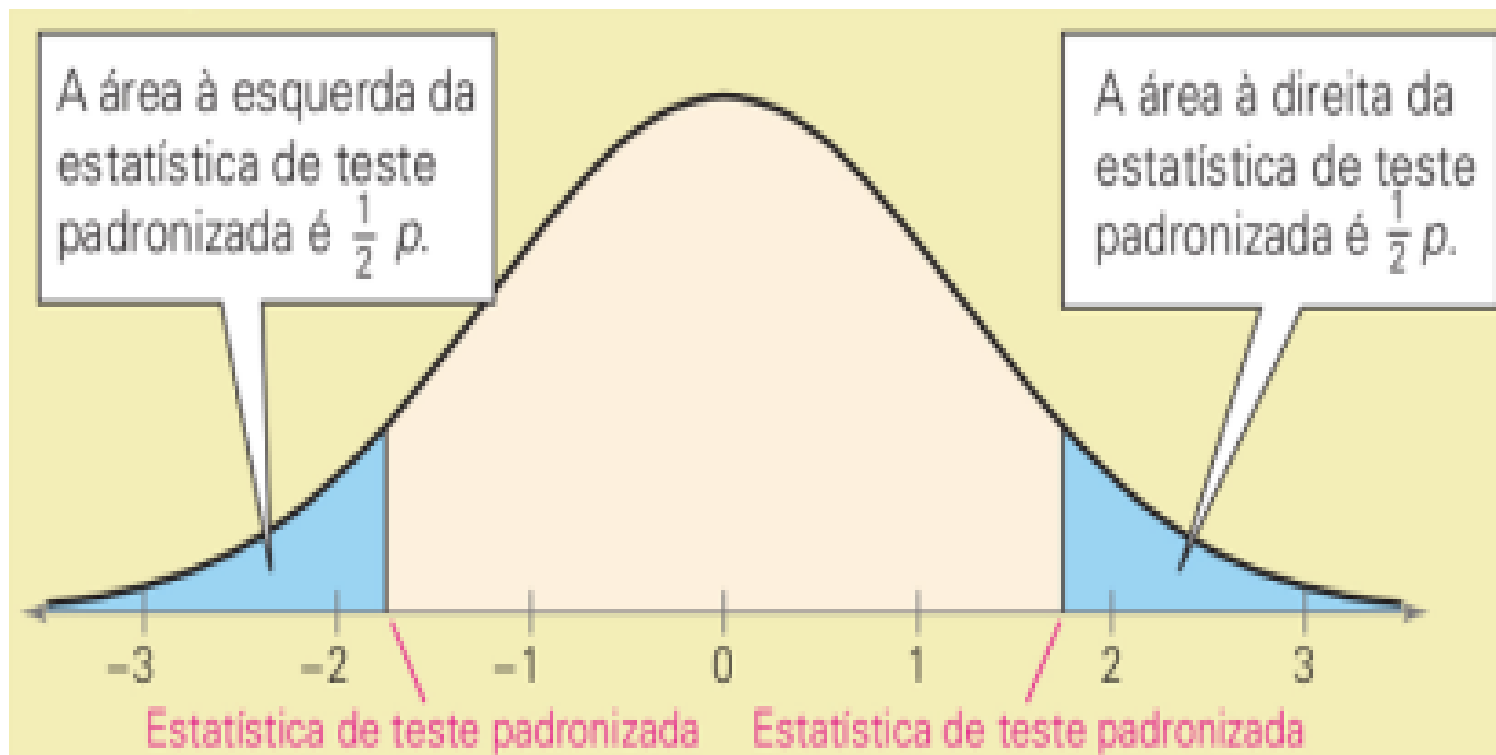


Fonte: Larson e Farber (2015, [n. p.]).



Testes estatísticos e valores p

Teste Bilateral



Fonte: Larson e Farber (2015, [n. p.]).

Teste de hipóteses, regressão linear simples e correlação

Correlação linear

Bloco 2

Rafaela Rodrigues Oliveira Amaro



Em busca de padrões

Uso de camisa versus vitória do time

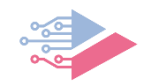


Fonte: Angela/ adobe.stock.com.

Valor do veículo versus imposto

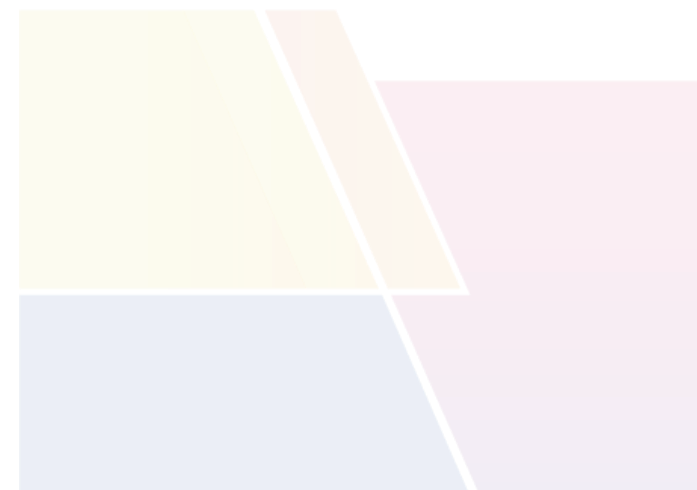


Fonte: Elena/ adobe.stock.com.



Definição e composição

- Correlação é o estudo do grau de associação entre variáveis.
- Na correlação é analisado se duas ou mais variáveis são independentes ou variam juntas.
- Variável dependente ou explanatória (x).
- Variável independente ou resposta (y).



Coeficiente de correlação de Pearson

Medida padronizada da correlação entre variáveis:

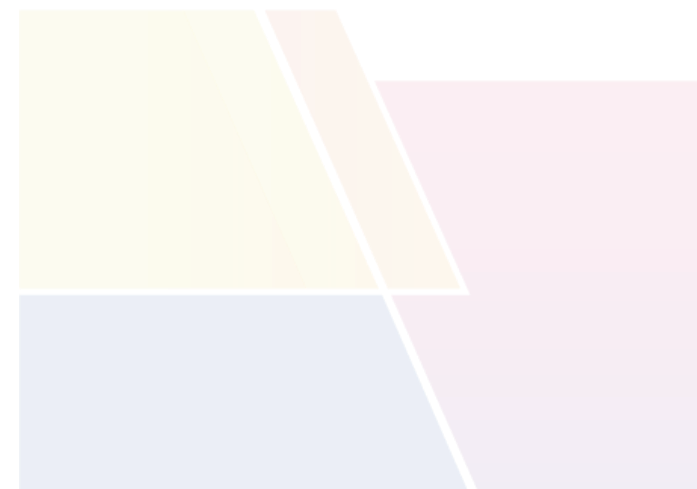
$$r = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{n \sum x^2 - (\sum x)^2} \sqrt{n \sum y^2 - (\sum y)^2}}$$

n = total de pares de dados.

$r = +1 \Rightarrow$ duas variáveis estão perfeitamente correlacionadas de forma positiva (se uma aumenta, a outra aumenta proporcionalmente).

$r = -1 \Rightarrow$ duas variáveis estão perfeitamente correlacionadas de forma negativa (se uma aumenta, a outra diminui proporcionalmente).

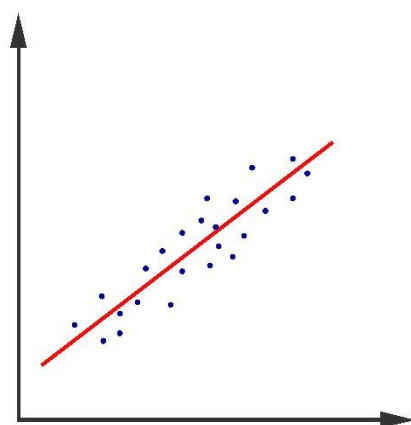
$r = 0 \Rightarrow$ ausência de relacionamento linear.



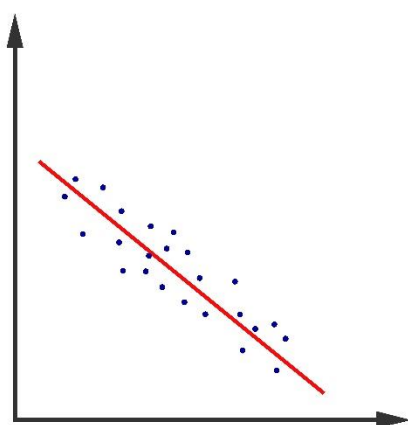


Representação gráfica

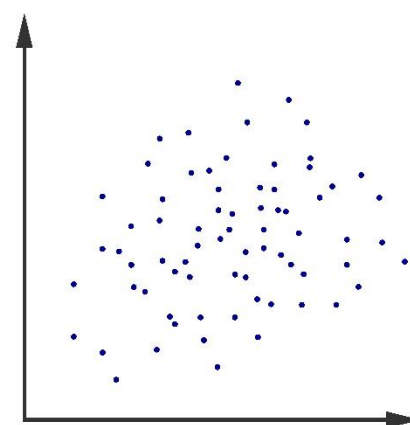
Diagrama de dispersão



Correlação positiva



Correlação negativa



Sem correlação

Fonte: Andrii/Adobe Stock.com.

O diagrama de dispersão indica a forma da relação entre as variáveis estudadas.



Exemplo: teste de aptidão em crianças

Com o objetivo de verificar se existe correlação positiva entre aptidão em matemática e aptidão em música, foi selecionado um grupo de crianças de oito a dez anos, que foram submetidas a dois testes de aptidão: um de matemática e um de música.

Teste de aptidão em crianças

Criança	Valores de aptidão em		Criança	Valores de aptidão em	
	Matemática	Música		Matemática	Música
1	60	80	7	48	79
2	58	62	8	72	88
3	73	70	9	75	54
4	51	83	10	83	82
5	54	62	11	62	64
6	75	92	12	52	69

Fonte: adaptado de Barbetta (2008, p. 286).

Exemplo: teste de aptidão em crianças

Exemplo

Matemática (x)	Música (y)	x^2	y^2	$x \cdot y$
60	80	3600	6400	4800
58	62	3364	3844	3596
73	70	5329	4900	5110
51	83	2601	6889	4233
54	62	2916	3844	3348
75	92	5625	8464	6900
48	79	2304	6241	3792
72	88	5184	7744	6336
75	54	5625	2916	4050
83	82	6889	6724	6806
62	64	3844	4096	3968
52	69	2704	4761	3588
$\sum x = 763$	$\sum y = 885$	$\sum x^2 = 49985$	$\sum y^2 = 66823$	$\sum x \cdot y = 56517$

$$r = \frac{12 \cdot 56517 - (763 \cdot 885)}{\sqrt{12 \cdot 49985 - (763)^2} \sqrt{12 \cdot 66823 - (885)^2}} = \frac{678204 - 675255}{\sqrt{17621} \cdot \sqrt{18651}} \cong 0,16$$

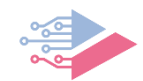
Fonte: elaborado pela autora.

Teste de hipóteses, regressão linear simples e correlação

Regressão linear

Bloco 3

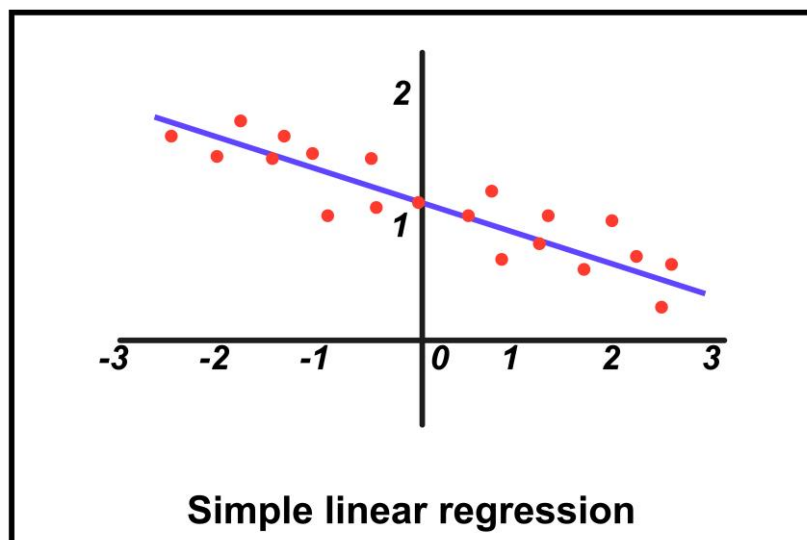
Rafaela Rodrigues Oliveira Amaro



Análise de regressão

- Determina como duas ou mais variáveis se relacionam.
- Obtém uma equação matemática, que caracteriza a relação entre duas ou mais variáveis. Por meio dela, é possível estimar e/ou mensurar valores da variável dependente.

Reta de regressão



Fonte: MakZin/ adobe.stock.com.



Reta de regressão

$$\hat{y} = mx + b$$

$$m = \frac{n \sum xy - (\sum x)(\sum y)}{n \sum x^2 - (\sum x)^2}$$

$$b = \underline{y} - m\underline{x}$$

$$b = \frac{\sum y}{n} - m \frac{\sum x}{n}$$

Onde:

- \underline{y} = média dos valores de y no conjunto de dados.
- \underline{x} = média dos valores de x e n é o número de pares de dados.



Exemplo: teste de aptidão em crianças

Com o objetivo de verificar se existe correlação positiva entre aptidão em matemática e aptidão em música, foi selecionado um grupo de crianças de oito a dez anos, que foram submetidas a dois testes de aptidão: um de matemática.

Matemática (x)	Música (y)
60	80
58	62
73	70
51	83
54	62
75	92
48	79
72	88
75	54
83	82
62	64
52	69

Exemplo

Do exercício para encontrar o coeficiente de correlação, temos que:

$$\sum x = 763 \quad \sum y = 885 \quad \sum x \cdot y = 56517 \quad \sum x^2 = 49985$$

$$\hat{y} = mx + b$$

$$\bullet \quad m = \frac{n \sum xy - (\sum x)(\sum y)}{n \sum x^2 - (\sum x)^2} = \frac{12 \cdot 56517 - (763 \cdot 885)}{12 \cdot 49985 - (763)^2} = \frac{2949}{17651} \cong 0,17$$

$$\bullet \quad b = \frac{\sum y}{n} - m \frac{\sum x}{n} = \frac{885}{12} - 0,17 \cdot \frac{763}{12} \cong 62,94$$

$$\text{Assim, encontramos: } \hat{y} = 0,17x + 62,94.$$

Fonte: elaborado pela autora.



Correlação e regressão linear, no Excel

4/5 – Desenvolva o conteúdo aqui

Para ter acesso ao recurso interativo, assista à videoaula.



Semelhanças e diferenças

Correlação
linear.

Regressão
linear.



Teste de hipóteses, regressão linear simples e correlação

Teoria em prática

Bloco 4

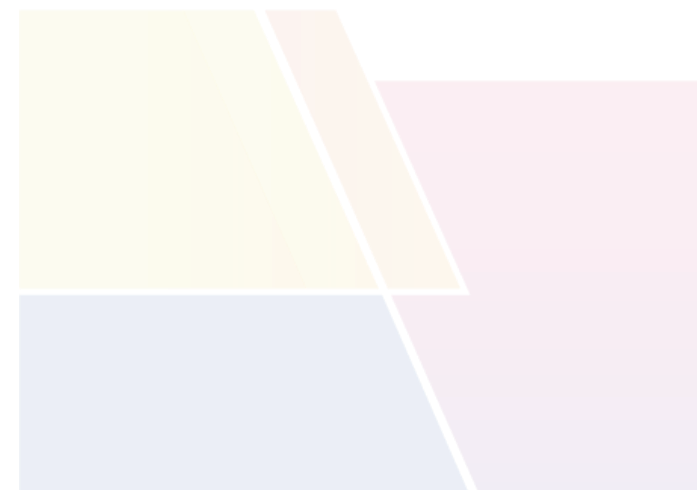
Rafaela Rodrigues Oliveira Amaro

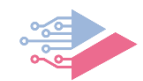


Refleta sobre a seguinte situação

Você é colaborador(a) de uma indústria que coleta, periodicamente, amostras aleatórias de 500 unidades em sua linha de produção para verificação da qualidade. As peças da amostra recebem uma classificação como defeituosa ou não defeituosa. De acordo com as diretrizes da empresa, o processo produtivo deve ser revisto caso haja evidência de que existam mais de 1,5% de peças defeituosas.

Considere que, na última amostra, foram encontradas nove peças defeituosas. Você será o(a) responsável por avaliar se o processo precisa ser revisto ou não. Para isso, use um nível de significância de 1%.





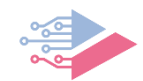
Norte para a resolução

1) Estabelecer a hipótese nula (H_0) e a hipótese alternativa (H_1) .

2) Identificar o nível de significância α .

3) Calcular a estatística de teste: se $n \leq 30$, teste z, caso contrário, $n > 30$, teste *t-student*.

4) Decidir, usando a evidência x , ao nível de significância α , se aceita ou rejeita a hipótese nula.



Norte para a resolução

Como estamos procurando evidências de que a proporção de peças defeituosas é maior que 1,5% do total, será necessária a utilização de um teste unilateral à direita, definido pelas seguintes hipóteses:

- **Hipótese nula** $\Rightarrow H_0: p = 1,5\%$;
- **Hipótese alternativa** $\Rightarrow H_1: p > 1,5\%$.

A amostra é composta por quinhentas peças, logo, $n = 500$, dessas nove foram classificadas como defeituosas. Assim, temos a seguinte proporção: $\hat{p} = \frac{9}{500} = 0,018$.





Norte para a resolução

É possível adotar os critérios para uma aproximação normal ($n \leq 30$), por isso, a necessidade em encontrar uma estatística de teste z , dada por:

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 \cdot (1 - p_0)}{n}}} = \frac{0,018 - 0,015}{\sqrt{\frac{0,015 \cdot (1 - 0,015)}{500}}} = \frac{0,003}{0,00543} = 0,5518$$

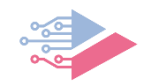
De acordo com teste unilateral à esquerda, com um nível de significância de $\alpha = 1\% = 0,01$, temos o valor crítico $z_c = 2,575$

Resolução

Nível de confiança	α	z_c
90%	0,1	1,645
95%	0,05	1,96
99%	0,01	2,575

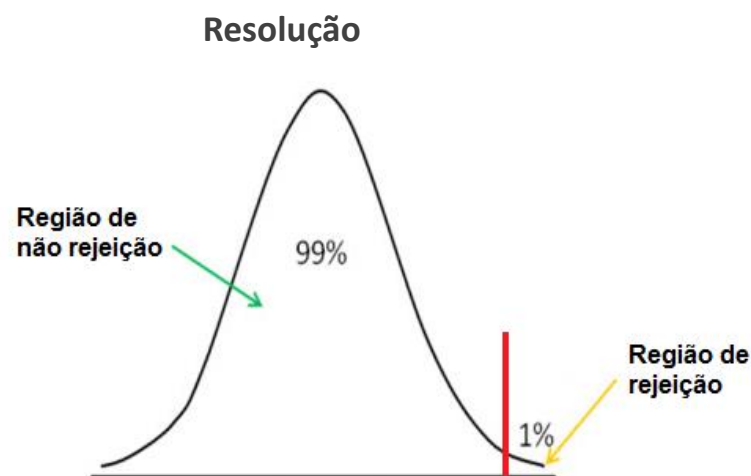
Fonte: elaborado pela autora.





Norte para a resolução

Como a amostra corresponde a estatística de teste $z = 0,5518$ verifica-se sua posição na área de aceitação da hipótese.



Fonte: adaptado de <https://www.learningaboutelectronics.com/Artigos/Calculadora-teste-de-hipotese-estatistica.php#resposta>

Logo, podemos concluir que não será necessário um processo de revisão no processo de verificação da qualidade das peças, já que não há evidências de que mais que 1,5% dos itens verificados sejam defeituosos, considerando um nível de significância de 1%.

Teste de hipóteses, regressão linear simples e correlação

Consolidando o aprendizado

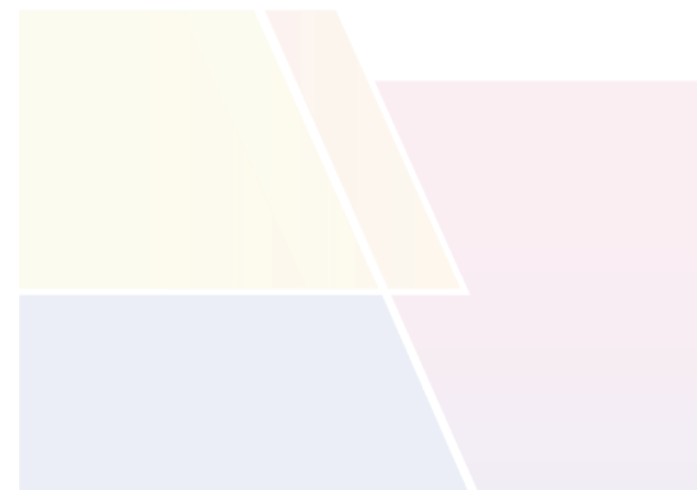
Bloco 5

Rafaela Rodrigues Oliveira Amaro



Consolidando o aprendizado

- Teste de hipóteses: hipótese nula e hipótese alternativa.
- Testes estatísticos: z (distribuição normal) e t (t-student).
- Nível de confiança versus nível de significância.
- Conclusão teste de hipótese.
- Correlação linear.
- Regressão linear.
- Correlação linear e regressão linear, no Excel.





Quiz



O teste de hipóteses é uma metodologia estatística, que auxilia na tomada de decisões sobre uma ou mais populações, tendo como base dados obtidos pela amostra. Por isso, é possível verificar se:

A

Os dados amostrais são verdadeiros ou falsos coletados na pesquisa.

B

Os dados amostrais são coerentes à população estudada na pesquisa.

C

Os dados amostrais trazem evidência para sustentar ou não a hipótese nula.

D

Os dados amostrais trazem evidência para sustentar ou não a hipótese alternativa.

Quiz



O teste de hipóteses é uma metodologia estatística, que auxilia na tomada de decisões sobre uma ou mais populações, tendo como base dados obtidos pela amostra. Por isso, é possível verificar se:

A

Os dados amostrais são verdadeiros ou falsos coletados na pesquisa.

B

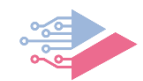
Os dados amostrais são coerentes à população estudada na pesquisa.

C

Os dados amostrais trazem evidência para sustentar ou não a hipótese nula.

D

Os dados amostrais trazem evidência para sustentar ou não a hipótese alternativa.

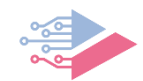


Quiz – Resolução

Resposta correta: **c) Os dados amostrais trazem evidência para sustentar ou não a hipótese nula.**

Na dinâmica de um teste de hipóteses, com base em dados amostrais, uma hipótese nula é estruturada. Assim como base no nível de significância, que define uma área de aceitação e de rejeição que é limitada por uma estatística de teste (teste z ou teste t-student), a hipótese nula é aceita ou rejeitada.





Leitura Fundamental

Prezado estudante, as indicações a seguir podem estar disponíveis em algum dos parceiros da nossa Biblioteca Virtual (faça o login por meio do seu AVA), e outras podem estar disponíveis em sites acadêmicos (como o SciELO), repositórios de instituições públicas, órgãos públicos, anais de eventos científicos ou periódicos científicos, todos acessíveis pela internet.

Isso não significa que o protagonismo da sua jornada de autodesenvolvimento deva mudar de foco. Reconhecemos que você é a autoridade máxima da sua própria vida e deve, portanto, assumir uma postura autônoma nos estudos e na construção da sua carreira profissional.

Por isso, nós o convidamos a explorar todas as possibilidades da nossa Biblioteca Virtual e além! Sucesso!





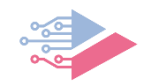
Indicação de leitura 1

Os testes estatísticos podem robustecer o entendimento do pesquisador acerca das questões de investigação por ele definidas. Partindo dessa afirmação, o autor defende que cabe ao investigador, em função do objeto investigado, decidir pela usabilidade ou não de testes estatísticos. Assim, no decorrer deste artigo, definições estatísticas são retomadas e exemplificadas.

Referência:

LAPA, L. Testes estatísticos: breves reflexões. **Reflexões em torno de metodologias de investigação: recolha de dados**, [s. l.], v. 2, p. 73-86, 2021.





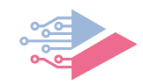
Indicação de leitura 2

Este artigo científico apresenta o conteúdo de regressão linear quando aplicado a Engenharia Civil, especificamente aos custos inerentes a construção de edifícios.

Referência:

MAUÉS, F. C. A.; MELO, K. P. de; LEÃO, C. B. de O. *et al.*
Estimativa de custos paramétricos de construção de edifícios usando modelo de regressão linear. **Gestão & Tecnologia de Projetos**, São Carlos, v. 17, n. 2, p. 19-37, 2022.





Referências

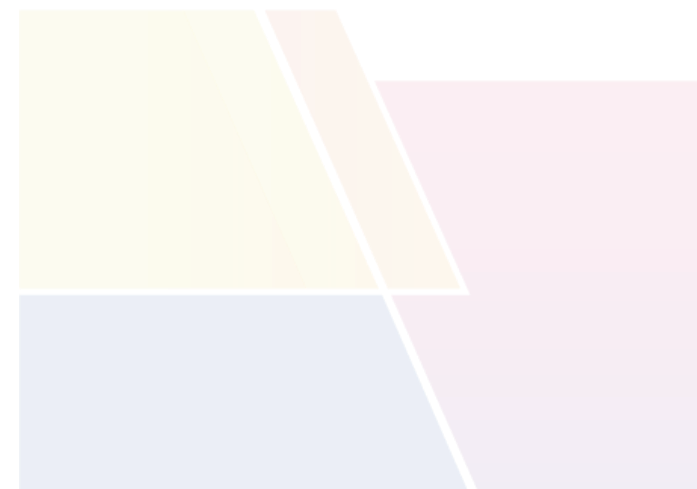
BARBETTA, P. A. **Estatística aplicada às ciências sociais**. Florianópolis: UFSC, 2008.

BUSSAB, W. O.; MORETTIN, P. A. **Estatística básica**. São Paulo: Saraiva 2009.

DA CUNHA, S. B.; CARVAJAL, S. R. **Estatística Básica – a arte de trabalhar com dados**. São Paulo: Elsevier Brasil, 2009.

LAPA, L. Testes estatísticos: breves reflexões. **Reflexões em torno de metodologias de investigação: recolha de dados**, [s. l.], v. 2, p. 73-86, 2021.

MAUÉS, F. C. A.; MELO, K. P. de; LEÃO, C. B. de O. *et al.* Estimativa de custos paramétricos de construção de edifícios usando modelo de regressão linear. **Gestão & Tecnologia de Projetos**, São Carlos, v. 17, n. 2, p. 19-37, 2022. Disponível em: <https://www.revistas.usp.br/gestaodeprojetos/article/view/168355>. Acesso em: 22 jul. 2024.





Bons estudos!