Mathtrace – manual

V. Estatística

Índice

V. Estatística

41. Estatística

41. Estatística

Os dados escrevem-se como data(…): valores separados por ponto e vírgula, espaços ou mudanças de linha; a vírgula decimal vale como no resto (data(1,5 2,5 3)). Uma tabela – uma linha por observação, uma coluna por variável – escreve-se linha a linha separadas por ponto e vírgula; a primeira linha pode ter os nomes das variáveis. Mais cómodo: a tecla de grelha do separador f(x) abre uma tabela para preencher, colar da área de transferência (células do Excel ou do Google Sheets) ou carregar de um ficheiro CSV, e o botão «Como dados» insere-a para estatística – mesmo uma tabela maior do que a grelha.

Com uma variável, um chip escolhe o que calcular. Estatística descritiva: média aritmética, mediana, moda, quartis (medianas das metades inferior e superior), amplitude e amplitude interquartil, variância e desvio padrão amostrais (n − 1) e populacionais (n), coeficiente de variação e valores atípicos (fora de $Q_1 - 1{,}5 \cdot IQR$ e $Q_3 + 1{,}5 \cdot IQR$), com diagrama de caixa. Frequências e histograma: a tabela de frequências absolutas, relativas e acumuladas – por valores para inteiros com poucos valores, senão por classes (o seu número pela regra de Sturges). Pontuações z: cada valor padronizado pela média e pelo desvio padrão dos dados, com a pontuação T.

Duas variáveis (duas colunas) dão correlação e regressão: somas, covariância, o coeficiente de correlação de Pearson r com a força da relação em palavras, a reta de regressão pelos mínimos quadrados e o coeficiente de determinação r^2, com diagrama de dispersão. Três ou mais variáveis recebem a sua estatística descritiva lado a lado com diagramas de caixa.

Inferência estatística a partir de dados: o intervalo de confiança da média (chip “Intervalo de confiança da média”, 90, 95 ou 99 %) pela distribuição t; o teste t (chip “Teste t”) – uma variável contra o valor testado μ_0, duas variáveis como amostras independentes (teste de Welch, ou com variâncias iguais) ou valores pareados (diferenças dos pares); o teste χ^2 (chip “Teste χ^2”) – uma coluna de frequências como teste de aderência à distribuição uniforme, uma tabela como teste de independência com o V de Cramér. Sob os chips introduz-se o nível de significância α, a alternativa (bilateral, unilateral à esquerda ou à direita) e μ_0. A resolução: hipóteses, estatística de teste, valor e região críticos, valor p e decisão; o gráfico mostra a distribuição com a região crítica sombreada. Duas amostras independentes de tamanhos diferentes escrevem-se como colunas com nome – data(A = 23 25 28; B = 20 22 21 25).

Comparar grupos: o chip «Teste de postos (Mann-Whitney, Wilcoxon, Kruskal-Wallis)» compara dois grupos pelo teste de Mann-Whitney (com «Pareadas», valores emparelhados pelo teste de Wilcoxon) e três ou mais grupos pelo teste de Kruskal-Wallis – pelos postos, sem pressupor distribuição normal; com amostras pequenas sem empates, o valor p é exato. O chip «ANOVA (análise de variância)» compara as médias dos grupos pela análise de variância de um fator – tabela ANOVA, teste F e tamanho do efeito η^2. Escreva os grupos em colunas, de qualquer comprimento (colunas com nome); o gráfico mostra os diagramas de caixa e a distribuição da estatística. No teste χ^2 de uma coluna de

frequências, o campo «Proporções esperadas» aceita proporções esperadas próprias – 1:2:1, 0,25 0,5 0,25 ou em percentagem.

A regressão de duas variáveis tem sob o seu chip a escolha do modelo: linear (com o teste de significância da correlação), quadrática (equações normais de mínimos quadrados), exponencial y = a·e^(bx), potência y = a·x^b e logarítmica y = a + b·ln x (por linearização com logaritmos) – com o coeficiente de determinação R² e o gráfico da curva ajustada. O formulário “Estatística e psicometria” tem ainda o intervalo de confiança e testes t a partir de valores resumidos (tamanho, média, desvio padrão) – incluindo o teste z com σ conhecido e o teste t de duas amostras.

A distribuição normal está no formulário «Estatística e psicometria» (secção «Aplicações» do ecrã inicial, área Estatística): as probabilidades P(X < x), P(X > x), P(a < X < b) e o quantil – o valor abaixo do qual fica uma proporção dada. A resolução padroniza, z = (x − μ)/σ, e usa a função de distribuição Φ, um z negativo como 1 − Φ(|z|) como nas tabelas; o gráfico mostra a densidade com a área colorida.

Como introduzir

`data(3; 5; 7; 7; 9)`	Uma variável – os valores com ponto e vírgula, espaços ou um por baixo do outro.
`data(x y; 1 52; 2 58; 3 61)`	Uma tabela linha a linha, a primeira linha com os nomes das variáveis (opcional).
`data(x = 1 2 3 4; y = 2 4 5 4)`	Colunas com nome, cada uma com os seus valores.

Exemplos

Estatística descritiva de oito valores.

Enunciado: `data(2; 4; 4; 4; 5; 5; 7; 9)`

Resultado

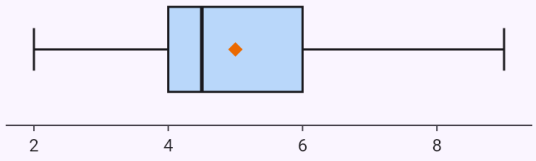
$$\bar{x} = 5$$
$$\tilde{x} = 4{,}5$$
$$s \approx 2{,}138$$

Passos

- Número de valores $n = 8$, dados ordenados: $2,\ 4,\ 4,\ 4,\ 5,\ 5,\ 7,\ 9$
- Média aritmética: $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{40}{8} = 5$
- A mediana é o valor central dos dados ordenados: $\tilde{x} = \frac{x_{(4)} + x_{(5)}}{2} = \frac{4+5}{2} = 4{,}5$
- Moda (o valor mais frequente, 3×): $\hat{x} = 4$
- Os quartis são as medianas das metades inferior e superior: $Q_1 = 4, \quad Q_2 = \tilde{x} = 4{,}5, \quad Q_3 = 6$
- Amplitude e amplitude interquartil
  $R = x_{max} - x_{min} = 9 - 2 = 7$
  $IQR = Q_3 - Q_1 = 6 - 4 = 2$
- Soma dos quadrados dos desvios em relação à média: $\sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum x_i^2 - n\bar{x}^2 = 32$
- Variância – amostral s² e populacional σ²
  $s^2 = \frac{1}{n-1}\sum (x_i - \bar{x})^2 = \frac{32}{7} \approx 4{,}571$
  $\sigma^2 = \frac{1}{n}\sum (x_i - \bar{x})^2 = \frac{32}{8} = 4$
- Desvio padrão
  $s = \sqrt{s^2} = \sqrt{4{,}571} \approx 2{,}138$
  $\sigma = \sqrt{4} = 2$
- Coeficiente de variação: $v = \frac{s}{\bar{x}} = \frac{2{,}138}{5} \approx 0{,}4276 = 42{,}8\,\%$
- Barreiras para valores atípicos $Q_1 - 1{,}5\,IQR = 1, \quad Q_3 + 1{,}5\,IQR = 9$ – nenhum valor fica fora.
- A variância amostral s² divide por n − 1 e estima a variância da população – a escolha habitual. A populacional σ² divide por n, para dados que são toda a população.
- Atenção: há poucos dados, as estimativas (sobretudo da variância e dos quartis) são incertas.

Gráfico

Diagrama de caixa (mediana, quartis, valores atípicos; ◆ média)



O valor atípico 31 no diagrama de caixa.

Enunciado: `data(12 15 14 13 16 15 14 31)`

Resultado

$$\bar{x} = 16{,}25$$
$$\tilde{x} = 14{,}5$$
$$s \approx 6{,}089$$

Passos

- Número de valores $n = 8$, dados ordenados: $12,\ 13,\ 14,\ 14,\ 15,\ 15,\ 16,\ 31$
- Média aritmética: $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{130}{8} = 16{,}25$
- A mediana é o valor central dos dados ordenados: $\tilde{x} = \frac{x_{(4)} + x_{(5)}}{2} = \frac{14 + 15}{2} = 14{,}5$
- Moda (o valor mais frequente, 2×): $\hat{x} = 14,\ 15$
- Os quartis são as medianas das metades inferior e superior: $Q_1 = 13{,}5, \quad Q_2 = \tilde{x} = 14{,}5, \quad Q_3 = 15{,}5$
- Amplitude e amplitude interquartil
  $R = x_{max} - x_{min} = 31 - 12 = 19$
  $IQR = Q_3 - Q_1 = 15{,}5 - 13{,}5 = 2$
- Soma dos quadrados dos desvios em relação à média: $\sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum x_i^2 - n\bar{x}^2 = 259{,}5$
- Variância – amostral s² e populacional σ²
  $s^2 = \frac{1}{n-1}\sum (x_i - \bar{x})^2 = \frac{259{,}5}{7} \approx 37{,}07$
  $\sigma^2 = \frac{1}{n}\sum (x_i - \bar{x})^2 = \frac{259{,}5}{8} \approx 32{,}44$
- Desvio padrão
  $s = \sqrt{s^2} = \sqrt{37{,}07} \approx 6{,}089$
  $\sigma = \sqrt{32{,}44} \approx 5{,}695$
- Coeficiente de variação: $v = \frac{s}{\bar{x}} = \frac{6{,}089}{16{,}25} \approx 0{,}3747 = 37{,}5\,\%$
- Os valores atípicos ficam fora das barreiras $Q_1 - 1{,}5\,IQR = 10{,}5, \quad Q_3 + 1{,}5\,IQR = 18{,}5$: 31
- A variância amostral s² divide por n − 1 e estima a variância da população – a escolha habitual. A populacional σ² divide por n, para dados que são toda a população.
- Atenção: há poucos dados, as estimativas (sobretudo da variância e dos quartis) são incertas.

Gráfico

Diagrama de caixa (mediana, quartis, valores atípicos; ◆ média)

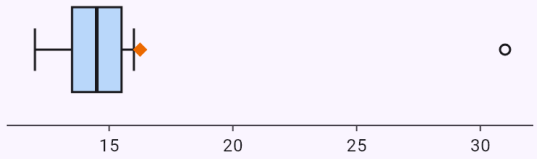


Tabela de frequências de notas e histograma.

Enunciado: `data(1 2 2 3 3 3 3 4 4 5 1 3 2 4 3)` Frequências e histograma

Resultado

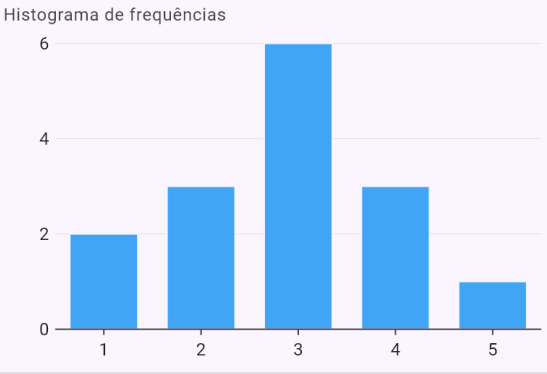
$$n = 15$$
$$\hat{x} = 3$$

Passos

- Número de valores $n = 15$, dados ordenados: $1,\ 1,\ 2,\ 2,\ 2,\ 3,\ 3,\ 3,\ 3,\ 3,\ 3,\ 4,\ 4,\ 4,\ 5$
- Tabela de frequências – absoluta n, relativa f, acumuladas N e F

x_i	n_i	f_i	N_i	F_i
1	2	0,133	2	0,133
2	3	0,200	5	0,333
3	6	0,400	11	0,733
4	3	0,200	14	0,933
5	1	0,067	15	1,000
$\sum$	15	1		

Gráfico



Dados contínuos agrupados em classes.

Enunciado: `data(12.5 14.1 15.8 17.2 18.9 21.3 22.0 23.7 25.1 26.4 28.8 31.2)`
Frequências e histograma

Resultado

$$n = 12$$
$$k = 5, \quad h = 5$$

Passos

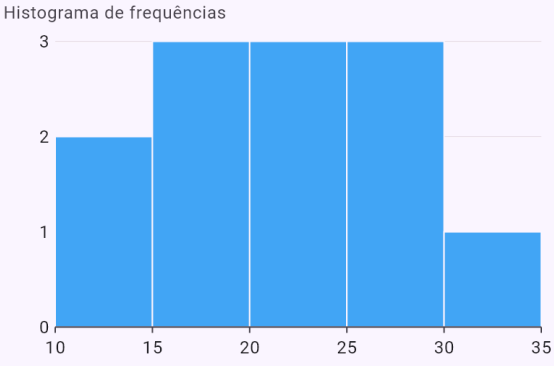
- Número de valores $n = 12$, dados ordenados $12{,}5,\ 14{,}1,\ 15{,}8,\ 17{,}2,\ 18{,}9,\ 21{,}3,\ 22,\ 23{,}7,\ 25{,}1,\ 26,$
- Há muitos valores – agrupamo-los em 5 classes (regra de Sturges)

$$k \approx 1 + \log_2 n = 1 + \log_2 12 \approx 4{,}585 \Rightarrow k = 5$$
$$h \approx \frac{R}{k} = \frac{18{,}7}{5} \approx 3{,}74 \Rightarrow h = 5$$

- Tabela de frequências das classes (x o ponto médio)

I_j	x_j	n_j	f_j	N_j	F_j
$[10,\ 15)$	12,5	2	0,167	2	0,167
$[15,\ 20)$	17,5	3	0,250	5	0,417
$[20,\ 25)$	22,5	3	0,250	8	0,667
$[25,\ 30)$	27,5	3	0,250	11	0,917
$[30,\ 35]$	32,5	1	0,083	12	1,000
$\sum$		12	1		

Gráfico



Pontuações z e T dos valores.

Enunciado: `data(2 4 4 4 5 5 7 9)` Pontuações z

Resultado

$$\bar{x} = 5$$
$$s \approx 2{,}138$$

Passos

- Número de valores $n = 8$, dados ordenados: $2,\ 4,\ 4,\ 4,\ 5,\ 5,\ 7,\ 9$
- Média aritmética: $\bar{x} = \dfrac{40}{8} = 5$
- Desvio padrão

$$s = \sqrt{\frac{1}{n-1}\sum (x_i - \bar{x})^2} = \sqrt{\frac{32}{7}} \approx 2{,}138$$

- Padronizamos cada valor

$$z_i = \frac{x_i - \bar{x}}{s} = \frac{x_i - 5}{2{,}138}$$
$$T_i = 50 + 10 z_i$$

- Pontuações z e T dos valores

x_i	z_i	T_i
2	−1,40	36,0
4	−0,47	45,3
4	−0,47	45,3
4	−0,47	45,3
5	0,00	50,0
5	0,00	50,0
7	0,94	59,4
9	1,87	68,7

- As pontuações z usam a média e o desvio padrão destes dados. Para comparar com uma população usam-se a média e o DP da norma (modelo «Conversão de pontuações padronizadas»).

Correlação e reta de regressão – horas de estudo e pontos no teste.

Enunciado: `data(x y; 1 52; 2 58; 3 61; 4 70; 5 74; 6 79)`

Resultado

$$r \approx 0{,}9937$$

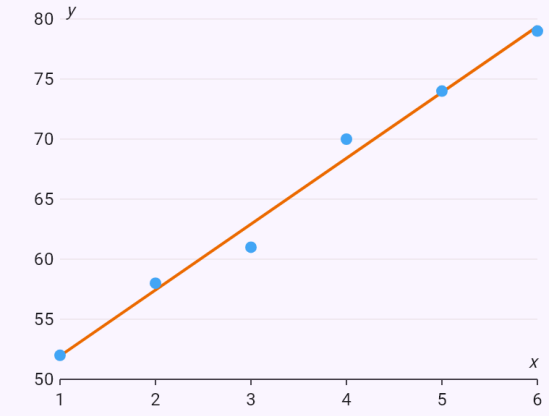
$$\hat{y} = 46{,}47 + 5{,}486x$$

Passos

- Número de valores $n = 6$, dados ordenados
  $[1,\ 52],\ [2,\ 58],\ [3,\ 61],\ [4,\ 70],\ [5,\ 74],\ [6,\ 79]$
- Somas e médias
  $\sum x_i = 21$
  $\sum y_i = 394$
  $\sum x_i^2 = 91$
  $\sum y_i^2 = 26\,410$
  $\sum x_i y_i = 1475$
  $\bar{x} = 3{,}5$
  $\bar{y} = 65{,}67$
- Desvio padrão
  $$s_x = \sqrt{\frac{\sum x_i^2 - n\bar{x}^2}{n-1}} \approx 1{,}871$$
  $$s_y = \sqrt{\frac{\sum y_i^2 - n\bar{y}^2}{n-1}} \approx 10{,}33$$
- Covariância
  $$s_{xy} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{n-1} = \frac{1475 - 6 \cdot 3{,}5 \cdot 65{,}67}{5} = 19{,}2$$
- Coeficiente de correlação de Pearson:
  $r = \frac{s_{xy}}{s_x s_y} = \frac{19{,}2}{1{,}871 \cdot 10{,}33} \approx 0{,}9937$ – a relação é muito forte, positiva (y cresce com x).
- Coeficientes da reta de regressão pelos mínimos quadrados
  $b = \frac{s_{xy}}{s_x^2} = \frac{19{,}2}{3{,}5} \approx 5{,}486$
  $a = \bar{y} - b\bar{x} = 65{,}67 - 5{,}486 \cdot 3{,}5 \approx 46{,}47$
- Reta de regressão: $\hat{y} = 46{,}47 + 5{,}486x$
- Coeficiente de determinação $r^2 = 0{,}9937^2 \approx 0{,}9874$ – a reta explica $98{,}7\,\%$ da variância de y.
- Teste da hipótese nula ρ = 0 (distribuição t com 4 graus de liberdade) – a correlação é estatisticamente significativa.
  $$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0{,}9937\sqrt{4}}{\sqrt{1-0{,}9937^2}} \approx 17{,}73$$
  $p < 0{,}0001 \le \alpha = 0{,}05$
- A correlação descreve só uma relação linear: não diz o que causa o quê e pode não captar uma relação não linear. A reta de regressão só deve usar-se no intervalo dos x medidos.

Gráfico

Diagrama de dispersão com a reta de regressão



Três variáveis lado a lado.

Enunciado: `data(A B C; 12 15 9; 14 18 11; 11 16 10; 15 19 12; 13 17 10)`

Resultado

$$\bar{x}_{\mathrm{A}} = 13$$
$$\bar{x}_{\mathrm{B}} = 17$$
$$\bar{x}_{\mathrm{C}} = 10{,}4$$

Passos

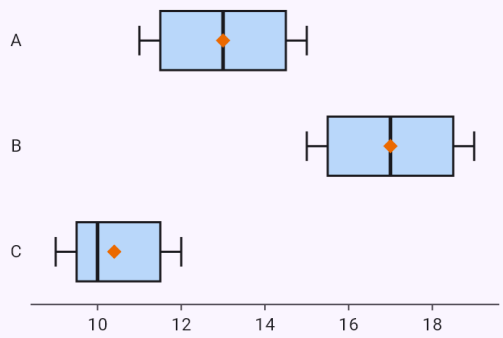
- Estatística descritiva das variáveis

	A	B	C
n	5	5	5
$\bar{x}$	13	17	10,4
$\tilde{x}$	13	17	10
Q_1	11,5	15,5	9,5
Q_3	14,5	18,5	11,5
s	1,581	1,581	1,14
σ	1,414	1,414	1,02
x_{min}	11	15	9
x_{max}	15	19	12

- A variância amostral s² divide por n − 1 e estima a variância da população – a escolha habitual. A populacional σ² divide por n, para dados que são toda a população.

Gráfico

Diagrama de caixa (mediana, quartis, valores atípicos; ◆ média)



Intervalo de confiança de 95 % da média de dez medições.

Enunciado: `data(102 98 110 105 99 107 101 104 96 108)` Intervalo de confiança da média

Resultado

$$95\,\%: \ [99{,}71,\ 106{,}3]$$

Passos

- Tamanho, média e desvio padrão amostral:
  $n = 10, \quad \bar{x} = 103, \quad s \approx 4{,}595$
- O erro padrão: $SE = \dfrac{s}{\sqrt{n}} = \dfrac{4{,}595}{\sqrt{10}} \approx 1{,}453$
- O quantil para a confiança escolhida: $t_{0{,}975}(9) \approx 2{,}262$
- O intervalo de confiança
  $\bar{x} \pm t \cdot SE = 103 \pm 2{,}262 \cdot 1{,}453 = 103 \pm 3{,}287 \Rightarrow [99,$
- Confiança de 95 % significa: se a amostragem fosse repetida muitas vezes, cerca de 95 % dos intervalos calculados assim conteriam a média verdadeira μ. Não significa que μ esteja neste intervalo concreto com probabilidade de 95 %.
- Para uma amostra pequena (n < 30), a distribuição t supõe que os dados vêm de uma distribuição aproximadamente normal. Com dados muito assimétricos ou valores atípicos, um teste não paramétrico é mais adequado.

A média difere de 100? Teste t de uma amostra.

Enunciado: `data(102 98 110 105 99 107 101 104 96 108)` Teste t

Resultado

$$t \approx 2{,}065$$

$$p \approx 0{,}06895 > \alpha = 0{,}05$$

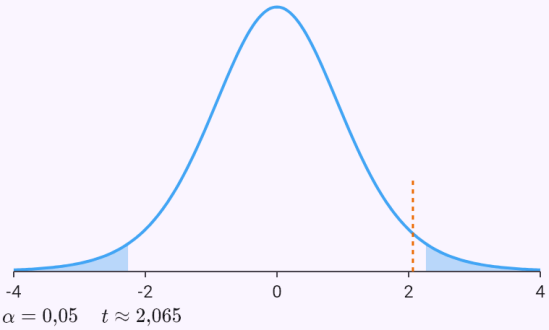
Passos

- Tamanho, média e desvio padrão amostral:
  $n = 10, \quad \bar{x} = 103, \quad s \approx 4{,}595$
- A hipótese nula e a alternativa:
  $H_0 : \mu = 100, \quad H_1 : \mu \neq 100$
- O erro padrão: $SE = \dfrac{s}{\sqrt{n}} = \dfrac{4{,}595}{\sqrt{10}} \approx 1{,}453$
- A estatística de teste – sob H_0 segue uma distribuição t com 9 graus de liberdade:
  $t = \dfrac{\bar{x} - \mu_0}{SE} = \dfrac{103 - 100}{1{,}453} \approx 2{,}065$
- O valor crítico ao nível α: $t_{0{,}975}(9) \approx 2{,}262$, a região crítica $|t| \geq 2{,}262$
- O valor p: $p = 2\,P(T \geq 2{,}065) \approx 0{,}06895$
- Como $p \approx 0{,}06895 > \alpha = 0{,}05$, não rejeitamos a hipótese nula – os dados não fornecem evidência suficiente contra ela.
- O valor p é a probabilidade de um resultado pelo menos tão extremo quanto o observado se H_0 for verdadeira – não a probabilidade de H_0 ser verdadeira. Não rejeitar H_0 não significa que H_0 seja verdadeira; a significância estatística nada diz sobre o tamanho ou a importância prática de uma diferença.

- Para uma amostra pequena (n < 30), a distribuição t supõe que os dados vêm de uma distribuição aproximadamente normal. Com dados muito assimétricos ou valores atípicos, um teste não paramétrico é mais adequado.

Gráfico

Distribuição t – sombreada a região crítica, tracejado o valor da estatística



$\alpha = 0{,}05 \quad t \approx 2{,}065$

Duas amostras independentes de tamanhos diferentes – teste t de Welch.

Enunciado: `data(A = 23 25 28 30 26 27; B = 20 22 21 25 19 24 23)` Teste t

Resultado

$$t \approx 3{,}503$$

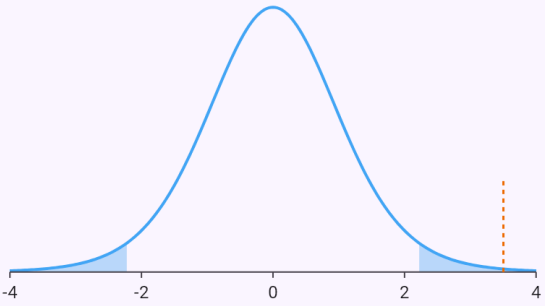
$$p \approx 0{,}005546 \leq \alpha = 0{,}05$$

Passos

- Tamanho, média e desvio padrão amostral
  $n_1 = 6,\ \bar{x}_1 = 26{,}5,\ s_1 \approx 2{,}429$
  $n_2 = 7,\ \bar{x}_2 = 22,\ s_2 \approx 2{,}16$
- A hipótese nula e a alternativa:
  $H_0 : \mu_1 = \mu_2, \quad H_1 : \mu_1 \neq \mu_2$
- O erro padrão
  $$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{2{,}429^2}{6} + \frac{2{,}16^2}{7}} \approx 1{,}285$$
- Os graus de liberdade de Welch-Satterthwaite
  $$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}} \approx 10{,}18$$
- A estatística de teste – sob H_0 segue uma distribuição t com $10{,}18$ graus de liberdade:
  $$t = \frac{\bar{x}_1 - \bar{x}_2}{SE} = \frac{26{,}5 - 22}{1{,}285} \approx 3{,}503$$
- O valor crítico ao nível α: $t_{0{,}975}(10{,}18) \approx 2{,}223$, a região crítica $|t| \geq 2{,}223$
- O valor p: $p = 2\,P(T \geq 3{,}503) \approx 0{,}005546$
- Como $p \approx 0{,}005546 \leq \alpha = 0{,}05$, rejeitamos a hipótese nula – o resultado é estatisticamente significativo.
- O valor p é a probabilidade de um resultado pelo menos tão extremo quanto o observado se H_0 for verdadeira – não a probabilidade de H_0 ser verdadeira. Não rejeitar H_0 não significa que H_0 seja verdadeira; a significância estatística nada diz sobre o tamanho ou a importância prática de uma diferença.
- O teste de Welch não supõe variâncias iguais dos grupos; os seus graus de liberdade v não precisam ser inteiros.

Gráfico

Distribuição t – sombreada a região crítica, tracejado o valor da estatística



$\alpha = 0{,}05 \quad t \approx 3{,}503$

Teste t pareado – medições antes e depois.

Enunciado: `data(x1 x2; 72 70; 75 73; 68 69; 80 76; 77 74; 74 71)` Teste t · Pareadas

Resultado

$$t \approx 3{,}081$$

$$p \approx 0{,}02743 \leq \alpha = 0{,}05$$

Passos

- As diferenças dos pares d (coluna 1 − coluna 2): $2,\ 2,\ -1,\ 4,\ 3,\ 3;\ n = 6,\quad \bar{d} \approx 2{,}167,\quad s_d \approx 1{,}722$
- A hipótese nula e a alternativa: $H_0 : \mu_d = 0,\quad H_1 : \mu_d \neq 0$
- O erro padrão: $SE = \dfrac{s}{\sqrt{n}} = \dfrac{1{,}722}{\sqrt{6}} \approx 0{,}7032$
- A estatística de teste – sob H_0 segue uma distribuição t com 5 graus de liberdade: $t = \dfrac{\bar{d} - \mu_0}{SE} = \dfrac{2{,}167 - 0}{0{,}7032} \approx 3{,}081$
- O valor crítico ao nível α: $t_{0{,}975}(5) \approx 2{,}571$, a região crítica $|t| \geq 2{,}571$
- O valor p: $p = 2\,P(T \geq 3{,}081) \approx 0{,}02743$
- Como $p \approx 0{,}02743 \leq \alpha = 0{,}05$, rejeitamos a hipótese nula – o resultado é estatisticamente significativo.
- O valor p é a probabilidade de um resultado pelo menos tão extremo quanto o observado se H_0 for verdadeira – não a probabilidade de H_0 ser verdadeira. Não rejeitar H_0 não significa que H_0 seja verdadeira; a significância estatística nada diz sobre o tamanho ou a importância prática de uma diferença.
- Para uma amostra pequena (n < 30), a distribuição t supõe que os dados vêm de uma distribuição aproximadamente normal. Com dados muito assimétricos ou valores atípicos, um teste não paramétrico é mais adequado.
- O teste t pareado é um teste t de uma amostra para as diferenças (coluna 1 − coluna 2) – cada par pertence ao mesmo sujeito (p. ex. antes e depois).

Gráfico

Distribuição t – sombreada a região crítica, tracejado o valor da estatística



$\alpha = 0{,}05 \quad t \approx 3{,}081$

O dado é honesto? Teste χ² de aderência.

Enunciado: `data(8 12 9 11 6 14)` Teste χ²

Resultado

$$\chi^2 = 4{,}2$$
$$p \approx 0{,}521 > \alpha = 0{,}05$$

Passos

- Hipótese nula: as 6 categorias são igualmente prováveis ($H_0: p_1 = \cdots = p_6 = \frac{1}{6}$); alternativa: pelo menos uma difere.
- As frequências esperadas sob H_0 $E_i = \frac{N}{k} = \frac{60}{6} = 10$ e as contribuições das categorias

i	O_i	E_i	$\frac{(O_i - E_i)^2}{E_i}$
1	8	10	0,4
2	12	10	0,4
3	9	10	0,1
4	11	10	0,1
5	6	10	1,6
6	14	10	1,6
$\sum$	60	60	4,2

- A estatística de teste – sob H_0 segue uma distribuição χ² com 5 graus de liberdade

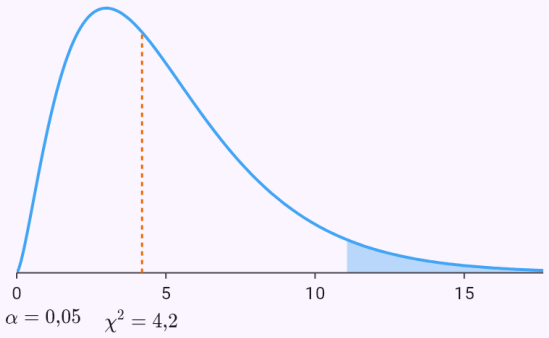
$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = 0{,}4 + 0{,}4 + 0{,}1 + 0{,}1 + 1{,}6 + 1$$

- O valor crítico ao nível α: $\chi^2_{0{,}95}(5) \approx 11{,}07$, a região crítica $\chi^2 \geq 11{,}07$
- O valor p: $p = P(\chi^2 \geq 4{,}2) \approx 0{,}521$
- Como $p \approx 0{,}521 > \alpha = 0{,}05$, não rejeitamos a hipótese nula – os dados não fornecem evidência suficiente contra ela.

- O valor p é a probabilidade de um resultado pelo menos tão extremo quanto o observado se H_0 for verdadeira – não a probabilidade de H_0 ser verdadeira. Não rejeitar H_0 não significa que H_0 seja verdadeira; a significância estatística nada diz sobre o tamanho ou a importância prática de uma diferença.
- O teste χ² trabalha com frequências (números de casos), não com porcentagens ou médias.

Gráfico

Distribuição χ² – sombreada a região crítica, tracejado o valor da estatística



Teste de independência numa tabela 2×2.

Enunciado: `data(20 30; 25 15)` Teste χ²

Resultado

$$\chi^2 = 4{,}5$$
$$p \approx 0{,}03389 \leq \alpha = 0{,}05$$
$$V \approx 0{,}2236$$

Passos

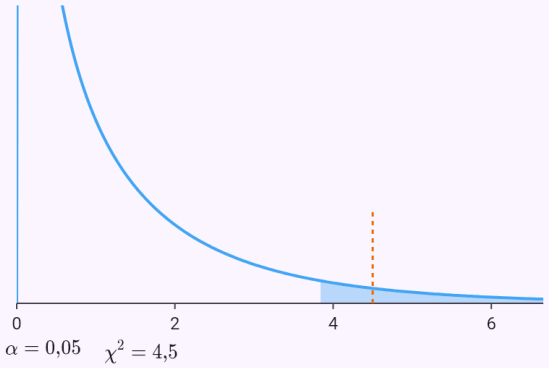
- Hipótese nula: a variável das linhas e a das colunas são independentes; alternativa: são dependentes.
- As frequências esperadas sob independência $E_{ij} = \frac{n_{i\cdot}\, n_{\cdot j}}{N}$, entre parênteses as observadas

	1	2	∑
1	25 (20)	25 (30)	50
2	20 (25)	20 (15)	40
∑	45	45	90

- A estatística de teste – sob H_0 segue uma distribuição χ² com 1 graus de liberdade: $\chi^2 = \sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = 1 + 1 + 1{,}25 + 1{,}25 = 4{,}5$
- O valor crítico ao nível α: $\chi^2_{0{,}95}(1) \approx 3{,}841$, a região crítica $\chi^2 \geq 3{,}841$
- O valor p: $p = P(\chi^2 \geq 4{,}5) \approx 0{,}03389$
- Como $p \approx 0{,}03389 \leq \alpha = 0{,}05$, rejeitamos a hipótese nula – o resultado é estatisticamente significativo.
- O valor p é a probabilidade de um resultado pelo menos tão extremo quanto o observado se H_0 for verdadeira – não a probabilidade de H_0 ser verdadeira. Não rejeitar H_0 não significa que H_0 seja verdadeira; a significância estatística nada diz sobre o tamanho ou a importância prática de uma diferença.
- A força da associação – V de Cramér (0 nenhuma, 1 total) $V = \sqrt{\frac{\chi^2}{N(\min(r,c)-1)}} = \sqrt{\frac{4{,}5}{90 \cdot 1}} \approx 0{,}2236$
- O teste χ² trabalha com frequências (números de casos), não com porcentagens ou médias.

Gráfico

Distribuição χ² – sombreada a região crítica, tracejado o valor da estatística



Dois grupos pelo teste de Mann-Whitney – amostras pequenas, valor p exato.

Enunciado: `data(A = 23 25 28 30 26 27; B = 20 22 21 25 19 24 23)`

Teste de postos (Mann-Whitney, Wilcoxon, Kruskal-Wallis)

Resultado

$$U = 3$$
$$z \approx -2{,}579$$
$$p \approx 0{,}009922 \leq \alpha = 0{,}05$$

Passos

- Hipótese nula: os dois grupos vêm da mesma distribuição, alternativa: os valores de um tendem a ser maiores

  H_0: $F_{\mathrm{A}} = F_{\mathrm{B}}$
  H_1: $F_{\mathrm{A}} \neq F_{\mathrm{B}}$

- Ordenamos juntos os valores dos dois grupos e numeramo-los pelo posto R (valores iguais recebem a média dos seus postos)

x	R	
19	1	B
20	2	B
21	3	B
22	4	B
23	5,5	A
23	5,5	B
24	7	B
25	8,5	A
25	8,5	B
26	10	A
27	11	A
28	12	A
30	13	A

- Tamanhos dos grupos e somas dos postos: $n_1 = 6, \quad n_2 = 7, \quad R_1 = 60, \quad R_2 = 31$

- A estatística U de Mann-Whitney – o número de pares em que o valor de um grupo tem posto maior que o do outro; testa-se a menor

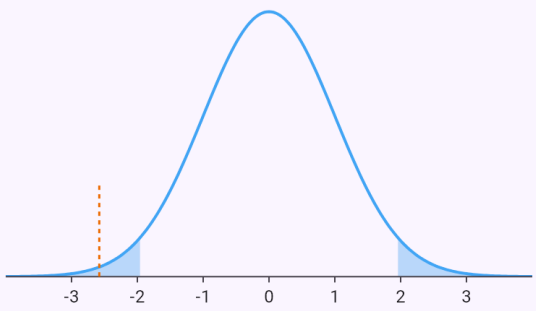
$$U_1 = R_1 - \frac{n_1(n_1+1)}{2} = 60 - \frac{6 \cdot 7}{2} = 39$$
$$U_2 = n_1 n_2 - U_1 = 6 \cdot 7 - 39 = 3$$
$$U = \min(U_1, U_2) = 3$$

- A estatística de teste – sob H_0 segue a distribuição normal padrão: $z = \dfrac{U - \frac{n_1 n_2}{2}}{\sigma_U} = \dfrac{3 - 21}{6{,}981} \approx -2{,}579$

- O valor crítico ao nível α: $z_{0{,}975} \approx 1{,}96$, a região crítica $|z| \geq 1{,}96$
- O valor p: $p = 2\,P(Z \geq 2{,}579) \approx 0{,}009922$
- Como $p \approx 0{,}009922 \leq \alpha = 0{,}05$, rejeitamos a hipótese nula – o resultado é estatisticamente significativo.
- O valor p é a probabilidade de um resultado pelo menos tão extremo quanto o observado se H_0 for verdadeira – não a probabilidade de H_0 ser verdadeira. Não rejeitar H_0 não significa que H_0 seja verdadeira; a significância estatística nada diz sobre o tamanho ou a importância prática de uma diferença.
- O tamanho do efeito r (orientação: 0,1 pequeno, 0,3 médio, 0,5 grande): $r = \dfrac{|z|}{\sqrt{N}} = \dfrac{2{,}579}{\sqrt{13}} \approx 0{,}7152$
- O teste de Mann-Whitney é um teste de postos: não pressupõe distribuição normal e serve também para dados ordinais e valores atípicos. Compara as distribuições inteiras – com formas semelhantes, é um teste de diferença de medianas.

Gráfico

Distribuição normal padrão – sombreada a região crítica, tracejado o valor da estatística



$\alpha = 0{,}05 \quad z \approx -2{,}579$

Valores emparelhados antes e depois pelo teste de Wilcoxon.

Enunciado: `data(pred po; 72 70; 75 73; 68 69; 80 76; 77 74; 74 71; 79 72)`

Teste de postos (Mann-Whitney, Wilcoxon, Kruskal-Wallis) · Pareadas

Resultado

$$T = 1$$
$$z \approx -2{,}205$$
$$p \approx 0{,}02743 \leq \alpha = 0{,}05$$

Passos

- Hipótese nula: a mediana das diferenças é zero (sem mudança), alternativa: não é: $H_0 : \tilde{d} = 0, \quad H_1 : \tilde{d} \neq 0$
- As diferenças dos pares d (1.ª coluna − 2.ª coluna): $2,\ 2,\ -1,\ 4,\ 3,\ 3,\ 7$; as diferenças nulas omitem-se: $n = 7$
- Ordenamos as diferenças não nulas pelo tamanho (sem sinal) e numeramo-las pelo posto R; o sinal fica com o posto

d_i	R	
-1	1	$-$
2	$2{,}5$	$+$
2	$2{,}5$	$+$
3	$4{,}5$	$+$
3	$4{,}5$	$+$
4	6	$+$
7	7	$+$

- As somas dos postos das diferenças positivas e negativas
  $T^+ = 27$
  $T^- = 1$
  $T = \min\left(T^+, T^-\right) = 1$
- A estatística de teste – sob H_0 segue a distribuição normal padrão: $z = \dfrac{T - \frac{n(n+1)}{4}}{\sigma_T} = \dfrac{1 - 14}{5{,}895} \approx -2{,}205$
- O valor crítico ao nível α: $z_{0{,}975} \approx 1{,}96$, a região crítica $|z| \geq 1{,}96$
- O valor p: $p = 2\,P(Z \geq 2{,}205) \approx 0{,}02743$
- Como $p \approx 0{,}02743 \leq \alpha = 0{,}05$, rejeitamos a hipótese nula – o resultado é estatisticamente significativo.
- O valor p é a probabilidade de um resultado pelo menos tão extremo quanto o observado se H_0 for verdadeira – não a probabilidade de H_0 ser verdadeira. Não rejeitar H_0 não significa que H_0 seja verdadeira; a significância estatística nada diz sobre o tamanho ou a importância prática de uma diferença.
- O tamanho do efeito r (orientação: 0,1 pequeno, 0,3 médio, 0,5 grande): $r = \dfrac{|z|}{\sqrt{n}} = \dfrac{2{,}205}{\sqrt{7}} \approx 0{,}8335$
- O teste de Wilcoxon é a versão por postos do teste t emparelhado: não pressupõe diferenças normais e tem em conta o sinal e o tamanho pelos postos.

Gráfico

Distribuição normal padrão – sombreada a região crítica, tracejado o valor da estatística



$\alpha = 0{,}05 \quad z \approx -2{,}205$

Três grupos pelo teste de Kruskal-Wallis.

Enunciado: `data(A = 12 15 14 11 13; B = 16 18 17 15 19; C = 13 14 12 16 15)`

Teste de postos (Mann-Whitney, Wilcoxon, Kruskal-Wallis)

Resultado

$$H \approx 8{,}263$$
$$p \approx 0{,}01606 \leq \alpha = 0{,}05$$

Passos

- Hipótese nula: todos os grupos vêm da mesma distribuição, alternativa: pelo menos um difere
  $H_0 : F_1 = \cdots = F_3$
  $H_1 : \exists\, i, j : F_i \neq F_j$

- Ordenamos juntos os valores de todos os grupos e numeramo-los pelo posto R (valores iguais recebem a média dos seus postos)

x	R	
11	1	A
12	2,5	A
12	2,5	C
13	4,5	A
13	4,5	C
14	6,5	A
14	6,5	C
15	9	A
15	9	B
15	9	C
16	11,5	B
16	11,5	C
17	13	B
18	14	B
19	15	B

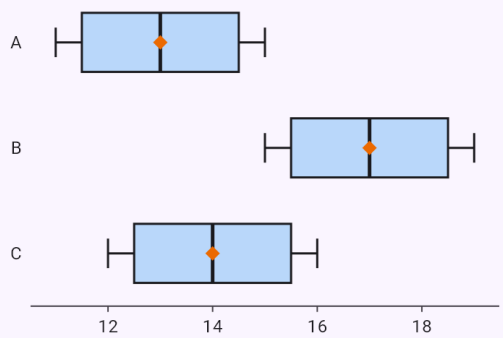
- A soma dos postos de cada grupo
  $R_{\mathrm{A}} = 23{,}5 \ (n = 5)$
  $R_{\mathrm{B}} = 62{,}5 \ (n = 5)$
  $R_{\mathrm{C}} = 34 \ (n = 5)$

- A estatística de teste – sob H_0 segue uma distribuição χ^2 com 2 graus de liberdade
  $$H = \frac{12}{N(N+1)} \sum \frac{R_j^2}{n_j} - 3(N+1) = \frac{12}{15 \cdot 16} \left(\frac{23{,}5^2}{5} \right.$$
  $$C = 1 - \frac{\sum (t^3 - t)}{N^3 - N} = 1 - \frac{48}{3360} \approx 0{,}9857$$
  $$H_c = \frac{H}{C} = \frac{8{,}145}{0{,}9857} \approx 8{,}263$$

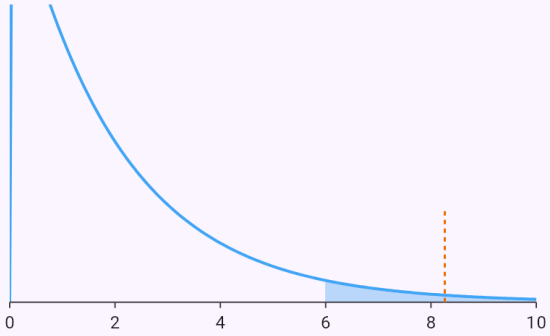
- O valor crítico ao nível α: $\chi^2_{0{,}95}(2) \approx 5{,}991$, a região crítica $\chi^2 \geq 5{,}991$
- O valor p: $p = P(\chi^2 \geq 8{,}263) \approx 0{,}01606$
- Como $p \approx 0{,}01606 \leq \alpha = 0{,}05$, rejeitamos a hipótese nula – o resultado é estatisticamente significativo.
- O valor p é a probabilidade de um resultado pelo menos tão extremo quanto o observado se H_0 for verdadeira – não a probabilidade de H_0 ser verdadeira. Não rejeitar H_0 não significa que H_0 seja verdadeira; a significância estatística nada diz sobre o tamanho ou a importância prática de uma diferença.
- O tamanho do efeito ε² – a parte da variância dos postos explicada pelos grupos:
  $$\varepsilon^2 = \frac{H}{N-1} = \frac{8{,}263}{14} \approx 0{,}5902$$
- O teste de Kruskal-Wallis é a versão por postos da ANOVA de um fator. Diz apenas que alguns grupos diferem – quais, mostra-o a comparação por pares (Mann-Whitney com correção de Bonferroni).

Gráfico

Diagrama de caixa (mediana, quartis, valores atípicos; ◆ média)



Distribuição χ² – sombreada a região crítica, tracejado o valor da estatística



$\alpha = 0{,}05$ $\chi^2 \approx 8{,}263$

As médias de três grupos pela ANOVA de um fator.

Enunciado: `data(A = 12 15 14 11 13; B = 16 18 17 15 19; C = 13 14 12 16 15)`

ANOVA (análise de variância)

Resultado

$$F \approx 8{,}667$$
$$p \approx 0{,}004687 \leq \alpha = 0{,}05$$
$$\eta^2 \approx 0{,}5909$$

Passos

- Hipótese nula: as médias de todos os grupos são iguais, alternativa: pelo menos duas diferem

  $H_0 \colon \mu_1 = \cdots = \mu_3$

  $H_1 \colon \exists\, i, j \colon \mu_i \neq \mu_j$

- Tamanhos, médias e desvios padrão dos grupos e a média global $\bar{x} \approx 14{,}67$

	n_j	$\bar{x}_j$	s_j
A	5	13	1,581
B	5	17	1,581
C	5	14	1,581
$\sum$	15	14,67	

- Análise de variância – soma de quadrados entre grupos (B) e dentro deles (W), graus de liberdade e quadrados médios

  $SS_B = \sum n_j (\bar{x}_j - \bar{x})^2$

  $SS_W = \sum \sum (x_{ij} - \bar{x}_j)^2$

	SS	df	MS	F
B	43,33	2	21,67	8,667
W	30	12	2,5	
$\sum$	73,33	14		

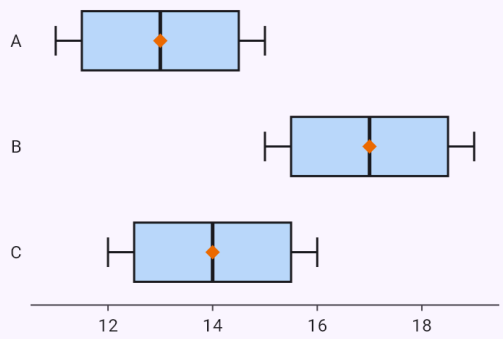
- A estatística de teste – sob H_0 segue a distribuição F com 2 e 12 graus de liberdade:

  $$F = \frac{MS_B}{MS_W} = \frac{21{,}67}{2{,}5} \approx 8{,}667$$

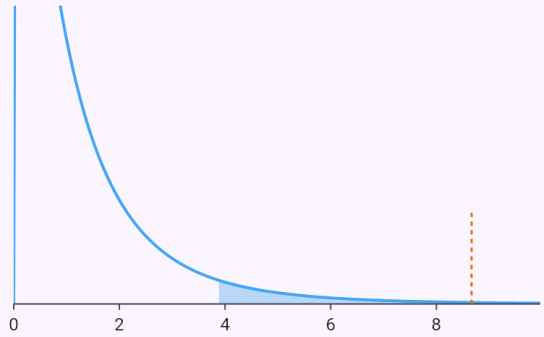
- O valor crítico ao nível α: $F_{0{,}95}(2;\, 12) \approx 3{,}885$, a região crítica $F \geq 3{,}885$
- O valor p: $p = P(F \geq 8{,}667) \approx 0{,}004687$
- Como $p \approx 0{,}004687 \leq \alpha = 0{,}05$, rejeitamos a hipótese nula – o resultado é estatisticamente significativo.
- O tamanho do efeito η² – a parte da variância explicada pelas diferenças entre grupos (orientação: 0,01 pequeno, 0,06 médio, 0,14 grande): $\eta^2 = \frac{SS_B}{SS_T} = \frac{43{,}33}{73{,}33} \approx 0{,}5909$
- O valor p é a probabilidade de um resultado pelo menos tão extremo quanto o observado se H_0 for verdadeira – não a probabilidade de H_0 ser verdadeira. Não rejeitar H_0 não significa que H_0 seja verdadeira; a significância estatística nada diz sobre o tamanho ou a importância prática de uma diferença.
- A ANOVA pressupõe grupos independentes, uma distribuição aproximadamente normal em cada um e variâncias semelhantes. Se não for o caso (grupos pequenos, atípicos, dados ordinais), use o teste de Kruskal-Wallis.
- Rejeitar H_0 diz apenas que pelo menos duas médias diferem. Quais, mostra-o a comparação por pares (testes post hoc, p. ex. testes t com correção de Bonferroni).

Gráfico

Diagrama de caixa (mediana, quartis, valores atípicos; ◆ média)



Distribuição F – sombreada a região crítica, tracejado o valor da estatística



$\alpha = 0{,}05$ $F \approx 8{,}667$

Teste de ajustamento às proporções esperadas 1 : 2 : 1.

Enunciado: `data(30 50 20)` Teste χ² · Proporções esperadas 1:2:1

Resultado

$$\chi^2 = 2$$
$$p \approx 0{,}3679 > \alpha = 0{,}05$$

Passos

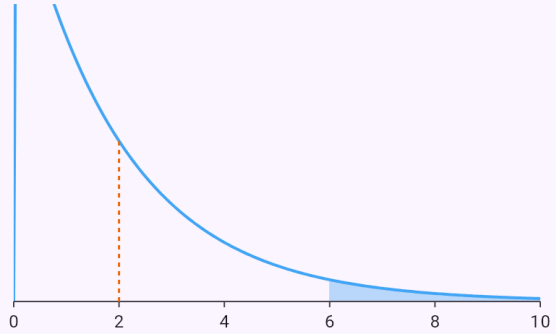
- Hipótese nula: as categorias têm as proporções dadas, alternativa: pelo menos uma difere:
  $H_0: p_1 = 0{,}25,\ p_2 = 0{,}5,\ p_3 = 0{,}25$
- As frequências esperadas pelas proporções dadas $E_i = N \cdot p_i, \quad N = 100$ e as contribuições das categorias

i	O_i	E_i	$\frac{(O_i - E_i)^2}{E_i}$
1	30	25	1
2	50	50	0
3	20	25	1
$\sum$	100	100	2

- A estatística de teste – sob H_0 segue uma distribuição χ² com 2 graus de liberdade:
  $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = 1 + 0 + 1 = 2$
- O valor crítico ao nível α: $\chi^2_{0{,}95}(2) \approx 5{,}991$, a região crítica $\chi^2 \geq 5{,}991$
- O valor p: $p = P(\chi^2 \geq 2) \approx 0{,}3679$
- Como $p \approx 0{,}3679 > \alpha = 0{,}05$, não rejeitamos a hipótese nula – os dados não fornecem evidência suficiente contra ela.
- O valor p é a probabilidade de um resultado pelo menos tão extremo quanto o observado se H_0 for verdadeira – não a probabilidade de H_0 ser verdadeira. Não rejeitar H_0 não significa que H_0 seja verdadeira; a significância estatística nada diz sobre o tamanho ou a importância prática de uma diferença.
- O teste χ² trabalha com frequências (números de casos), não com porcentagens ou médias.

Gráfico

Distribuição χ² – sombreada a região crítica, tracejado o valor da estatística



Regressão quadrática.

Enunciado: `data(x y; 1 2.1; 2 3.9; 3 9.2; 4 15.8; 5 25.1; 6 36.2)`

Correlação e regressão · Quadrática

Resultado

$$\hat{y} = 1{,}73 - 0{,}9354x + 1{,}116x^2$$

$$R^2 \approx 0{,}9997$$

Passos

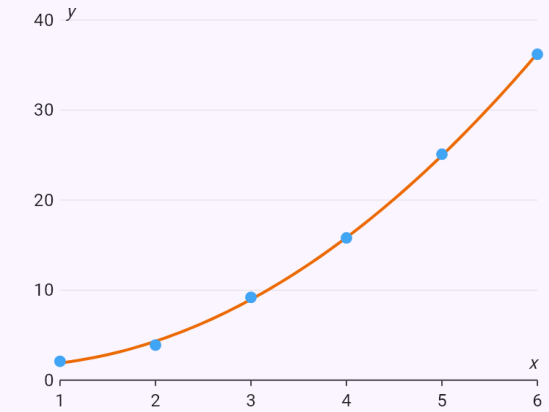
- Número de valores $n = 6$, dados ordenados
  $[1,\ 2{,}1],\ [2,\ 3{,}9],\ [3,\ 9{,}2],\ [4,\ 15{,}8],\ [5,\ 25{,}1],\ [6,\ 36{,}2]$
- Somas e médias
  $\sum x_i = 21$
  $\sum x_i^2 = 91$
  $\sum x_i^3 = 441$
  $\sum x_i^4 = 2275$
  $\sum y_i = 92{,}3$
  $\sum x_i y_i = 443{,}4$
  $\sum x_i^2 y_i = 2284$
- As equações normais de mínimos quadrados para ŷ = a + bx + cx²
  $$6a + 21b + 91c = 92{,}3$$
  $$21a + 91b + 441c = 443{,}4$$
  $$91a + 441b + 2275c = 2284$$
- A curva de regressão: $\hat{y} = 1{,}73 - 0{,}9354x + 1{,}116x^2$
- O coeficiente de determinação – a curva explica $99{,}97\,\%$ da variância de y.

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} = 1 - \frac{0{,}3012}{874{,}5} \approx 0{,}9997$$

- A curva de regressão só é razoável dentro do intervalo dos x medidos – fora dele (extrapolação) a estimativa pode ser muito imprecisa.

Gráfico

Diagrama de dispersão com a curva de regressão



Regressão exponencial por linearização.

Enunciado: `data(x y; 0 2.0; 1 2.7; 2 3.7; 3 5.0; 4 6.8; 5 9.1)`

Correlação e regressão · Exponencial

Resultado

$$\hat{y} = 2{,}002\, e^{0{,}3042x}$$

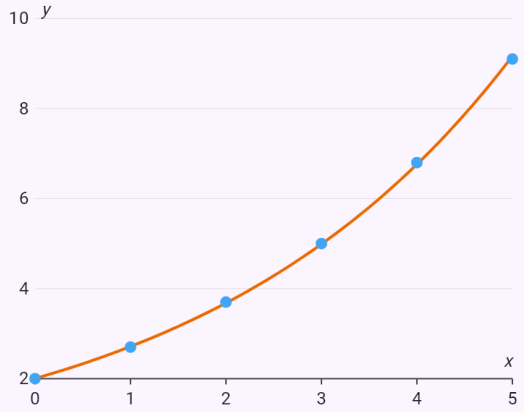
$$R^2 \approx 0{,}9998$$

Passos

- Número de valores $n = 6$, dados ordenados
  $[0,\ 2],\ [1,\ 2{,}7],\ [2,\ 3{,}7],\ [3,\ 5],\ [4,\ 6{,}8],\ [5,\ 9{,}1]$
- Aplicando logaritmos, o modelo torna-se uma reta nas variáveis (u, v): $y = a\, e^{bx} \Rightarrow \ln y = \ln a + bx$; os pares transformados
  $[0,\ 0{,}6931],\ [1,\ 0{,}9933],\ [2,\ 1{,}308],\ [3,\ 1{,}609],\ [4,\ 1{,}91$
- Coeficientes da reta de regressão pelos mínimos quadrados
  $$b = \frac{\sum (u_i - \bar{u})(v_i - \bar{v})}{\sum (u_i - \bar{u})^2} = \frac{5{,}324}{17{,}5} \approx 0{,}3042$$
  $$\ln a = \overline{\ln y} - b\bar{x} = 1{,}455 - 0{,}3042 \cdot 2{,}5 \approx 0{,}6943$$
- O coeficiente a pela transformação inversa:
  $a = e^{0{,}6943} \approx 2{,}002$
- A curva de regressão: $\hat{y} = 2{,}002\, e^{0{,}3042x}$
- O coeficiente de determinação – a curva explica $99{,}98\,\%$ da variância de y.
  $$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} = 1 - \frac{0{,}006563}{35{,}95} \approx 0{,}9998$$
- Os coeficientes são ajustados por mínimos quadrados aos dados transformados (logaritmos), não aos valores originais – o resultado pode diferir um pouco de uma regressão não linear em programas estatísticos.
- A curva de regressão só é razoável dentro do intervalo dos x medidos – fora dele (extrapolação) a estimativa pode ser muito imprecisa.

Gráfico

Diagrama de dispersão com a curva de regressão



Teste t unilateral à direita a partir de valores resumidos: n = 25, x̄ = 104, s = 15, μ_0 = 100.

Enunciado: `model(ttest; find=test[one]; n=25[one]; mean=104[one]; sd=15[one]; mu0=100[one];`

```
alpha=0.05[one]; alternative=greater; sigma=sample)
```

Resultado

$t \approx 1{,}333$

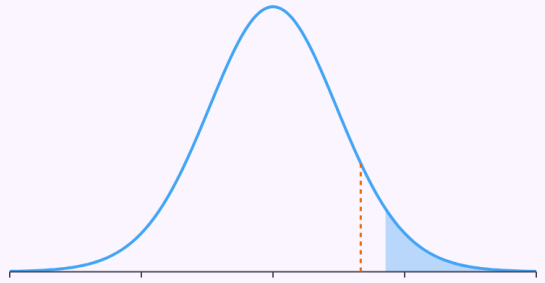
$p \approx 0{,}09747 > \alpha = 0{,}05$

Passos

- Procuramos t; dados:
  $n = 25, \quad \bar{x} = 104, \quad s = 15, \quad \mu_0 = 100, \quad \alpha = 0{,}05$
- O teste t verifica se a média populacional difere de um valor μ₀. A estatística t = (x̄ − μ₀)/(s/√n) segue, sob a hipótese nula, uma distribuição t com n − 1 graus de liberdade; o valor p indica quão provável seria um resultado assim ou mais extremo se H_0 fosse verdadeira. Com σ conhecido é um teste z. Equação do modelo: $t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$
- A hipótese nula e a alternativa:
  $H_0 : \mu = 100, \quad H_1 : \mu > 100$
- O erro padrão: $SE = \dfrac{s}{\sqrt{n}} = \dfrac{15}{\sqrt{25}} = 3$
- A estatística de teste – sob H_0 segue uma distribuição t com 24 graus de liberdade:
  $t = \dfrac{\bar{x} - \mu_0}{SE} = \dfrac{104 - 100}{3} \approx 1{,}333$
- O valor crítico ao nível α: $t_{0{,}95}(24) \approx 1{,}711$, a região crítica $t \geq 1{,}711$
- O valor p: $p = P(T \geq 1{,}333) \approx 0{,}09747$
- Como $p \approx 0{,}09747 > \alpha = 0{,}05$, não rejeitamos a hipótese nula – os dados não fornecem evidência suficiente contra ela.
- O valor p é a probabilidade de um resultado pelo menos tão extremo quanto o observado se H_0 for verdadeira – não a probabilidade de H_0 ser verdadeira. Não rejeitar H_0 não significa que H_0 seja verdadeira; a significância estatística nada diz sobre o tamanho ou a importância prática de uma diferença.
- Para uma amostra pequena (n < 30), a distribuição t supõe que os dados vêm de uma distribuição aproximadamente normal. Com dados muito assimétricos ou valores atípicos, um teste não paramétrico é mais adequado.

Gráfico

Distribuição t – sombreada a região crítica, tracejado o valor da estatística



$\alpha = 0{,}05 \quad t \approx 1{,}333$

Teste t de Welch de dois grupos a partir de valores resumidos.

Enunciado: `model(twosample; find=test[one]; n1=30[one]; mean1=52[one]; sd1=8[one];`

```
n2=28[one]; mean2=47[one]; sd2=10[one]; alpha=0.05[one]; variances=welch; alternative=two)
```

Resultado

$$t \approx 2{,}093$$
$$p \approx 0{,}04124 \leq \alpha = 0{,}05$$

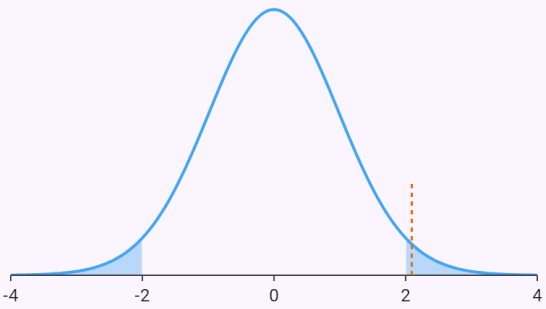
Passos

- Procuramos t; dados
  $n_1 = 30, \quad \bar{x}_1 = 52, \quad s_1 = 8, \quad n_2 = 28, \quad \bar{x}_2 = 47,$
- Comparação das médias de dois grupos independentes. O teste de Welch não supõe variâncias iguais e é hoje a opção recomendada; o teste t clássico com variância combinada supõe que ambos os grupos têm a mesma variância. Bastam os tamanhos, médias e desvios padrão dos dois grupos. Equação do modelo:
  $t = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$
- A hipótese nula e a alternativa:
  $H_0 : \mu_1 = \mu_2, \quad H_1 : \mu_1 \neq \mu_2$
- O erro padrão
  $$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{8^2}{30} + \frac{10^2}{28}} \approx 2{,}388$$
- Os graus de liberdade de Welch-Satterthwaite
  $$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}} \approx 51{,}71$$
- A estatística de teste – sob H_0 segue uma distribuição t com $51{,}71$ graus de liberdade:
  $$t = \frac{\bar{x}_1 - \bar{x}_2}{SE} = \frac{52 - 47}{2{,}388} \approx 2{,}093$$
- O valor crítico ao nível α: $t_{0{,}975}(51{,}71) \approx 2{,}007$, a região crítica $|t| \geq 2{,}007$
- O valor p: $p = 2\,P(T \geq 2{,}093) \approx 0{,}04124$

- Como $p \approx 0{,}04124 \leq \alpha = 0{,}05$, rejeitamos a hipótese nula – o resultado é estatisticamente significativo.
- O valor p é a probabilidade de um resultado pelo menos tão extremo quanto o observado se H_0 for verdadeira – não a probabilidade de H_0 ser verdadeira. Não rejeitar H_0 não significa que H_0 seja verdadeira; a significância estatística nada diz sobre o tamanho ou a importância prática de uma diferença.
- O teste de Welch não supõe variâncias iguais dos grupos; os seus graus de liberdade v não precisam ser inteiros.

Gráfico

Distribuição t – sombreada a região crítica, tracejado o valor da estatística



$\alpha = 0{,}05 \quad t \approx 2{,}093$

Distribuição normal: P(X < 85) para N(100; 15²).

Enunciado: `model(normal; find=below[one]; mu=100[one]; sigma=15[one]; x=85[one])`

Resultado

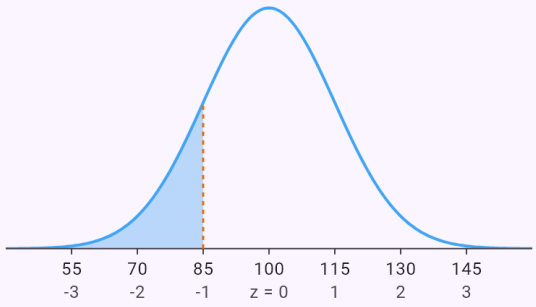
$$P(X < 85) \approx 0{,}1587 = 15{,}87\,\%$$

Passos

- Procuramos P; dados: $\mu = 100, \quad \sigma = 15, \quad x = 85$
- Muitas grandezas (altura, desempenho num teste, erro de medida) variam simetricamente em torno da média μ, tanto mais raramente quanto mais longe dela – é a distribuição normal (de Gauss) N(μ; σ²) com desvio padrão σ. As probabilidades calculam-se padronizando, z = (x − μ)/σ, com a função de distribuição Φ da normal padrão N(0; 1): P(X < x) = Φ(z). Um quantil é o valor abaixo do qual fica a proporção p. Equação do modelo
  $X \sim N(\mu,\, \sigma^2)$
  $Z = \dfrac{X - \mu}{\sigma} \sim N(0,\, 1)$
- Isolamos a grandeza procurada: $z = \dfrac{x - \mu}{\sigma}$
- Substituímos: $z = \dfrac{85 - 100}{15} = -1$
- Substituímos
  $P(X < 85) = \Phi(-1) = (1 - \Phi(1)) \approx (1 - 0{,}8413) \approx 0$

Gráfico

Densidade normal, a área colorida é a probabilidade procurada



$P \approx 0{,}1587 \quad x = 85$

A proporção de pessoas entre 160 e 185 cm com μ = 170, σ = 8.

Enunciado: `model(normal; find=between[one]; mu=170[one]; sigma=8[one]; a=160[one];`

b=185[one])

Resultado

$$P(160 < X < 185) \approx 0{,}864 = 86{,}4\,\%$$

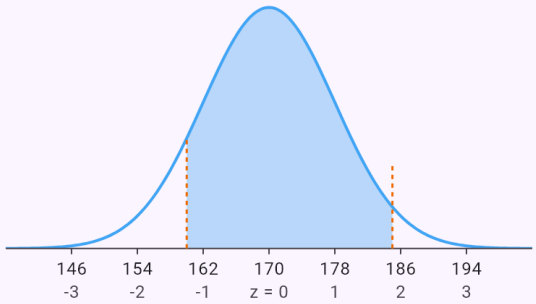
Passos

- Procuramos P; dados:
  $\mu = 170, \quad \sigma = 8, \quad a = 160, \quad b = 185$
- Muitas grandezas (altura, desempenho num teste, erro de medida) variam simetricamente em torno da média μ, tanto mais raramente quanto mais longe dela – é a distribuição normal (de Gauss) N(μ; σ²) com desvio padrão σ. As probabilidades calculam-se padronizando, z = (x − μ)/σ, com a função de distribuição Φ da normal padrão N(0; 1): P(X < x) = Φ(z). Um quantil é o valor abaixo do qual fica a proporção p. Equação do modelo
  $X \sim N(\mu,\, \sigma^2)$
  $Z = \dfrac{X - \mu}{\sigma} \sim N(0,\, 1)$
- Isolamos a grandeza procurada:
  $z_a = \dfrac{a - \mu}{\sigma}, \quad z_b = \dfrac{b - \mu}{\sigma}$
- Substituímos:
  $z_a = \dfrac{160 - 170}{8} = -1{,}25, \quad z_b = \dfrac{185 - 170}{8} = 1{,}875$
- Substituímos

$$P(160 < X < 185) = \Phi(1{,}875) - \Phi(-1{,}25) \approx 0{,}9696$$

Gráfico

Densidade normal, a área colorida é a probabilidade procurada



$P \approx 0{,}864 \quad a \quad b$

Um quantil: a pontuação abaixo da qual fica 95 % da população.

Enunciado: `model(normal; find=quantile[one]; mu=100[one]; sigma=15[one]; p=95[pct])`

Resultado

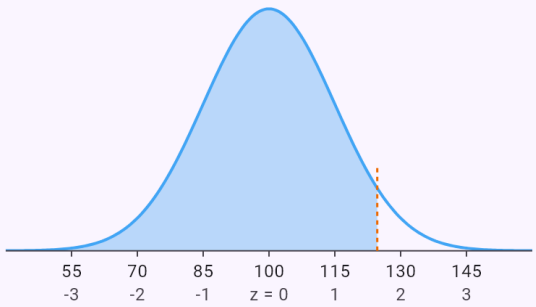
$$x \approx 124{,}7$$

Passos

- Procuramos x; dados: $\mu = 100, \quad \sigma = 15, \quad p = 95\,\%$
- Convertemos as unidades: $p = 95\,\% = 0{,}95$
- Muitas grandezas (altura, desempenho num teste, erro de medida) variam simetricamente em torno da média μ, tanto mais raramente quanto mais longe dela – é a distribuição normal (de Gauss) N(μ; σ²) com desvio padrão σ. As probabilidades calculam-se padronizando, z = (x − μ)/σ, com a função de distribuição Φ da normal padrão N(0; 1): P(X < x) = Φ(z). Um quantil é o valor abaixo do qual fica a proporção p. Equação do modelo
  $X \sim N(\mu,\, \sigma^2)$
  $Z = \dfrac{X - \mu}{\sigma} \sim N(0,\, 1)$
- Isolamos a grandeza procurada
  $P(X < x) = \Phi\left(\dfrac{x - \mu}{\sigma}\right) = p \Rightarrow x = \mu + z_p\sigma, \quad z_p =$
- Substituímos
  $z_{0{,}95} = \Phi^{-1}(0{,}95) \approx 1{,}645, \quad x = 100 + 1{,}645 \cdot 15 \approx 1$

Gráfico

Densidade normal, a área colorida é a probabilidade procurada



$p = 0{,}95 \quad x \approx 124{,}7$

Limitações conhecidas

- Os quartis são as medianas das metades (o método escolar); os programas estatísticos (Excel, SPSS) usam outras fórmulas e com poucos dados podem obter valores um pouco diferentes.
- A comparação por pares depois da ANOVA ou do teste de Kruskal-Wallis (post hoc) e a ANOVA de vários fatores ainda não existem – compare os pares por um teste t ou de Mann-Whitney com correção de Bonferroni; as regressões não lineares calculam-se por linearização.
- Os dados introduzem-se só à mão, pela grelha ou por um ficheiro CSV – o app ainda não reconhece uma tabela numa foto.