Dados bivariados — dispersão, correlação e reta de regressão

10.º Ano | Matemática B | 4 páginas | 12 exercícios

Competências que desenvolve: Interpretar um diagrama de dispersão (forma, direção e força), classificar o coeficiente de correlação linear, determinar e usar a reta de regressão com tecnologia, distinguir correlação de causalidade e prever dentro do intervalo dos dados.

Revisão rápida

- **Par de dados:** cada unidade estatística dá (x, y) — x é a variável *explanatória* (independente) e y a *resposta* (dependente).
- **Diagrama de dispersão:** ler *forma* (linear ou não), *direção* (positiva/negativa) e *força* (fraca/forte) da nuvem de pontos.
- **Coeficiente r ∈ [−1, 1]:** o sinal dá a direção · perto de ±1 = associação forte · perto de 0 = fraca. Só mede associação *linear*.
- **Reta de regressão** (mínimos quadrados): $\hat{y} = a \cdot x + b$, calculada com tecnologia — serve para prever *dentro* do intervalo dos dados. **Correlação não implica causalidade.**

Exercícios

1. Bivariado ou univariado? a) (idade, altura) __ b) horas de estudo semanais __ c) (horas de estudo, nota final) __ — em (a), qual é a variável explanatória e qual a resposta?

2. Nuvem de pontos: (1, 2), (2, 4), (3, 5), (4, 6), (5, 8). a) forma __ b) direção __ c) força __ d) que valor de r esperarias (sem calcular)? __

3. Classifica cada r: −0,92 → __ · 0,15 → __ · −0,45 → __ · 0,99 → __ — que valores de r são *impossíveis*? __

4. V/F (justifica): «Se r = 0,85 entre duas variáveis, então uma *causa* a outra.» __ — apresenta um exemplo em que pode existir correlação sem causalidade.

© 2026 Estudantices

5. Reta de regressão $\hat{y} = 2{,}1x + 3{,}4$. a) declive __ e o que ele significa? (1 frase) b) ordenada na origem __ c) previsão para x = 10 __ d) previsão para x = 100 __ — é fiável? Porquê? (1 frase)

6. Método dos mínimos quadrados: a) que reta é que ele escolhe? (1 frase) b) porque é chamado «dos mínimos quadrados»? c) que ferramentas tecnológicas podes usar para a obter? (duas)

7. Tabela: x = 2, 4, 6, 8 · y = 5, 9, 13, 17. a) a relação é linear? __ b) qual é a equação que liga x a y? __ c) valor de y para x = 5 __

8. Pontos: (1, 3), (2, 5), (3, 7), (4, 9) — todos alinhados. a) equação da reta __ b) valor de r __ c) se afastarmos um ponto da reta, o que acontece a |r|? (1 frase)

9. Um estudo dá r = 0,02 entre duas variáveis. a) pode existir mesmo relação entre elas? __ b) descreve, em 1 frase, uma situação em que a associação existe mas não é linear.

© 2026 Estudantices

10. Algoritmo (pensamento computacional):

```
def reta(x):
    return 2.1 * x + 3.4
print(round(reta(10), 1))
```

a) o que é impresso __ b) o que faz a linha `def reta(x):`? (1 frase) __ c) se o declive passar a 2.5, o que é impresso? __

11. Correlação ≠ causalidade: numa cidade, durante o verão, à medida que aumentam as vendas de gelado aumentam também os afogamentos na praia. a) que tipo de relação descreve este exemplo? __ b) porque é que não se pode concluir que o gelado causa afogamentos? (1 frase) __ c) que variável em causa poderia explicar ambos? __

12. Produção: vais apresentar a relação entre horas de estudo e nota final numa turma. Em 3 frases, indica: o gráfico a usar, o valor esperado de r (com justificação) e um cuidado a ter na interpretação.

© 2026 Estudantices

Soluções — Dados bivariados

1. a) **bivariado** b) **univariado** c) **bivariado** — em (a): explanatória = **idade**, resposta = **altura** (estuda-se como a altura varia com a idade).

2. a) **linear** b) **positiva** (y cresce com x) c) **forte** — os pontos estão quase alinhados d) r próximo de **+1** (≈ 0,99).

3. −0,92 → **negativa forte** · 0,15 → **praticamente nula (fraca)** · −0,45 → **negativa moderada/fraca** · 0,99 → **positiva muito forte**. Impossíveis: **r > 1 ou r < −1**.

4. F — (modelo) r mede associação, não mecanismo: pode haver uma terceira variável (ex.: tempo disponível) a explicar ambas, ou a relação ser acidental. Sem experimentação/controle não se conclui causalidade.

5. a) a = **2,1** — por cada unidade que x aumenta, y aumenta em média 2,1 b) **3,4** c) 2,1 × 10 + 3,4 = **24,4** d) **244,4** — **não**: está fora do intervalo dos dados (extrapolação); a reta só é fiável dentro do intervalo observado.

6. a) (modelo) a reta que está «mais perto» de todos os pontos ao mesmo tempo b) porque minimiza a **soma dos quadrados** das distâncias verticais entre os pontos e a reta — quadrar torna positivos os desvios negativos c) (modelo) folha de cálculo (funções de regressão linear), calculadora gráfica ou AGD.

7. a) **sim** — y aumenta 4 sempre que x aumenta 2 (declive 2) b) **y = 2x + 1** (2 × 2 + 1 = 5 ✓) c) 2 × 5 + 1 = **11**.

8. a) **y = 2x + 1** (3 = 2 × 1 + 1 ✓) b) **r = 1** — associação linear perfeita e positiva c) |r| **abaixa de 1** — os pontos afastam-se da linha e a associação deixa de ser perfeita.

9. a) **sim** — (modelo) r ≈ 0 só diz que não há associação *linear* b) uma relação em forma de U (a variável cresce até certo ponto e depois desfaz), que anula a correlação linear.

10. a) `24.4` b) (modelo) Define uma função chamada `reta` com parâmetro `x`, para reutilizar o cálculo ŷ = 2,1x + 3,4 chamando-a com qualquer valor c) `28.4`.

11. a) **correlação positiva** b) (modelo) Os dois fenómenos são provocados pela mesma causa — a estação do ano — sem que um provoque o outro c) a **estação do verão** (maior calor e mais tempo na praia).

12. (modelo) Usaria um diagrama de dispersão com horas de estudo no eixo x e nota final no y. Esperaria r positivo e forte, porque mais horas de estudo tendem a acompanhar notas mais altas. Cuidado: a associação não prova que estudar *causa* a nota (podem intervir fatores como apoio em casa) e as previsões só valem dentro do intervalo dos dados observados.

© 2026 Estudantices