

Oplossingenblad — Proefexamen Statistiek

voor de ouder: volledige uitwerking + socratische hulpvragen + veelgemaakte fouten

Hoe gebruik je dit oplossingenblad?

- **Tijdens het examen:** laat je dochter eerst zelfstandig werken (50 min).
- **Stokt ze?** Gebruik de **socratische hulpvragen** per vraag, van algemeen naar specifiek — zo verklar je het antwoord *niet*.
- **Bij correctie:** tussenstappen (formule, rangnummer, invullen) tellen volwaardig mee.
- **Eenheid of eindantwoord vergeten?** Trek licht af (typisch $\frac{1}{2}$ punt).
- **Veelgemaakte fouten** staan onderaan elke oplossing — gebruik ze als checklist.

Score-overzicht

Vraag	Topic	Punten	Behaald
V1	Begrippen	3	___ / 3
V2	Frequentietabel	5	___ / 5
V3	Centrummaten	5	___ / 5
V4	Kwartielen & spreiding	5	___ / 5
V5	Variantie & standaardafw.	4	___ / 4
V6	Boxplot lezen	4	___ / 4
V7	Boxplot tekenen	4	___ / 4
V8	Verdeling & misleiding	3	___ / 3
V9	Spreidingsdiagram	4	___ / 4
V10	Correlatiecoëfficiënt	3	___ / 3
	TOTAAL	40	___ / 40

Vraag 1. [Begrippen] (3 ptn)

Na elke rit van de Ronde van Frankrijk worden 20 renners (verspreid over de ploegen) gecontroleerd op doping.

- a) Geef de populatie en de steekproef.
- b) Wat is de onderzochte variabele?
- c) Is 'het aantal gewonnen ritten van een renner' een kwantitatieve of kwalitatieve variabele? Discreet/continu of nominaal/ordinaal? Verklaar.

Volledige uitwerking

- a) **Populatie:** alle renners van de Ronde van Frankrijk. **Steekproef:** de 20 gecontroleerde renners.
- b) **Variabele:** wel/niet positief op de dopingcontrole (de testuitslag).
- c) Het aantal gewonnen ritten is een **kwantitatieve, discrete** variabele: het is een telling (gehele getallen) waarmee je kunt rekenen.

Socratische hulpvragen — gebruik in volgorde, stop zodra ze weer verder kan:

- 1. Wat is het verschil tussen de hele groep en het deel dat je echt onderzoekt?
- 2. Welk kenmerk meet/noteer je bij elke renner?
- 3. Kun je met 'aantal ritten' rekenen? Is het een telling of een meting?

Veelgemaakte fouten (instinkers)

- Populatie en steekproef omwisselen.
- Bij c) 'discreet' en 'continu' verwarren (telling = discreet).

Vraag 2. [Frequentietabel] (5 ptn)

Een onderzoek levert onderstaande (onvolledige) frequentietabel op. Vul de kolommen **caf_i**, **rf_i** (op 3 decimalen) en **rf_i%** aan. (n = 75)

x_i	af_i	rf_i	$rf_i \%$	caf_i
10	7			
20	12			
30	13			
40	20			
50	7			
60	16			
n	75	1	100 %	

a) Hoeveel waarnemingen zijn kleiner dan of gelijk aan 30?

b) Welk percentage van de waarnemingen is gelijk aan 40?

Volledige uitwerking

caf: 7 ; 19 ; 32 ; 52 ; 59 ; 75.

rf: 0,093 ; 0,160 ; 0,173 ; 0,267 ; 0,093 ; 0,213 (elke af / 75).

rf%: 9,3% ; 16,0% ; 17,3% ; 26,7% ; 9,3% ; 21,3%.

a) caf bij 30 = **32** waarnemingen zijn ≤ 30 .

b) rf% van 40 = $20/75 \cdot 100 = \mathbf{26,7\%}$.

Socratische hulpvragen — gebruik in volgorde, stop zodra ze weer verder kan:

1. Hoe bereken je een relatieve frequentie uit een absolute frequentie?
2. Wat betekent 'cumulatief'? Tel je op van boven naar onder?
3. a) In welke kolom lees je 'hoeveel zijn er ≤ 30 ' rechtstreeks af?

Veelgemaakte fouten (instinkers)

- rf delen door het aantal rijen (6) i.p.v. door n = 75.
- caf fout optellen (vergeet niet cumulatief op te tellen).
- rf% vergeten (rf $\times$ 100).

Vraag 3. [Centrumsmaten] (5 ptn)

De resultaten van een toets (op 10) in een klas staan in de tabel. Bereken het **gemiddelde**, de **modus** en de **mediaan**.

resultaat x_i	4	5	6	7	8	9
af_i	2	3	8	5	4	1

Volledige uitwerking

$n = 2+3+8+5+4+1 = 23$.

Gemiddelde: $\bar{x} = (2 \cdot 4 + 3 \cdot 5 + 8 \cdot 6 + 5 \cdot 7 + 4 \cdot 8 + 1 \cdot 9) / 23$
 $= (8+15+48+35+32+9) / 23 = 147/23 \approx 6,39$.

Modus: de waarde met de grootste af = **6** (komt 8 keer voor).

Mediaan: rangnummer = $(23+1)/2 = 12$. caf: $4 \rightarrow 2, 5 \rightarrow 5, 6 \rightarrow 13$. Plaats 12 valt bij waarde 6 $\rightarrow$ **Me = 6**.

Socratische hulpvragen — gebruik in volgorde, stop zodra ze weer verder kan:

1. Hoeveel leerlingen zijn er in totaal (n)? Tel de frequenties op.
2. Bij het gemiddelde: waardoor moet je delen — door 6 of door n?
3. Mediaan: bereken eerst het rangnummer. Gebruik dan de caf om de plaats te vinden.

Veelgemaakte fouten (instinkers)

- Gemiddelde delen door 6 (aantal verschillende waarden) i.p.v. door 23.
- Modus = 8 antwoorden (dat is de frequentie, niet de waarde).
- Mediaan zonder rangnummer/caf gokken.

Vraag 4. [Kwartielen & spreiding] (5 ptn)

Een reeks meetwaarden (al gesorteerd): **3 5 6 7 8 10 12 13 15 16 18**. Bepaal Q1, Q2 (= Me) en Q3, de variatiebreedte R en de interkwartielafstand IKA.

Volledige uitwerking

$n = 11$ (oneven). rangnummer Me = $(11+1)/2 = 6 \rightarrow$ **Me = Q2 = 10** (6de waarde).

Q1 = mediaan van de eerste helft (incl. Me): 3, 5, 6, 7, 8, 10. rangnummer = $(6+1)/2 = 3,5 \rightarrow Q1 = (6+7)/2 = 6,5$.

Q3 = mediaan van de tweede helft (incl. Me): 10, 12, 13, 15, 16, 18. $Q3 = (13+15)/2 = 14$.

R = max – min = $18 - 3 = 15$.

IKA = $Q3 - Q1 = 14 - 6,5 = 7,5$.

Socratische hulpvragen — gebruik in volgorde, stop zodra ze weer verder kan:

1. Hoeveel waarnemingen zijn er? Even of oneven? Wat betekent dat voor de mediaan?
2. Bij oneven n: reken je Q1 en Q3 inclusief of exclusief de mediaan?
3. Hoe bereken je R en IKA als je min, max, Q1 en Q3 hebt?

Veelgemaakte fouten (instinkers)

- De rij niet eerst sorteren (hier al gesorteerd — maar check altijd!).
- Bij oneven n de mediaan vergeten mee te nemen in elke helft.
- R en IKA verwisselen.

Vraag 5. [Variantie & standaardafw.] (4 ptn)

Op 5 dagen werd de middagtemperatuur gemeten: 16° , 18° , 20° , 22° , 24° . Bereken de variantie σ^2 en de standaardafwijking σ (op 2 decimalen).

Volledige uitwerking

Gemiddelde: $\bar{x} = (16+18+20+22+24)/5 = 100/5 = 20^\circ$.

Afwijkingen: $-4, -2, 0, 2, 4$.

Kwadraten: $16, 4, 0, 4, 16$.

Variantie: $\sigma^2 = (16+4+0+4+16)/5 = 40/5 = 8$.

Standaardafwijking: $\sigma = \sqrt{8} \approx 2,83^\circ$.

Socratische hulpvragen — gebruik in volgorde, stop zodra ze weer verder kan:

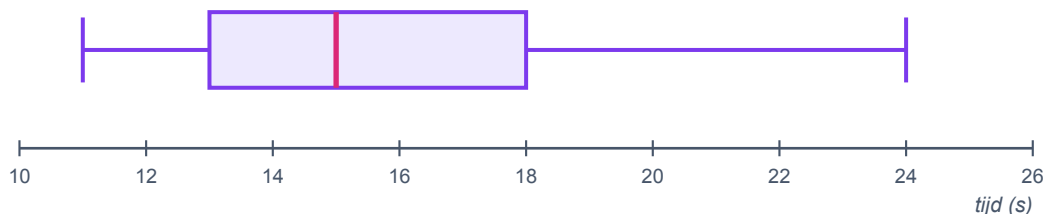
1. Wat is de eerste stap voor de variantie — welk getal heb je altijd eerst nodig?
2. Wat doe je met elke afwijking voordat je optelt? (denk aan het woord 'kwadratische')
3. Hoe ga je van variantie naar standaardafwijking?

Veelgemaakte fouten (instinkers)

- Vergeten te kwadrateren (gewoon de afwijkingen optellen geeft 0!).
- De wortel vergeten bij de standaardafwijking.
- Delen door 4 i.p.v. door $n = 5$.

Vraag 6. [Boxplot lezen] (4 ptn)

Hieronder zie je de boxplot van de 100m-sprinttijden van klas 3A (20 leerlingen).



- a) Bepaal de variatiebreedte R en de interkwartielafstand IKA .
- b) Hoeveel procent van de leerlingen loopt trager dan 18 s (= $Q3$)?
- c) Hoeveel leerlingen lopen sneller dan 15 s (= de mediaan)?

Volledige uitwerking

Afgelezen: $\min = 11$; $Q1 = 13$; $Me = 15$; $Q3 = 18$; $\max = 24$.

a) $R = \max - \min = 24 - 11 = 13$ s. $IKA = Q3 - Q1 = 18 - 13 = 5$ s.

b) Boven $Q3$ ligt **25%** van de leerlingen (trager dan 18 s).

c) Onder de mediaan ligt 50% $\rightarrow$ 50% van 20 = **10 leerlingen** lopen sneller dan 15 s.

Socratische hulpvragen — gebruik in volgorde, stop zodra ze weer verder kan:

1. Welke 5 kengetallen lees je rechtstreeks af op een boxplot?
2. Elk stuk van de boxplot bevat hoeveel procent van de gegevens?
3. c) 50% van de leerlingen — hoeveel leerlingen is dat als er 20 zijn?

Veelgemaakte fouten (instinkers)

- Aflezen van verkeerde waarden (let op de schaal).
- Vergeten dat elk stuk 25% is.
- Bij c) % niet omzetten naar een aantal leerlingen.

Vraag 7. [Boxplot tekenen] (4 ptn)

Tien leerlingen behaalden volgende punten (op 20): **8 10 11 12 13 13 14 15 17 19**. Bepaal de 5 kengetallen (min, Q1, Me, Q3, max) en schets de boxplot.

Volledige uitwerking

n = 10 (even). Min = 8 ; Max = 19.

Me: rangnummer = $(10+1)/2 = 5,5 \rightarrow \text{Me} = (13+13)/2 = 13$.

Q1: eerste helft (plaats 1–5): 8, 10, 11, 12, 13 $\rightarrow$ mediaan = 3de waarde = **11**.

Q3: tweede helft (plaats 6–10): 13, 14, 15, 17, 19 $\rightarrow$ mediaan = **15**.

Boxplot: rechthoek van 11 tot 15, streepje op 13, snorharen tot 8 en 19.

Socratische hulpvragen — gebruik in volgorde, stop zodra ze weer verder kan:

1. Is n even of oneven? Hoe bepaal je dan de mediaan?
2. Bij even n: hoe verdeel je de rij in twee helften voor Q1 en Q3?
3. Welke onderdelen heeft een boxplot, en wat stelt de rechthoek voor?

Veelgemaakte fouten (instinkers)

- Bij even n de twee helften fout afbakenen.
- De mediaan als verticaal streepje vergeten in de box.
- Snorharen niet tot het echte min/max trekken.

Vraag 8. [Verdeling & misleiding] (3 ptn)

Beantwoord beide deelvragen kort maar verklarend.

a) In een rechtsscheve verdeling: ligt het gemiddelde links of rechts van de mediaan? Verklaar.

b) Een nieuwsbericht toont een staafdiagram waarin 50% en 43% sterk lijken te verschillen. Hoe kan een grafiek dit verschil misleidend groot maken? Noem één manier.

Volledige uitwerking

a) Het gemiddelde ligt aan de kant van de staart. Bij een rechtsscheve verdeling ligt de staart rechts, dus het gemiddelde ligt **rechts** van de mediaan (de grote waarden in de staart trekken het gemiddelde omhoog).

b) Door de **verticale as niet bij 0 te laten beginnen** (bv. vanaf 38%) lijkt het verschil tussen de staven veel groter dan het is. (Andere voorbeelden: punten weglaten op de tijdas, of een niet-chronologische volgorde.)

Socratische hulpvragen — gebruik in volgorde, stop zodra ze weer verder kan:

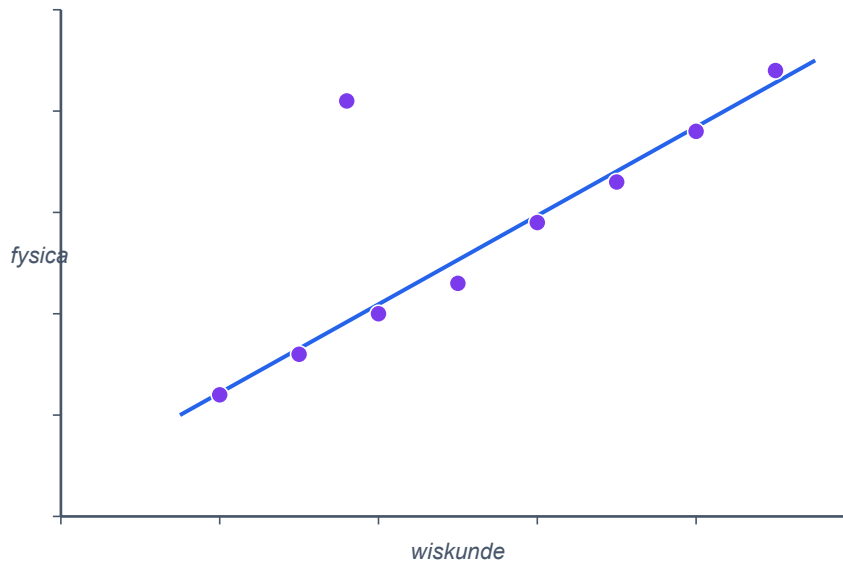
1. a) Aan welke kant van de verdeling ligt het gemiddelde altijd — de top of de staart?
2. b) Waar moet de verticale as normaal beginnen om eerlijk te zijn?
3. b) Welke andere trucs ken je nog (denk aan de tijdas)?

Veelgemaakte fouten (instinkers)

- Bij a) 'links' en 'rechts' verwisselen.
- Bij b) geen concrete manier noemen (gewoon 'de grafiek liegt' is te vaag).

Vraag 9. [Spreidingsdiagram] (4 ptn)

Het spreidingsdiagram hieronder toont voor 9 leerlingen het resultaat voor wiskunde (x) en voor fysica (y).



- Welk soort verband zie je? Beschrijf de richting.
- Is het verband sterk of zwak? Verklaar aan de hand van de puntenwolk.
- Welk punt zou je als een mogelijke uitschieter aanduiden?

Volledige uitwerking

- Een **positief** verband: grote x-waarden gaan samen met grote y-waarden. De trendlijn **stijgt**.
- De meeste punten liggen dicht bij de trendlijn → **sterk** positief lineair verband (op één punt na).
- Het punt linksboven (lage wiskunde, hoge fysica) ligt ver van de trendlijn en is de **uitschieter**.

Socratische hulpvragen — gebruik in volgorde, stop zodra ze weer verder kan:

- Gaan grote x-waarden samen met grote of met kleine y-waarden?
- Hoe dicht bij de rechte liggen, hoe ... het verband?
- Welk punt ligt duidelijk los van de algemene trend?

Veelgemaakte fouten (instinkers)

- Positief en negatief verband verwisselen (kijk naar de richting van de wolk).
- Sterkte verwarren met richting.
- De uitschieter op de trendlijn zoeken i.p.v. er ver vanaf.

Vraag 10. [Correlatiecoëfficiënt] (3 ptn)

Voor een dataset berekende een leerling met Excel de correlatiecoëfficiënt $r = 0,86$.

- a) Welk soort lineair verband leid je hieruit af?
- b) Een tweede leerling gebruikt dezelfde gegevens, maar zet x en y omgekeerd op de assen. Welke waarde krijgt zij voor r? Verklaar.
- c) Na het toevoegen van één extra (afwijkend) meetpunt daalt r naar 0,42. Wat besluit je hieruit over r?

Volledige uitwerking

- a) $r = 0,86$ ligt dicht bij +1 → een **sterk positief** lineair verband.
- b) r blijft **exact gelijk** (0,86). Het omwisselen van x en y heeft geen invloed op de correlatiecoëfficiënt.
- c) Één afwijkend punt doet r flink dalen (van sterk naar zwak). Besluit: **de correlatiecoëfficiënt is gevoelig voor uitschieters**.

Socratische hulpvragen — gebruik in volgorde, stop zodra ze weer verder kan:

1. Waar ligt r tussen, en wat betekent een waarde dicht bij 1?
2. b) Verandert er iets aan de samenhang als je enkel de assen verwisselt?
3. c) Wat gebeurt er met r door één punt dat ver van de rest ligt?

Veelgemaakte fouten (instinkers)

- Bij b) denken dat r van teken verandert of $1/r$ wordt.
- Bij a) 'sterk' en 'zwak' verwisselen (0,86 is sterk).