

Voor elk onderdeel is aangegeven hoeveel punten kunnen worden behaald. Antwoorden moeten *altijd* zijn voorzien van een berekening, toelichting of argumentatie.

MAX: 50 punten

OPGAVE 1 (8 PNT)

Op het Nicolaaslyceum doen alle leerlingen uit 4 havo mee aan een parkloop. Van de tijd die de leerlingen nodig hebben om de parkloop te voltooien is de verdelingskromme hiernaast gemaakt.

2p

a De verdelingskromme heeft twee toppen.

Geef hiervoor een mogelijke verklaring.

2p

b Hoe ligt bij deze verdeling de mediaan ten opzichte van de modus?

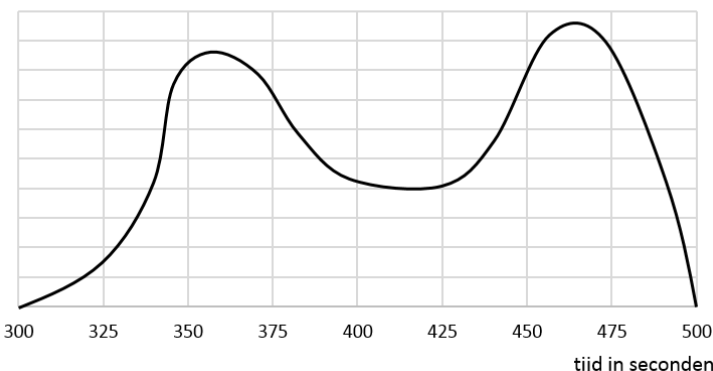
2p

c Schets de bijbehorende relatieve cumulatieve verdelingskromme.

2p

d Hieronder zie je de boxplots A, B en C.

aantal leerlingen



Welke van deze boxplots hoort bij de verdelingskromme? Licht toe.

OPGAVE 2 BLOEDDRUK (5 PNT)

Bij een onderzoek naar de gevolgen van het pilgebruik wordt onder andere de bloeddruk van vrouwen gemeten. Hieronder zie je gegevens (uitgedrukt in procenten) over de bloeddruk van vrouwen (in de leeftijdscategorie 25-34 jaar) die al meer dan een jaar de pil slikken (*users*) en vrouwen die dat niet doen (*non-users*).

Bloeddruk (in mm)	Non-user	User
Onder 90	1	
90-95	1	
95-100	5	4
100-105	11	5
105-110	11	10
110-115	17	15
115-120	18	17
120-125	11	15
125-130	9	12
130-135	7	10
135-140	4	5
140-145	2	4
145-150	2	2
150-155	1	1
Totaal (in procenten)	100	100

Onderzoek hoe groot het verschil in bloeddruk is tussen de users en de non-users.
Kies hiervoor een geschikte maat.

OPGAVE 3 (10 pnt)

- 3p **a** Van 45 pasgeboren Aziatische olifanten is het gemiddelde gewicht 96 kg en de standaardafwijking 12 kg.
Bereken het 95%-betrouwbaarheidsinterval voor het gewicht van pasgeboren Aziatische olifanten. Rond in het antwoord af op één decimaal.
- 3p **b** Van 32 Afrikaanse olifanten is het geboortegewicht in kg vastgesteld. Hieruit volgt dat het 95%-betrouwbaarheidsinterval voor het geboortegewicht $[99,7; 110,3]$ is.
Bereken de steekproefstandaardafwijking in gehele kg nauwkeurig.
- 4p **c** Voordat Leo een spreekbeurt over de Afrikaanse olifant houdt, laat hij alle leerlingen van zijn klas opschrijven hoeveel kg volgens hen het geboortegewicht is van een Afrikaanse olifant. Het blijkt dat de standaardafwijking van de gegeven antwoorden maar liefst 60 kg is.
Zo krijgt Leo het 95%-betrouwbaarheidsinterval $[56, 104]$ voor het geboortegewicht in kg.
Hoeveel leerlingen zitten er bij Leo in de klas?

OPGAVE 4 VLEES (11 PNT)

Er is een onderzoek gedaan of de mate waarin men vlees eet afhangt van de leeftijd. Daartoe heeft men onderscheid gemaakt tussen de leeftijden van 18 tot 30 jaar en van 30 jaar en ouder. Een van de vragen ging over de vleesconsumptie bij de warme maaltijd. Zie de tabel. Je ziet dat er 15 ondervraagden in de leeftijd 18-30 jaar waren die nooit vlees eten bij de warme maaltijd.

VLEESCONSUMPTIE BIJ DE WARME MAALTIJD

	18-30 jaar	30 jaar en ouder
nooit	15	16
minder dan eens per week	12	5
1 of 2 keer per week	25	28
3 of 4 keer per week	30	45
meer dan 4 keer per week	8	46

- 6p **a** Bepaal met behulp van een vuistregel op het formuleblad of het verschil tussen de leeftijden 18-30 jaar en 30 jaar en ouder wat betreft de vleesconsumptie bij de warme maaltijd groot, middelmatig of gering is.
- 5p **b** Onderzoek of je dezelfde conclusie trekt als je 'nooit' en 'minder dan eens per week' samenvoegt en ook '3 of 4 keer per week' en 'meer dan 4 keer per week', dus als je de categorieën 'nooit of vrijwel nooit', '1 of 2 keer per week' en 'vaker dan 2 keer per week' hebt.

OPGAVE 5, Schoolkinderen (1 + 2 + 2 + 3pnt)

Kinderen in een mammoetklas met meer dan 30 leerlingen, omdat de school moet bezuinigen? Niet alleen basisscholen zijn er ongelukkig mee. Ouders vrezen dat hun kroost te weinig aandacht krijgt. "We draaien de klok terug naar begin jaren negentig", stelt pedagoog Bas Levering. Tussen 1997 en 2002 daalde het gemiddeld aantal kinderen in een onderbouwklas van 23,7 naar 20,9 en kwamen er onderwijsassistenten. Uit een evaluatierapport blijkt dat kleinere groepen en meer handen in de klas de kwaliteit van het onderwijs in de onderbouw van het basisonderwijs hebben verbeterd.

a. Wat is hier de populatie?

Veronderstel dat de gegevens over 1997 gebaseerd zijn op een aselechte steekproef van 80 klassen en dat de standaardafwijking gelijk is aan 3.

b. Bereken het 95%-betrouwbaarheidsinterval voor de gemiddelde omvang van een onderbouwklas in 1997.

Veronderstel dat de gegevens over 2002 gebaseerd zijn op een aselechte steekproef van 120 klassen en dat de standaardafwijking gelijk is aan 4.

c. Bereken het 95%-betrouwbaarheidsinterval voor de gemiddelde omvang van een onderbouwklas in 2002. Overlapt dit 95% betrouwbaarheidsinterval met dat van 1997?

Veronderstel dat een onderzoeker met 95 procent zekerheid de gemiddelde omvang van een onderbouwklas op 1 decimaal nauwkeurig wil vaststellen.

Veronderstel verder dat de standaardafwijking gelijk is aan 4.

d. Bereken hoe groot de steekproef minimaal moet zijn om aan de gegeven eis te kunnen voldoen.

OPGAVE 6 BEVOLKING (8 PNT)

De tabel Bevolkingsprognose hieronder komt uit de CBS-publicatie Trends in Nederland 2016.

Bevolkingsprognose

	2016	2040	2060
	x 1 000		
Bevolking, 1 januari	16 974	18 108	18 175
jonger dan 20 jaar	3 815	3 916	3 825
20-39 jaar	4 166	4 248	4 395
40-64 jaar	5 909	5 142	5 183
65-79 jaar	2 336	3 166	2 762
80 jaar of ouder	749	1 637	2 010
Bevolking, jaarmutatie	113	11	8
waaronder			
levendgeborenen	173	186	193
overledenen	148	192	200
geboorte-overschot	25	-6	-7
immigratie	239	185	187
emigratie (incl. administratieve correcties)	151	168	171
migratiesaldo (incl. administratieve correcties)	88	17	15
	kinderen per vrouw		
Totaal vruchtbaarheidscijfer	1,66	1,75	1,75
	jaren		
Levensverwachting bij de geboorte			
mannen	79,9	84,0	86,8
vrouwen	83,3	87,5	90,3
	%		
Bevolking, 1 januari			
jonger dan 20 jaar	22,5	21,6	21,0
20-64 jaar	59,3	51,9	52,7
65 jaar of ouder	18,2	26,5	26,3

2p **a** Licht toe hoe je de jaarmutatie 113 die bij 2016 staat uit de andere gegevens kunt berekenen.

2p **b** Licht toe hoe je het percentage 22,5 dat bij 2016 staat uit de andere gegevens kunt berekenen.

De groene druk is het aantal 0-19-jarigen t.o.v. het aantal 20-64-jarigen.

De grijze druk is het aantal personen van 65 jaar en ouder t.o.v. het aantal 20-64-jarigen.

De demografische druk is de som van de groene druk en de grijze druk. De groene druk, de grijze druk en de demografische druk worden uitgedrukt in procenten.

4p **c** Bereken de groene druk, de grijze druk en de demografische druk in 2016 en in 2040 volgens deze prognose.

FORMULEBLAD

Betrouwbaarheidsintervallen

Het 95%-betrouwbaarheidsinterval voor de populatieproportie is $p \pm 2 \cdot \sqrt{\frac{p(1-p)}{n}}$, met p de steekproefproportie en n de steekproefomvang.

Het 95%-betrouwbaarheidsinterval voor het populatiegemiddelde is $\bar{X} \pm 2 \cdot \frac{S}{\sqrt{n}}$, met $\bar{X}$ het steekproefgemiddelde, n de steekproefomvang en S de steekproefstandaardafwijking.

Vuistregels bij de grootte van het verschil van twee groepen

2 × 2-kruistabel $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, met $\phi = \frac{ad - bc}{\sqrt{(a+b)(a+c)(b+d)(c+d)}}$

- als $\phi < -0,4$ of $\phi > 0,4$, dan zeggen we 'het verschil is groot',
- als $-0,4 \leq \phi < -0,2$ of $0,2 < \phi \leq 0,4$, dan zeggen we 'het verschil is middelmatig',
- als $-0,2 \leq \phi \leq 0,2$, dan zeggen we 'het verschil is gering'.

Maximale verschil in cumulatief percentage (max V_{cp}) (met steekproefomvang $n > 100$)

- als $\max V_{cp} > 40$, dan zeggen we 'het verschil is groot',
- als $20 < \max V_{cp} \leq 40$, dan zeggen we 'het verschil is middelmatig',
- als $\max V_{cp} \leq 20$, dan zeggen we 'het verschil is gering'.

Effectgrootte $E = \frac{\bar{X}_1 - \bar{X}_2}{\frac{1}{2}(S_1 + S_2)}$ met $\bar{X}_1$ en $\bar{X}_2$ de steekproefgemiddelden ($\bar{X}_1 \geq \bar{X}_2$), S_1

en S_2 de steekproefstandaardafwijkingen

- als $E > 0,8$, dan zeggen we 'het verschil is groot',
- als $0,4 < E \leq 0,8$, dan zeggen we 'het verschil is middelmatig',
- als $E \leq 0,4$, dan zeggen we 'het verschil is gering'.

Twee boxplots vergelijken

- als de boxen¹ elkaar niet overlappen, dan zeggen we 'het verschil is groot'
- als de boxen elkaar wel overlappen en een mediaan van een boxplot buiten de box van de andere boxplot ligt, dan zeggen we 'het verschil is middelmatig'
- in alle andere gevallen zeggen we 'het verschil is gering'.

¹ De 'box' is het interval vanaf het eerste kwartiel tot en met het derde kwartiel.

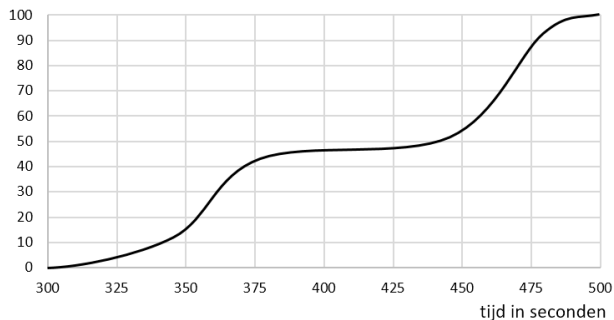
CORRECTIEBLAD

OPGAVE 1

totaal 8p

Basisvraag

- a de jongens lopen de parkloop gemiddeld sneller dan de meisjes, daardoor ontstaan twee toppen 2p
- b de mediaan ligt links van de modus 2p
- c relatieve cumulatieve frequentie 2p



- d boxplot C, want uit de verdelingskromme volgt dat Q1 ruim 350 is en de mediaan ruim 400

OPGAVE 2

Onderstaande opgave komt uit de SLO4 bundel.

Nodig: GRM, Lijsten invullen -> CALC-> 1 VAR STATS, en lees af gemiddelde en stddev.

Dat doe je voor de beide lijsten.

Dan heb je twee gemiddelden en twee stddev. Dus Effectgrootte.

Opgave 31

Bloeddruk is een kwantitatieve variabele, dus de effectgrootte gebruiken of de boxplots vergelijken. Het meest geschikt is de effectgrootte. Daarvoor eerst het gemiddelde en de standaardafwijking van beide groepen uitrekenen.

Neem aan: onder de 90 -> de waarde is 87,5

Dan kom je op:

	Non-user	User
Gemiddelde	117,1	120,6
Standaardafwijking	12,77	12,01

$E = 0,28$, dus het verschil is gering.

Merk op dat hier NIET sprake is van een proportie (getallen als fractie en dus tussen 0 en 1). Men spreekt hier over dingen, mensen, voorwerpen, etc.. Dus gebruik je die andere formule:

$\bar{X} \pm 2 \cdot \frac{S}{\sqrt{n}}$ Soms moet je die terugrekenen

Opgave 3**totaal 10p**

a $96 - 2 \cdot \frac{12}{\sqrt{45}} \approx 92,4$ en $96 + 2 \cdot \frac{12}{\sqrt{45}} \approx 99,6$ 2p

het 95%-betrouwbaarheidsinterval is $[92,4; 99,6]$ 1p

b $4 \cdot \frac{S}{\sqrt{32}} = 110,3 - 99,7$ ofwel $4 \cdot \frac{S}{\sqrt{32}} = 10,6$ 2p

oplossen van de vergelijking $4 \cdot \frac{S}{\sqrt{32}} = 10,6$ geeft $S \approx 15$ dus de steekproefstandaardafwijking is 15 kg 1p

c $4 \cdot \frac{60}{\sqrt{n}} = 104 - 56$ ofwel $4 \cdot \frac{60}{\sqrt{n}} = 48$ 2p

oplossen (met de GR) van de vergelijking $4 \cdot \frac{60}{\sqrt{n}} = 48$ geeft $n = 25$,

dus er zitten 25 leerlingen bij Leo in de klas 2p

Opgave 4**totaal 8p**

a het maken van een tabel met cum. perc. en V_{cp} 4p

Die kolommen ontbreken bij veel mensen.

	18-30 jaar		30 jaar en ouder		V_{cp}
	cum.	cum. perc.	cum.	cum. perc.	
nooit	15	16,7%	16	11,4%	5,3%
minder dan eens per week	27	30%	21	15%	15%
1 of 2 keer per week	52	57,8%	49	35%	22,8%
3 of 4 keer per week	82	91,1%	94	67,1%	24%
meer dan 4 keer per week	90	100%	140	100%	0%

$\max.V_{cp} = 24\%$ 1p

omdat $20\% < \max.V_{cp} \leq 40\%$ is het verschil middelmatig 1p

b het maken van een tabel met cum. perc. en V_{cp} bij deze situatie 3p

	18-30 jaar	30 jaar en ouder	V_{cp}
	cum. perc.	cum. perc.	
nooit of vrijwel nooit	46,7%	18,3%	28,4%
1 of 2 keer per week	57,8%	35%	22,8%
meer dan 2 keer per week	100%	100%	0%

$\max.V_{cp} = 28,4\%$, dus ook nu is het verschil middelmatig
je trekt dezelfde conclusie 1p

1p
1p

Opgave 5: schoolkinderen

- a. Onderbouwklassen in het basisonderwijs in Nederland.
- b. [23,03; 24,37].
- c. [20,17; 21,63], de intervallen overlappen niet.
- d. $n > 25600$.

Vraag (a) Pas op: Het ging hier NIET om kinderen, maar om klassen die een gemiddeld aantal kinderen hebben. Klassen dus.

Verder gewone intervalletjes. Som d terugrekenen met GRM of die kwadraattruc.

Opgave 6 totaal 8p

a geboorteoverschot + migratiesaldo = 25 + 88 = 113 2p

Variantantwoorden zijn soms goedgekeurd. Je kan op basis van de tabel ook alles narekenen.

b $\frac{3815}{16974} \times 100\% = 22,47\ldots\% \approx 22,5\%$ 2p

c 2016 groene druk = $\frac{22,5}{59,3} \times 100\% \approx 37,9\%$
 grijze druk = $\frac{18,2}{59,3} \times 100\% \approx 30,7\%$
 demografische druk = 37,9% + 30,7% = 68,6% 2p

2040 groene druk = $\frac{21,6}{51,9} \times 100\% \approx 41,6\%$
 grijze druk = $\frac{26,5}{51,9} \times 100\% \approx 51,1\%$
 demografische druk = 41,6% + 51,1% = 92,7% 2p