

Bij dit examen hoort een uitwerkbijlage.

Dit examen bestaat uit 21 vragen.

Voor dit examen zijn maximaal 75 punten te behalen.

Voor elk vraagnummer staat hoeveel punten met een goed antwoord behaald kunnen worden.

Als bij een vraag een verklaring, uitleg of berekening vereist is, worden aan het antwoord meestal geen punten toegekend als deze verklaring, uitleg of berekening ontbreekt.

Geef niet meer antwoorden (redenen, voorbeelden e.d.) dan er worden gevraagd. Als er bijvoorbeeld twee redenen worden gevraagd en je geeft meer dan twee redenen, dan worden alleen de eerste twee in de beoordeling meegeteld.



Vuistregels voor de grootte van het verschil van twee groepen

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

waarin a, b, c en d absolute aantallen zijn

- 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”

Maximaal verschil in cumulatief percentage ($\max V_{cp}$)

(met voor beide groepen een 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”

noot 1 De ‘box’ is het interval vanaf het eerste kwartiel tot en met het derde kwartiel.



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.



Appels met peren vergelijken

In figuur 1 staat een bericht dat op internet te lezen was over de appel- en perenoogst in Nederland. Hierin wordt gesproken over de **oogstraming** van 2021; dit is de schatting van de oogstopbrengst (in miljoen kilo) van dat jaar.

figuur 1

Meer appels en minder peren verwacht dan vorig jaar
Geplaatst op 27 juli 2021

Dit jaar zullen in Nederland naar verwachting meer appels worden geplukt dan vorig jaar, terwijl de totale perenoogst lager wordt geraamd dan vorig jaar. De Nederlandse oogstraming komt uit op 250 miljoen kilo appels en 325 miljoen kilo peren. Dit blijkt uit informatie van GroentenFruit Huis en NFO.

Er hangen meer appels aan de bomen dan in 2020. Hierdoor wordt de totale appeloogst (in miljoen kilo) 14 procent hoger ingeschat dan in 2020. Bij peren is het beeld juist andersom waardoor de perenoogst (in miljoen kilo) 19 procent lager wordt geraamd dan vorig jaar.

Neem aan dat een appel gemiddeld 150 gram weegt en een peer gemiddeld 225 gram. Op basis van de oogstraming kun je het geschatte aantal appels en het geschatte aantal peren in 2021 berekenen.

- 3p 1 Bereken het verschil tussen deze twee aantallen. Geef je antwoord in een geheel aantal miljoenen.

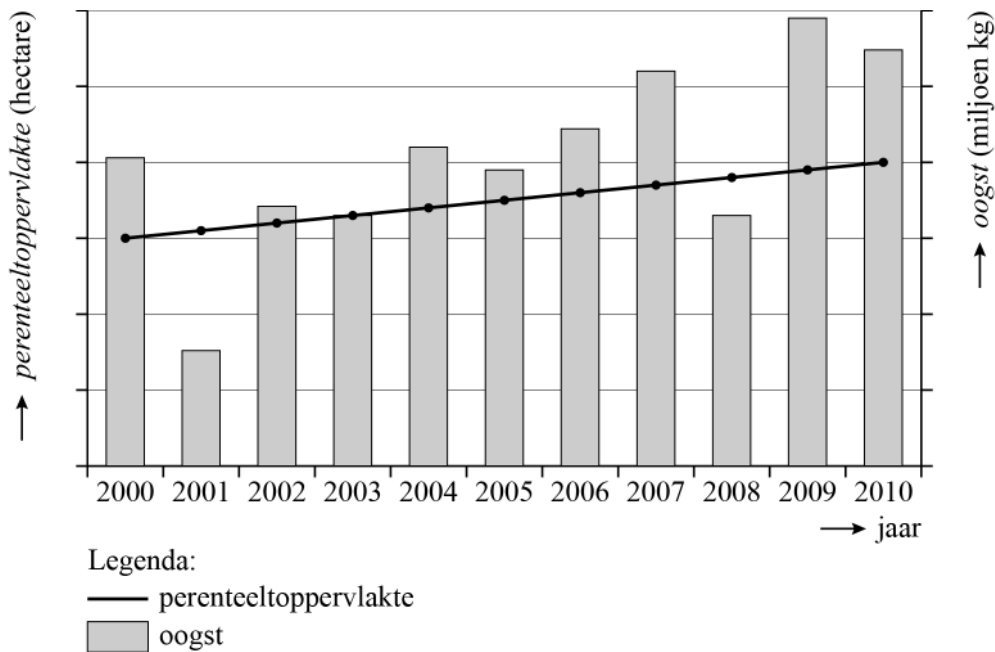
Op basis van het bericht kun je het totale gewicht berekenen dat in 2020 aan appels en peren samen is geoogst.

- 3p 2 Bereken dit totale gewicht. Geef je antwoord in een geheel aantal miljoenen kilogram.

In figuur 2 zie je voor de periode 2000-2010 het totaalgewicht aan peren dat in Nederland is geoogst. Ook zie je de oppervlakte waarop deze peren werden geteeld. Neem aan dat deze oppervlakte gedurende een kalenderjaar gelijk blijft.



figuur 2 perenteelt



- 3p 3 Beredeneer, zonder gebruik te maken van getallenvoorbeelden, met behulp van figuur 2 of de gemiddelde perenoogst in miljoen kg per hectare in 2008 groter of kleiner is dan in 2003.

De perenteeltoppervlakte neemt in de periode van 2000 tot en met 2015 lineair toe. We kunnen voor die periode de perenteeltoppervlakte beschrijven met de volgende formule:

$$T_p = 6000 + 200t$$

Hierin is T_p de perenteeltoppervlakte in hectare en t het aantal gehele jaren na 2000.

In de tabel staat de appelteeltoppervlakte in 2003 en 2015.

tabel

jaar	2003	2015
appelteeltoppervlakte T_a (in hectare)	10 296	7600

Ga ervan uit dat de appelteeltoppervlakte in de periode 2000-2015 lineair afneemt en dat deze oppervlakte gedurende een kalenderjaar steeds gelijk blijft.

- 5p 4 Bereken in welk jaar de perenteeltoppervlakte voor het eerst groter is dan de appelteeltoppervlakte.

Men verwachtte dat na 2015 de appelteeltoppervlakte exponentieel zou afnemen met 19,4% per vijf jaar.

- 4p 5 Bereken met hoeveel procent de appelteeltoppervlakte dan per jaar afneemt. Geef je antwoord in één decimaal.

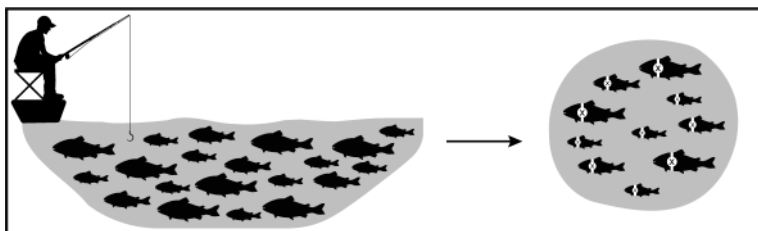


Vangen-hervangen

Het is niet eenvoudig het totaal aantal vissen in een meer te tellen. Met behulp van de statistische techniek 'vangen-hervangen' is het mogelijk een schatting te doen van het aantal vissen. Deze techniek werkt als volgt:

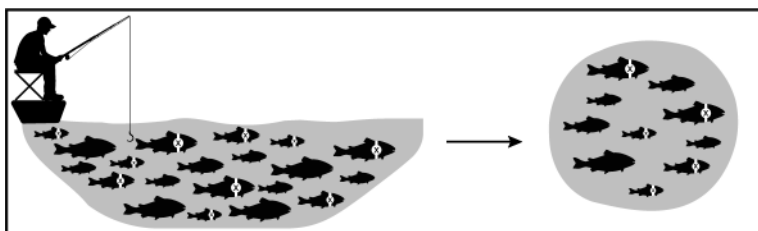
Vang een aantal vissen uit het meer. Dit is de eerste steekproef. Geef deze vissen een merkteken. Zie figuur 1.

figuur 1



Zet de gemerkte vissen vervolgens weer terug in het meer. Wacht een tijdje tot de gemerkte vissen zich over het hele meer verspreid hebben en vang nogmaals een aantal vissen uit het meer. Dit is de tweede steekproef. Zie figuur 2.

figuur 2



We kunnen nu veronderstellen dat de proportie gemerkte vissen in de tweede steekproef, gelijk is aan de proportie gemerkte vissen in het hele meer. In formulevorm:

$$\frac{R}{C} = \frac{M}{N} \quad \text{(formule 1)}$$

Hierbij is:

- R het aantal gemerkte vissen in de tweede steekproef;
- C het aantal gevangen vissen dat in de tweede steekproef;
- M het aantal gevangen, en dus gemerkte, vissen in de eerste steekproef;
- N het (geschatte) aantal vissen in het meer.



Biologen vangen in een meer bij een eerste steekproef 80 vissen, merken deze en zetten ze weer terug in het meer. Een paar dagen later wordt een tweede steekproef vissen gevangen. Hierin bevinden zich 6 gemerkte en 92 ongemerkte vissen.

- 3p 6 Bereken met behulp van formule 1 hoeveel vissen er in het meer zijn. Geef je antwoord in een geheel honderdtal.

We kunnen formule 1 herleiden tot een formule waarmee je eenvoudiger het aantal vissen N in het meer kunt berekenen:

$$N = \frac{M \cdot C}{R} \quad (\text{formule 2})$$

Biologen Adri en Fiona vangen samen een eerste steekproef, waarna zij alle 30 vissen merken die ze samen hebben gevangen. De gemerkte vissen zetten ze terug in het meer. Om de betrouwbaarheid van de schatting van het aantal vissen in het meer te vergroten, vangt elk van hen 180 vissen als tweede steekproef. Eerst vangt Adri 180 vissen en hij telt het aantal gemerkte vissen. Daarna zet hij de 180 vissen terug in het meer. Vervolgens vangt Fiona 180 vissen. In de steekproef van Fiona is er één vis meer gemerkt dan in de steekproef van Adri. Beiden berekenen met formule 2 het aantal vissen in het meer. Fiona's uitkomst blijkt 60 lager te zijn dan Adri's uitkomst.

- 4p 7 Bereken het aantal gemerkte vissen in de steekproef van Adri.

Er is een situatie mogelijk, waarbij je vissen in de eerste steekproef vangt en merkt, in een tweede steekproef weer vissen vangt, maar dat het niet lukt om met formule 1 of 2 het aantal vissen in het meer te schatten.

- 2p 8 Beschrijf deze situatie en leg uit waarom het in die situatie niet lukt om met formule 2 het aantal vissen in het meer te schatten.

Formule 1 en 2 werken in de praktijk minder goed als R klein is. Daarom is er met computersimulaties gezocht naar een formule die betere schattingen geeft. Uit die computersimulaties bleek dat bij kleine waarden van R , onderstaande formule betere schattingen geeft voor het aantal vissen in het meer dan formule 2.

$$N = \frac{(M+1) \cdot (C+1)}{R+1} - 1 \quad (\text{formule 3})$$

Het verschil in de schattingen van het aantal vissen volgens de formules 2 en 3 wordt groter naarmate er minder gemerkte vissen in een tweede steekproef worden gevangen.

We bekijken de situatie waarbij in de eerste steekproef 120 vissen zijn gevangen en gemerkt en in een tweede steekproef 150 vissen zijn gevangen.

- 4p 9 Bereken bij welke aantallen gemerkte vissen in een tweede steekproef de schatting volgens formule 2 meer dan 100 hoger is dan de schatting volgens formule 3.



In 2018 verschenen er berichten in de media over de (toilet)hygiëne op Nederlandse basisscholen en middelbare scholen. De berichten waren gebaseerd op drie onderzoeken onder respectievelijk leraren die werkzaam waren binnen het basis- of voortgezet onderwijs, leerlingen tussen de 5 en 16 jaar oud en ouders met een kind tussen de 5 en 16 jaar oud. In totaal namen 630 leraren, 452 leerlingen en 626 ouders deel aan het onderzoek. We nemen aan dat elk van de drie steekproeven aselekt en representatief was.

In de steekproef van de leerlingen gaven 194 van de 452 leerlingen aan regelmatig moeite te hebben met opletten tijdens de les, omdat zij op sommige momenten naar het toilet moeten, maar niet gaan.

- 4p **10** Bereken het 95%-betrouwbaarheidsinterval voor het percentage Nederlandse leerlingen tussen de 5 en 16 jaar oud, waarvoor het bovenstaande geldt. Geef de percentages in je antwoord in gehelen.

Van de 452 leerlingen in de steekproef zaten er 302 op de basisschool en 150 op de middelbare school. Precies de helft van de basisschoolleerlingen gaat op school niet naar het toilet. Bij de middelbare scholieren is dit 60%.

In deze steekproef is er dus een verschil tussen basisschoolleerlingen en middelbare scholieren in het deel dat op school niet naar het toilet gaat.

- 4p **11** Onderzoek met behulp van het formuleblad of dat verschil groot, middelmatig of gering is.

Veel ouders vermoeden dat een betere hygiëne op een school samengaat met een lager percentage ziekmeldingen onder de leerlingen van de betreffende school.

Neem aan dat de hygiëne van een school uitgedrukt wordt in een score tussen 0 en 100 waarbij geldt: hoe hoger de score, hoe beter de hygiëne van de betreffende school.

- 2p **12** Geef op de uitwerkbijlage een schets van een mogelijke puntenwolk die het vermoeden ondersteunt dat een betere hygiëne samengaat met een lager percentage ziekmeldingen.



De 630 leraren in de steekproef hebben gereageerd op de stelling of hun school over voldoende schoonmaakbudget beschikt om de hygiëne op school te kunnen waarborgen. Het resultaat staat in de tabel.

tabel

stelling: het schoonmaakbudget is voldoende						
reactie	weet niet	helemaal mee oneens	mee oneens	neutraal	mee eens	helemaal mee eens
percentage	27	11	27	12	19	4

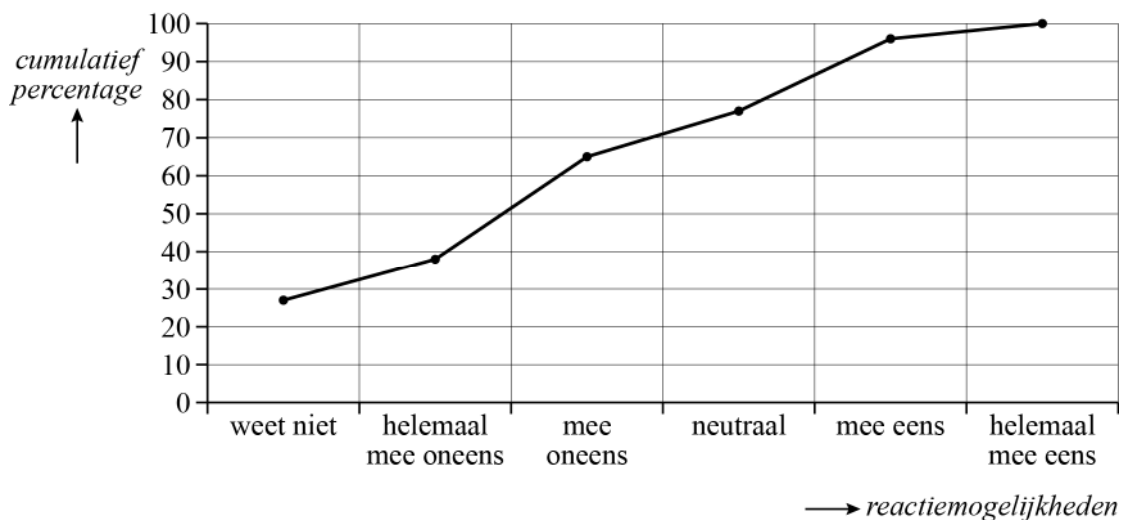
In de media werd naar aanleiding van dit onderzoek geconcludeerd dat ruim een derde van de leraren aangeeft dat het schoonmaakbudget te laag is om de hygiëne te kunnen waarborgen.

Kennelijk had men hiervoor de twee reactiemogelijkheden 'helemaal mee oneens' (11%) en 'mee oneens' (27%) samengevoegd.

- 2p **13** Had men op basis van deze tabel ook een uitspraak over het gemiddelde van de lerarenreacties kunnen doen? Licht je antwoord toe.

Een journalist maakte bij de gegevens uit de tabel een grafiek van het cumulatieve percentage. Zie de figuur.

figuur



De manier waarop de reactiemogelijkheid 'weet niet' is opgenomen in de grafiek kan misleidend worden genoemd.

- 3p **14** Leg uit waarom deze manier misleidend kan worden genoemd en geef aan wat een betere manier is om met deze reactiemogelijkheid om te gaan bij het maken van zo'n grafiek.



Printabonnement

Een copyshop biedt printabonnementen aan. Voor een vast bedrag per maand kun je een vast aantal pagina's afdrukken. Wil je meer pagina's afdrukken in die maand, dan moet je daarvoor extra betalen. Het aantal extra pagina's dat je afdruckt, wordt bijgehouden door de copyshop. In deze opgave bedoelen we met '(afdruk)kosten' telkens de kosten voor de klant, niet voor de copyshop.

Er zijn verschillende abonnementen. Zie de tabel.

tabel

Naam abonnement	Print Gratis	Print Soms	Print Regelmatig
vast bedrag per maand	€ 0	€ 3	€ 5
aantal gratis pagina's per maand	15	50	100
variabele kosten voor extra pagina's	€ 1 per 10 pagina's	€ 1 per 10 pagina's	€ 1 per 15 pagina's

Bijvoorbeeld: als je het Print Soms-abonnement hebt en in een maand 51 pagina's afdruckt, dan zijn de kosten voor die maand $3 + 1 = 4$ euro. Ook 52 pagina's kosten bij dit abonnement 4 euro.

Bij het afdrukken van 35 pagina's in een maand zijn de kosten met een Print Gratis-abonnement lager dan met een Print Soms-abonnement.

3p **15** Bereken hoeveel procent lager. Geef je antwoord in hele procenten.

Redouan en Jelle hebben beiden het Print Soms-abonnement voor een jaar afgesloten. In dat jaar drukken ze elk 600 pagina's af. Het is mogelijk dat Redouan voor dat jaar een ander bedrag kwijt is dan Jelle.

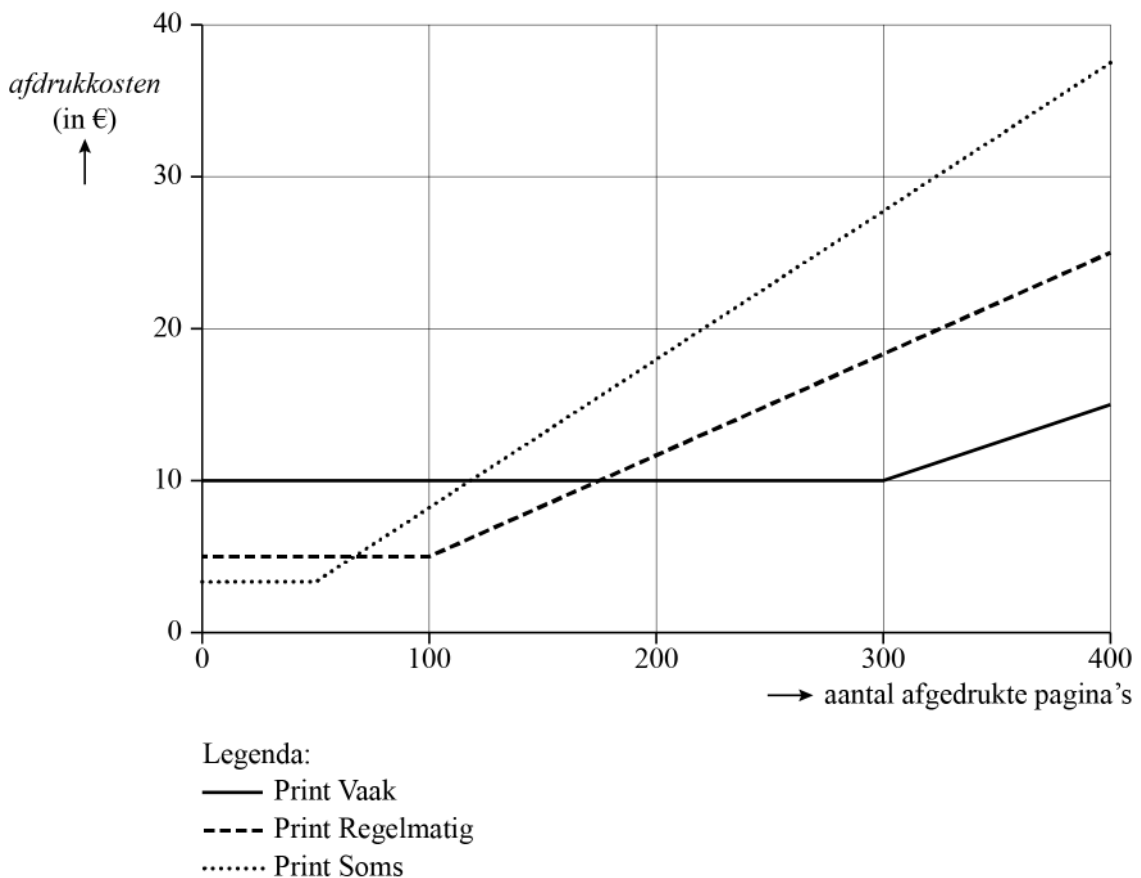
4p **16** Bereken het maximale verschil in euro's dat mogelijk is.



De copyshop besluit dat je voortaan het afdrukken van de extra pagina's niet meer per 10 of 15 pagina's betaalt, maar per stuk. Voor elk abonnement kunnen de kosten per afgedrukte pagina worden berekend met behulp van de variabele kosten uit de onderste rij van de tabel. Zo zijn bijvoorbeeld deze kosten bij het Print Gratis-abonnement 0,10 euro per pagina.

In de figuur zijn voor deze nieuwe situatie de afdrukkosten in een bepaalde maand voor de abonnementen Print Soms en Print Regelmatig weergegeven, afhankelijk van het aantal afgedrukte pagina's. Ook is de grafiek weergegeven van een nieuw abonnement dat de copyshop aanbiedt: Print Vaak.

figuur



Met behulp van de beschrijving van de nieuwe situatie en de figuur, zou je aan de tabel (van de oude situatie) een kolom kunnen toevoegen voor het abonnement Print Vaak.

- 4p 17 Bepaal voor het abonnement Print Vaak het vaste bedrag per maand, het aantal gratis pagina's per maand en het aantal extra pagina's dat je voor 1 euro kunt afdrukken. Licht je drie antwoorden toe.



In de rest van deze opgave bedoelen we met ‘afdrukkosten’ de afdrukkosten per maand voor de klant en met ‘pagina’s’ afgedrukte pagina’s per maand.

De afdrukkosten (in euro) voor de abonnementen Print Soms K_S en Print Regelmatig K_R worden gegeven door de volgende formules:

$$K_S = \begin{cases} 3 & \text{als } 0 \leq p < 50 \\ 0,1p - 2 & \text{als } p \geq 50 \end{cases} \quad (\text{formule 1})$$

$$K_R = \begin{cases} 5 & \text{als } 0 \leq p < 100 \\ \frac{1}{15}p - 1\frac{2}{3} & \text{als } p \geq 100 \end{cases} \quad (\text{formule 2})$$

Hierin is p het aantal pagina’s.

Voor kleine aantallen pagina’s zijn de afdrukkosten bij het abonnement Print Soms lager dan bij Print Regelmatig.

- 3p 18 Bereken met behulp van de formules voor K_S en K_R vanaf welk aantal pagina’s de afdrukkosten bij het abonnement Print Regelmatig lager zijn dan bij Print Soms.

Voor de **afdrukkosten per pagina** G_R voor het abonnement Print Regelmatig geldt:

$$G_R = \frac{K_R}{p} \quad (\text{formule 3})$$

Hierin is p het aantal pagina’s, waarbij p minstens 100 is.

Formule 3 kan in de volgende vorm worden geschreven:

$$G_R = a - \frac{b}{p} \quad (\text{formule 4})$$

Hierin zijn a en b positieve getallen en p is het aantal pagina’s, waarbij minstens 100 is.

- 3p 19 Bepaal de waarden van a en b in formule 4, zonder gebruik te maken van getallenvoorbeelden voor p . Geef deze waarden in twee decimalen.

De afdrukkosten per pagina G kunnen voor elk abonnement, als het aantal pagina’s groter is dan het aantal gratis pagina’s, beschreven worden met een formule van de vorm:

$$G = c - \frac{d}{p} \quad (\text{formule 5})$$

Hierin zijn c en d positieve getallen.

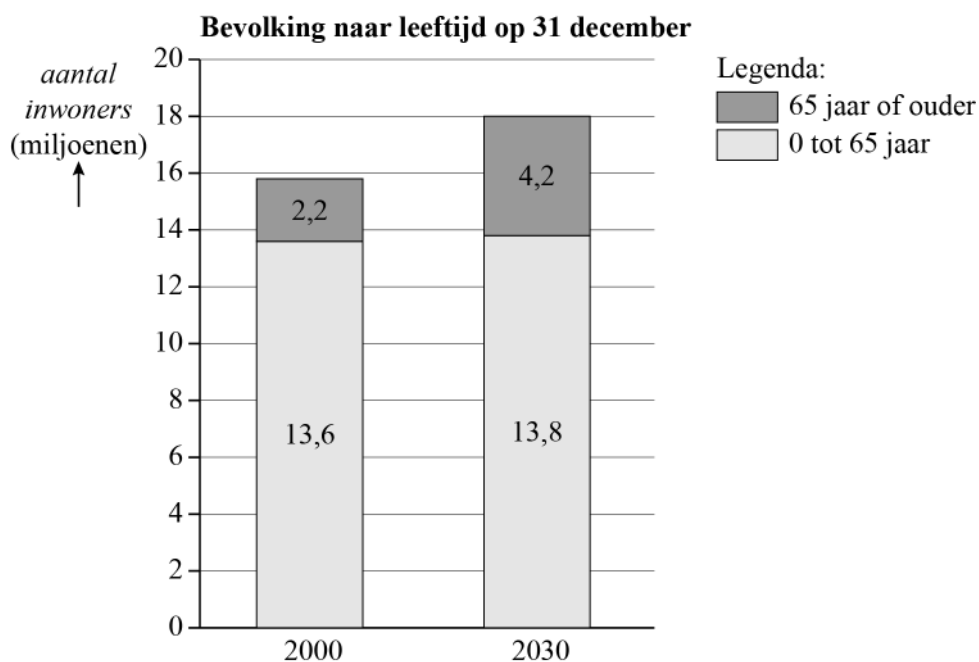
- 3p 20 Beredeneer met behulp van formule 5 of de afdrukkosten per pagina dan afnemen of toenemen als het aantal pagina’s toeneemt.



65-plussers

Het CBS (Centraal Bureau voor de Statistiek) publiceerde eind 2018 een artikel met prognoses over het aantal inwoners in Nederland waaruit bleek dat dit aantal minstens tot 2030 zal blijven toenemen. Ook beweerde het CBS in dit artikel dat in 2030 bijna een kwart van de bevolking 65-plus (65 jaar of ouder) is. Zie de figuur. De gegevens gelden telkens op 31 december van het jaar.

figuur



Op basis van gegevens van de jaren 2000 tot en met 2018 en prognoses van 2019 tot en met 2030 verwacht men dat de totale bevolking in de periode 2000-2030 nagenoeg lineair zal groeien, terwijl het aantal 65-plussers exponentieel zal groeien. In deze opgave gaan we ervan uit dat deze lineaire en exponentiële groei zich beide na 2030 voortzetten.

- 9p 21 Onderzoek in welk jaar er op 31 december meer dan een kwart van de bevolking 65-plusser is.

Bronvermelding

Een opsomming van de in dit examen gebruikte bronnen, zoals teksten en afbeeldingen, is te vinden in het bij dit examen behorende correctievoorschrift.

