Antwoorden en uitwerkingen reader statistiek voor de minor

Inhoudsopgave

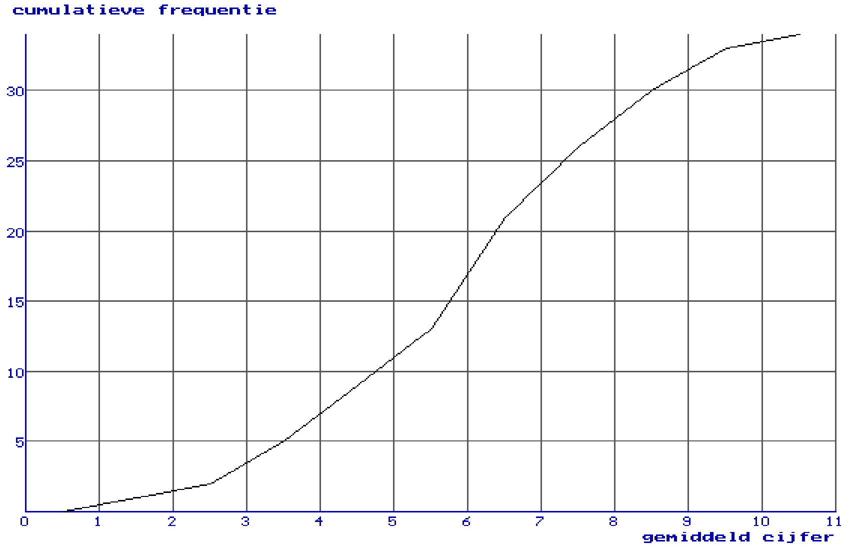
Laatst bijgewerkt op dinsdag 25 februari 2020

Hoofdstuk 1

Opgave 1

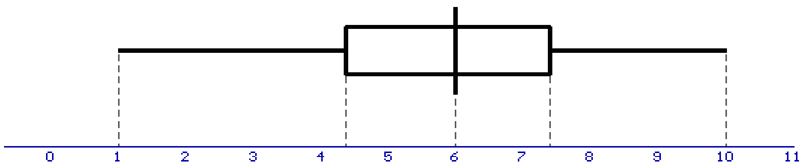
a. m=5,9 en s=2,1

b.



c. Mediaan=6, q1=4,4 en q3=7,4

d.



e. De modus is 6.

Opgave 2

a. $P(\text{minstens 1 rood}) = 1 - \frac{6}{10} \cdot \frac{5}{9} \cdot \frac{4}{8} = \frac{5}{6}$

b. $P(\text{3 rood}) = \frac{4}{10} \cdot \frac{3}{9} \cdot \frac{2}{8} = \frac{1}{30}$

c. $P(A \mid B) = \frac{P(A \text{ en } B)}{P(B)} = \frac{\frac{1}{30}}{\frac{1}{30}} = 1$ en $P(B \mid A) = \frac{P(A \text{ en } B)}{P(A)} = \frac{\frac{1}{30}}{\frac{5}{6}} = \frac{1}{25}$

d. Nee, $P(A \text{ en } B) \neq P(A) \cdot P(B)$

Opgave 3

a. $P(\text{minstens 1 rood}) = 1 - \left(\frac{6}{10}\right)^3 = \frac{98}{125}$

b. $P(3\ \text{rood}) = \left(\frac{4}{10}\right)^3 = \frac{8}{125}$

c. $P(A \mid B) = \frac{P(A\ \text{en}\ B)}{P(B)} = 1\ \ \text{en}\ P(B \mid A) = \frac{P(A\ \text{en}\ B)}{P(A)} = \frac{\frac{8}{125}}{\frac{98}{125}} = \frac{4}{49}$

d. Nee, $P(A\ \text{en}\ B) \neq P(A) \cdot P(B)$

Opgave 4

a. $\frac{9 \cdot 8 \cdot 7}{3!} = 84$

b.

CDA-PvdA-VVD:	4·2·2=16
CDA-PvdA-GL:	4·2·1=8
CDA-VVD-GL:	4·2·1=8
PvdA-VVD-GL:	2·2·1=4
In **totaal**:	36 mogelijkheden

c. P(minstens twee van dezelfde partij) =

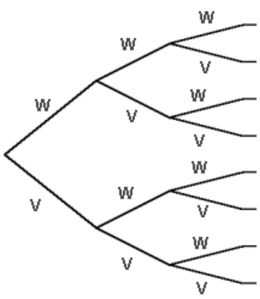
1 – P(allemaal verschillend)= $1 - \frac{36}{84} = 1 - \frac{3}{7} = \frac{4}{7} \approx 0{,}571$

Opgave 5

We noemen de kans om van **pa** te winnen **p** en de kans om van **ma** te winnen **m**.

De kans om te winnen:

P(winnen bij pa-ma-pa) = p·m + (1 – p)·m·p = p·m·(2 – p)
P(winnen bij ma-pa-ma)= m·p + (1 – m)·p·m = p·m·(2 – m)



Omdat **p<m** geldt:
P(winnen bij pa-ma-pa) > P(winnen bij ma-pa-ma)
Dus: **pa-ma-pa**

En dat is anders dan je zou verwachten, denk ik…☺

Opgave 6

a.

X : aantal passagiers dat op komt dagen

$p = 0,975$

$n = 120$

$P(X > 116) = 1 - P(X \le 116) = 1 - 0,35275 = 0,647$

b.

X : aantal passagiers dat op komt dagen

$p = 0,975$

$n = 120$

$P(X < 116) = P(X \le 115) = 0,183$

Opgave 7

$$\frac{\binom{16}{0}\binom{36}{4}}{\binom{52}{4}}\cdot 0+\frac{\binom{16}{1}\binom{36}{3}}{\binom{52}{4}}\cdot 10+\frac{\binom{16}{2}\binom{36}{2}}{\binom{52}{4}}\cdot 20+\frac{\binom{16}{3}\binom{36}{1}}{\binom{52}{4}}\cdot 30+\frac{\binom{16}{4}\binom{36}{0}}{\binom{52}{4}}\cdot 40\approx 12,31$$

Of

$$\sum_{k=0}^{4}\frac{\binom{16}{k}\binom{36}{4-k}}{\binom{52}{4}}\cdot 10k$$

Opgave 8

$$P(\text{meer dan 40 euro}) = 1-\left(\frac{1}{2}+\left(\frac{1}{2}\right)^2+\left(\frac{1}{2}\right)^3+\left(\frac{1}{2}\right)^4\right)\approx 0,063$$

Of veel handiger: $P(\text{meer dan 40 euro})=\left(\frac{1}{2}\right)^4 \approx 0,063$

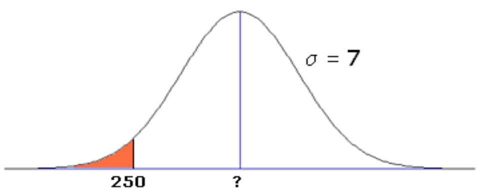
Opgave 9

a.

$\Phi(z) = 0,05 \Rightarrow z = -1,64$

$z = \dfrac{250-\mu}{7} = -1,64$

$\mu \approx 261,5$



b.

$\Phi(z) = 0,05 \Rightarrow z = -1,64$

$z = \dfrac{250-256}{\sigma} = -1,64$

$\sigma \approx 3,66$

Opgave 10

$E(A+B) = 75 + 50 = 125$

$\sigma(A+B) = \sqrt{3^2 + 2^2} = \sqrt{13} \approx 3,6$

$A + B \sim \text{Normaal}(125; 2,6)$

Gevraagd : $P(X > 130)$

Met de GR : normalcdf$(130, \infty, 125, 3.6) \rightarrow 0,0828$

$P(X > 130) = 0,0828$

Opgave 11

a. We willen 'aantonen' dat $\mu > 100$. Dit doen we door uit te gaan van $\mu = 100$. Als dat niet is vol te houden, verwerpen we H_0 en nemen we H_1 aan. We gaan dus uit van het tegendeel van wat we willen aantonen.

b.

Onder H_0 geldt :

$X \sim$ normaal verdeeld met $\mu = 100$ en $\sigma = \dfrac{15}{\sqrt{25}} = 3$

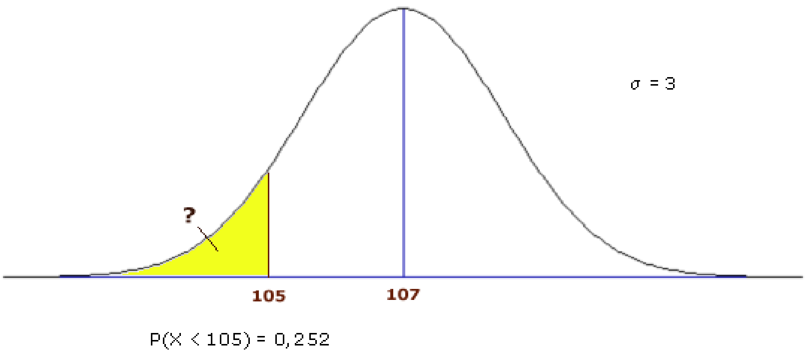
$\alpha = 0,05$

$P(X > k) = 0,05$

$k \approx 104,93$

Dus bij 105 of hoger verwerpen we H_0.

c.



Opgave 12

a. X: aantal weeffouten per rol

X~ Poissonverdeeld met $\lambda = 1$

$P(X \geq 2) = 1 - P(X \leq 1) \approx 1 - 0,736 = 0,264$

b. P(A kwaliteit) = 0,736

X: aantal rollen met A-kwaliteit

X~ binomiaal verdeeld met n=40 en p=0,736

$P(X \geq 30) = 1 - P(X \leq 29) \approx 1 - 0,499 = 0,501$

Hoofdstuk 2

Opgave 1

a. Herhalingscombinatie: aantal $= \binom{7-1+2}{2} = \binom{8}{2} = 28$

b. Permutatie: aantal=6! = 720

c. Combinatie: aantal=$\binom{15}{7} = 6.435$

d. Rangschikking met herhaling: $10^4 = 10.000$

Opgave 2

a. aantal $= \dfrac{14!}{7!\cdot 4!\cdot 3!}$

b. $P(2 \text{ rood, } 3 \text{ wit en } 4 \text{ blauw}) = \dfrac{9!}{2!\cdot 3!\cdot 4!}\cdot\left(\dfrac{2}{45}\right)^2\cdot\left(\dfrac{3}{45}\right)^3\cdot\left(\dfrac{4}{45}\right)^4 \approx 0{,}00000005$

Opgave 2

a. 8! = 40.320

b. 6 · 3! · 5! = 4.320

c. 5 · 4 · 6! = 14.400

Opgave 3

Met de formule "$\text{kans} = \dfrac{\text{aantal gunstige uitkomsten}}{\text{aantal mogelijke uitkomsten}}$" geeft dit P(3 kop)=$\dfrac{\binom{10}{3}}{2^{10}} = \dfrac{15}{128}$.

Met de binomiale verdeling: P(3 kop)=$\binom{10}{3}\cdot\left(\frac{1}{2}\right)^3\cdot\left(\frac{1}{2}\right)^7 = \frac{15}{128}$.

Opgave 4

Van de overige 8 vragen moet Jan dus gokken met 50% kans om goed te gokken. Hij moet nog minimaal 3 vragen goed hebben.

X: aantal vragen goed van de 8
X~binomiaal verdeeld met:
n=8
p=$\frac{1}{2}$
Gevraagd: $P(X \geq 3)$
Oplossing: $P(X \geq 3) = 1 - P(X \leq 2) \approx 1 - 0{,}1445 = 0{,}8555$

Opgave 5

In onderstaande tabellen zie je steeds in geel de mogelijkheden dat LINKS wint van BOVEN. In blauw BOVEN wint van LINKS.

		B					
		3	3	3	3	3	3
	0						
	0						
A	4						
	4						
	4						
	4						

		C					
		2	2	2	2	6	6
	3						
	3						
B	3						
	3						
	3						
	3						

		D					
		1	1	1	5	5	5
	2						
	2						
C	2						
	2						
	6						
	6						

		A					
		0	0	4	4	4	4
	1						
	1						
D	1						
	5						
	5						
	5						

...en dat is toch verwonderlijk! Kennelijk geldt hier niet wat bij getallen geldt: $a>b$ en $b>c \Rightarrow a>c$

Opgave 6

I.

$$\text{De kans } P = \frac{\text{aantal gunstige uitkomsten}}{\text{aantal mogelijke uitkomsten}}$$

Aantal gunstige uitkomsten is $3\cdot 8\cdot 7\cdot 6=1008$
Aantal mogelijke uitkomsten is $14\cdot 13\cdot 12=2184$
P(2 witte ballen)=$^{1008}/_{2184}={}^{6}/_{13}$.

II.

Er zijn eigenlijk twee soorten ballen: witte en niet-witte. Wat is nu de kans op (precies) twee witte ballen? Er zijn nu verschillende mogelijkheden: de eerste 2 ballen kunnen wit zijn en de derde niet, de eerste en de derde kunnen wit zijn, enzovoort....
Om de kans uit te rekenen kun je kijken naar één zo'n volgorde. We nemen maar wit, wit, niet-wit.
We kijken naar P(w,w,n), dus de kans op precies die volgorde!

P(w,w,n) = $^{8}/_{14}\cdot{}^{7}/_{13}\cdot{}^{6}/_{12} = {}^{2}/_{13}$

Vervolgens kijk je hoeveel verschillende volgordes je kunt maken met twee witte en een niet-witte.
Dit kan op 3 verschillende manieren, dus de kans op precies twee witte ballen is $3\cdot{}^{2}/_{13} = {}^{6}/_{13}$

$$P(\textit{twee witte ballen}) = \overbrace{3}^{\text{aantal verschillende volgorden}}\cdot\overbrace{\frac{8}{14}\cdot\frac{7}{13}\cdot\frac{6}{12}}^{\text{kans op één bepaalde volgorde}} = \frac{6}{13}$$

III.

Je kunt dit soort vragen ook oplossen door naar combinaties te kijken. Op hoeveel manieren kan je 2 witte knikkers uit de vaas halen.

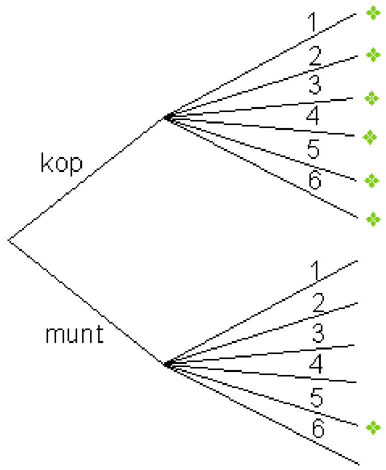
$$P(2\ \text{witte}) = \frac{\binom{8}{2}\cdot\binom{6}{1}}{\binom{14}{3}} \approx 0{,}462$$

Opgave 7

a. De optelregel geldt alleen als twee kansen disjunct zijn, dat wil zeggen: als A gebeurt, gebeurt B niet en als B gebeurt dan gebeurt A niet. In dit geval klopt dat dus niet, want je kan best 'kop' gooien en ook 'vijf'.

b. Blijft de vraag: wat is het dan wel!?

Oplossing 1
Teken een boomdiagram en tel het aantal takken, die aan de voorwaarde voldoen! We zien: **7** takken van de **12** voldoen, dus de kans op kop of vijf is $^7/_{12}$.



Oplossing 2
P(A of B)=P(A)+P(B)-P(A en B)
P(A of B)=$^1/_2+^1/_6-^1/_{12}=^7/_{12}$

Oplossing 3
Maak een kanstabel!

		vijf		
		ja	nee	
kop	ja			**1/2**
	nee			**1/2**
		1/6	**5/6**	**1**

Zijn de gebeurtenissen 'gooi kop' en 'gooi vijf' onafhankelijk? Antwoord: ja, dus:

		vijf		
		ja	nee	
kop	ja	**1/12**	**5/12**	1/2
	nee	**1/12**	5/12	1/2
		1/6	**5/6**	**1**

P(kop of vijf)=$^7/_{12}$

c. Zie oplossing 2

Opgave 8

a. P(A)=$\frac{1}{2}$, P(B)=$\frac{1}{6}$, P(A en B)=$\frac{1}{6}$ en P(A of B)=$\frac{1}{2}$.
b. Klopt!
c. P(A) · P(B)=$\frac{1}{12}$ en dat is niet hetzelfde als P(A en B). A en B zijn **niet** onafhankelijk.
d. P(A|B)=1 en P(B|A)=$\frac{1}{3}$

Opgave 9

a. P(A)=$\frac{5}{12}$, P(B)=$\frac{1}{6}$ en P(A en B)=$\frac{1}{12}$

b. P(A of B) = P(A) + P(B) - P(A en B)= $\frac{5}{12}$ + $\frac{1}{6}$ - $\frac{1}{12}$ = $\frac{1}{2}$

c. Nee, want P(A en B) $\neq$ 0.

d. Nee, want P(A) · P(B) $\neq$ P(A en B)

Opgave 10

Gebruik $\mathrm{P(A|B)} = \frac{\mathrm{P(A\ en\ B)}}{\mathrm{P(B)}}$. Er geldt: $0{,}30 = \frac{0{,}20}{\mathrm{P(B)}}$ $\Rightarrow$ P(B)=$\frac{2}{3}$

Opgave 11

Als A en B onafhankelijk zijn dan geldt P(A en B)=P(A) · P(B). Er geldt: P(A en B)=0,21

Opgave 12

1126 $\longrightarrow$ 12 manieren

1135 $\longrightarrow$ 12 manieren

1144 $\longrightarrow$ 6 manieren

1225 $\longrightarrow$ 12 manieren

1234 $\longrightarrow$ 24 manieren

1333 $\longrightarrow$ 4 manieren

2224 $\longrightarrow$ 4 manieren

2233 $\longrightarrow$ 6 manieren

Er zijn 80 ‘manieren’ om 10 ogen te gooien. P(10 gooien met 4 dobbelstenen) =$\frac{80}{6^4} = \frac{5}{81}$

Hoofdstuk 3

Opgave 1

a. P(na 5 keer 6)=$\left(\frac{5}{6}\right)^4 \cdot \frac{1}{6} = \frac{625}{7776}$

b. $P(X>3)=1-\left(P(X=1)+P(X=2)+P(X=3)\right)=1-\left(\frac{1}{6}+\frac{5}{6}\cdot\frac{1}{6}+\left(\frac{5}{6}\right)^2\cdot\frac{1}{6}\right)=\frac{125}{216}$

Opgave 2

a. X: aantal studenten dat op komt dagen

X is binomiaal verdeeld met p=0,9 en n=55.

Gevraagd: $P(X>50)$

Oplossing: $P(X>50)=1-P(X\leq 50)\approx 1-0,655=0,345$

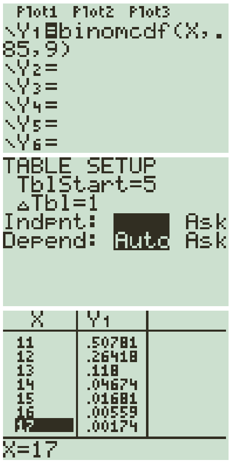
b. X:aantal toetsen dat nodig is

X is binomiaal verdeeld met p=0,9 en n=55

Gevraagd: $P(X\leq 49)$

Oplossing: $P(X\leq 49)=0,476$

Opgave 3



X: aantal mensen dat bereikt wordt

X is binomiaal verdeeld met p=0,85 en n=onbekend.

Er geldt: $P(X\geq 10)>0,90$

Gevraagd: Wat is n?

Oplossing:

$P(X\geq 10)>0,90$

$1-P(X\leq 9)>0,90$

$P(X\leq 9)<0,10$

$n=14$

Opgave 4

a. $P(3\text{ mannen})=\dfrac{\binom{20}{3}\binom{14}{2}}{\binom{34}{5}}\approx 0,373$

b. P(minstens 3 mannen)=$\dfrac{\binom{20}{3}\binom{14}{2}}{\binom{34}{5}}+\dfrac{\binom{20}{4}\binom{14}{1}}{\binom{34}{5}}+\dfrac{\binom{20}{5}}{\binom{34}{5}}\approx 0,672$

Opgave 5

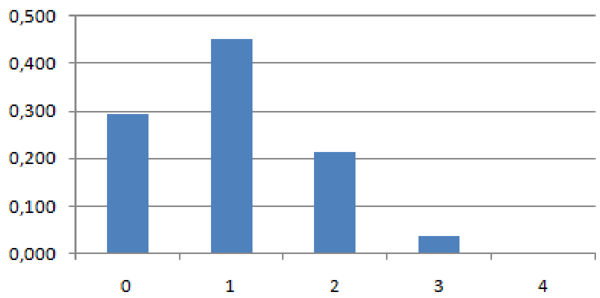
$$P(X=0)=\frac{\binom{4}{0}\binom{28}{8}}{\binom{32}{8}}\approx 0{,}295$$

$$P(X=1)=\frac{\binom{4}{1}\binom{28}{7}}{\binom{32}{8}}\approx 0{,}450$$

$$P(X=2)=\frac{\binom{4}{2}\binom{28}{6}}{\binom{32}{8}}\approx 0{,}215$$

$$P(X=3)=\frac{\binom{4}{3}\binom{28}{5}}{\binom{32}{8}}\approx 0{,}037$$

$$P(X=4)=\frac{\binom{4}{4}\binom{28}{4}}{\binom{32}{8}}\approx 0{,}002$$



Opgave 6

1 huis per 2000 geeft voor een gemeente met 6000 huizen: $\lambda = 3$

$$P(X=4)=e^{-3}\cdot\frac{3^4}{4!}\approx 0{,}168$$

Opgave 7

a. Wegens de symmetrie in de grafiek van de normale verdeling is het gevraagde percentage gelijk aan het percentage dat korter is dan 20 cm, dus 5%

b.

$$\left.\begin{array}{l}\mu = 25 \text{ en } \sigma = s\\ P(X<20)=0{,}05\\ z=-1{,}64\end{array}\right\}\Rightarrow -1{,}64=\frac{20-25}{s}\Rightarrow s\approx 3{,}05$$

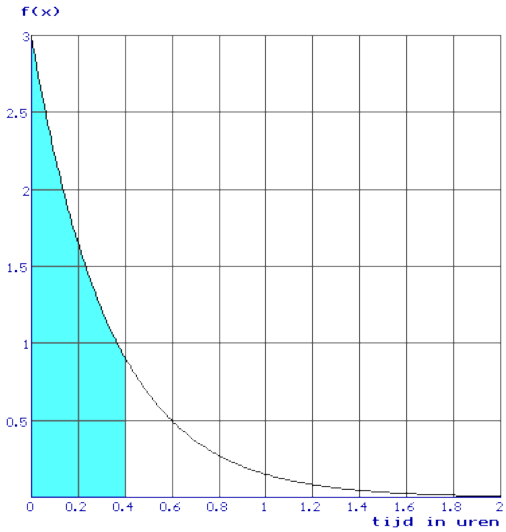
c.

- De gevraagde kans is $P(140 < X < 170 \mid \mu = 145 \text{ en } \sigma = 15)$
- Het gebruik van een geschikte functie op de GR
- Het antwoord 0,58

d.

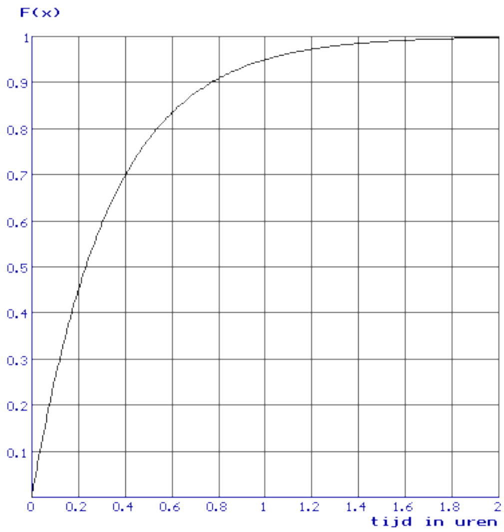
- Als er bij 100 bomen a kleine bomen zijn, zijn er $100 - a$ grote
- Dit geeft de vergelijking $10a + 15(100 - a) = 1300$
- De oplossing hiervan is $a = 40$
- 40 van de 100 bomen, dus 40%, moet als klein worden verkocht
- $P(X < g \mid \mu = 145 \text{ en } \sigma = 15) = 0{,}4$
- Dit geeft $g \approx 141{,}2 \approx 141$ cm

Opgave 8



24 minuten komt overeen met 0,4 uur en $\lambda = 3$.

$$P(X > 0,4) = 1 - P(X \leq 0,4) = 1 - \int_0^{0,4} 3e^{-3t}\,dt \approx 0,301$$



Dat was andere koek! ☺

Hoofdstuk 4

Opdracht 1

a. De populatie is de verzameling deeltijdstudenten van jaar 1. Het steekproefkader is de lijst in Osiris van het vakproject en mijn steekproef bestaat uit de 5 willekeurige studenten van de lijst.

b. Er zijn nogal wat bezwaren:
   1. Niet elke deeltijdstudent van jaar 1 doet mee met het vakproject.
   2. Niet elke deeltijdstudent van jaar 1 schrijft zich in Osiris in.
   3. De steekproef is (vast) niet groot genoeg (zie vervolg).
   4. …

c. Dit is een binomiaal kansprobleem:

   X:aantal studenten dat voor de klas staat

   X~binomiaal verdeeld met p=0,2 en n=5

   $H_0 : p = 0,2$

   $H_1 : p \neq 0,2$

   Onder H_0:$P(X \geq 3)=1-P(X \leq 2) \approx 0,058$

   0,058 is groter dan 0,025 ($\alpha = 0,05$)

   Er is geen reden om H_0 te verwerpen.

Opdracht 2

- “Het zit hem er natuurlijk in, dat er veel meer mensen zeggen dat de bussen vol zijn, omdat er nou eenmaal meer mensen in een volle bus zitten!”
  http://hhofstede.nl/misbruik/busritten.htm
- Rekenvoorbeeld: Neem aan dat je twee bussen hebt, één bus met 1 persoon en één bus met 29 personen. De gemiddelde busbezetting is $\frac{1+29}{2} = 15$. Maar volgens de manier van de opdracht is de gemiddelde busbezetting gelijk aan: $\frac{1 \cdot 1 + 29 \cdot 29}{30} \approx 28$

Opdracht 3

$$46 - 2.58 \cdot \frac{5}{\sqrt{100}} < \mu < 46 + 2.58 \cdot \frac{5}{\sqrt{100}}$$

$$44.71 < \mu < 47.29$$

Opdracht 4

a.

$$20.142 - 1.96 \cdot \frac{0.1}{\sqrt{25}} < \mu < 20.142 + 1.96 \cdot \frac{0.1}{\sqrt{25}}$$

$$20.103 < \mu < 20.181$$

b.

$$20.142 - 2.58 \cdot \frac{0.1}{\sqrt{25}} < \mu < 20.142 + 2.58 \cdot \frac{0.1}{\sqrt{25}}$$

$$20.090 < \mu < 20.194$$

Opdracht 5

$z = 1.96$

$\sigma = 4$

$a = 1$

$$n \geq \frac{1.96^2 \cdot 4^2}{1^2} \approx 62$$

Opdracht 6

Met de functie **tcdf(lower, upper, df)** kan je de oppervlakte onder de grafiek bepalen van ‘linkergrens’ tot ‘rechtergrens’ met ‘het aantal vrijheidsgraden’.

Zo zou de opdracht tcdf(2.015,9E99,5) gelijk moeten zijn aan 0,05.
En wat denk je?

```
tcdf(2.015,9E99,
5)
     .0500030842
```

Klopt!

Opdracht 7

$\bar{x} = 212$ en $s_X = 18.688$

We passen de t-verdeling toe met 95% betrouwbaarheid. Het aantal vrijheidsgraden is v=8 en de rechteroverschrijdingskans is 0,025.
Uit de tabel: $t_8 = 2.306$

$$\bar{x} - t \cdot \frac{s}{\sqrt{n}} < \mu < \bar{x} + t \cdot \frac{s}{\sqrt{n}}$$

$$212 - 2.306 \cdot \frac{18.688}{\sqrt{9}} < \mu < 212 + 2.306 \cdot \frac{18.688}{\sqrt{9}}$$

$$197.6 < \mu < 226.4$$

Opdracht 8

$\bar{x} = 86$ en $s_X = 4.123$

We passen de t-verdeling toe met 99% betrouwbaarheid. Het aantal vrijheidsgraden is v=8 en de rechteroverschrijdingskans is 0,005.
Uit de tabel: $t_8 = 3.355$

$$\bar{x} - t \cdot \frac{s}{\sqrt{n}} < \mu < \bar{x} + t \cdot \frac{s}{\sqrt{n}}$$

$$86 - 3.355 \cdot \frac{4.123}{\sqrt{9}} < \mu < 86 + 3.355 \cdot \frac{4.123}{\sqrt{9}}$$

$$81.39 < \mu < 90.61$$

Hoofdstuk 5

Opdracht 1

X: aantal bacteriën

H_0: $\mu(X)=900$

H_1: $\mu(X)>900$, rechtszijdige toets met $\alpha=0{,}01$

Toetsingsgrootheid: $\bar{X}$, het gemiddeld aantal bacteriën in de steekproef

$\sigma(X)$ is onbekend, dus puntschatten met s_X

$$\left.\begin{array}{l} \bar{x}=940 \\ s_X \approx 21{,}34 \\ \left.\begin{array}{l}\alpha=0{,}01 \\ v=9\end{array}\right\} \Rightarrow t=2{,}821 \end{array}\right\} \Rightarrow x=900+2{,}821\cdot\frac{21{,}34}{\sqrt{10}}\approx 919$$

$$Z=\{x \mid x>919\}$$

Het gemiddelde ligt in het kritieke gebied. We verwerpen H_0 en nemen H_1 aan. Er zitten waarschijnlijk gemiddeld meer dan 900 bacteriën in een monster.

Opgave 2

X: aantal personen waarbij het middel niet werkt

H_0: $p\leq 0{,}01$

H_1: $p>0{,}01$

X~binomiaal verdeeld met n=200 en p=0,01 (onder H_0)

$\mu(X)=0{,}01\cdot 200=2$

Omdat n groot is en p klein mogen we dit ook benaderen met de poissonverdeling.

X~poisson verdeeld met $\lambda=2$

$P(X \geq 6) < 0{,}01$ dus het kritieke gebied is $\{X \mid X \geq 6\}$.

We hadden gevonden X=8, dus we verwerpen H_0 en nemen H_1 aan. De bewering van de fabrikant is waarschijnlijk niet juist.

Opgave 3

X: aantal keren 'zes gooien'.

H_0: $p=\frac{1}{6}$

H_1: $p\neq\frac{1}{6}$

X~binomiaal verdeeld met n=60 en $p=\frac{1}{6}$ (onder H_0)

$P(X\geq k)<0{,}025 \Rightarrow P(X\leq k-1)>0{,}95 \Rightarrow k-1=16 \Rightarrow k=17$

Z={X | X ≥ 17}. We hadden gevonden X=13. Dat ligt niet in het kritieke gebied. We verwerpen H_0 niet. Er is dus geen reden om aan te nemen dat de dobbelsteen onzuiver is.

Opgave 4

Observed

	CDA	D'66	PvdA	VVD	Totaal
Voor de Euro	96	84	106	154	440
Tegen de Euro	154	38	100	128	420
Totaal	250	122	206	282	**860**

Expected

	CDA	D'66	PvdA	VVD	Totaal
Voor de Euro	127,9	62,4	105,4	144,3	440
Tegen de Euro	122,1	59,6	100,6	137,7	420
Totaal	250	122	206	282	**860**

$\chi^2 \approx 32{,}94$

't Is een 2 bij 4 tabel, dus het aantal vrijheidsgraden v = (2-1)(4-1)=3

Met α=0,05 aflezen in de tabel geeft $g_{0,95} = 7{,}81$

$Z=\{\chi^2 \mid \chi^2>7{,}81\}$

We hadden gevonden χ^2=32,94. Dat ligt in het kritieke gebied. We verwerpen H_0 en nemen H_1 aan. We kunnen concluderen dat het voor- of tegen de euro zijn afhangt van de politieke voorkeur.

Opgave 5

Observed

Aantal ogen	1	2	3	4	5	6
Waargenomen frequentie	7	10	12	9	9	13

Expected

Aantal ogen	1	2	3	4	5	6
Waargenomen frequentie	10	10	10	10	10	10

$$\chi^2 = \frac{(7\text{-}10)^2}{10} + \frac{(10\text{-}10)^2}{10} + \frac{(12\text{-}10)^2}{10} + \frac{(9\text{-}10)^2}{10} + \frac{(9\text{-}10)^2}{10} + \frac{(13\text{-}10)^2}{10}$$

χ^2= 0,9 + 0 + 0,4 + 0,1 + 0,1 + 0,9 = 2,4

Het aantal vrijheidsgraden is 5.

Met α=0,05 aflezen in de tabel geeft $g_{0,95} = 11{,}07$

De gevonden waarde ligt niet in het kritieke gebied. Er is geen reden om H_0 te verwerpen. Er is derhalve geen reden om aan te nemen dat de dobbelsteen onzuiver is.

Opgave 6

X~norm(μ=?, σ=10) met $\sigma(\bar{X})=\frac{\sigma(X)}{\sqrt{n}}=\frac{10}{\sqrt{10}}=\sqrt{10}$

Y~norm(μ=?, σ=$\sqrt{30}$) met $\sigma(\bar{Y})=\frac{\sigma(Y)}{\sqrt{m}}=\frac{\sqrt{30}}{\sqrt{5}}=\sqrt{6}$

H_0: X-Y=0
H_1: X-Y≠0, tweezijdig met α=0,05
Toetsingsgrootheid V=X-Y

V~normaal verdeeld met $\mu(\bar{V})$=0 en $\sigma(\bar{V})=\sqrt{\sigma^2(\bar{X})+\sigma^2(\bar{Y})}=\sqrt{10+6}=4$

Bij α=0,05 (tweezijdig) hoort een z-score van z=1,96

$Z=\langle\leftarrow,0-1{,}96\cdot 4\rangle\cup\langle 0+1{,}96\cdot 4,\rightarrow\rangle$

$Z=\langle\leftarrow,-7{,}84\rangle\cup\langle 7{,}84,\rightarrow\rangle$

We hadden gevonden V=-5. Dat ligt niet in het kritiek gebied. Er is geen reden om H0 te verwerpen.
De variabelen kunnen dezelfde verwachtingswaarde hebben.

Opgave 7

H_0: $\mu_1-\mu_2=0$
H_1: $\mu_1-\mu_2\neq 0$, tweezijdig met α=0,05
Toetsingsgrootheid $V=\bar{X}_1-\bar{X}_2$
Onder H_0 geldt: $\mu(V)=0$ en σ(V) is onbekend, maar $\sigma(X_1)$ en $\sigma(X_2)$ zijn wel gelijk.

We weten wel dat $\sigma(V)=\sqrt{\frac{s_{X_1}^2}{11}+\frac{s_{X_2}^2}{7}}$. $s_{X_1}^2$ en $s_{X_2}^2$ kun je dan (pooled) schatten met s_p^2.

$$s_p^2=\frac{(11-1)s_{X_1}^2+(7-1)s_{X_2}^2}{11-1+7-1}=\frac{10s_{X_1}^2+6s_{X_2}^2}{16}$$

$s_{X_1}^2=30{,}2$

$s_{X_2}^2=31$

$$s_p^2=\frac{10\cdot 30{,}2+6\cdot 31}{16}\approx 30{,}5$$

$$s_V=\sqrt{\frac{30{,}5}{11}+\frac{30{,}5}{7}}\approx 2{,}67$$

De t-toets met α=0,05 (tweezijdig) en v=16 geeft t=2,120

$0-t\cdot s_V<\bar{V}<0+t\cdot s_V$

$0-2{,}12\cdot 2{,}67<\bar{V}<0+2{,}12\cdot 2{,}67$

$-5{,}66<\bar{V}<5{,}66$

$Z=\langle\leftarrow;-5{,}66\rangle\cup\langle 5{,}66;\rightarrow\rangle$

We hadden gevonden V=40-30=10. We verwerpen H_0 en accepteren H_1.
De gemiddelde inkomens zijn ongelijk.

Opgave 8

H_0: μ_1-μ_2=0

H_1: μ_1-$\mu_2 \neq 0$, tweezijdig met α=0,05

Toetsingsgrootheid $V = \overline{X}_1 - \overline{X}_2$

Onder H_0 geldt: $\mu(V) = 0$ en σ(V) is onbekend en $\sigma(X_1)$ en $\sigma(X_2)$ zijn ongelijk.

Met je GR:

$s^2_{X_1} = 30,2$

$s^2_{X_2} = 31$

$$s_V = \sqrt{\frac{s^2_{X_1}}{11} + \frac{s^2_{X_2}}{7}} = \sqrt{\frac{30,2}{11} + \frac{31}{7}} \approx 2,68\,.$$

De t-toets met α=0,05 (tweezijdig) en v=16 geeft t=2,120

$0 - t \cdot s_V < \overline{V} < 0 + t \cdot s_V$

$0 - 2,12 \cdot 2,68 < \overline{V} < 0 + 2,12 \cdot 2,68$

$-5,68 < \overline{V} < 5,68$

$Z = \langle \leftarrow; -5,68 \rangle \cup \langle 5,68; \rightarrow \rangle$

We hadden gevonden V=40-30=10. We verwerpen H_0 en accepteren H_1.
De gemiddelde inkomens zijn ongelijk.

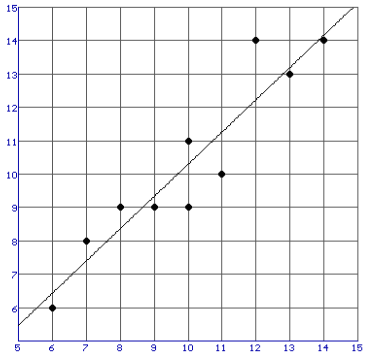
En dat maakt kennelijk weinig uit… het zou nog een idee zijn om te kijken naar twee steekproeven waarbij de standaarddeviatie sterk verschillen (minimaal een factor 4). Maar misschien doen we dat nog wel…

Hoofdstuk 6

Opdracht 1

Variabele	Nominaal	Ordinaal	Interval	Verhouding
Geslacht	X			
Leeftijd in jaren				X
Lengte in meter				X
Gewicht in kg				X
Hoogst afgeronde opleiding		X		
Soort rijbewijs	X			
Aantal kinderen				X
Gemiddeld aantal glazen alcohol per week				X
Aantal jaren in Nederland				X
IQ			X	

Opdracht 2



lijn door (5;5,6) en (14;14,1)

$$rc = \frac{14{,}1-5{,}6}{14-5} = \frac{8{,}5}{9} \approx 0{,}94$$

Vul (5;5,6) in: $Y = b + 0{,}94X$

$5{,}6 = b + 0{,}94 \cdot 5 \Rightarrow b \approx 0{,}88$

$Y = 0{,}88 + 0{,}94 \cdot X$

Opdracht 3

$$\left.\begin{array}{l} a = \dfrac{10 \cdot 1088 - 100 \cdot 103}{10 \cdot 1060 - 100^2} \approx 0{,}966667 \\ b = \dfrac{103 - 0{,}966667 \cdot 100}{10} \approx 0{,}63333 \end{array}\right\} \Rightarrow \text{Klopt!}$$

Opdracht 4

$$\left.\begin{array}{l} a = \dfrac{7 \cdot 4717 - 49 \cdot 560}{7 \cdot 427 - 49^2} \approx 9{,}49 \\ b = \dfrac{560 - 9{,}49 \cdot 49}{7} \approx 13{,}58 \end{array}\right\} \Rightarrow Y = 13{,}58 + 9{,}48 \cdot X$$

Opdracht 5

$$r = \frac{10 \cdot 1088 - 100 \cdot 103}{\sqrt{\left(10 \cdot 1060 - 100^2\right)\left(10 \cdot 1125 - 103^2\right)}} \approx 0{,}935$$

Opdracht 6

X	0	2	4	1	3	5	6	21
Y	2	3	6	1	4	5	5	26
X^2	0	4	16	1	9	25	36	91
Y^2	4	9	36	1	16	25	25	116
XY	0	6	24	1	12	25	30	98

$$r = \frac{7 \cdot 98 - 21 \cdot 26}{\sqrt{\left(7 \cdot 91 - 21^2\right)\left(7 \cdot 116 - 26^2\right)}} \approx 0{,}857$$

Opdracht 7

$r \approx 0{,}857492...$

Opdracht 8

a.

```
LinReg
 y=ax+b
 a=.1561959654
 b=3.855907781
 r²=.921533526
 r=.9599653775
```

b.

X	Y_1
60	13.228
61	13.384
62	13.54
63	13.696
64	13.852
65	14.009
66	14.165

X=65

Dat is 14

Opdracht 9

	1	2	3	4	5	6
Hoogst genoten onderwijs	VO	VO	MBO	HBO	HBO	UNIVERSITEIT
Rangnummers	1	2	3	4	5	6
Rangnummers na correctie	1,5	1,5	3	4,5	4,5	6
Inkomen	32.000	44.000	36.000	37.500	42.000	41.000
Rangnummers inkomen	1	6	2	3	5	4
d_i	0,5	-4,5	1	1,5	-0,5	2
d_i^2	0,25	20,25	1	2,25	0,25	4

$$\sum d_i^2 = 0{,}25 + 20{,}25 + 1 + 2{,}25 + 0{,}25 + 4 = 28$$

$$r_S = 1 - \frac{6\sum d_i^2}{n^3 - n} = 1 - \frac{6 \cdot 28}{6^3 - 6} \approx 0{,}2$$

Opdracht 10

In de verste verte niet!

Opdracht 11

Nou net niet...☺

Opdracht 12

Observed

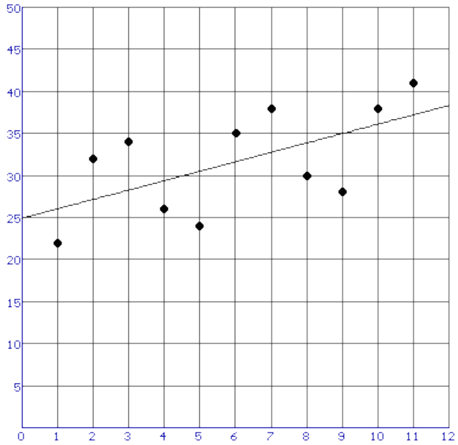
	Zelfstandig	Bij ouders	Totaal
Man	16	24	40
Vrouw	12	8	20
	28	32	60

Expected

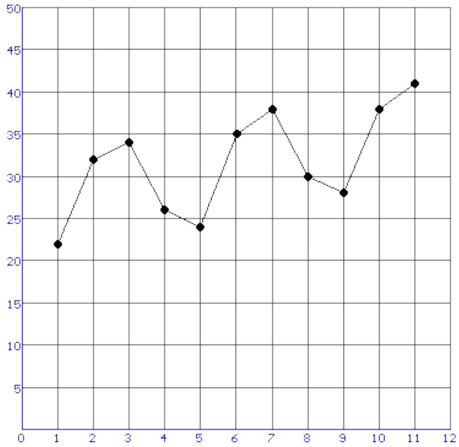
	Zelfstandig	Bij ouders	Totaal
Man	$18\frac{2}{3}$	$21\frac{1}{3}$	40
Vrouw	$9\frac{1}{3}$	$10\frac{2}{3}$	20
	28	32	60

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} \approx 2{,}143 \qquad \phi = \sqrt{\frac{\chi^2}{n}} = \sqrt{\frac{2{,}143}{60}} \approx 0{,}189 \text{ en } V = \sqrt{\frac{\chi^2}{n(k-1)}} = \sqrt{\frac{2{,}143}{60 \cdot 1}} \approx 0{,}189$$

Opdracht 13

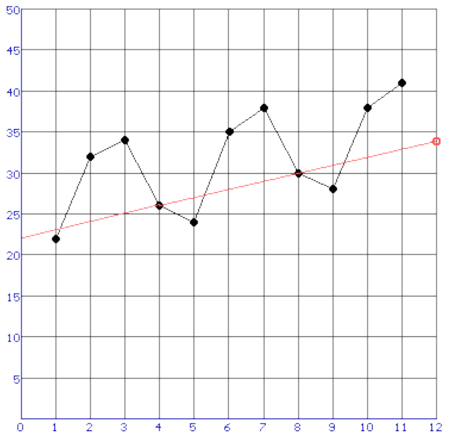


Hiermee zou je kunnen voorspellen dat de omzet in de IV-de periode van 1996 ongeveer 38,3 zal zijn.
Maar 'erg waarschijnlijk' lijkt me dat niet!



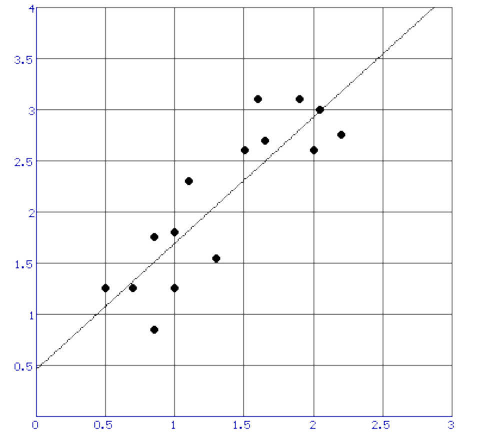
Er valt iets op! Dit lijkt wel een periodiek verschijnsel! Weliswaar neemt de totale omzet per jaar toe, maar er is duidelijk sprake van een seizoengebonden omzet!

Dus voor een voorspelling van de omzet in periode IV in 1996 zou je beter alleen naar de punten in de verschillende IV-de periodes kunnen kijken:



In dat geval lijkt me de voorspelling 'de omzet in periode IV in 1996 is ongeveer 34' realistischer dan de voorspelling hierboven.

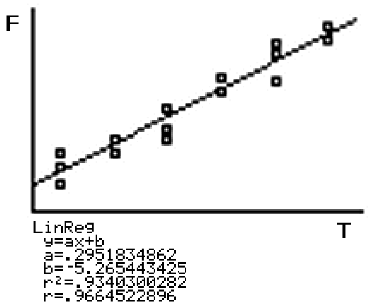
Opdracht 14

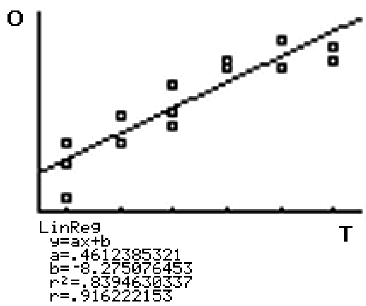


a. r ≈ 0,87

b. Wat dacht je van 'Frisdrank veroorzaakt verkeersdrukte' of 'Verkeersdeelneming maakt dorstig'? Een beetje onzin waarschijnlijk...

c.





d. Het blijkt dat er sprake is van een derde factor. Er bestaat een samenhang tussen temperatuur en het nuttigen van frisdrank. Er bestaat ook een samenhang tussen temperatuur en verkeersdrukte (allemaal naar 't strand). Maar de samenhang tussen frisdrank en verkeersdrukte is onzin... een klassieke valkuil...

Vrij vertaald van http://www.hhofstede.nl/misbruik/derdefactor.htm

Polio wordt niet veroorzaakt door een bacterie maar door een virus.