



**STATISTIEK** VOOR HET SECUNDAIR ONDERWIJS

Verklarende statistiek

7. Betrouwbaarheidsintervallen voor proporties

*Werktekst voor de leerling*

Prof. dr. Herman Callaert

Hans Bekaert  
Cecile Goethals  
Lies Provoost  
Marc Vancaudenberg

# Betrouwbaarheidsintervallen voor proporties

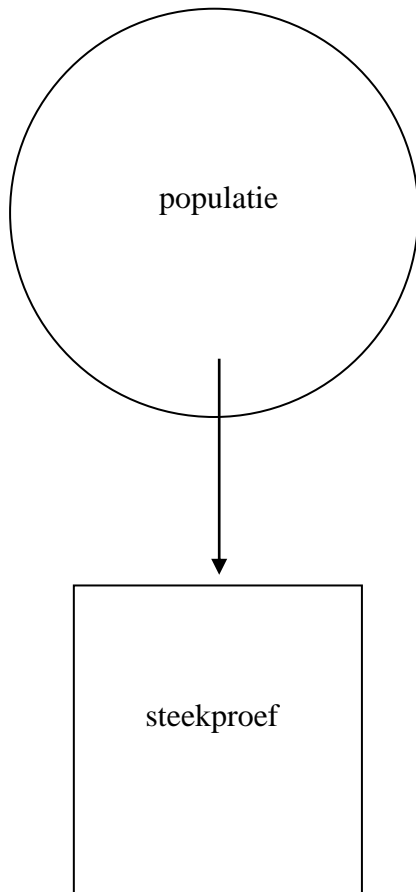
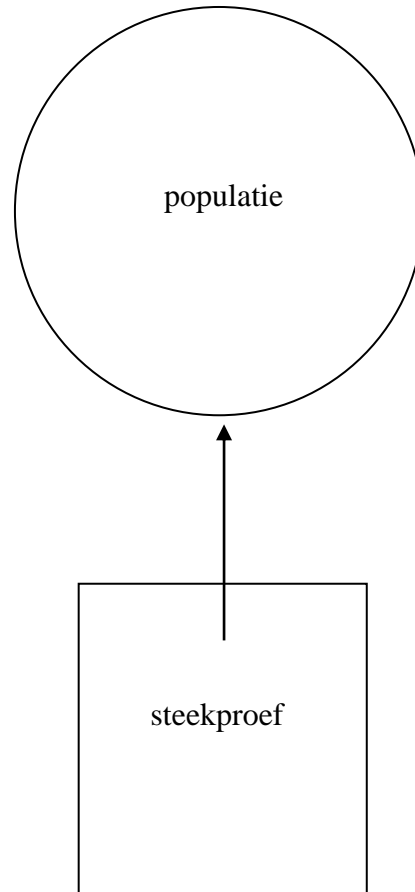
<b>1. Verklarende statistiek .....</b>	<b>2</b>
<b>2. Een populatieparameter schatten .....</b>	<b>3</b>
<b>3. Puntschatten .....</b>	<b>3</b>
<b>4. Intervalschatten .....</b>	<b>4</b>
4.1. Hoe het werkt, in theorie .....	4
4.1.1. Een voorbeeld .....	5
4.1.2. Een kansmodel voor intervallen .....	7
4.2. Hoe het werkt, in de praktijk .....	7
4.2.1. Een voorbeeld .....	8
4.2.2. Een kansmodel voor betrouwbaarheidsintervallen .....	9
4.2.3. Een experiment .....	10
4.3. Observeerbaarheid .....	13
<b>5. Hoe werken betrouwbaarheidsintervallen? .....</b>	<b>14</b>
5.1. Het Rad van Fortuin .....	14
5.2. Niet zeker, en dan? .....	15
5.3. Betrouwbaarheid is .....	15
<b>6. Betrouwbaarheid en precisie .....</b>	<b>16</b>
6.1. De betrouwbaarheid veranderen .....	16
6.2. Wat is een foutenmarge? .....	17
6.3. Een voorbeeld uit de media .....	19
<b>7. Hoe groot moet de steekproef zijn? .....</b>	<b>20</b>
7.1. De theorie .....	20
7.2. De praktijk .....	21
7.2.1. Externe informatie .....	21
7.2.2. Speel op veilig .....	22
7.2.3. Exit polls bij nipte verkiezingen .....	22
7.2.4. Risicobaby's .....	23
<b>8. Voor niets gaat de zon op .....</b>	<b>25</b>
8.1. Drie basisgrootheden .....	25
8.2. Samenvatting .....	27

## 1. Verklarende statistiek

Bij de studie van kansmodellen heb je geleerd hoe je uit het kansmodel van de populatie allerlei eigenschappen kan afleiden voor het kansmodel van het steekproefgemiddelde of van de steekproefproportie.

In de verklarende statistiek ga je omgekeerd te werk. Uit de informatie in een steekproef probeer je iets te weten te komen over een kenmerk van de populatie, zoals het populatiegemiddelde of de populatieproportie.

kansmodellen

verklarende  
statistiek

## 2. Een populatieparameter schatten

Het gemiddelde geboortegewicht van alle kinderen die in 2003 in Vlaanderen geboren zijn ken je niet omdat je de volledige populatie van al die geboortegewichten niet hebt. Maar dit gemiddelde geboortegewicht bestaat wel. Je noteert het met  $\mu$ . Om dezelfde reden weet je ook niet wat de echte proportie meisjes is bij al die geboorten. Maar ook die bestaat en zij wordt genoteerd door  $\pi$ . Grootheden zoals  $\mu$  en  $\pi$  zijn populatieparameters. Het zijn **vaste getallen**. Zij geven een bepaalde eigenschap van een populatie weer. Maar je kent ze niet.

Een populatieparameter zoals de populatieproportie  $\pi$  kan je proberen te schatten op basis van wat je in je steekproef hebt gezien. Je kan dat op verschillende manieren doen. Je kan één getal (één punt) aangeven, namelijk de proportie “successen” die jij in je steekproef vindt. Dat noem je puntschatten. Je kan ook werken met intervallen. Dat noem je intervalschatten.

## 3. Puntschatten

Je trekt een steekproef (EAS) van grootte  $n = 100$  uit de geboorten van het jaar 2003 en je hebt daar 53 meisjes tussen. Jij moet nu zeggen wat de proportie meisjes is van alle kinderen die in 2003 zijn geboren. Wat ga je doen?

Begin met de grootheden van het vraagstuk duidelijk te beschrijven, met de juiste notatie.

### Opdracht 1

---

De vraag gaat over de proportie meisjes van alle geboorten in 2003. Gaat die vraag dan over een populatieproportie of over een steekproefproportie? Wat is de juiste notatie?

Jij hebt 53 meisjes in je steekproef. Wat is dan je gevonden steekproefproportie? Hoe noteer je die? Hoort bij jouw uitkomst ook een kansmodel? Hoe heet dat en met welke notatie duid je dat aan?

Om op de vraag over de proportie meisjes van 2003 te antwoorden heb je eigenlijk niet veel keuze. De enige informatie die je hebt is je eigen steekproef. Daarin heb je 53 % meisjes gevonden. Wat jij nu kan zeggen is dat, volgens jou, er een proportie van 53 % meisjes geboren is in 2003. Wat zou je anders wel zeggen en hoe zou je dat verantwoorden?

Inderdaad, je hebt nu het beste gedaan wat je kon doen. Maar toch geeft dat niet zo'n goed gevoel. Waarom? Wat heeft je medeleerling, die uit dezelfde populatie ook een steekproef van grootte  $n = 100$  heeft getrokken, gevonden? En als je niet hetzelfde hebt, wie heeft dan “gelijk”?



In de verklarende statistiek kan je “je gelijk” niet halen op basis van een uitkomst. Je moet gaan kijken naar het onderliggende kansmodel dat die uitkomst genereert.

Hier heb je gewerkt met een kansmodel  $\hat{P}$  dat nu eens een te grote en dan eens een te kleine proportie  $\hat{p}$  oplevert, maar waarbij je weet dat, bij heel veel herhalingen, de gevonden proporties  $\hat{p}$  “gemiddeld” op de populatieproportie  $\pi$  vallen want  $E(\hat{P}) = \pi$ . Maar bij een puntschatting kan je niet aangeven hoever je van de echte  $\pi$  terechtkomt en met welke kans. Om daar iets meer over te weten moet je van een punt overstappen op een interval.

*Nota.*

- Om de populatieproportie  $\pi$  te schatten werk je met het kansmodel  $\hat{P}$ . Dat noem je een schatter van  $\pi$ . Een toevallige uitkomst  $\hat{p}$  noem je een schatting.
- Als het gemiddelde van een schatter samenvalt met de populatieparameter die hij schat dan spreek je over een **onvertekende** schatter. Zo is  $\hat{P}$  een onvertekende schatter van  $\pi$  want  $E(\hat{P}) = \pi$ . En  $\bar{X}$  is een onvertekende schatter van  $\mu$  want  $E(\bar{X}) = \mu$ .

## 4. Intervalschatten

Bij een intervalschatting voor  $\pi$  moet je het kansmodel van de steekproefproportie  $\hat{P}$  kennen. Om te starten onderstel je dat de steekproef groot genoeg is zodat je de normale benadering kan gebruiken.

### 4.1. Hoe het werkt, in theorie

Je weet van vroeger dat elke normale met 95 % kans niet verder van zijn gemiddelde valt dan 1.96 standaardafwijkingen. Als de steekproefproportie  $\hat{P}$  zich (benaderend) gedraagt als een normale dan kan je voor  $\hat{P}$  schrijven dat:

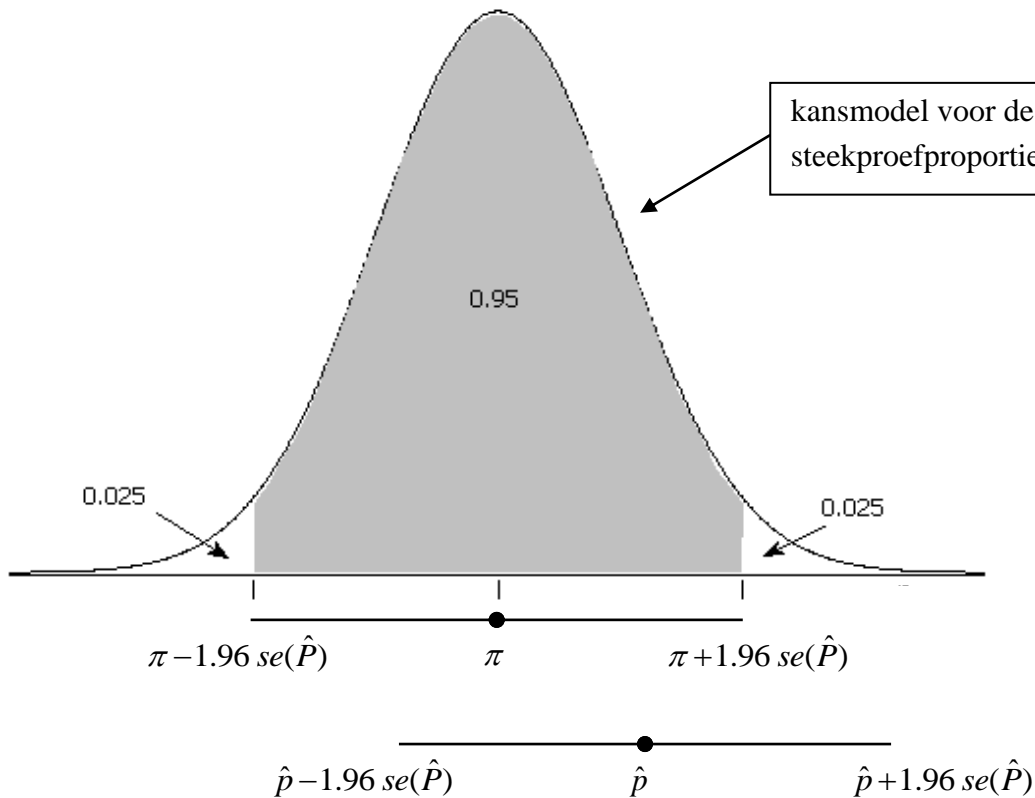
$$P(\pi - 1.96 \text{ se}(\hat{P}) \leq \hat{P} \leq \pi + 1.96 \text{ se}(\hat{P})) = 0.95$$

$\hat{P}$  komt met 95 % kans terecht in het interval  $[\pi - 1.96 \text{ se}(\hat{P}); \pi + 1.96 \text{ se}(\hat{P})] = \pi \pm 1.96 \text{ se}(\hat{P})$ .

Als jij nu een steekproef trekt en jouw steekproefproportie  $\hat{p}$  uitreken dan kan je rond je gevonden  $\hat{p}$  een interval maken van de vorm  $\hat{p} \pm 1.96 \text{ se}(\hat{P}) = [\hat{p} - 1.96 \text{ se}(\hat{P}); \hat{p} + 1.96 \text{ se}(\hat{P})]$ . Een belangrijke eigenschap van het interval dat je op die manier opstelt ontdek je in de volgende opdracht.

### Opdracht 2

Wanneer bevat het interval  $\hat{p} \pm 1.96 se(\hat{P})$  de populatieproportie  $\pi$ ? Wanneer niet? Waar moet je daarvoor met je steekproefproportie  $\hat{p}$  zijn terechtgekomen? Kijk goed naar figuur 1. Daar vind je het antwoord. Schrijf dit antwoord nu ook op in je eigen woorden.

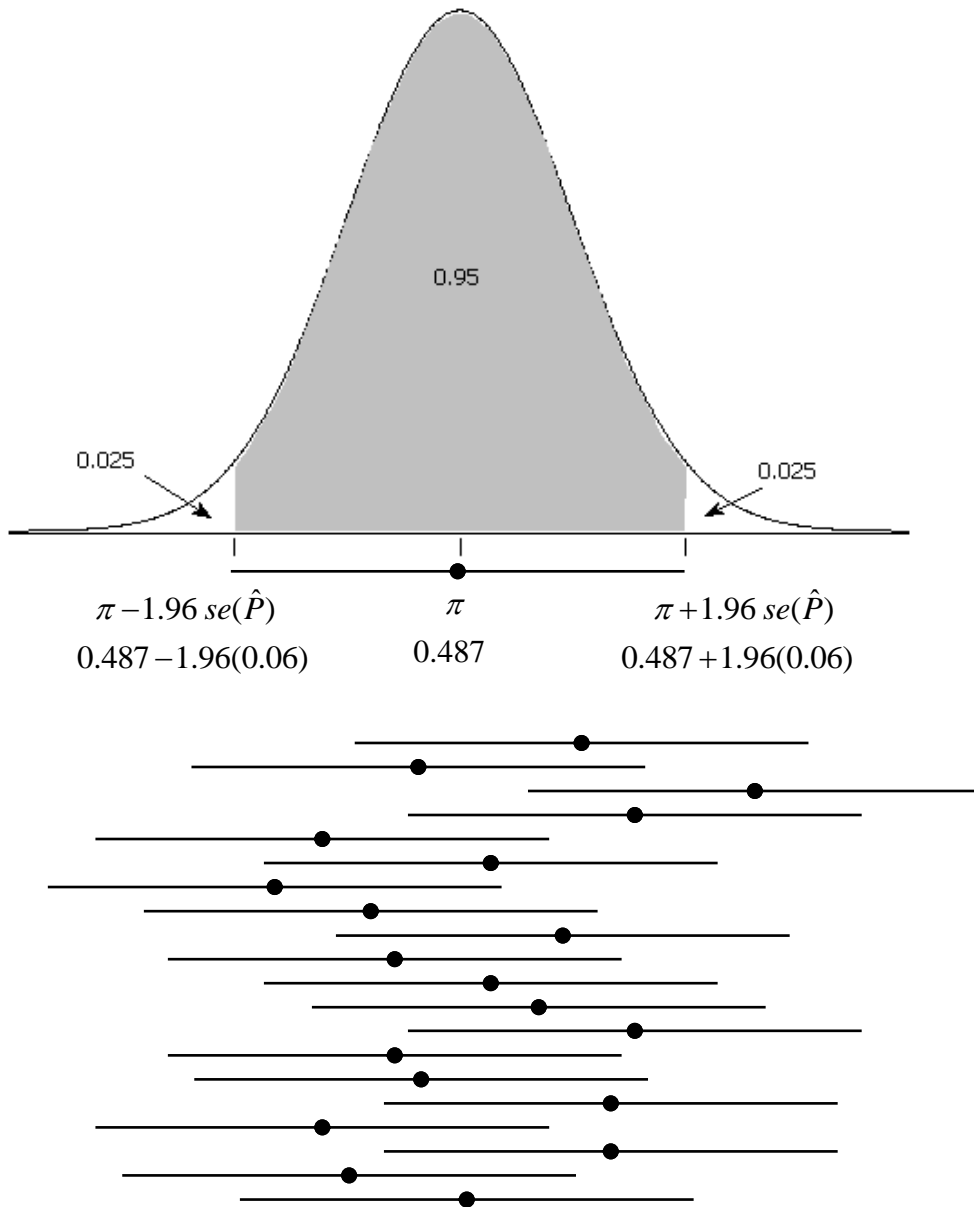


Figuur 1

#### 4.1.1. Een voorbeeld

Werk met een steekproef van grootte  $n = 70$  uit de populatie geboorten van het jaar 2003 en kijk naar de proportie meisjes daarin. Als je mag aannemen dat de populatieproportie gelijk is aan  $\pi = 0.487$  dan is  $se(\hat{P}) = \frac{\sqrt{\pi(1-\pi)}}{\sqrt{n}} = \frac{\sqrt{0.487(1-0.487)}}{\sqrt{70}} \cong 0.06$ .

Onderstel nu eens dat 20 leerlingen elk een steekproef van grootte  $n = 70$  trekken en noteren welke proportie  $\hat{p}$  zij daarin vinden. Daarna maken zij elk een interval van de vorm  $[\hat{p} - 1.96 se(\hat{P}); \hat{p} + 1.96 se(\hat{P})]$ . In dit voorbeeld is dit dus  $[\hat{p} - 1.96 (0.06); \hat{p} + 1.96 (0.06)]$ . Als je voor elk van die leerlingen hun interval tekent dan zou je een resultaat kunnen krijgen zoals figuur 2.



Figuur 2

	gevonden proportie $\hat{p}$	gevonden interval $[\hat{p} - 1.96(0.06); \hat{p} + 1.96(0.06)]$
leerling 1	$37/70 \cong 0.529$	[ 0.411 ; 0.647 ]
leerling 2	$32/70 \cong 0.457$	[ 0.339 ; 0.575 ]
leerling 3	$44/70 \cong 0.629$	[ 0.511 ; 0.747 ]
....	....	....
....	....	....

Tabel 1

De derde leerling is met zijn proportie in de rechterstaart terechtgekomen, voorbij het punt  $0.487 + 1.96(0.06)$ . Hij heeft daarna, zoals alle andere leerlingen, zijn interval opgesteld. Maar dat interval bevat de echte populatieproportie  $\pi$  niet want  $0.487$  ligt niet in  $[0.511; 0.747]$ . Bij alle andere leerlingen ligt  $0.487$  wel in het door hen opgestelde interval. Zij hebben allemaal een interval dat de populatieproportie  $\pi$  wel bevat.

Figuur 2 is een grafische voorstelling van wat je “in the long run” verwacht. Als je miljoenen dergelijke intervallen zou opstellen dan zou ongeveer 95 % van die intervallen de populatieproportie  $\pi$  bevatten en 5 % niet.

#### 4.1.2. Een kansmodel voor intervallen

Figuur 2 toont dat elke leerling een eigen uitkomst heeft na het trekken van zijn steekproef. Die uitkomst is “een interval”.

Nu kan je terug de basisvraag stellen. Wat zal er gebeuren als ik een steekproef zal trekken en ook zo’n interval zal maken? Welk interval zal ik dan hebben?

Je weet dat het enige goede antwoord op zo’n vraag gegeven wordt door een kansmodel. Hier is dat een model dat je vertelt welke intervallen je allemaal kan uitkomen en met welke kansen.

Als je een toevallig gevonden interval van een leerling hebt voorgesteld door  $[\hat{p} - 1.96(0.06); \hat{p} + 1.96(0.06)]$  dan is het nogal logisch dat je het kansmodel voor die intervallen met hoofdletters opschrijft, namelijk als  $[\hat{P} - 1.96(0.06); \hat{P} + 1.96(0.06)]$ .



Denk goed na over dit kansmodel, want het is de eerste keer in je leven dat je zoiets tegenkomt. Je bent ondertussen al gewoon geraakt om te werken met modellen die zeggen op welke manier **getallen** tot jou komen. Maar nu heb je een model dat zegt op welke manier **intervallen** tot jou komen.

Het model  $[\hat{P} - 1.96(0.06); \hat{P} + 1.96(0.06)]$  genereert intervallen waarbij je 95 % kans hebt dat er een interval tot jou zal komen dat de populatieproportie  $\pi$  bevat. Dat komt omdat  $\hat{P}$  zelf met kans 95 % terechtkomt in  $[\pi - 1.96(0.06); \pi + 1.96(0.06)]$ . Voor elke waarde  $\hat{p}$  die daar terechtkomt kan je het bijhorende interval  $[\hat{p} - 1.96(0.06); \hat{p} + 1.96(0.06)]$  maken. Dat is dan een interval dat  $\pi$  bevat, dat heb je op figuur 1 geleerd.

#### 4.2. Hoe het werkt, in de praktijk

Tot nu toe heb je gezien hoe je na het trekken van een steekproef een interval van de vorm  $[\hat{p} - 1.96(0.06); \hat{p} + 1.96(0.06)]$  kon maken. Dat is niet moeilijk want na het trekken van de steekproef ken je jouw geobserveerde proportie  $\hat{p}$ . Bij de eerste leerling was dat  $\hat{p} = 0.529$  en zo kwam die aan het interval  $[\hat{p} - 1.96(0.06); \hat{p} + 1.96(0.06)] = [0.411; 0.647]$ . Dat is een gekend interval, met gekend beginpunt en gekend eindpunt.



Vanwaar het getal 1.96 komt bij die berekening is duidelijk. Dat is het kritisch punt van de standaard normale verdeling wanneer je in het centrum een kans van 95 % wil hebben. Maar waar komt die 0.06 vandaan? Dat is de standaardfout van de steekproefproportie en de formule daarvoor is

$$se(\hat{P}) = \frac{\sqrt{\pi(1-\pi)}}{\sqrt{n}}$$

Als je met een steekproef van grootte  $n = 70$  werkt dan kan je de noemer al invullen. Maar wat doe je met de teller? Daar staat de populatieproportie  $\pi$  in. Daarvoor heb je tot nu toe 0.487 ingevuld, maar eigenlijk mag dat niet. Als je de populatieproportie  $\pi$  aan het schatten bent dan ken je ze niet. Wat moet je dan invullen voor die teller?

De oplossing bestaat erin dat je de populatieproportie  $\pi$  in de formule van  $se(\hat{P})$  vervangt door de proportie  $\hat{p}$  die jij in je steekproef hebt gevonden.

### 4.2.1. Een voorbeeld

De eerste leerling had een proportie  $\hat{p} = 0.529$ . Zij kent de populatieproportie  $\pi$  niet en kan dus het interval

$$\left[ \hat{p} - 1.96 \frac{\sqrt{\pi(1-\pi)}}{\sqrt{n}} ; \hat{p} + 1.96 \frac{\sqrt{\pi(1-\pi)}}{\sqrt{n}} \right]$$

niet uitrekenen. Maar als zij  $\pi$  vervangt door  $\hat{p}$  dan krijgt zij het interval  $\left[ \hat{p} - 1.96 \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} ; \hat{p} + 1.96 \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} \right]$ . Dat kan zij wel volledig uitrekenen, daarvoor moet zij  $\pi$  niet kennen. Voor haar levert dat [ 0.412 ; 0.646 ] .

Een interval dat op deze manier verkregen wordt is **een 95 % betrouwbaarheidsinterval voor  $\pi$** .

De GRM gebruikt deze “praktijk”-formule om betrouwbaarheidsintervallen uit te rekenen. Dat gaat als volgt.

Druk [STAT], loop naar TEST en dan naar beneden naar 1-PropZInt en druk [ENTER]. Vul het scherm in zoals

EDIT CALC TESTS 5:1-PropZTest... 6:2-PropZTest... 7:ZInterval... 8:TInterval... 9:2-SampZInt... 0:2-SampTInt... 1:1-PropZInt...	1-PropZInt x:37 n:70 C-Level:.95 Calculate	1-PropZInt (.41163,.64551) P=.5285714286 n=70
------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------	--------------------------------------------------------

aangegeven, loop naar Calculate en druk [ENTER]. Je kan nu de gevonden  $\hat{p}$  en het 95 % betrouwbaarheidsinterval voor  $\pi$  gewoon aflezen.

De resultaten van enkele andere leerlingen staan in tabel 2.

	gevonden proportie $\hat{p}$	gevonden interval $\left[ \hat{p} - 1.96 \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} ; \hat{p} + 1.96 \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} \right]$
leerling 1	$37 / 70 \cong 0.529$	[ 0.412 ; 0.646 ]
leerling 2	$32 / 70 \cong 0.457$	[ 0.340 ; 0.574 ]
leerling 3	$44 / 70 \cong 0.629$	[ 0.515 ; 0.742 ]
....	....	....

Tabel 2

## 4.2.2. Een kansmodel voor betrouwbaarheidsintervallen

De intervallen die de leerlingen hebben opgesteld zijn van de vorm

$$\left[ \hat{p} - 1.96 \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} ; \hat{p} + 1.96 \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} \right].$$

Het kansmodel dat deze intervallen genereert stel je voor met hoofdletters. Dat wordt dus

$$\left[ \hat{P} - 1.96 \frac{\sqrt{\hat{P}(1-\hat{P})}}{\sqrt{n}} ; \hat{P} + 1.96 \frac{\sqrt{\hat{P}(1-\hat{P})}}{\sqrt{n}} \right].$$

Dit kansmodel genereert uitkomsten die intervallen zijn. Zij worden betrouwbaarheidsintervallen genoemd. Als je met dit kansmodel heel veel intervallen zou maken dan zullen ongeveer 95 % van die intervallen de populatieproportie  $\pi$  bevatten en 5 % van die intervallen zullen dat niet doen.

In het kansmodel voor betrouwbaarheidsintervallen  $\left[ \hat{P} - 1.96 \frac{\sqrt{\hat{P}(1-\hat{P})}}{\sqrt{n}} ; \hat{P} + 1.96 \frac{\sqrt{\hat{P}(1-\hat{P})}}{\sqrt{n}} \right]$  zie je het getal 1.96 staan. Dat wijst erop dat je met de normale benadering werkt.

Om de normale benadering te mogen gebruiken bestaat er een criterium en in dat criterium komt de populatieproportie  $\pi$  voor. Die ken je niet. Ook hier bestaat de oplossing erin dat je  $\pi$  vervangt door de proportie  $\hat{p}$  die jij in je steekproef vindt. Het aangepaste criterium zegt dat er moet voldaan zijn

aan  $\begin{cases} n \hat{p} \geq 15 \\ n(1-\hat{p}) \geq 15 \end{cases}$ . Bemerk dat  $n\hat{p}$  = aantal successen in je steekproef. Inderdaad,

$$n \hat{p} = n \frac{\text{aantal successen}}{\text{steekproefgrootte}} = n \frac{\text{aantal successen}}{n} = \text{aantal successen}.$$

En  $n(1-\hat{p})$  is niets anders dan het aantal mislukkingen in je steekproef.

Je moet in je steekproef dus minstens 15 successen en 15 mislukkingen hebben om de normale benadering te mogen gebruiken

Wat moet je doen als er niet voldaan is aan dat criterium voor de normale benadering? Dan maak je gebruik van een recent ontdekte eigenschap. Die zegt dat je 2 moet optellen bij het gevonden aantal successen en 2 optellen bij het gevonden aantal mislukkingen. Dat betekent dus ook dat je 4 moet optellen bij de steekproefgrootte  $n$ . En deze nieuwe getallen gebruik je in hetzelfde model als vroeger, met het kritisch punt uit de normale verdeling.

- Als je bij een steekproef van grootte 30 slechts 3 successen gevonden hebt dan werk je met  $x=5$  voor het aantal successen en  $n=34$  voor de steekproefgrootte.

Zo krijg je  $[0.03 ; 0.27]$  als 95 % betrouwbaarheidsinterval voor  $\pi$ . De steekproefproportie  $\hat{p}=0.147$  die je GRM in zo'n geval aangeeft is dan niet meer de succesproportie  $\hat{p}=3/30=0.10$  die jij echt in je steekproef hebt gevonden.

```
1-PropZInt
x:5
n:34
C-Level:.95
```

```
1-PropZInt
(.02801,.2661)
P=.1470588235
n=34
```

Kansmodel voor een 95 % betrouwbaarheidsinterval  
voor de populatieproportie  $\pi$

Als ik in mijn steekproef minstens 15 successen en minstens 15 mislukkingen vind, dan werk ik met het kansmodel

$$\left[ \hat{p} - 1.96 \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} ; \hat{p} + 1.96 \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} \right] \quad (*)$$

In het andere geval tel ik 2 op bij het gevonden aantal successen en tel ik ook 2 op bij het gevonden aantal mislukkingen. Bij de steekproefgrootte  $n$  tel ik dus 4 op. Deze nieuwe getallen gebruik ik dan in (\*).

Tabel 3

Nota.

- Bij het rapporteren van je betrouwbaarheidsinterval voor  $\pi$  ga je als volgt te werk. Als je toevallig een steekproef getrokken hebt waarvoor de uitgerekenede formule een negatief resultaat geeft voor de ondergrens van je betrouwbaarheidsinterval dan maak je daar een nul van. Een proportie kan immers nooit negatief zijn. Eenzelfde opmerking geldt als je voor de bovengrens een getal vindt dat groter is dan 1. Daar maak je een 1 van.
- Je GRM doet geen controle op het criterium voor de normale benadering  $\begin{cases} n \hat{p} \geq 15 \\ n(1-\hat{p}) \geq 15 \end{cases}$ .  
Dat moet je dus vooraf zelf doen en eventueel aangepaste getallen in je GRM inbrengen.

### 4.2.3. Een experiment

Als je kijkt naar de intervallen die de leerlingen gevonden hebben dan zijn die van de vorm  $\left[ \hat{p} - 1.96 \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} ; \hat{p} + 1.96 \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} \right]$ . Sommige van deze intervallen bevatten de populatieproportie  $\pi$  maar sommige ook niet. Je verwacht dat, “in the long run” er ongeveer 95 op de 100 intervallen  $\pi$  wel zullen bevatten. Je kan dat eens experimenteel nagaan op een eenvoudig voorbeeld.

Om te weten of de gevonden betrouwbaarheidsintervallen de echte populatieproportie  $\pi$  te pakken hebben moet je natuurlijk weten wat  $\pi$  is. Je start dus met een 0–1 populatie die je kent en dan doe je alsof je ze niet kent zodat je  $\frac{\sqrt{\pi(1-\pi)}}{\sqrt{n}}$  moet vervangen door  $\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}$ .

Werk nu met een 0–1 populatie  $X$  die een 60 % succeskans heeft. Die ziet er zo uit.

$x$	0	1
$P(X=x)$	0.4	0.6

Kansmodel voor een 0–1 populatie  $X$  met succeskans  $\pi = 0.6$

Tabel 4

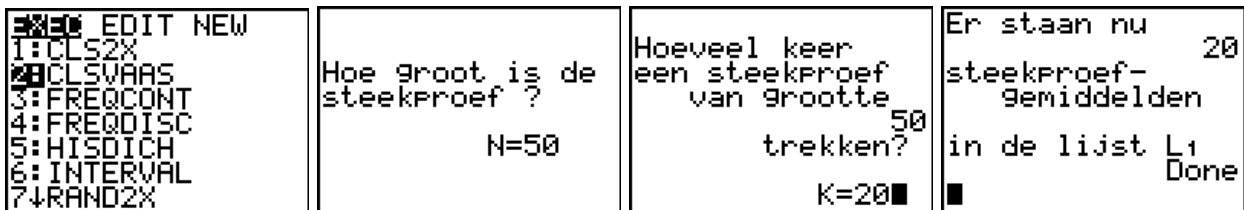
Op basis van een steekproef van grootte  $n = 50$  stel je nu een 95 % betrouwbaarheidsinterval op voor  $\pi$ . Je moet hierbij gebruik maken van je gevonden steekproefproportie  $\hat{p}$  want je onderstelt dat je  $\pi$  niet kent. Dit betekent dat je het interval  $\left[ \hat{p} - 1.96 \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}; \hat{p} + 1.96 \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} \right]$  moet berekenen waarbij je bovendien  $\hat{p}$  en  $n$  moet aanpassen als je niet minstens 15 successen en 15 mislukkingen in je steekproef hebt gevonden.

### Opdracht 3

Werk samen in groepjes van 5 leerlingen. Elke leerling werkt eerst afzonderlijk en daarna breng je alle resultaten samen.

Trek 20 keer een steekproef van grootte  $n = 50$  en stel telkens een 95 % betrouwbaarheidsinterval op voor  $\pi$ . Gebruik daarvoor je GRM op de volgende manier.

Vertaal tabel 4 in een vaasmodel en zet dit vaasmodel in de lijst [L5]. Druk dan **[PRGM]** kies CLSVAAS en druk 2 keer **[ENTER]**. Vul in dat je met een steekproef van grootte  $n = 50$  wil werken en dat je 20 keer zo'n steekproef wil trekken. Het programma heeft ongeveer 1 minuut nodig.




Bemerk dat het programma zegt dat er 20 steekproefgemiddelden in de lijst [L1] staan. Jij trekt hier steekproeven uit een 0–1 populatie en dan krijgt het steekproefgemiddelde de naam steekproefproportie. Je krijgt dus als resultaat in [L1] de proportie successen  $\hat{p}$  die in elk van die 20 steekproeven gevonden zijn. Daarmee moet je nu 20 keer een betrouwbaarheidsinterval opstellen. Dat doe je als volgt.

Druk **[PRGM]**, kies INTERVAL en druk 2 keer **[ENTER]**. Tik 95 wanneer er naar de betrouwbaarheid in percent wordt gevraagd en druk **[ENTER]**. Vul in dat je telkens een steekproef van grootte  $n = 50$  hebt getrokken en druk **[ENTER]**.

Betrouwbaarheid in Percent bv B=95 of B=80 B=95■	Hoe groot was de steekproef? N=50■	De populatie- proportie $\pi$ Tik de waarde van $\pi$ met $0 < \pi < 1$ $\pi$ onbekend = 2 en druk ENTER ?0.6■
-----------------------------------------------------------	------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------

Tik 0.6 om de waarde van  $\pi$  in te brengen en druk **[ENTER]**. Tik 1(=ja) om de intervallen te tekenen en druk **[ENTER]**. Bekijk de intervallen en druk dan **[ENTER]**. Je GRM zegt nu hoeveel van die 20 intervallen  $\pi = 0.6$  bevatten. In dit voorbeeld zijn er dat 19 van de 20. Hoeveel zijn er dat bij jou? Schrijf jouw getal in tabel 5.

Wil je 20 betrouwbaarheids intervallen tekenen ? Tik 0=neen 1=ja en druk ENTER ?1	Vervolgs=ENTER 	19 van in totaal 20 betrouwbaarheids intervallen bevatten $\pi$ . Druk ENTER	Er staan nu 20 betrouwbaarheids intervallen in [ L2 , L3 ] Done
--------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------

De 20 intervallen staan in [L2] (ondergrens) en [L3] (bovengrens). Je kan die daar gaan bekijken met **[STAT]** en 1:Edit....

Vat nu de resultaten van alle leerlingen van je groep samen.

	aantal betrouwbaarheidsintervallen die $\pi = 0.6$ wel bevatten	aantal betrouwbaarheidsintervallen die $\pi = 0.6$ niet bevatten
leerling 1		
leerling 2		
leerling 3		
leerling 4		
leerling 5		
TOTAAL		

Tabel 5

Hoeveel percent van alle opgestelde intervallen bevat de populatieproportie?  
“In the long run” zou dat ongeveer 95 % moeten zijn.

### 4.3. Observeerbaarheid

Uiteindelijk wil je na het trekken van je steekproef niet terecht komen op een interval waar nog een onbekende  $\pi$  in staat. De succesproportie  $\pi$  in de populatie kan je nooit observeren, de succesproportie  $\hat{p}$  in je getrokken steekproef zie je wel.

Dat je in de praktijk overstapt van  $\frac{\sqrt{\pi(1-\pi)}}{\sqrt{n}}$  op  $\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}$  heeft dus gewoon te maken met “observeerbaarheid”.

Bij elke grootheid kan je kijken of zij aan het toeval onderhevig is en of zij observeerbaar is. Je kan daarbij de volgende omschrijvingen gebruiken.

**Observeerbaar** : als je na het trekken van een steekproef de steekproefresultaten in je grootheid invult dan zal er in je grootheid geen enkel onbekend element meer staan.

**Aan het toeval onderhevig** : als je de volgende keer juist dezelfde procedure toepast (uit dezelfde populatie een steekproef van dezelfde grootte trekt), dan zal je voor je grootheid heel waarschijnlijk een andere waarde vinden dan vorige keer.

Grootheid	Observeerbaar	Aan het toeval onderhevig	Verklaring
$\pi$	Neen	Neen	De populatieproportie is een vast getal dat je niet kent.
$\hat{p}$	Ja	Ja	De steekproefproportie is aan het toeval onderhevig maar na de steekproef ken jij jouw proportie.
$n$	Ja	Neen	De steekproefgrootte wordt vooraf vastgelegd en is gekend.
$\left[ \hat{P} - 1.96 \frac{\sqrt{\pi(1-\pi)}}{\sqrt{n}} ; \hat{P} + 1.96 \frac{\sqrt{\pi(1-\pi)}}{\sqrt{n}} \right]$	Neen	Ja	Na de steekproef kan je jouw waarde $\hat{p}$ van $\hat{P}$ invullen maar $\pi$ blijft onbekend.
$\left[ \hat{P} - 1.96 \frac{\sqrt{\hat{P}(1-\hat{P})}}{\sqrt{n}} ; \hat{P} + 1.96 \frac{\sqrt{\hat{P}(1-\hat{P})}}{\sqrt{n}} \right]$	Ja	Ja	Na de steekproef kan je jouw waarde $\hat{p}$ van $\hat{P}$ invullen en dan ken je het interval.

Tabel 6

## 5. Hoe werken betrouwbaarheidsintervallen?

In de verklarende statistiek is er iets wat jij volledig onder controle hebt en er is ook iets waar jij helemaal geen vat op hebt. Het eerste is het model en het tweede is de uitkomst. *Je leert statistiek om te weten met welk model je moet werken en om dit model te kunnen aanpassen aan je eigen eisen.* Daarna gebruik je dat model. Op basis van jouw toevallige steekproefresultaten heb jij dan een toevallige uitkomst van jouw model gevonden. Wat die uitkomst is heb jij niet in handen. Dat is nu eenmaal zo, want zo werkt statistiek.

### 5.1. Het Rad van Fortuin

Om goed het verschil tussen “het kansmodel voor betrouwbaarheidsintervallen” en “jouw betrouwbaarheidsinterval” te zien, maken we een vergelijking met een rad van fortuin.

Om te beginnen mag jij aan de quizmaster vragen om het rad van fortuin in een aantal gelijke sectoren te verdelen (bijvoorbeeld in 100 sectoren) en om in elke sector een kaartje te hangen. Op dat kaartje moet een interval staan met begin- en eindpunt tussen nul en één.

Verder mag jij zeggen hoeveel intervallen de echte populatieproportie  $\pi$  moeten bevatten. Als jij graag 95 kaartjes hebt met intervallen die  $\pi$  bevatten dan neemt de quizmaster 100 kaartjes, 95 “goede” en 5 “slechte”, en hangt die in willekeurige volgorde in de 100 sectoren. Hij vertelt je niet welke de “goede” en welke de “slechte” kaartjes zijn. Daarna verdwijnt hij ... voorgoed.

Nu ben jij aan de beurt.

VOORALEER je aan het rad draait kan je zeggen dat je een kans van 95 % hebt om een interval te vinden dat de echte  $\pi$  bevat. Inderdaad, op 95 van de 100 kaartjes staat een “goed” interval en op de andere 5 kaartjes staat een “slecht” interval.

Dan draai je aan het rad en na een tijdje stopt het bij één bepaalde sector. Het kaartje dat in die sector hangt moet jij nemen. Dat is jouw betrouwbaarheidsinterval.

Wat kan je nu zeggen, NADAT je aan het rad gedraaid hebt en jouw kaartje hebt genomen? Jij staat daar nu met dat kaartje in je hand, waarop een begin- en eindpunt van één interval staat.

Wat de echte populatieproportie  $\pi$  is, weet je niet. Je zal die ook nooit kennen want alleen de quizmaster kende die en hij is verdwenen. Wat je (hopelijk) wel weet is dat de populatieproportie  $\pi$  een vast getal is, niet observeerbaar maar helemaal niet aan het toeval onderhevig.

Het vaste interval dat nu op je kaartje staat bevat het vaste getal  $\pi$  wel of het bevat  $\pi$  niet. Daar kan je 100 % zeker van zijn. Maar je kan geen kansuitspraken meer maken over feiten die al gebeurd zijn.



Zeggen dat  $\pi$  met kans 95 % in jouw interval valt is fout. De populatieproportie  $\pi$  is een vast getal en dat “valt” niet ergens “met een of andere kans”.

Zeggen dat jouw betrouwbaarheidsinterval met kans 95 % de populatieproportie  $\pi$  bevat is al even fout. Jij hebt nu een vast interval en dat interval bevat  $\pi$  wel, of het bevat  $\pi$  niet.

## 5.2. Niet zeker, en dan?

Je hebt nu een zogenaamd 95 % betrouwbaarheidsinterval en wat je daar met 100 % zekerheid over weet is dat het ofwel “goed” is ofwel “slecht”. Frustrerend?

Eigenlijk moet je de dingen wat anders bekijken, in hun totale context. ACHTERAF heb jij inderdaad één interval, dat ofwel  $\pi$  bevat ofwel niet. Maar je kan tegelijk zeggen dat jij met een model gewerkt hebt waarbij jij VOORAF een kans van 95 % had om een “goed” interval te vinden. En nu “hoop” je dat jouw interval een “goed” interval is. Zekerheid heb je niet.

Wat doe je daarna?

Voor het vervolg van je studie ga je er vanuit dat je kan verder werken in de onderstelling dat  $\pi$  één van de waarden van jouw interval is. Voor jou zijn dat “aannemelijke waarden” voor  $\pi$ .

## 5.3. Betrouwbaarheid is ...

In elke taal zoekt men naar een goed woord om uit te leggen wat er gebeurt als je een betrouwbaarheidsinterval opstelt.

In het Engels spreekt men over een 95 % “*confidence interval*”. Leerling 1 zou in de Verenigde Staten leren dat zij moet zeggen: “op basis van mijn steekproef kan ik 95 % *confident* zijn dat bij de geboorten van het jaar 2003 de proportie meisjes ligt tussen 0.432 en 0.628”. Of ze zou ook mogen zeggen: “ik heb 95 % *confidence* dat het interval [ 0.432 ; 0.628 ] de proportie meisjes bij de geboorten van 2003 te pakken heeft”.

Op het centrale examen voor statistiek (in de USA) gebruiken veel leerlingen inderdaad één van beide zinnen als er een vraag is over betrouwbaarheidsintervallen. Zij schrijven zo’n zin letterlijk op, op zo’n manier dat je het gevoel krijgt dat ze die zin van buiten geleerd hebben. Misschien komt dit omdat zij niet weten wat er nu juist met het woord *confidence* bedoeld wordt. In hun dagelijkse spreektaal komt “*I have 95 % confidence that...*” niet voor.

Ook in het Nederlands is het een probleem.

Zeggen dat je een interval hebt gevonden dat met **95% betrouwbaarheid** (en dus NIET “met 95 % kans”) de populatieproportie  $\pi$  bevat blijft een moeilijke uitspraak. Je moet die uitspraak telkens dubbel interpreteren. De analogie met een rad van fortuin kan je daarbij helpen. Het begrip “95 % betrouwbaarheid” bestaat uit 2 delen: “95 %” en “betrouwbaarheid”.

1. De “95 %” verwijst naar het model waarmee je gewerkt hebt, of naar het rad van fortuin, waarbij je er VOORAF zelf hebt kunnen voor zorgen dat je 95 % kans had op een “goed” interval.
2. Het woord “betrouwbaarheid” verwijst ernaar dat je met dat rad van fortuin al gespeeld hebt. ACHTERAF zit je daar met jouw interval. Het kan een “goed” of een “slecht” zijn.

En dan werk je verder met wat volgens jou “aannemelijke waarden” voor de populatieproportie  $\pi$  zijn. De waarden waar je *vertrouwen* (= *confidence*) in hebt. Dat zijn de waarden die jij in jouw interval gevonden hebt. Iets anders heb je niet. Zekerheid ook niet.

Als voor het opstellen van een betrouwbaarheidsinterval aan jou gevraagd wordt om “**te werken met een betrouwbaarheid van 95 %**” dan wordt daarmee bedoeld dat jij moet werken met een MODEL dat met KANS 95 % een “goed” interval naar jou **zal** sturen. Een UITKOMST uit dat model noem je **daarna** een 95 % BETROUWBAARHEIDSINTERVAL.



## 6. Betrouwbaarheid en precisie

Wanneer je beslist om met een vooraf gekozen steekproefgrootte  $n$  te werken (bijvoorbeeld met  $n = 70$ ) dan ligt nog niet alles vast. Tot nu toe heb je alleen **95 %** betrouwbaarheidsintervallen bestudeerd, maar je kan ook met een andere betrouwbaarheid werken.

### 6.1. De betrouwbaarheid veranderen

Het model voor betrouwbaarheidsintervallen in tabel 3 werkt met

$$\left[ \hat{P} - 1.96 \frac{\sqrt{\hat{P}(1-\hat{P})}}{\sqrt{n}} ; \hat{P} + 1.96 \frac{\sqrt{\hat{P}(1-\hat{P})}}{\sqrt{n}} \right] \quad (*)$$

Dit model genereert intervallen die (gemiddeld) 95 keer op de 100 de populatieproportie  $\pi$  bevatten. Bij het opstellen van het model heb je de normale benadering gebruikt. Het getal 1.96 ken je heel goed uit de studie van normale kansmodellen. Je weet dat elk normaal model met kans 95 % niet verder dan 1.96 standaardafwijkingen van zijn gemiddelde valt. Voor de standaard normale  $Z$  schrijf je dat als  $P(-1.96 \leq Z \leq 1.96) = 0.95$ . Dit is de reden waarom je het getal **1.96** in (\*) ziet verschijnen. Daarom ook is (\*) een model voor een **95 %** betrouwbaarheidsinterval.

Als je nu met een systeem wil werken dat kleinere of grotere betrouwbaarheid oplevert, dan hoef je alleen maar je kennis over normale kansmodellen te gebruiken. Je weet bijvoorbeeld dat  $P(-1.645 \leq Z \leq 1.645) = 0.90$  en dus is

$$\left[ \hat{P} - 1.645 \frac{\sqrt{\hat{P}(1-\hat{P})}}{\sqrt{n}} ; \hat{P} + 1.645 \frac{\sqrt{\hat{P}(1-\hat{P})}}{\sqrt{n}} \right] \quad (**)$$

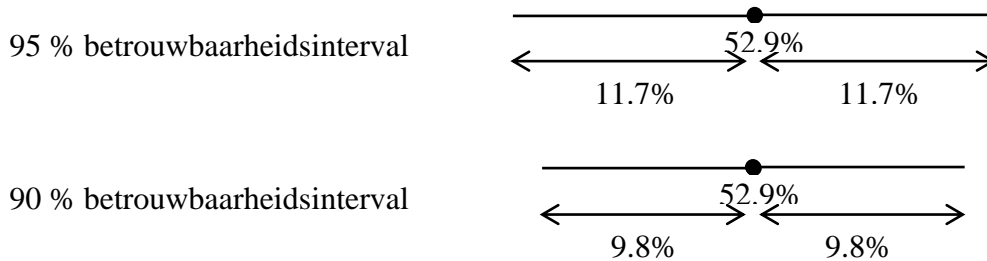
een model dat 90 % betrouwbaarheidsintervallen voor  $\pi$  genereert. Moeilijker is het niet.

### 6.2. Wat is een foutenmarge?

Leerling 1 heeft in haar steekproef van 70 geboorten 37 meisjes gevonden. Dat is een proportie  $\hat{p} \cong 0.529$ . Zij gebruikt het model (\*) en vindt  $[0.412 ; 0.646]$  voor haar 95 % betrouwbaarheidsinterval voor  $\pi$ . Je kan dat interval ook schrijven als  $0.529 \pm 0.117$ . In percenten uitgedrukt is dat  $52.9 \% \pm 11.7 \%$

Zij wil ook weten wat er gebeurt als zij met een 90 % betrouwbaarheid zou werken. Daarom vult zij haar gevonden proportie in in het model (\*\*) en krijgt een 90 % betrouwbaarheidsinterval dat eruit ziet als  $[0.430 ; 0.627]$ . Dat interval kan je schrijven als  $0.529 \pm 0.098$  of als  $52.9 \% \pm 9.8 \%$

<pre>1-PropZInt x:37 n:70 C-Level:.95 Calculate</pre>	<pre>1-PropZInt (.41163,.64551) P=.5285714286 n=70</pre>
<pre>1-PropZInt x:37 n:70 C-Level:.90 Calculate</pre>	<pre>1-PropZInt (.43043,.62671) P=.5285714286 n=70</pre>



Figuur 3

Je merkt dat, bij een gevonden proportie  $\hat{p}$ , een 90 % betrouwbaarheidsinterval korter is dan een 95 % betrouwbaarheidsinterval. Dat kon je vooraf weten. Bij (\*) gebruik je 1.96 en bij (\*\*) 1.645.

De halve lengte van een betrouwbaarheidsinterval, dat is dus het getal dat na  $\pm$  staat, noem je de **foutenmarge**. Hoe korter het interval hoe preciezer de informatie. Maar daarvoor betaal je een prijs. Je moet dan met een model werken waarbij de kans groter is op een “slecht” interval. Bij (\*\*) is die kans 10 % en bij (\*) is die maar 5 %.

Wat je zopas gezien hebt is algemeen waar.

Een grotere betrouwbaarheid gaat samen met een grotere foutenmarge. Je hebt dan minder precieze informatie (want een langer interval).

Omgekeerd, als je bereid bent om een model te gebruiken dat met minder kans een “goed” interval genereert (bijvoorbeeld met kans 90 % in plaats van met kans 95 %) dan krijg je in ruil meer precieze informatie (want een korter interval).

Je moet dus zelf de voor- en nadelen afwegen wanneer je beslist met welk model je wil werken. Daarvoor moet je de context van het onderzoek kennen.

### Opdracht 4

---

Wat is de foutenmarge van een 80 % betrouwbaarheidsinterval voor  $\pi$  bij 37 gevonden successen in een steekproef van grootte  $n = 70$ ? Verwacht je een preciezer (en dus korter) interval te vinden dan bij een 95 % betrouwbaarheidsinterval? Waarom? Zoek nu dat interval en zeg wat de foutenmarge is. Toon je redenering en je berekeningen.

### Opdracht 5

---

Met welk rad van fortuin werk je als je, zonder  $\pi$  te kennen, zelf zo'n rad zou moeten maken dat een 100 % betrouwbaarheidsinterval oplevert? Wat moet je dan op de kaartjes schrijven die je aan dat rad hangt? Wat gebeurt er in de statistiek als je "zekerheid" (100 % betrouwbaarheid) wil hebben?

### 6.3. Een voorbeeld uit de media

Een krant schrijft dat er bij 4 op de 10 van de 1301 ondervraagde gezinnen 's ochtends samen wordt ontbeten. Er staat ook bij dat het onderzoek een foutenmarge van 2.6 % heeft. Als je zoiets in een krant leest dan mag je ervan uitgaan dat zij met het model (\*) gewerkt hebben.

In dit onderzoek is de gevonden steekproefproportie  $\hat{p} = 0.4$  (4 op de 10 gezinnen) zodat een 95 % betrouwbaarheidsinterval voor  $\pi$  ( $\pi$  is de proportie van **alle** Vlaamse gezinnen die samen ontbijten) gelijk is aan

$$\hat{p} \pm 1.96 \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} = 0.4 \pm 1.96 \frac{\sqrt{0.4(1-0.4)}}{\sqrt{1301}} = 0.4 \pm 0.0266 \cong 40\% \pm 2.6\%$$



**Datum :** 13/01/2005

**Titel :** Vlaamse gezinnen streven gezond ontbijt na

Vlaamse gezinnen weten over het algemeen goed wat gezond ontbijten is en proberen daar zo consequent mogelijk naar te handelen.

Vooraf moeders eten en drinken wat zij denken dat bij een evenwichtig ontbijt hoort. Dat blijkt uit een onderzoek van de Gezinsbond over het ontbijtgedrag van hun leden.

In vier op de tien van de 1.301 ondervraagde gezinnen wordt samen ontbeten. Vaders ontbijten wel vaker in hun eentje. Oudere kinderen eten vaker alleen dan hun jongere broertjes of zusjes.

Vlaams minister van Volksgezondheid en Gezin Inge Vervotte benadrukt dat een gezellig ontbijt bijdraagt tot een positieve gezinscultuur. Ze nodigt de Gezinsbond uit om samen met andere betrokkenen concrete acties op te zetten om gezonde voeding nog meer te stimuleren.

Het onderzoek van de Gezinsbond heeft een foutenmarge van 2,6 procent. Hoger opgeleiden en lagere schoolkinderen zijn oververtegenwoordigd.

#### Opdracht 6

Sommige mensen die dit artikel lezen denken als volgt. In die studie waren er 40 % gezinnen die verklaarden dat zij samen ontbijten, maar de onderzoekers zijn er niet zeker van of er in de totale populatie van alle gezinnen ook juist 40 % zijn die samen ontbijten. Daarom hebben zij een foutenmarge bepaald. Dat is de maximale fout (te groot of te klein) die er kan opzitten. De populatieproportie is misschien niet exact 40 % maar zij ligt zeker tussen 37.4 % en 42.6 %.

Wat denk jij van deze uitspraak? Zeg ook welke conclusie jij zelf uit het bericht in die krant trekt. Doe dit met een zo nauwkeurig mogelijk woordgebruik.

## 7. Hoe groot moet de steekproef zijn?

Bij elk “echt” statistisch onderzoek is de vraag naar de grootte van de te trekken steekproef zeer belangrijk. Het gaat niet alleen over tijd en geld van de onderzoekers. Dikwijls zijn bij een onderzoek ook anderen betrokken, zoals patiënten bij de speurtocht naar nieuwe geneesmiddelen.

### 7.1. De theorie

Uit tabel 3 weet je dat  $\left[ \hat{p} - 1.96 \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}; \hat{p} + 1.96 \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} \right]$  het kansmodel is dat 95 % betrouwbaarheidsintervallen voor de populatieproportie  $\pi$  genereert.

Als jij nu een steekproef trekt en daarin een proportie  $\hat{p}$  successen hebt, dan stel jij jouw 95 % betrouwbaarheidsinterval op. Dat is dan  $\left[ \hat{p} - 1.96 \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}; \hat{p} + 1.96 \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} \right]$  of, wat hetzelfde is,  $\hat{p} \pm 1.96 \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}$ . Je kan natuurlijk ook  $\hat{p} \pm z \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}$  schrijven als je met een algemeen kritisch punt  $z$  uit de standaard normale wil werken.



De foutenmarge  $z \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}$  die je hier ziet bestaat uit 3 grootheden:

- **de betrouwbaarheid** die bepaalt wat  $z$  is. Wil je 95 % betrouwbaarheid dan is  $z = 1.96$ . Werk je met 90 % betrouwbaarheid dan is  $z = 1.645$ .
- **de gevonden steekproefproportie**  $\hat{p}$  die in de teller voorkomt
- **de steekproefgrootte**  $n$  die in de noemer voorkomt

Als jij zegt dat je een 95 % betrouwbaarheidsinterval wil maken met een foutenmarge van 3 % dan hoef je dat alleen maar in te vullen in

$$\text{foutenmarge} = z \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}$$

of in

$$\sqrt{n} = \frac{z \sqrt{\hat{p}(1-\hat{p})}}{\text{foutenmarge}} \rightarrow n = \frac{z^2 \hat{p}(1-\hat{p})}{(\text{foutenmarge})^2} \quad (***)$$

Voor jouw vraag is de grootte  $n$  van de te trekken steekproef gelijk aan

$$n = \frac{z^2 \hat{p}(1-\hat{p})}{(\text{foutenmarge})^2} \rightarrow n = \frac{1.96^2 \hat{p}(1-\hat{p})}{(0.03)^2}$$

Zie je het probleem hier? Je probeert te bepalen hoe groot de steekproef moet zijn die je **zal** trekken. En om te weten hoe groot die moet zijn zou je al de resultaten van die steekproef moeten hebben, want je hebt  $\hat{p}$  nodig! Dat is de wereld op zijn kop.

## 7.2. De praktijk

In de praktijk moet je proberen een goede (benaderende) waarde te gebruiken voor de succesproportie  $\hat{p}$  die je verwacht te zullen vinden in je steekproef. Dat kan op 2 manieren. Ofwel gebruik je externe informatie, ofwel speel je op veilig.

### 7.2.1. Externe informatie

Externe informatie kan je halen uit vorige onderzoeken, uit de literatuur, uit pilootstudies enz. Als deze informatie betrouwbaar is en als je geen grote veranderingen verwacht tegenover wat er vroeger gevonden werd, dan kan je die externe informatie gebruiken om te bepalen met welke steekproefgrootte  $n$  je nu moet werken.

Bij externe informatie zet je een dubbele stap. Externe informatie geeft je immers een (benaderend) idee over de succesproportie **in de populatie**, dus over  $\pi$ . En dan redeneer je als volgt. Als ik uit zo'n populatie trek, dan verwacht ik in mijn steekproef een succesproportie  $\hat{p}$  die niet te ver afligt van  $\pi$ . En dan gebruik je wat je weet over  $\pi$  als een benaderende waarde voor  $\hat{p}$  in (\*\*\*)

Op [http://aps.vlaanderen.be/statistiek/publicaties/stat\\_Publicaties\\_vrind2006.htm](http://aps.vlaanderen.be/statistiek/publicaties/stat_Publicaties_vrind2006.htm) vind je heel wat informatie over Vlaanderen in de publicatie "VRIND 2006". Van alle leerlingen in de tweede en derde graad secundair onderwijs zitten er ongeveer 31 % in het TSO. Dat is een cijfer voor heel Vlaanderen voor de periode 1994-2005. Jouw vraag is nu: "hoeveel percent leerlingen zitten er dit jaar in mijn provincie in het TSO"? Je wil dit snel weten, op basis van een steekproef en zonder te moeten wachten tot de volledige inschrijvingslijsten van alle scholen gepubliceerd zijn.

Als je mag onderstellen dat jouw provincie niet veel verschilt van de algemene trend in Vlaanderen dan betekent dit dat er daar globaal ook ongeveer 31 % leerlingen voor TSO kiezen. Dit is dus een vermoeden over de populatieproportie  $\pi$  in jouw studie (jouw populatie bestaat uit alle leerlingen 2<sup>e</sup> + 3<sup>e</sup> graad in jouw provincie).

Als in jouw provincie (bij benadering) 31 % van de leerlingen voor TSO kiezen dan verwacht je dat je ook ongeveer 31 % "successen" (= leerlingen die voor TSO kiezen) in je steekproef zal vinden. In de formule (\*\*\*) vervang je dan  $\hat{p}$  door 31 %.

Om een 95 % betrouwbaarheidsinterval op te stellen voor de proportie leerlingen die dit jaar in jouw provincie TSO volgen heb je de volgende steekproefgrootte nodig als je wil werken met een foutenmarge van "ongeveer" 3 %. Uit (\*\*\*) volgt dat je

$$n = \frac{1.96^2 \hat{p}(1-\hat{p})}{(0.03)^2} \cong \frac{1.96^2 (0.31)(1-0.31)}{(0.03)^2} \cong 914$$

leerlingen moet trekken.

*Nota*

- Bij het bepalen van de steekproefgrootte rond je af naar boven.

## 7.2.2. Speel op veilig

Soms heb je geen externe informatie of is die niet betrouwbaar of niet stabiel. Hoe moet je dan iets over de te verwachten  $\hat{p}$  te weten komen?

In de formule (\*\*\*) om  $n$  te bepalen staat niet  $\hat{p}$  maar wel  $\hat{p}(1-\hat{p})$ . Wat  $\hat{p}$  zal zijn weet je niet, maar je weet wel dat elke proportie een getal is tussen 0 en 1. Je zou dus eens kunnen onderzoeken wat de mogelijke waarden zijn van  $\hat{p}(1-\hat{p})$  terwijl je  $\hat{p}$  laat lopen van 0 tot 1. Met je GRM is dat geen probleem.

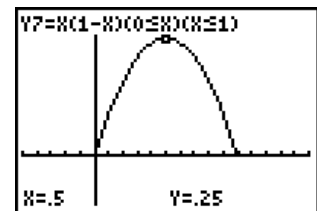
Indien nodig druk je eerst  $\boxed{2nd}$   $\boxed{DRAW}$  en  $1:ClrDraw$ .

Pas  $\boxed{WINDOW}$  aan zoals aangegeven. Druk dan  $\boxed{Y=}$ . Zorg ervoor dat alle functies af staan. Loop naar een functie die nog vrij is, bijvoorbeeld  $Y_7=$  en tik dan  $\boxed{X,T,\theta,n}$   $\boxed{1 - \boxed{X,T,\theta,n}}$  samen met de grenzen waar je met  $\boxed{2nd}$   $\boxed{TEST}$  de gepaste ongelijkheidstekens invult. Druk dan  $\boxed{TRACE}$ .

Terwijl  $\hat{p}$  van 0 naar 1 loopt zie je dat  $\hat{p}(1-\hat{p})$  een stuk van een parabool beschrijft. De maximale waarde van  $\hat{p}(1-\hat{p})$  is 0.25 en die wordt bereikt voor  $\hat{p} = 0.5$ .

```
WINDOW
Xmin=-.5
Xmax=1.5
Xscl=.1
Ymin=-.1
Ymax=.3
Yscl=1
Xres=1
```

```
Plot1 Plot2 Plot3
Y6=X^2*normalPd
f(X)
Y7=X(1-X)(0≤X)(<
X≤1)
Y8=
Y9=
Y0=
```



De redenering is nu als volgt. Voor een 95 % betrouwbaarheidsinterval met

een foutenmarge van 0.03 heb je een steekproefgrootte  $n$  nodig die gelijk is aan  $n = \frac{1.96^2 \hat{p}(1-\hat{p})}{(0.03)^2}$ .

Wat  $\hat{p}(1-\hat{p})$  is weet je niet maar het kan nooit groter zijn dan 0.25. Dat gebeurt voor  $\hat{p} = 0.5$ .

Vervang dus  $\hat{p}$  door 0.5 en werk met  $n = \frac{1.96^2 (0.5)(1-0.5)}{(0.03)^2}$ . Je steekproef is dan zeker groot genoeg en je foutenmarge is zeker nooit meer dan wat er gevraagd was.

Bemerk dat je een steekproef van  $n = \frac{1.96^2 (0.5)(1-0.5)}{(0.03)^2} \cong 1068$  leerlingen uit je provincie moet trekken als je de externe informatie niet vertrouwt en dus “op veilig” speelt.

## 7.2.3. Exit polls bij nipte verkiezingen

In 2000 namen Hillary Clinton en Rick Lazio het tegen elkaar op om U.S. senator te worden in de staat New York. Soms zijn dergelijke verkiezingen vooraf al duidelijk en wint de favoriete kandidaat met een voorsprong van meer dan 40 % op zijn tegenstrever. Maar bij Hillary was dat helemaal niet zo duidelijk en de peilingen vooraf voorspelden voor haar een nipte overwinning, zoiets rond de 55 % van de stemmen.

Op de dag van de verkiezingen wil je op basis van een exit poll snel kunnen zeggen wie er gewonnen heeft. Daarom probeer je met een zo klein mogelijke steekproef te werken, want dat gaat het snelst. Je stelt voor om gebruik te maken van de externe informatie (Hillary haalt waarschijnlijk 55 %). Je stelt ook voor om de foutenmarge wat groter te nemen, zoals 5 %.

Beide voorstellen zijn hier niet aan te raden.

## Opdracht 7

Kan je, vooraleer verder te lezen, redenen bedenken waarom beide voorstellen hier niet aan te raden zijn? Welke? Denk aan wat je in de praktijk verneemt op radio en TV wanneer men probeert te voorspellen wie er zal winnen bij nipte verschillen tussen partijen (of politici).

- *De vermelde 55 % is twijfelachtig. Dat getal schommelde ook nogal wat in de weken die aan de verkiezingen voorafgingen, afhankelijk van de krant of de TV-zender die de peiling had laten uitvoeren. Soms was het 52 % en een andere keer 56 %. Je kan dus maar beter “op veilig” spelen en 0.50 invullen voor  $\hat{p}$ .*
- *Een foutenmarge van 5 % is in deze situatie af te raden. Als je bijvoorbeeld bij je exit poll een overwinning voor Hillary vindt van 53 % dan zit je met een betrouwbaarheidsinterval dat er uitziet als [ 48 % ; 58 % ]. Je “aannemelijke waarden” vertellen je dan dat Hillary zowel kan verloren hebben als gewonnen. Dat helpt niet veel. Werk dus liever met een kleine foutenmarge, zoals 2 %*

Voor een 95 % betrouwbaarheidsinterval met een foutenmarge van ongeveer 2 % moet je werken

met een steekproef van grootte  $n = \frac{1.96^2 (0.5)(1-0.5)}{(0.02)^2} = 2401$ .

Het onderzoeksbureau dat op 7 november 2000 (de dag van de verkiezingen) de exit poll uitvoerde werkte met 2232 kiezers waarvan er 1244 voor Hillary Clinton hadden gestemd. Zij vonden  $55.7\% \pm 2.1\%$  of [ 53.6 % ; 57.8 % ] voor hun 95 % betrouwbaarheidsinterval. Alle “aannemelijke waarden” voor het percent van alle kiezers dat voor Hillary heeft gestemd liggen hier boven de 50 % en dus was de TV er als de kippen bij om te zeggen dat Hillary de verkiezingen had gewonnen.

<pre>1-PropZInt x:1244 n:2232 C-Level:.95 Calculate</pre>	<pre>1-PropZInt (.53674, .57795) P=.5573476703 n=2232</pre>
-----------------------------------------------------------	-------------------------------------------------------------

Na het tellen van alle 6.2 miljoen stembiljetten bleek dat 56.0 % kiezers op Hillary Clinton had gestemd.

### 7.2.4. Risicobaby's

Kinderen met een te laag geboortegewicht lopen groter risico op allerlei complicaties. In de geneeskunde spreekt men over “laag geboortegewicht” wanneer een kind *minder dan 2.5 kg* weegt. Met de term “zeer laag geboortegewicht” bedoelt men gewichten beneden 1.5 kg bij de geboorte. Dan heb je echt een baby met verhoogd risico.

In Vlaanderen zijn er gelukkig weinig kinderen die bij de geboorte minder dan 1.5 kg wegen. Dat zijn er 8 à 9 per duizend. Dit cijfer komt uit een medisch tijdschrift van 5 jaar geleden. Maar is dat ook de voorbije 3 jaar zo gebleven? Je wil voor het percent risicobaby's die de voorbije 3 jaar in Vlaanderen zijn geboren een 95 % betrouwbaarheidsinterval opstellen. Hoe pak je dat aan?



- **Een formule blindelings toepassen... is meestal geen goed idee.**

Je zou als volgt kunnen redeneren. Ik heb betrouwbare externe informatie om te gissen wat ik ongeveer voor  $\hat{p}$  zal vinden in mijn steekproef. Volgens dat medisch tijdschrift is de proportie zo'n 9 op duizend. Verder werk ik met de klassieke 95 % betrouwbaarheid en met de klassieke 3 % foutenmarge. Ik moet dus werken met een steekproef van grootte

$$n = \frac{1.96^2 (0.009)(1-0.009)}{(0.03)^2} \cong 39$$

STOP.... en DENK !

Om de zaak op de spits te drijven zou je eens kunnen onderstellen dat er maar 9 risicobaby's per honderdduizend geboorten zijn. Dan moet je slechts een steekproef trekken van grootte

$$n = \frac{1.96^2 (0.00009)(1-0.00009)}{(0.03)^2} \cong 1.$$

Is dat niet fantastisch?

- **Even ernstig**

Als je vooraf een vermoeden hebt dat de populatieproportie  $\pi$  heel klein is, let dan goed op de foutenmarge waarmee je wil werken. Bij het voorbeeld van de risicobaby's is die totaal zinloos.

Als het waar is dat er ongeveer 8 à 9 risicobaby's per 1000 geboorten zijn, dan betekent een "ernstige vermeerdering" dat er bijvoorbeeld 12 per duizend risicobaby's zouden zijn, en een gevoelige vermindering heb je als er in plaats van 9 slechts 5 à 6 per duizend zijn. Wat heb je dan aan een interval met een foutenmarge van 3 % of dus van 30 per duizend? Dat is absurd!

In de context van deze studie is een zinvolle foutenmarge 1 à 2 per duizend. Als je die gebruikt dan verhoogt de benodigde steekproef drastisch, namelijk tot  $n = \frac{1.96^2 (0.009)(1-0.009)}{(0.002)^2} \cong 8566$ . Je

hebt bovendien nu ook dat, voor een verwachte  $\hat{p}$  van ongeveer 9 per 1000,

$$\begin{cases} n \hat{p} \cong 8566(0.009) \cong 77 \\ n(1-\hat{p}) \cong 8566(0.991) \cong 8488 \end{cases} \text{ wat twee keer ruimschoots meer dan 15 is.}$$

## 8. Voor niets gaat de zon op

Meer precieze informatie krijg je uit een smaller interval. Daarvoor heb je een kleinere foutenmarge nodig. Die krijg je niet zomaar. Je moet daarvoor inboeten op de betrouwbaarheid, ofwel heb je meer observaties nodig. Alles heeft zijn prijs.

### 8.1. Drie basisgrootheden

Hoe de drie basisgrootheden op elkaar inwerken ga je nu in detail bekijken. Je begint telkens met een concreet voorbeeld en daaruit leer je hoe het algemene principe werkt.

Als je “op veilig” speelt dan is het verband tussen “de benodigde steekproefgrootte  $n$ ”, “de betrouwbaarheid” en “de foutenmarge” gegeven door:

$$n = \frac{(z)^2 (0.5)(1-0.5)}{(\text{foutenmarge})^2}$$

Hierbij is  $z$  het kritisch punt van de standaard normale dat door de gewenste betrouwbaarheid wordt vastgelegd. Wil je bijvoorbeeld werken met een betrouwbaarheid van 95 % dan is  $z = 1.96$ . Bij een betrouwbaarheid van 90 % neem je  $z = 1.645$  en bij 99 % hoort  $z = 2.58$ .

### Opdracht 8

Bekijk nu wat er gebeurt in verschillende situaties en geef een interpretatie in woorden.

1. Houd eerst even de betrouwbaarheid vast. Kies in dit voorbeeld eens een betrouwbaarheid van 90 % en werk met  $z = 1.645$ . Zoek dan de benodigde steekproefgrootte als je op veilig speelt en hoogstens een foutenmarge van 5 % wil hebben. Doe dit ook voor andere foutenmarges en vul de volgende tabel in. Zie je een bepaald patroon verschijnen? Kan je dat verklaren?

Bij een betrouwbaarheid van 90 % is:	
maximale foutenmarge	steekproefgrootte
5 %	
2 %	
1 %	

Tabel 7

Is de volgende uitspraak waar: “als je wil dat de foutenmarge maar half zo groot is dan moet je werken met een steekproef die dubbel zo groot is”? Wat is je advies voor mensen die met een zeer kleine foutenmarge willen werken? Bedenk dat elke extra observatie die je nodig hebt tijd en geld kost.

2. Je kan ook zeggen dat je hoogstens een foutenmarge van 3 % wil en dan zoeken wat de benodigde steekproefgrootte is bij een betrouwbaarheid van 90 %. Doe dit ook voor andere betrouwbaarheid en vul de volgende tabel in. Werk telkens met de formule waar je op veilig speelt. Kan je verklaren wat hier gebeurt?

Bij een maximale foutenmarge van 3 % is:	
betrouwbaarheid	steekproefgrootte
90 %	
95 %	
99 %	

Tabel 8

3. Als je vooraf zegt dat je alleen maar tijd en geld hebt voor een steekproef van grootte  $n = 400$ , hoe groot zal dan de foutenmarge ten hoogste zijn bij een betrouwbaarheid van 90 % ? Vul dat in in de volgende tabel en werk met verschillende waarden voor de betrouwbaarheid. Maak gebruik van het feit dat  $\sqrt{\hat{p}(1-\hat{p})} \leq \sqrt{0.5(1-0.5)} = 0.5$ . Welk patroon ontdek je hier?

Bij een vaste steekproefgrootte $n = 400$ is:	
betrouwbaarheid	foutenmarge
90 %	
95 %	
99 %	

Tabel 9

## 8.2. Samenvatting

In de onderstaande tabel staat per kolom een bepaalde situatie beschreven. Het zijn de situaties die je zopas in je opdracht hebt bestudeerd. Zo zegt bijvoorbeeld kolom A: als je de steekproefgrootte niet wil veranderen maar je wil wel een grotere betrouwbaarheid dan zal de foutenmarge vergroten en krijg je dus een minder precies interval.

	A	B	C
betrouwbaarheid	↗	↗	vast
foutenmarge	↗	vast	↘
steekproefgrootte	vast	↗	↗

Tabel 10

### Opdracht 9

---

Schrijf nu ook voluit wat kolom B en C zeggen