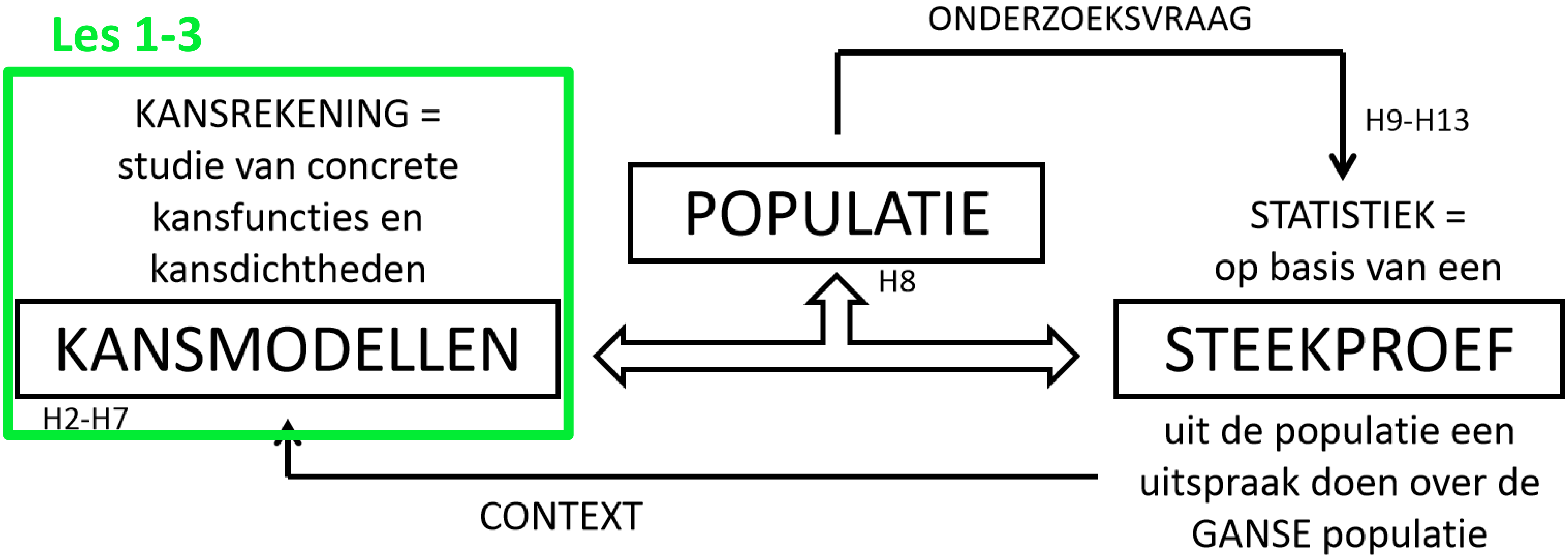
Even herhalen...



Formularium

Les 1:

BASISREGELS:

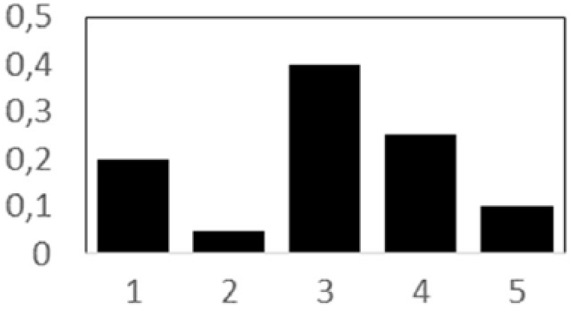
- Regel van Laplace: $P(A) = \frac{\#\text{ gunstige gevallen}}{\#\text{ mogelijke gevallen}}$
- Complementregel: $P(A^c) = 1 - P(A)$
- Somregel voor 2 gebeurtenissen: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- Productregel: $P(A \cap B) = P(B) \cdot P(A|B) = P(A) \cdot P(B|A)$
- Totale waarschijnlijkheid: $P(B) = \sum_i P(B \cap A_i)$ met alle A_i onderling disjunct en $\sum_i A_i = U$
- Regel van Bayes: $P(A_i|B) = \frac{P(B|A_i) \cdot P(A_i)}{P(B)}$

Formularium

Les 2:

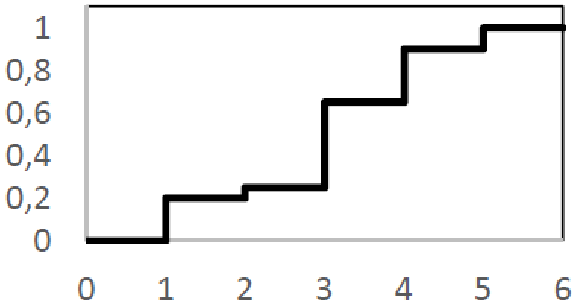
Discrete kansvariabele vs. continue kansvariabele

kansfunctie: $p_i = P(X = x_i)$



$$\sum_i p_i = 1$$

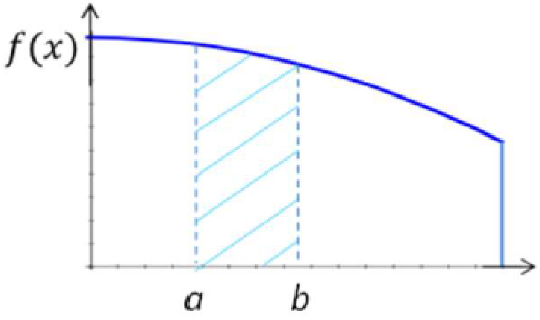
cumulatieve verdelingsfunctie $F_X(x) = P(X \leq x)$



$$\mu_X = E(X) = \sum_i x_i \cdot p_i$$

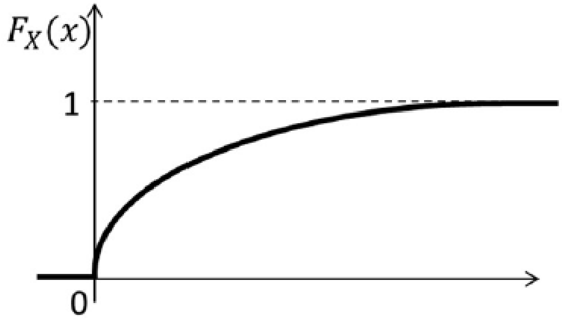
$$\sigma_X^2 = E((X - \mu_X)^2) = E(X^2) - (E(X))^2$$

kansdichtheid $f(x)$ met $P(a \leq X \leq b) = \int_a^b f(x)\,dx$



$$\int_{-\infty}^{+\infty} f(x)\,dx = 1$$

cumulatieve verdelingsfunctie $F_X(x) = \int_{-\infty}^{x} f(x)\,dx$



$$\mu_X = E(X) = \int_{-\infty}^{+\infty} x\,f(x)\,dx$$

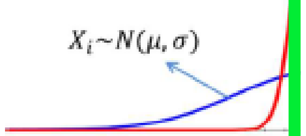
$$\sigma_X^2 = \int_{-\infty}^{+\infty} (x - \mu_X)^2 f(x)\,dx = E(X^2) - (E(X))^2$$

Fo

Les 2-3:

LINEAIRE COMBINATIES: Als $X_1, X_2, \ldots, X_n$ onderling **onafhankelijke** normaal verdeel
met verwachting μ_i en standaardafwijking σ_i voor i = 1, 2, …, n dan is $Y = a_1X_1$
ook normaal verdeeld met verwachting $\mu = a_1\mu_1 + a_2\mu_2 + \cdots + a_n\mu_n + b$
$\sigma = \sqrt{a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \cdots + a_n^2\sigma_n^2}$. *(opmerking: hierbij mogen de constanten a_i en b*

STEEKPROEFGEMIDDELDE: Als $X_1, X_2, \ldots, X_n$ onderling **onafhankelijke** normaal verdeelde toevalsvariabelen zijn met verwachting μ en standaardafwijking σ voor i = 1, 2, …, n dan is $\overline{X_n} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$.



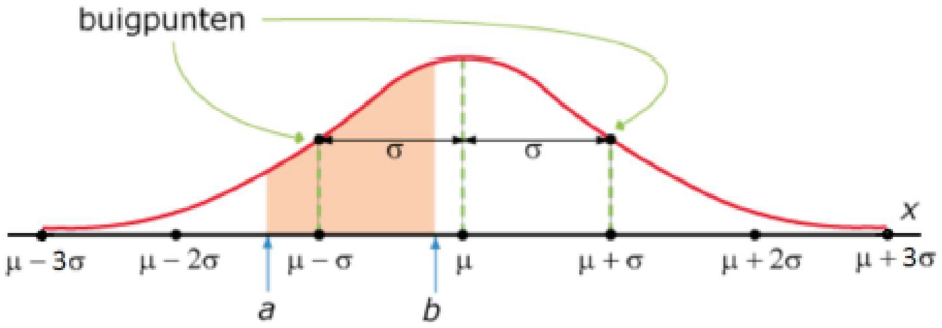
NORMALE BENADERING: Als $X \sim \mathrm{B}(n, p)$, dan is X bij benadering $N(np, \sqrt{np(1-p)}$
Als $X \sim \mathrm{Poi}(\lambda)$, dan is X bij benadering $N(\lambda, \sqrt{\lambda})$ als λ > 10.

CENTRALE LIMIETSTELLING: Voor een grote steekproef ($n \geq 30$) kan de verdeling van de som $S_n = X_1 + \cdots + X_n$ benaderd worden door een normale verdeling met verwachting $n\mu_X$ en standaardafwijking $\sqrt{n}\,\sigma_X$.

STEEKKAART NORMALE KANSVERDELING

KANSDICHTHEIDSFUNCTIE: $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ ← rekentoestel: normPdf(x, μ, σ)

NOTATIE: $X \sim N(\mu, \sigma)$ → E(X) = μ, Var(X) = σ²



KANSVERDELINGSFUNCTIE: $P(X \leq x) = \int_{-\infty}^{x} f_X(t)dt$ ← rekentoestel: normCdf(-∞, x, μ, σ)

KANS MET ONDER- en BOVENGRENS: $P(a \leq X \leq b)$ = normCdf(a, b, μ, σ)

BELANGRIJKSTE σ-NIVEAUS: Voor elke normale kansverdeling $X \sim N(\mu, \sigma)$ geldt

$P(\mu - \sigma \leq X \leq \mu + \sigma) = 68{,}27\%$ $\qquad P(\mu - 1.645\sigma \leq X \leq \mu + 1.645\sigma) \approx 90\%$
$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = 95{,}45\%$ $\qquad P(\mu - 1.96\sigma \leq X \leq \mu + 1.96\sigma) \approx 95\%$
$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = 99{,}73\%$ $\qquad P(\mu - 2.576\sigma \leq X \leq \mu + 2.576\sigma) \approx 99\%$

p-KWANTIELEN (0 ≤ p ≤ 1): zoek c zodat $P(X \leq c) = p$ ← rekentoestel: invNorm(p, μ, σ)

LINK met **STANDAARDNORMALE VERDELING:** Als $X \sim N(\mu, \sigma)$, dan heeft de gestandaardiseerde kansvariabele $Z = \frac{X-\mu}{\sigma}$ (=Z-score) een normale verdeling N(0,1) met verwachting 0 en standaardafwijking 1.

Formularium

Les 3:

Wanneer een kansexperiment bestaat uit n *onafhankelijke* deelexperimenten met elk **dezelfde** kans op succes (p) spreken we van een **binomiaal kansexperiment**. De kansvariabele X = het aantal keer succes, volgt dan de binomiale kansverdeling: $X \sim B(n, p)$.

De bijhorende kansfunctie: $P(X = i) = \binom{n}{i} p^i (1-p)^{n-i}$ $\leftarrow$ rekentoestel: binomPdf(n,p,i)

Voor een binomiale verdeling geldt: $\mu_X = np$ en $\sigma_X = \sqrt{np(1-p)}$.

Formularium

Les 3:

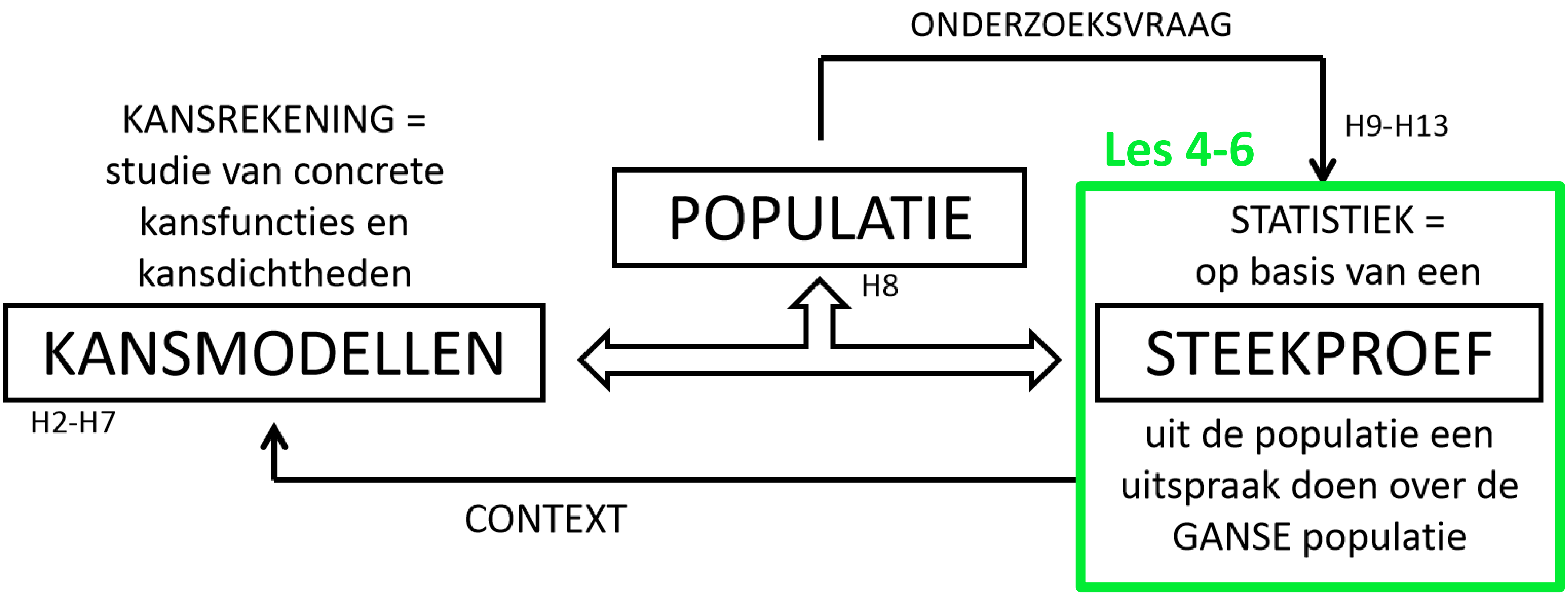
Noem X het aantal gebeurtenissen in een continu interval (bvb. een bepaalde afstand of een tijdsinterval).

Wanneer het **verwachte** aantal gebeurtenissen λ **evenredig** is met de lengte van het interval en bovendien het aantal gebeurtenissen die optreden in 2 niet-overlappende (deel)intervallen *onafhankelijk* zijn van elkaar, dan is de kansvariabele X **Poisson** verdeeld met gemiddelde λ: $X \sim Poi(\lambda)$.

De bijhorende kansfunctie: $P(X = k) = \dfrac{e^{-\lambda}\lambda^k}{k!}$ ← rekentoestel: poissPdf(λ,k)

Voor een Poisson verdeling geldt: $\mu_X = \lambda$ en $\sigma_X = \sqrt{\lambda}$.

Vanaf nu ‘echte’ statistiek!



Statistiek: les 4

Betrouwbaarheidsinterval en hypothesetoetsen voor het gemiddelde met σ gekend

Sabine Bertho sabine.bertho@kuleuven.be
Giovanni Vanroelen giovanni.vanroelen@uhasselt.be

De gezamenlijke opleiding industrieel ingenieur is een initiatief van UHasselt en KU Leuven.

Yoghurtpotjes...

Aselecte steekproef uit een populatie

- n onafhankelijke, willekeurige trekkingen $X_1, X_2, \ldots, X_n$ uit een populatie X
- Elke X_i heeft dezelfde kans(dichtheids)functie als de populatie X
- De steekproef moet representatief zijn voor de populatie

Yoghurtpotjes in het rekentoestel

Gemiddelde en standaardafwijking van de volledige populatie berekenen (menu – 6 – 3 – …):

3 → $\mathrm{mean}(massa)$ 153.102

9 → $\mathrm{stDevPop}(massa)$ 2.96215

Steekproef nemen in het rekentoestel

Kies een willekeurig getal

Zonder teruglegging

RandSeed 1572 *Gereed*

$steekproef1 := \mathrm{randSamp}(massa, 100, 1)$

$\{154.77, 153.61, 155.78, 151.48, 149.72, 159.32, 151.$ ▸

DataYoghurt.tns

	A massa	B	C
=			
1	150.1		
2	152.46		
3	149.31		
4	152.79		
5	151.96		
6	147.07		
7	152.7		
8	146.52		
9	154.11		
10	154.67		

G5

Yoghurtpotjes in het rekentoestel

Gemiddelde en standaardafwijking van de volledige populatie berekenen:

mean(*massa*)	153.102
stDevPop(*massa*)	2.96215

DataYoghurt.tns

	A massa	B steekproef1	C
=			
1	150.1	154.77	
2	152.46	153.61	
3	149.31	155.78	
4	152.79	151.48	
5	151.96	149.72	
6	147.07	159.32	
7	152.7	151.68	
8	146.52	149.95	
9	154.11	153.2	
10	154.67	152.47	
11	150.54	157.86	

Steekproef nemen in het rekentoestel

Kies een willekeurig getal

RandSeed 1572	*Gereed*
steekproef1:=randSamp(*massa*,100,1)	
{154.77,153.61,155.78,151.48,149.72,159.32,151.▸	

UHASSELT KU LEUVEN

Steekproefgemiddelde

Definitie 8.3 *Het* steekproefgemiddelde *is gelijk aan*

$$\boxed{\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.}$$

Yoghurtpotjes:

mean(*massa*)	153.102
mean(*steekproef1*)	152.785

	A massa	B steekproef1	C
=			
1	150.1	154.77	
2	152.46	153.61	
3	149.31	155.78	
4	152.79	151.48	
5	151.96	149.72	
6	147.07	159.32	
7	152.7	151.68	
8	146.52	149.95	
9	154.11	153.2	
10	154.67	152.47	
11	150.54	157.86	

Eigenschappen van het steekproefgemiddelde (déjà vu)

Eigenschap 8.4 (Gemiddelde en variantie van $\overline{X}$) *Het gemiddelde van het steekproefgemiddelde is gelijk aan het populatiegemiddelde:*

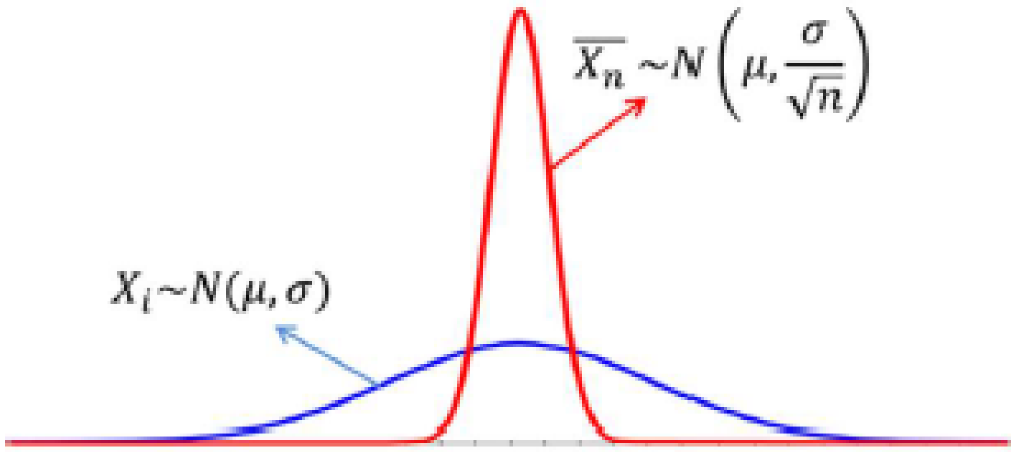
Vervelende maar belangrijke zin!

$$E(\overline{X}) = \mu$$

De variantie van het steekproefgemiddelde is gelijk aan de populatievariantie gedeeld door n:

$$\mathrm{Var}(\overline{X}) = \sigma^2_{\overline{X}} = \frac{\sigma^2}{n}$$

STEEKPROEFGEMIDDELDE: Als $X_1, X_2, \ldots, X_n$ onderling **onafhankelijke** normaal verdeelde toevalsvariabelen zijn met verwachting μ en standaardafwijking σ voor $i = 1, 2, \ldots, n$ dan is $\overline{X_n} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$.



Puntschatting van het gemiddelde

Eigenschap 8.4 (Gemiddelde en variantie van $\overline{X}$) *Het gemiddelde van het steekproef-gemiddelde is gelijk aan het populatiegemiddelde:*

Vervelende maar belangrijke zin!

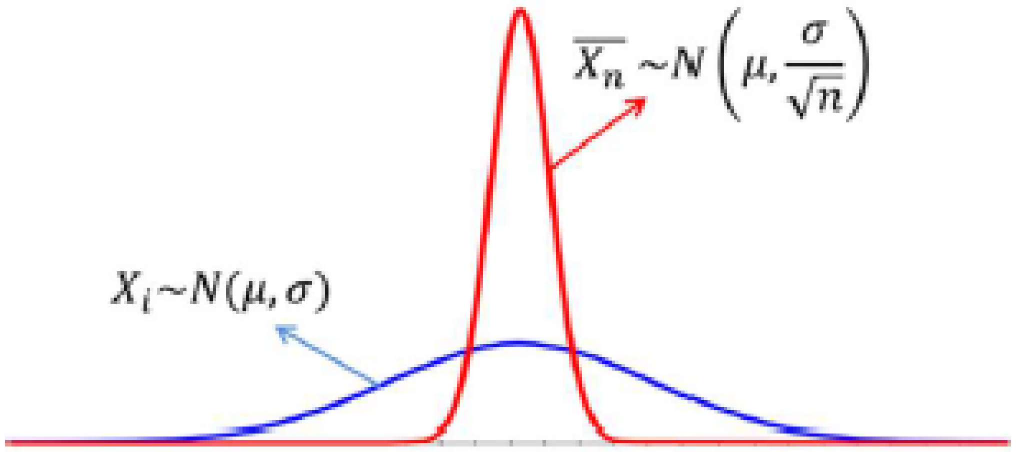
$$\mathrm{E}(\overline{X}) = \mu$$

De variantie v … *gedeeld door n:*

$\overline{X}$ is een zuivere schatter voor μ

$$\mathrm{Var}(\overline{X}) = \sigma^2_{\overline{X}} = \frac{\ \ }{n}$$

STEEKPROEFGEMIDDELDE: Als $X_1, X_2, \ldots, X_n$ onderling **onafhankelijke** normaal verdeelde toevalsvariabelen zijn met verwachting μ en standaardafwijking σ voor i = 1, 2, …, n dan is $\overline{X_n} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$.



UHASSELT KU LEUVEN

Notaties: afspraken

X = eigenschap die we bestuderen (bvb. lengte, gewicht, …)
= kansvariabele met een al dan niet gekende verdeling

μ = populatiegemiddelde

σ = populatiestandaardafwijking

$\bar{X}$ = kansvariabele steekproefgemiddelde

$\bar{x}$ = steekproefgemiddelde voor 1 bepaalde steekproef

Algemeen: hoofdletter = kansvariabele
kleine letter = 1 uitkomst

Intervalschatten

Hoe dicht ligt $\bar{x}$ bij μ?

Verdeling van $\bar{X}$?

Als $X \sim N(\mu, \sigma)$ dan $\bar{X} \sim N(\mu, \sigma/\sqrt{n})$

Maar wat als $X \sim ?$

Centrale limietstelling

Eigenschap 8.6 Centrale Limietstelling *Voor een grote steekproef ($n \geq 30$) kan de kansdichtheidsfunctie van het steekproefgemiddelde $\overline{X}$ benaderd worden door een normale kansdichtheidsfunctie. Dit wil zeggen*

We nemen een steekproef van 100 yoghurt potjes

$$\overline{X} \sim N(\mu, \sigma/\sqrt{n}) \quad \textit{voor } n \geq 30$$

Hoe groter de steekproeven zijn die je neemt, hoe meer de verdeling van $\overline{X}$ zal gelijken op een normale verdeling. Merk op dat de steekproefgrootte de variantie $\mathrm{Var}(\overline{X}) = \sigma^2_{\overline{X}} = \sigma^2/n$ beïnvloedt.

CENTRALE LIMIETSTELLING: Voor een grote steekproef ($n \geq 30$) kan de verdeling van de som $S_n = X_1 + \cdots + X_n$ benaderd worden door een normale verdeling met verwachting $n\mu_X$ en standaardafwijking $\sqrt{n}\,\sigma_X$.

Intervalschatten

$$\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

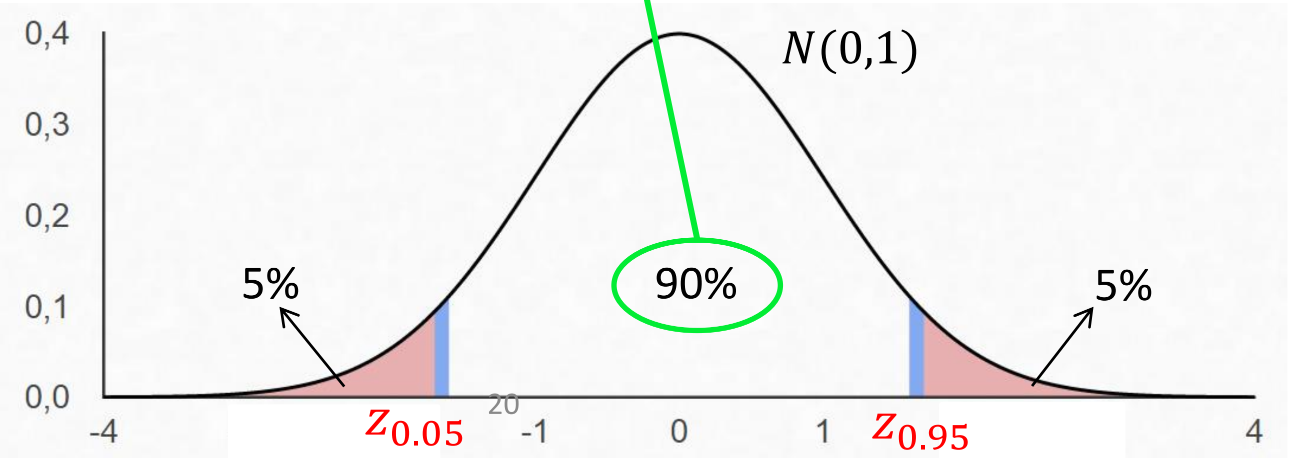
$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

$$\mathrm{P}\left(z_{0.05} < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < z_{0.95}\right) = 0.90$$

invNorm(0.95,0,1)

Deze les: σ gekend
Volgende les: σ niet gekend

We gaan hier op zoek naar een 90% betrouwbaarheidsinterval



Intervalschatten

$$P\left(z_{0.05} < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < z_{0.95}\right) = 0.90$$

$$P\left(z_{0.05}\cdot\frac{\sigma}{\sqrt{n}} < \bar{X}-\mu < z_{0.95}\cdot\frac{\sigma}{\sqrt{n}}\right) = 0.90$$

$$P\left(-z_{0.95}\cdot\frac{\sigma}{\sqrt{n}} < \bar{X}-\mu < z_{0.95}\cdot\frac{\sigma}{\sqrt{n}}\right) = 0.90$$

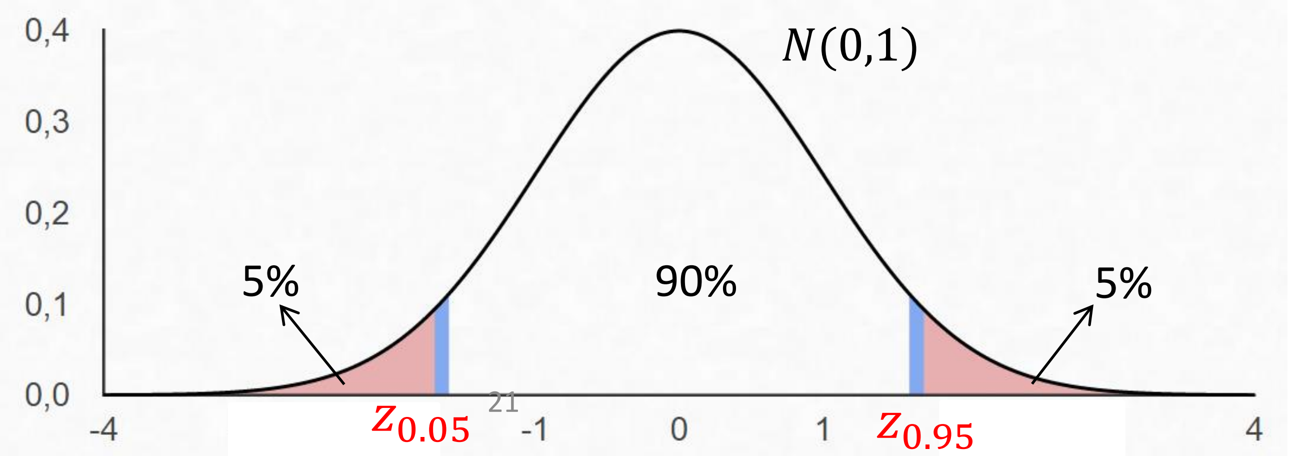
$$P\left(z_{0.95}\cdot\frac{\sigma}{\sqrt{n}} > \mu-\bar{X} > -z_{0.95}\cdot\frac{\sigma}{\sqrt{n}}\right) = 0.90$$

$$P\left(\underline{\bar{X}+z_{0.95}\cdot\frac{\sigma}{\sqrt{n}}} > \mu > \underline{\bar{X}-z_{0.95}\cdot\frac{\sigma}{\sqrt{n}}}\right) = 0.90$$

Er is 90% kans dat μ tussen deze grenzen ligt.

90% betrouwbaarheidsinterval:

$$\left[\bar{X}-z_{0.95}\cdot\frac{\sigma}{\sqrt{n}}, \bar{X}+z_{0.95}\cdot\frac{\sigma}{\sqrt{n}}\right]$$



Intervalschatten

$$P\left(z_{0.05} < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < z_{0.95}\right) = 0.90$$

$$P\left(z_{0.05} \cdot \frac{\sigma}{\sqrt{n}} < \bar{X} - \mu < z_{0.95} \cdot \frac{\sigma}{\sqrt{n}}\right) = 0.90$$

$$P\left(-z_{0.95} \cdot \frac{\sigma}{\sqrt{n}} < \bar{X} - \mu < z_{0.95} \cdot \frac{\sigma}{\sqrt{n}}\right) = 0.90$$

$$P\left(z_{0.95} \cdot \frac{\sigma}{\sqrt{n}} > \mu - \bar{X} > -z_{0.95} \cdot \frac{\sigma}{\sqrt{n}}\right) = 0.90$$

$$P\left(\bar{X} + z_{0.95} \cdot \frac{\sigma}{\sqrt{n}} > \mu > \bar{X} - z_{0.95} \cdot \frac{\sigma}{\sqrt{n}}\right) = 0.90$$

90% betrouwbaarheidsinterval:

$$\left[\bar{X} - z_{0.95} \cdot \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{0.95} \cdot \frac{\sigma}{\sqrt{n}}\right]$$

Foutmarge

en $z_{0.95} = \text{invNorm}(0.95,0,1) = 1.64485$

Yoghurtpotjes:

- $\sigma = 2.96215$ (bepaald uit hele populatie, dia 13)
- $n = 100$
- Foutmarge is dus $1.64485 \cdot \frac{2.96215}{\sqrt{100}} = 0.48723$

Intervalschatten

90% betrouwbaarheidsinterval op basis van de steekproef in dia 13-15:

$$[152.785 - 0.48723, 152.785 + 0.48723]$$

$$= [152.298, 153.272]$$

Hier ligt μ (=153.102) in het 90% BI.

MAAR…

RandSeed 94103 *Gereed*

steekproef2:=randSamp(*massa*,100,1)
{149.92,154.92,159.89,153.96,145.52,152.11,155.11,1…

mean(*steekproef2*) 152.604

Voor deze steekproef niet! (reken zelf uit)

90% betrouwbaarheidsinterval:

$$\left[\bar{X} - z_{0.95} \cdot \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{0.95} \cdot \frac{\sigma}{\sqrt{n}}\right]$$

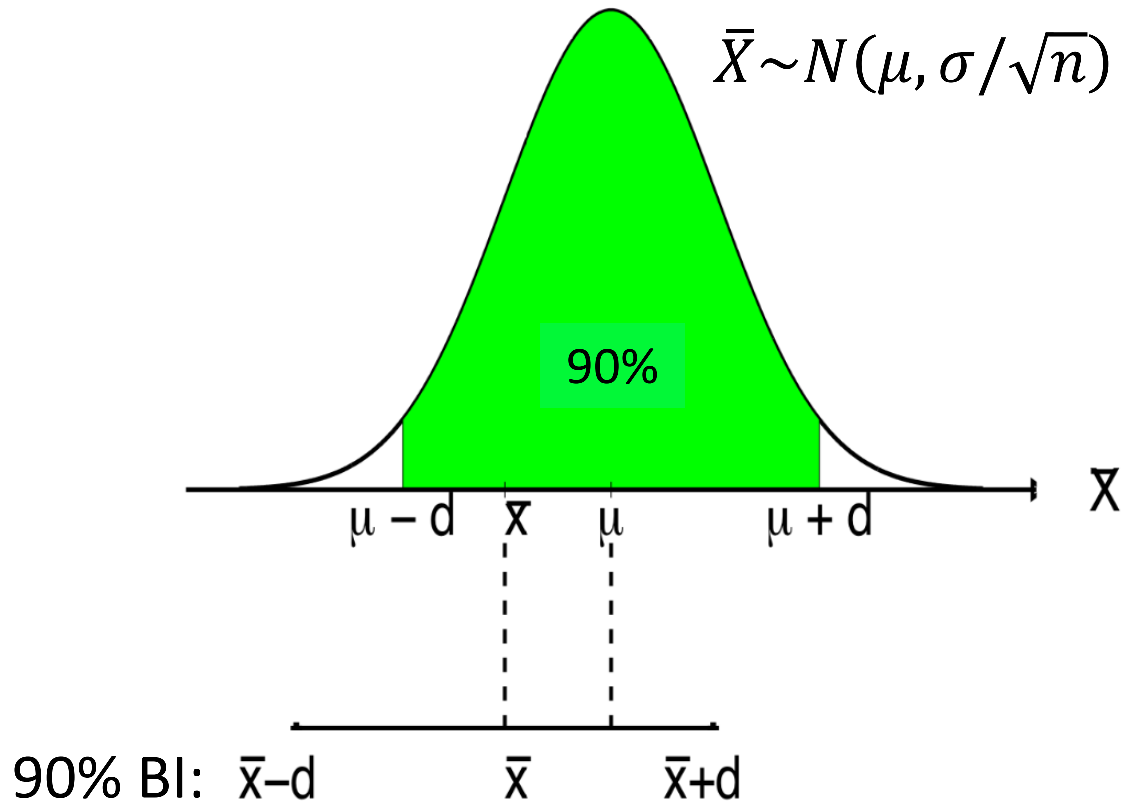
Foutmarge

en $z_{0.95} = \text{invNorm}(0.95,0,1) = 1.64485$

Yoghurtpotjes:

- $\sigma = 2.96215$ (bepaald uit hele populatie, dia 13)
- $n = 100$
- Foutmarge is dus $1.64485 \cdot \frac{2.96215}{\sqrt{100}} = 0.48723$

90% betrouwbaarheidsinterval: grafisch

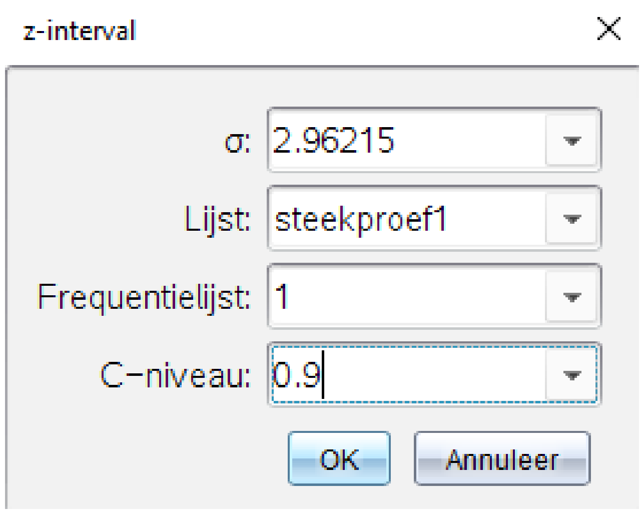


Een betrouwbaarheidsinterval is voor elke steekproef anders. Met een beetje pech is een 90% betrouwbaarheidsinterval 100% fout!

Intervalschatten met de TI-Nspire

Yoghurtpotjes:

- Betrouwbaarheidsinterval via: menu – 6: Statistieken – 6: Betrouwbaarheidsintervallen – 1: z-interval – Kies Gegevensinvoermethode: Gegevens (omdat we hier een lijst met getallen hebben)



zInterval 2.96215,*steekproef1*,1,0.9: *stat.results*

$$\begin{bmatrix} \text{"Titel"} & \text{"z-interval"} \\ \text{"CLower"} & 152.298 \\ \text{"CUpper"} & 153.272 \\ \text{"}\bar{x}\text{"} & 152.785 \\ \text{"ME"} & 0.48723 \\ \text{"sx := sn-1x"} & 2.90615 \\ \text{"n"} & 100. \\ \text{"}\sigma\text{"} & 2.96215 \end{bmatrix}$$

Betrouwbaarheidsinterval voor μ met σ gekend: samenvatting

90% betrouwbaarheidsinterval:

$$\left[\bar{X} - z_{0.95} \cdot \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{0.95} \cdot \frac{\sigma}{\sqrt{n}}\right]$$

en $z_{0.95} = \mathrm{invNorm}(0.95{,}0{,}1) = 1.64485$

Opmerking 2: M.b.v. deze teststatistieken kan je ook een B.I. voor μ opstellen → $\bar{x} \pm$ kwantiel x $\frac{\mathrm{STDEV}}{\sqrt{n}}$

UHASSELT KU LEUVEN

Betrouwbaarheidsinterval voor µ met σ gekend: samenvatting

90% betrouwbaarheidsinterval:

$$\left[\bar{X} - z_{0.95} \cdot \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{0.95} \cdot \frac{\sigma}{\sqrt{n}}\right]$$

en $z_{0.95} = \mathrm{invNorm}(0.95,0,1) = 1.64485$

Opmerking 2: M.b.v. deze teststatistieken kan je ook een B.I. voor µ opstellen → $\bar{x} \pm$ kwantiel x $\frac{\mathrm{STDEV}}{\sqrt{n}}$

Foutmarge

De foutmarge wordt kleiner als de steekproefgrootte *n* groter wordt!

Betrouwbaarheidsintervallen voor het gemiddelde

Hoeveel yoghurtpotjes moeten we wegen zodat de foutmarge kleiner wordt dan 0,3g?

90% betrouwbaarheidsinterval:

$$\left[\bar{X} - z_{0.95} \cdot \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{0.95} \cdot \frac{\sigma}{\sqrt{n}}\right]$$

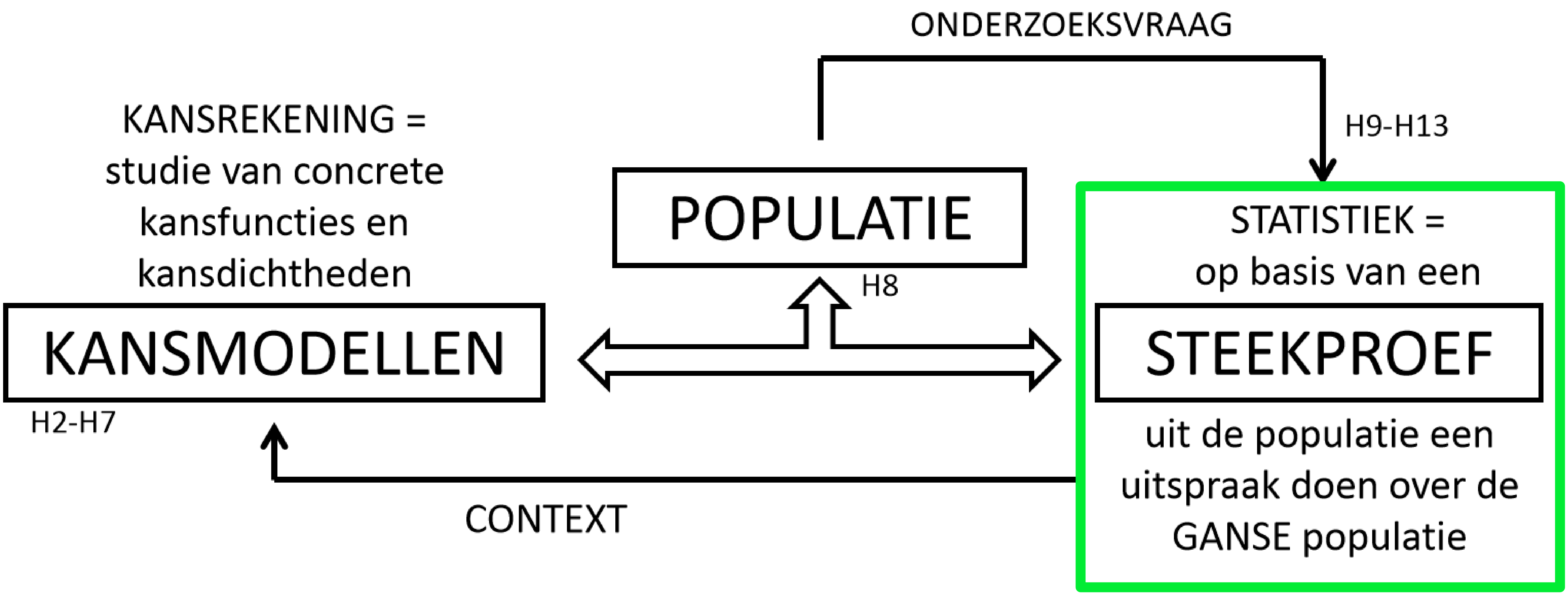
en $z_{0.95} = \text{invNorm}(0.95,0,1) = 1.64485$

$$1.64485 \cdot \frac{2.96215}{\sqrt{n}} < 0.3$$

$$\sqrt{n} > 1.64485 \cdot \frac{2.96215}{0.3}$$

$n > \left(1.64485 \cdot \frac{2.96215}{0.3}\right)^2$ = 263,77 dus minstens 264 metingen.

Even herhalen...



Stappenplan hypothesetoetsen

Zie formularium

ALGEMEEN STAPPENPLAN voor het toetsen van hypothesen

1. **Formuleer de nulhypothese (H_0) en de alternatieve hypothese (H_1).**
   Dikwijls is H_0 te interpreteren als een zeer neutrale, behoudende uitspraak.

2. **Bepaal het significantieniveau α waarop je de toets wil uitvoeren.**
   De kans op een type I-fout (d.w.z. : H_0 verwerpen terwijl dit niet mag) zal nu ten hoogste α zijn.

3. **Bereken** (op basis van een steekproef van grootte *n*) **de experimentele waarde t van een teststatistiek T waarvan de kansverdeling onder H_0 volledig gekend is** (al dan niet benaderend).

4. Hier heb je de keuze: **bepaal het kritisch gebied** (= verwerpingsgebied) **van je hypothesetoets** OF **bereken de p-waarde** (= overschrijdingskans) **van je toets**. Meer concreet:

   a. Voor een **linker-één-staart** toets: vind de grootste t_0 waarvoor $P(T \leq t_0) \leq \alpha$.
      Het kritisch gebied wordt dan gegeven door $\{x: x \leq t_0\}$ ←→ p-waarde = $P(T \leq t)$.
   b. Voor een **rechter-één-staart** toets: vind de kleinste t_0 waarvoor $P(T \geq t_0) \leq \alpha$.
      Het kritisch gebied wordt dan gegeven door $\{x: x \geq t_0\}$ ←→ p-waarde = $P(T \geq t)$.
   c. Voor een **twéé-staart** toets: vind de grootste t_1 waarvoor $P(T \leq t_1) \leq \alpha/2$ en de kleinste t_2 waarvoor $P(T \geq t_2) \leq \alpha/2$. Het kritisch gebied is dan: $\{x: x \leq t_1 \text{ en } x \geq t_2\}$
      ←→ de p-waarde in geval de kansverdeling van T **symmetrisch** is: $p = 2\,P(T \leq t)$ of $p = 2\,P(T \geq t)$ naargelang t zich in de linker- of rechterstaart van de verdeling van T bevindt.

5. **Formuleer je besluit**. Dit wil zeggen: verwerp H_0 op het significantieniveau α, indien de experimentele waarde t in het kritisch gebied ligt van je toets **OF** indien de p-waarde ≤ α.
   Zo niet, besluit dat je op basis van je steekproef H_0 niet kunt tegenspreken ("kan behouden").

Voorbeeld hypothesetoets: oef 16 p44

16. Een machine produceert metalen staafjes. De lengte van deze staafjes is normaal verdeeld met een vooraf gekende standaardafwijking van 2,0 cm. De gemiddelde lengte heeft als normwaarde 20,0 cm. Er kunnen storingen optreden waardoor het gemiddelde lager wordt, zonder dat de spreiding verandert.
    Ter controle neemt men regelmatig steekproeven van 5 stuks en meet hiervan de lengte. Bij een steekproef vond men een gemiddelde lengte van 18,6 cm.

    a. Toets de hypothese dat de instelwaarde van de machine nog steeds de normwaarde is. Neem als significantieniveau 2% en trek een conclusie op basis van een gepast significantiegebied (= verwerpingsgebied).

Nulhypothese en alternatieve hypothese

Een hypothesetoets begint met het opstellen van twee hypothesen: de *nulhypothese* H_0 en de *alternatieve hypothese* H_1. De alternatieve hypothese kan verschillende vormen hebben.

$$\begin{array}{lcl} H_0 & : & \mu = \mu_0 \\ H_1 & : & \mu > \mu_0 \end{array} \qquad \begin{array}{lcl} H_0 & : & \mu = \mu_0 \\ H_1 & : & \mu < \mu_0 \end{array} \qquad \begin{array}{lcl} H_0 & : & \mu = \mu_0 \\ H_1 & : & \mu \neq \mu_0 \end{array}$$

STAP 1

1. **Formuleer de nulhypothese (H_0) en de alternatieve hypothese (H_1).**
   Dikwijls is H_0 te interpreteren als een zeer neutrale, behoudende uitspraak.

Nulhypothese en alternatieve hypothese

16. Een machine produceert metalen staafjes. De lengte van deze staafjes is normaal verdeeld met een vooraf gekende standaardafwijking van 2,0 cm. De gemiddelde lengte heeft als normwaarde 20,0 cm. Er kunnen storingen optreden waardoor het gemiddelde lager wordt, zonder dat de spreiding verandert.

    Ter controle neemt men regelmatig steekproeven van 5 stuks en meet hiervan de lengte. Bij een steekproef vond men een gemiddelde lengte van 18,6 cm.

Nulhypothese

$H_0 : \mu = \mu_0 = 20{,}0 \text{ cm}$

In een normale situatie is de gemiddelde lengte 20,0 cm

Alternatieve hypothese:

$H_1 : \mu < \mu_0 = 20{,}0 \text{ cm}$

Als er storingen optreden kan het gemiddelde verlagen

Fouten van de eerste en de tweede soort

16. Een machine produceert metalen staafjes. De lengte van deze staafjes is normaal verdeeld met een vooraf gekende standaardafwijking van 2,0 cm. De gemiddelde lengte heeft als normwaarde 20,0 cm. Er kunnen storingen optreden waardoor het gemiddelde lager wordt, zonder dat de spreiding verandert.

    Ter controle neemt men regelmatig steekproeven van 5 stuks en meet hiervan de lengte.

    Bij een steekproef vond men een gemiddelde lengte van 18,6 cm.

STAP 2

2. **Bepaal het significantieniveau α waarop je de toets wil uitvoeren.**
   De kans op een type I-fout (d.w.z. : H_0 verwerpen terwijl dit niet mag) zal nu ten hoogste α zijn.

Wat is dit?

Fouten van de eerste en de tweede soort

Wat zijn de mogelijke conclusies van de hypothese toets?

	Je zegt dat H_0 waar is	Je zegt dat H_0 niet waar is
H_0 is waar	Juiste conclusie	Type I fout
H_0 is niet waar	Type II fout	Juiste conclusie

Ergste fout!

Het kan op 2 manieren fout gaan

$$\mathrm{P}(\text{fout van de eerste soort}) = \mathrm{P}(H_0 \text{ verwerpen} | H_0 \text{ waar})$$

= we zeggen dat H_0 niet waar is, terwijl dit in werkelijkheid wel zo is = ergste fout

$$\mathrm{P}(\text{fout van de tweede soort}) = \mathrm{P}(H_0 \text{ aanvaarden} | H_0 \text{ niet waar})$$

= we zeggen dat H_0 waar is, terwijl dit in werkelijkheid niet zo is

Fouten van de eerste en de tweede soort

Voorbeeld: ons rechtssysteem

H_0: beschuldigde is onschuldig

H_1: beschuldigde is schuldig

	Je zegt dat H_0 waar is	Je zegt dat H_0 niet waar is
H_0 is waar	onschuldige is vrij	onschuldige moet naar de gevangenis
H_0 is niet waar	schuldige mag vrij rondlopen	schuldige moet naar de gevangenis

Ergste fout!

Fouten van de eerste en de tweede soort

16. Een machine produceert metalen staafjes. De lengte van deze staafjes is normaal verdeeld met een vooraf gekende standaardafwijking van 2,0 cm. De gemiddelde lengte heeft als normwaarde 20,0 cm. Er kunnen storingen optreden waardoor het gemiddelde lager wordt, zonder dat de spreiding verandert.

Ter controle neemt men regelmatig steekproeven van 5 stuks en meet hiervan de lengte. Bij een steekproef vond men een gemiddelde lengte van 18,6 cm.

Nulhypothese

$H_0 : \mu = \mu_0 = 20{,}0 \text{ cm}$

Alternatieve hypothese:

$H_1 : \mu < \mu_0 = 20{,}0 \text{ cm}$

$$\mathrm{P}(\text{fout van de eerste soort}) = \mathrm{P}(H_0 \text{ verwerpen} | H_0 \text{ waar})$$

Fout van de eerste soort: we concluderen dat de instelwaarde verlaagd is terwijl deze in werkelijkheid nog altijd 20,0 cm is.

Fouten van de eerste en de tweede soort

16. Een machine produceert metalen staafjes. De lengte van deze staafjes is normaal verdeeld met een vooraf gekende standaardafwijking van 2,0 cm. De gemiddelde lengte heeft als normwaarde 20,0 cm. Er kunnen storingen optreden waardoor het gemiddelde lager wordt, zonder dat de spreiding verandert.
Ter controle neemt men regelmatig steekproeven van 5 stuks en meet hiervan de lengte. Bij een steekproef vond men een gemiddelde lengte van 18,6 cm.

Nulhypothese

$H_0 : \mu = \mu_0 = 20{,}0 \text{ cm}$

Alternatieve hypothese:

$H_1 : \mu < \mu_0 = 20{,}0 \text{ cm}$

$$\mathrm{P}(\text{fout van de tweede soort}) = \mathrm{P}(H_0\,\text{aanvaarden} | H_0\,\text{niet waar})$$

Fout van de tweede soort: we concluderen dat de instelwaarde nog altijd 20,0 cm is terwijl deze in werkelijkheid verlaagd is.

Fouten van de eerste en de tweede soort

16. Een machine produceert metalen staafjes. De lengte van deze staafjes is normaal verdeeld met een vooraf gekende standaardafwijking van 2,0 cm. De gemiddelde lengte heeft als normwaarde 20,0 cm. Er kunnen storingen optreden waardoor het gemiddelde lager wordt, zonder dat de spreiding verandert.

    Ter controle neemt men regelmatig steekproeven van 5 stuks en meet hiervan de lengte.

    Bij een steekproef vond men een gemiddelde lengte van 18,6 cm.

We moeten ervoor zorgen dat de kans op een fout van de eerste soort (= ergste fout) zo klein mogelijk is. We leggen deze kans vast vóórdat we aan de hypothesetoets beginnen als het significantieniveau α.

In de oefeningen of op het examen is deze altijd gegeven.

Fouten van de eerste en de tweede soort

16. Een machine produceert metalen staafjes. De lengte van deze staafjes is normaal verdeeld met een vooraf gekende standaardafwijking van 2,0 cm. De gemiddelde lengte heeft als normwaarde 20,0 cm. Er kunnen storingen optreden waardoor het gemiddelde lager wordt, zonder dat de spreiding verandert.
Ter controle neemt men regelmatig steekproeven van 5 stuks en meet hiervan de lengte. Bij een steekproef vond men een gemiddelde lengte van 18,6 cm.

    a. Toets de hypothese dat de instelwaarde van de machine nog steeds de normwaarde is. Neem als significantieniveau 2% en trek een conclusie op basis van een gepast significantiegebied (= verwerpingsgebied).

STAP 2

2. **Bepaal het significantieniveau α waarop je de toets wil uitvoeren.**
   De kans op een type I-fout (d.w.z. : H_0 verwerpen terwijl dit niet mag) zal nu ten hoogste α zijn.

Teststatistiek?

STAP 3

3. **Bereken** (op basis van een steekproef van grootte *n*) **de experimentele waarde t van een test-statistiek T waarvan de kansverdeling onder H_0 volledig gekend is** (al dan niet benaderend).

X = lengte staafjes
$\bar{X}$ = steekproefgemiddelde voor 5 staafjes

Stel dat H_0 waar is:

Het gemiddelde van de populatie

dan is $\mu_X = 20{,}0$ en $\sigma_X = 2{,}0$
en dus $\mu_{\bar{X}} = 20{,}0$ en $\sigma_{\bar{X}} = 2{,}0/\sqrt{5}$

Het gemiddelde van het steekproefgemiddelde

Teststatistiek:

$$\boxed{T = \bar{X} \sim N\left(20\,, \frac{2}{\sqrt{5}}\right)}$$

UHASSELT KU LEUVEN

Teststatistiek!

STAP 3

3. **Bereken** (op basis van een steekproef van grootte *n*) **de experimentele waarde t van een test-statistiek T waarvan de kansverdeling onder H_0 volledig gekend is** (al dan niet benaderend).

TESTSTATISTIEKEN T VOOR HET GEMIDDELDE

1. Toetsen op het gemiddelde μ van 1 groep

   a. σ vooraf gekend: z-toets → $T = \bar{X} \sim N(\mu, \sigma/\sqrt{n})$ OF $T = Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0{,}1)$

Experimentele waarde van de teststatistiek:

$t = \bar{x} = 18.6$

$$T = \bar{X} \sim N\left(20\,, \frac{2}{\sqrt{5}}\right)$$

Ter controle neemt men regelmatig steekproeven van 5 stuks en meet hiervan de lengte.

Bij een steekproef vond men een gemiddelde lengte van 18,6 cm.

Kritisch gebied = verwerpingsgebied

Even wachten, komt zo
dadelijk aan bod

STAP 4

4. Hier heb je de keuze: **bepaal het kritisch gebied** (= verwerpingsgebied) **van je hypothesetoets** OF **bereken de p-waarde** (= overschrijdingskans) **van je toets**. Meer concreet:

We vertrekken van:

$\alpha =$ maximaal risico dat je wil nemen op een fout van de eerste soort $= 0.02$

$\mathrm{P}(H_0 \text{ verwerpen} | H_0 \text{ waar})$

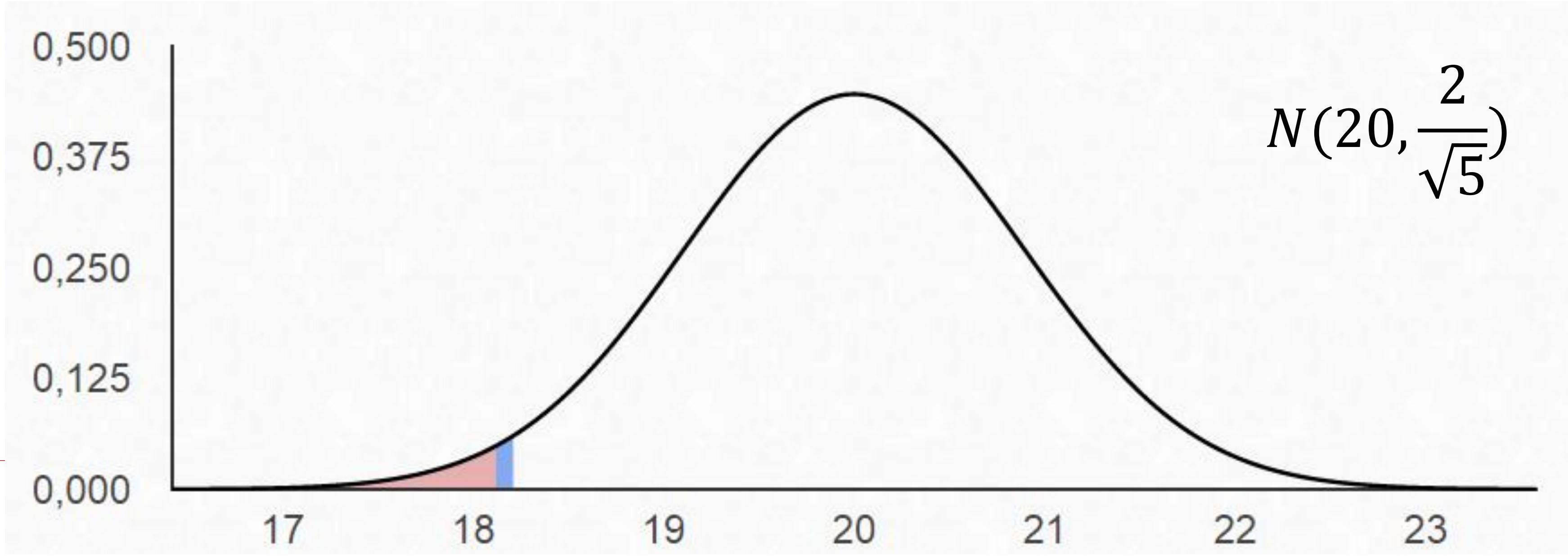
Kritisch gebied = verwerpingsgebied

$\alpha =$ maximaal risico dat je wil nemen op een fout van de eerste soort $= 0.02$

$\mathrm{P}(H_0 \text{ verwerpen} | H_0 \text{ waar})$

Stel dat $\mathrm{H_0}$ waar is: dan is $\mu_{\bar{X}} = 20{,}0$ en $\sigma_{\bar{X}} = 2{,}0/\sqrt{5}$ of ook $\bar{X} \sim N(20, \frac{2}{\sqrt{5}})$

Deze normale verdeling kunnen we tekenen:



Kritisch gebied = verwerpingsgebied

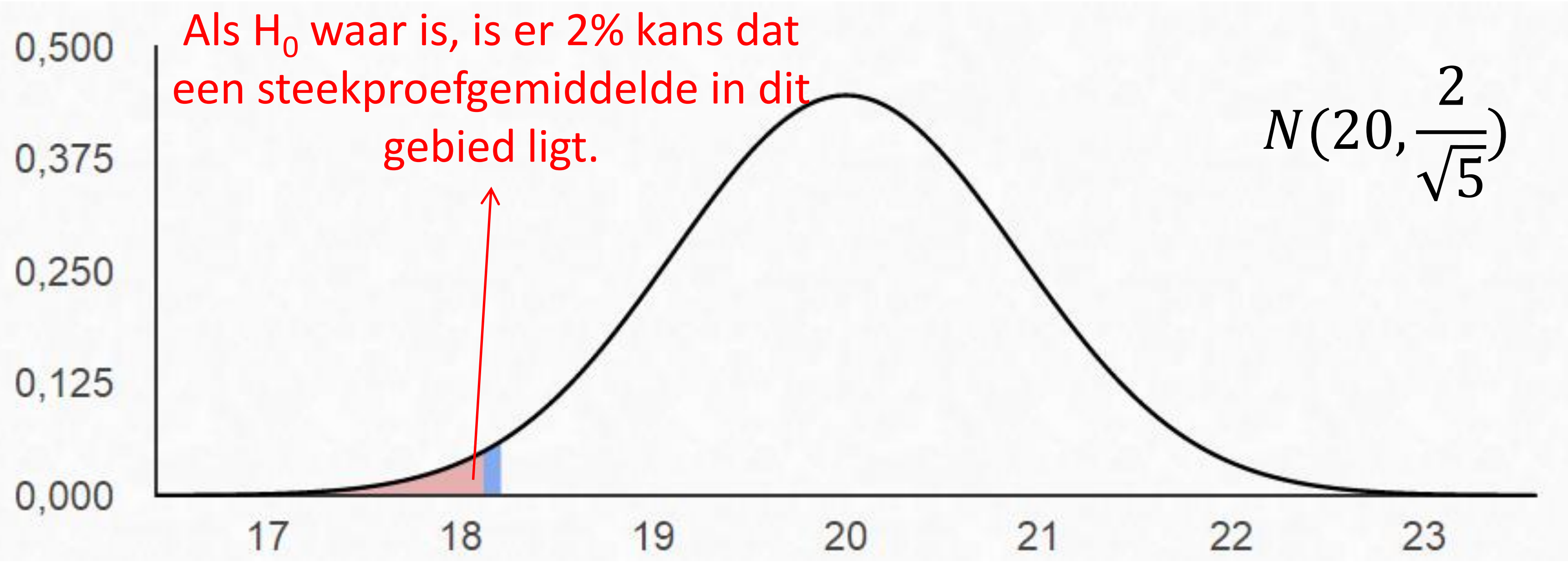
$\alpha =$ maximaal risico dat je wil nemen op een fout van de eerste soort

$= \mathrm{P}(H_0 \text{ verwerpen} | \underline{H_0 \text{ waar}})$

Stel dat $\underline{\text{H}_0 \text{ waar is}}$: dan is $\mu_{\bar{X}} = 20{,}0$ en $\sigma_{\bar{X}} = 2{,}0/\sqrt{5}$ of ook $\bar{X} \sim N(20, \frac{2}{\sqrt{5}})$

H_1 : $\mu < \mu_0$ en $\alpha = 0{,}02$

We tekenen links in de kansdichtheid een staartje met een oppervlakte van 2%. Dit is het verwerpingsgebied. Als het steekproef-gemiddelde in dit gebied ligt, is het laag genoeg om H_0 te verwerpen.



Kritisch gebied = verwerpingsgebied

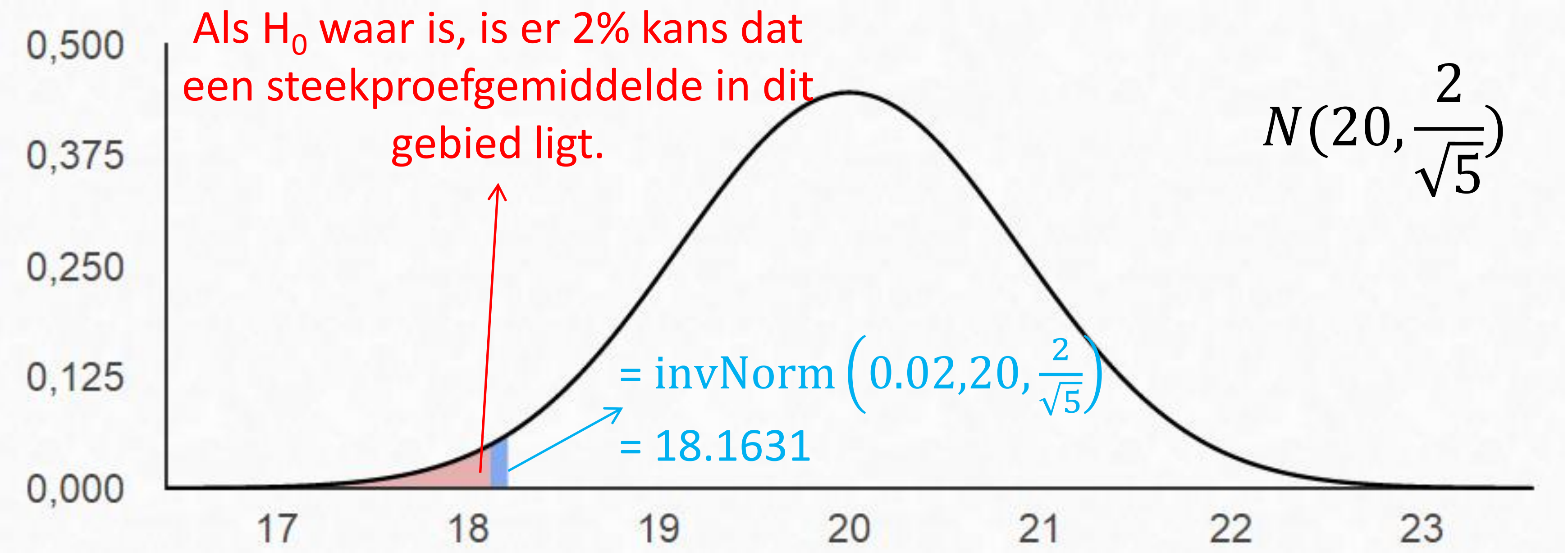
$\alpha =$ maximaal risico dat je wil nemen op een fout van de eerste soort

$= \mathrm{P}(H_0 \text{ verwerpen} | H_0 \text{ waar})$

Stel dat H_0 waar is: dan is $\mu_{\bar{X}} = 20{,}0$ en $\sigma_{\bar{X}} = 2{,}0/\sqrt{5}$ of ook $\bar{X} \sim N(20, \frac{2}{\sqrt{5}})$

Kritisch gebied = $]-\infty, 18.1631]$

We kunnen de grens-waarde van het verwerpingsgebied berekenen met het inverse normaal commando in het rekentoestel.



Besluit

STAP 5

16. Een machine produceert metalen staafjes. De lengte van deze staafjes is normaal

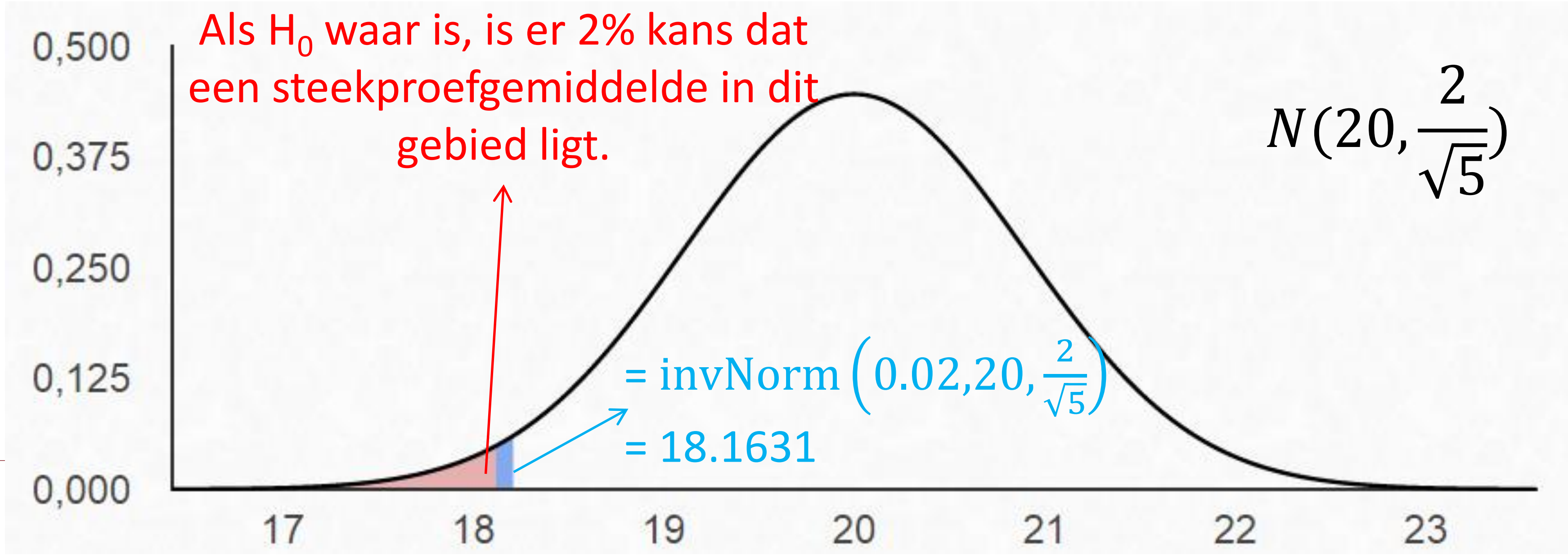
5. **Formuleer je besluit**. Dit wil zeggen: verwerp H_0 op het significantieniveau α, indien de experimentele waarde t in het kritisch gebied ligt van je toets **OF** indien de p-waarde ≤ α.
Zo niet, besluit dat je op basis van je steekproef H_0 niet kunt tegenspreken ("kan behouden").

Ter controle neemt men regelmatig steekproeven van 5 stuks en meet hiervan de lengte.

Bij een steekproef vond men een gemiddelde lengte van 18,6 cm. $= \bar{x} = t$

Deze waarde ligt niet in het verwerpingsgebied, maar gewoon in het aanvaardingsgebied

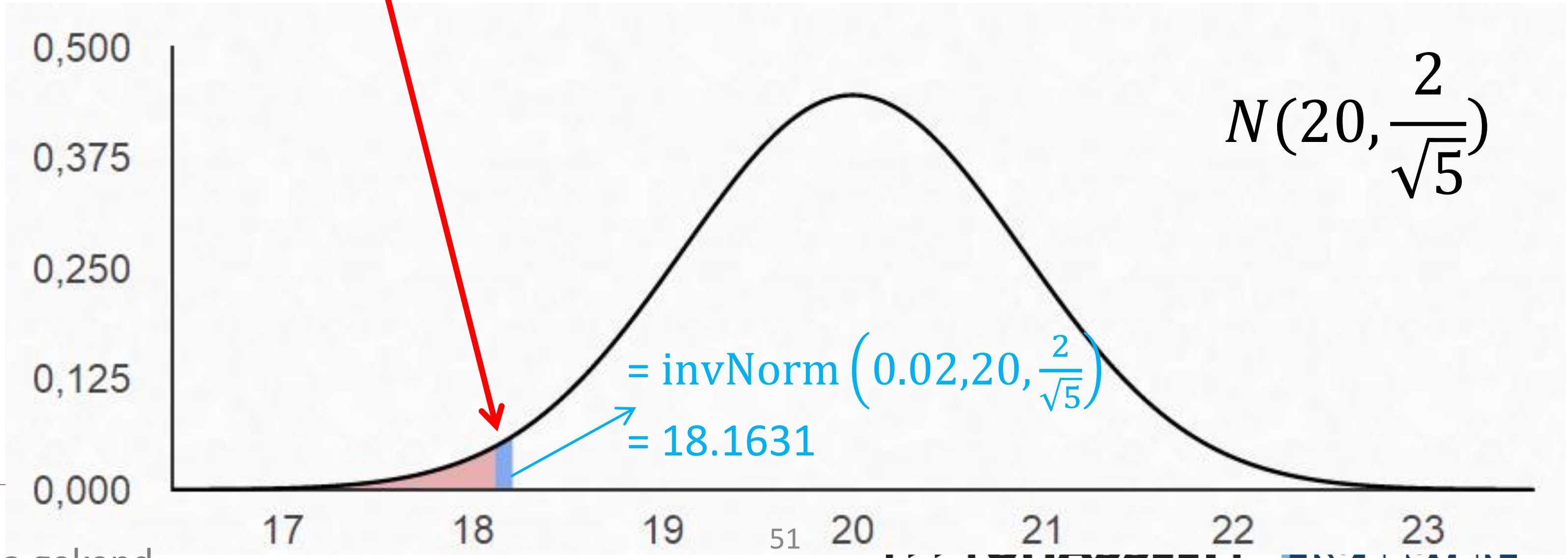
De steekproef had een gemiddelde van 18.6 cm. Dit ligt in het aanvaardingsgebied.
We aanvaarden hier dus de nulhypothese H_0 en concluderen dat de instelwaarde nog altijd 20.0 cm is.



Kritische grens = grenswaarde van het kritisch gebied

16. Een machine produceert metalen staafjes. De lengte van deze staafjes is normaal verdeeld met een vooraf gekende standaardafwijking van 2,0 cm. De gemiddelde lengte heeft als normwaarde 20,0 cm. Er kunnen storingen optreden waardoor het gemiddelde lager wordt, zonder dat de spreiding verandert.
Ter controle neemt men regelmatig steekproeven van 5 stuks en meet hiervan de lengte. Bij een steekproef vond men een gemiddelde lengte van 18,6 cm.
   b. Bij welke gemiddelde lengte van een steekproef van 5 stuks zou de nulhypothese van vraag a) verworpen worden (neem opnieuw α = 2%).

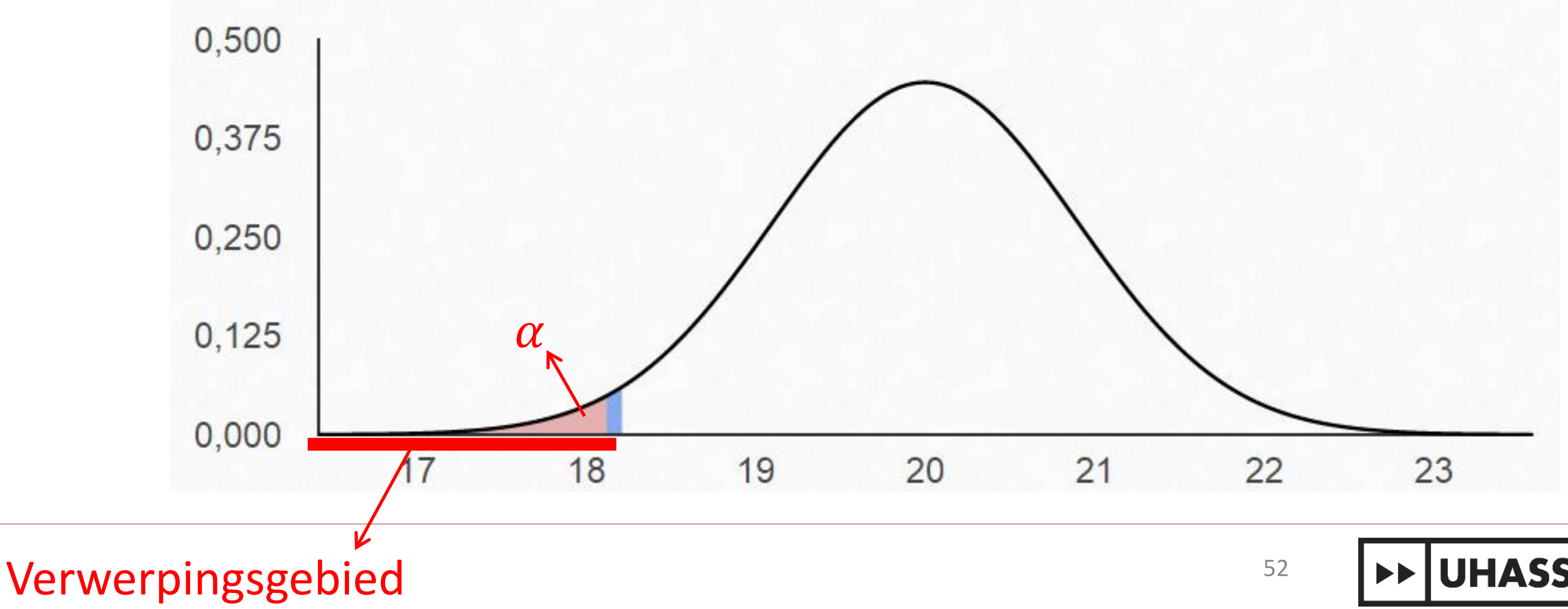
Bij een steekproef met een gemiddelde lengte van 18,16 cm of minder wordt de nulhypothese verworpen.



Verschillende soorten kritische gebieden: linkseenzijdige toets

Een hypothesetoets begint met het opstellen van twee hypothesen: de *nulhypothese* H_0 en de *alternatieve hypothese* H_1. De alternatieve hypothese kan verschillende vormen hebben.

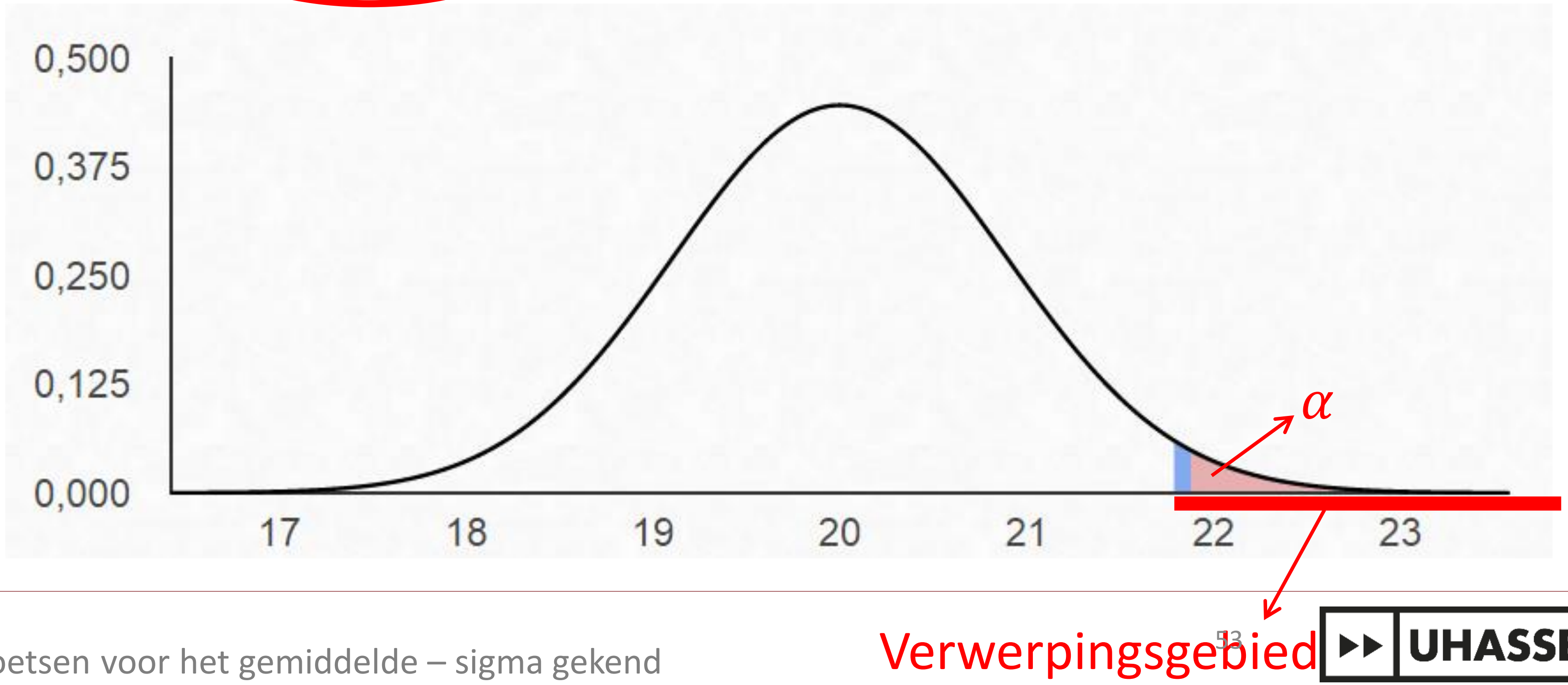
$$\begin{array}{lll} H_0 & : & \mu = \mu_0 \\ H_1 & : & \mu > \mu_0 \end{array} \qquad \begin{array}{lll} H_0 & : & \mu = \mu_0 \\ H_1 & : & \mu < \mu_0 \end{array} \qquad \begin{array}{lll} H_0 & : & \mu = \mu_0 \\ H_1 & : & \mu \neq \mu_0 \end{array}$$



Verschillende soorten kritische gebieden: rechtseenzijdige toets

Een hypothesetoets begint met het opstellen van twee hypothesen: de *nulhypothese* H_0 en de *alternatieve hypothese* H_1. De alternatieve hypothese kan verschillende vormen hebben.

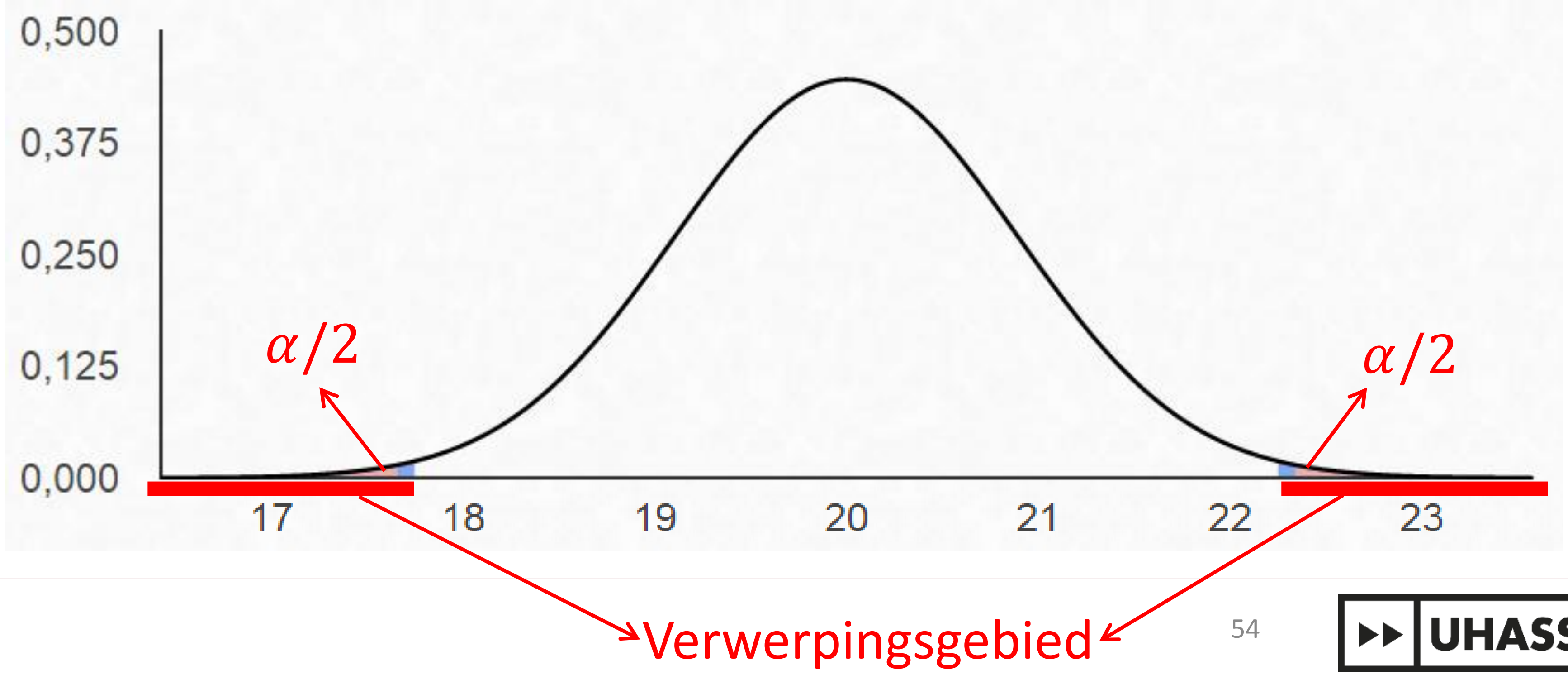
$$H_0 : \mu = \mu_0 \qquad H_0 : \mu = \mu_0 \qquad H_0 : \mu = \mu_0$$
$$H_1 : \mu > \mu_0 \qquad H_1 : \mu < \mu_0 \qquad H_1 : \mu \neq \mu_0$$



Verschillende soorten kritische gebieden: tweezijdige toets

Een hypothesetoets begint met het opstellen van twee hypothesen: de *nulhypothese* H_0 en de *alternatieve hypothese* H_1. De alternatieve hypothese kan verschillende vormen hebben.

$$\begin{array}{lcl} H_0 & : & \mu = \mu_0 \\ H_1 & : & \mu > \mu_0 \end{array} \qquad \begin{array}{lcl} H_0 & : & \mu = \mu_0 \\ H_1 & : & \mu < \mu_0 \end{array} \qquad \begin{array}{lcl} H_0 & : & \mu = \mu_0 \\ H_1 & : & \mu \neq \mu_0 \end{array}$$



Kans op type I fout?

16. Een machine produceert metalen staafjes. De lengte van deze staafjes is normaal verdeeld met een vooraf gekende standaardafwijking van 2,0 cm. De gemiddelde lengte heeft als normwaarde 20,0 cm. Er kunnen storingen optreden waardoor het gemiddelde lager wordt, zonder dat de spreiding verandert.
Ter controle neemt men regelmatig steekproeven van 5 stuks en meet hiervan de lengte. Bij een steekproef vond men een gemiddelde lengte van 18,6 cm.
    c. Hoe groot is de kans op een fout van de **eerste** soort bij de net uitgevoerde hypothesetoets?

Deze hebben we vooraf vastgelegd op 2%

= significantieniveau α

p-waarde

Interpreteer deze zin als volgt: stel dat we ons niets zouden aantrekken van het significantie-niveau α = 0,02 en op basis van het lage steekproefgemiddelde direct besluiten: we verwerpen de nulhypothese.

16. Een machine produceert metalen staafjes. De lengte van deze staafjes is normaal verdeeld met een vooraf gekende standaardafwijking van 2,0 cm. De gemiddelde lengte heeft als normwaarde 20,0 cm. Er kunnen storingen optreden waardoor het gemiddelde lager wordt, zonder dat de spreiding verandert.

Ter controle neemt men regelmatig steekproeven van 5 stuks en meet hiervan de lengte. Bij een steekproef vond men een gemiddelde lengte van 18,6 cm.

d. Stel dat je de nulhypothese zou verwerpen op basis van deze steekproef. Wat is dan de kans dat je een verkeerde conclusie neemt?

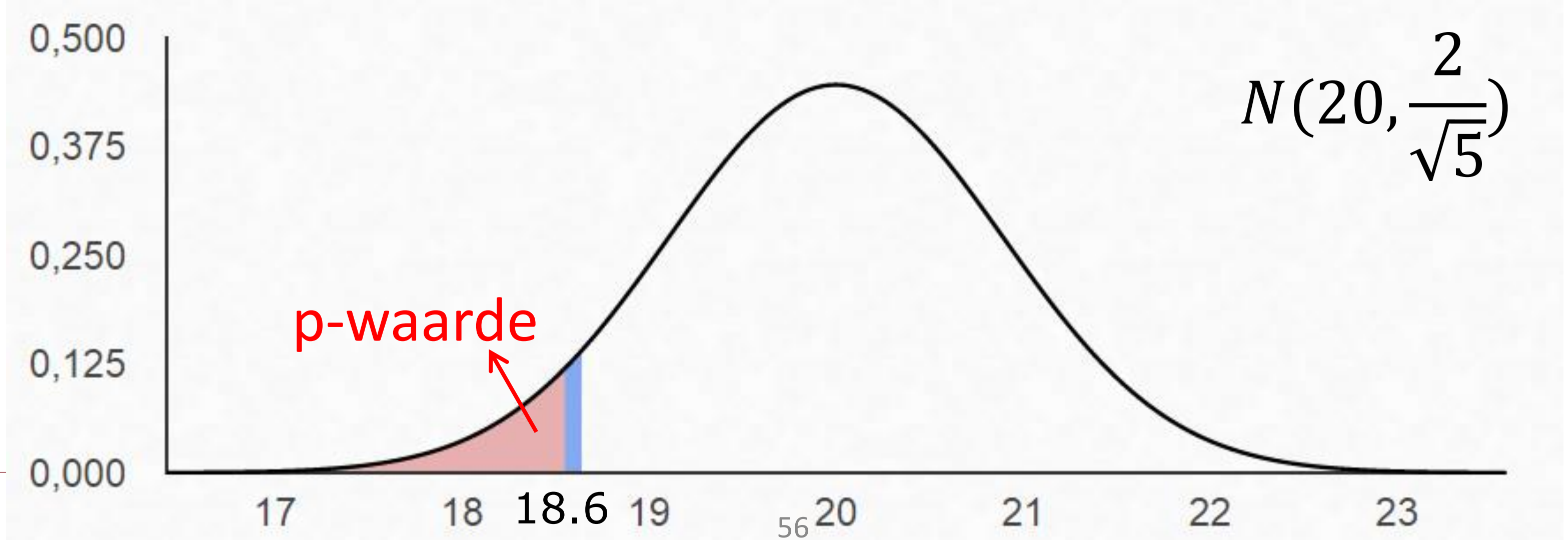
P(verkeerde conclusie)

$$= P(H_0 \text{ verwerpen} | H_0 \text{ waar})$$
$$= P(\bar{X} \leq 18.6 \mid \mu = 20)$$

We verwerpen als $\bar{X} = 18.6$ of minder

Als H_0 waar is dan is $\mu = 20$

$$= \text{normCdf}(-\infty, 18.6, 20, \frac{2}{\sqrt{5}})$$
$$= 0.058762$$
$$= \text{p-waarde}$$



p-waarde

p-waarde = de kans op dezelfde of een nog extremere uitkomst als de nulhypothese waar is

5. **Formuleer je besluit.** Dit wil zeggen: verwerp H_0 op het significantieniveau α, indien de experimentele waarde t in het kritisch gebied ligt van je toets **OF** indien de p-waarde ≤ α.
   Zo niet, besluit dat je op basis van je steekproef H_0 niet kunt tegenspreken ("kan behouden").

Zéér kleine p-waarde ($\leq \alpha$) :

$\Longrightarrow$ Zéér kleine kans op een extremere uitkomst/steekproef

$\Longrightarrow$ Deze steekproef was dus extreem (t.o.v. H_0)

$\Longrightarrow$ Het is gerechtvaardigd om H_0 te verwerpen

In het voorbeeld is de p-waarde nog niet klein genoeg om te verwerpen! ($0.058762 > 0.02$)

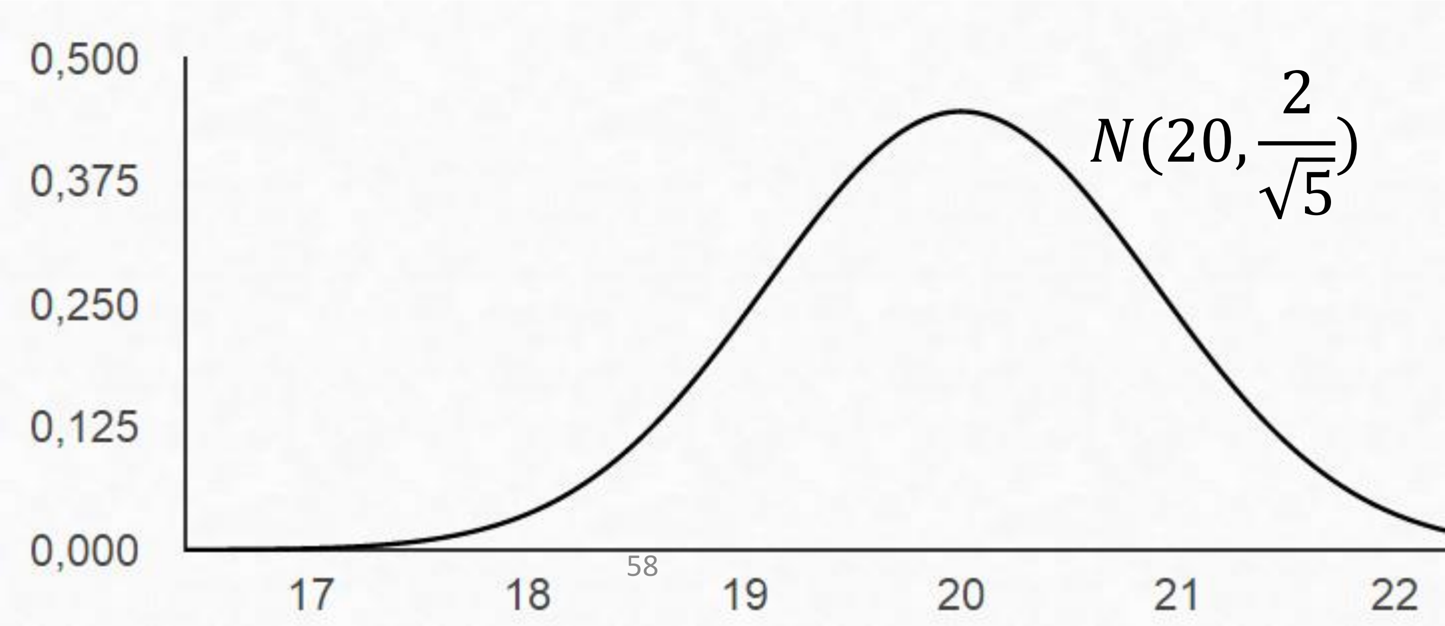
Fout van de tweede soort

e. Bereken in deze zelfde context de kans op een fout van de **tweede** soort onder de veronderstelling dat de instelwaarde 19.0 cm is.

$$= \mathrm{P}(H_0 \text{ aanvaarden} | H_0 \text{ niet waar})$$

De nulhypothese is nu niet waar dus we hebben een alternatief nodig waarvoor we de kansdichtheid kunnen opstellen.

$$\bar{X} \sim N\left(19, \frac{2}{\sqrt{5}}\right)$$



Fout van de tweede soort

e. Bereken in deze zelfde context de kans op een fout van de **tweede** soort onder de veronderstelling dat de instelwaarde 19.0 cm is.

$\boxed{= \mathrm{P}(H_0 \text{ aanvaarden} | H_0 \text{ niet waar})}$

$\bar{X} \sim N(19, \frac{2}{\sqrt{5}})$

$$\mathrm{P}(\text{type II fout}) = \beta$$

$$= P(H_0 \text{ aanvaarden} \mid H_0 \text{ niet waar})$$

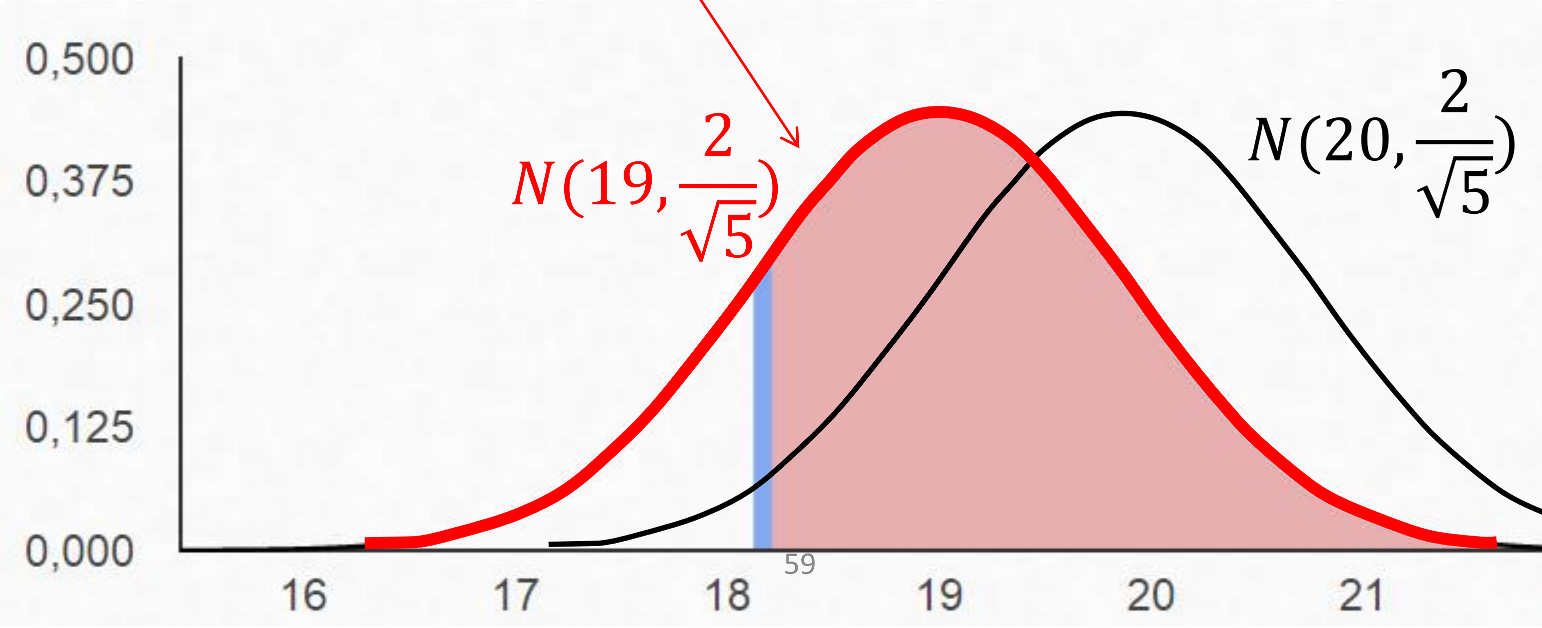
$$= P(H_0 \text{ aanvaarden} \mid H_1 \text{ waar})$$

$$= P(\bar{X} > 18.1631 \mid \mu = 19.0)$$

We aanvaarden H_0 als $\bar{x}$ in het aanvaardingsgebied ligt

$$= \mathrm{normCdf}\left(18.1631, \infty, 19, \frac{2}{\sqrt{5}}\right)$$

$$= 0.825282$$



Onderscheidingsvermogen 1-β

Het **onderscheidingsvermogen** van een toets is:

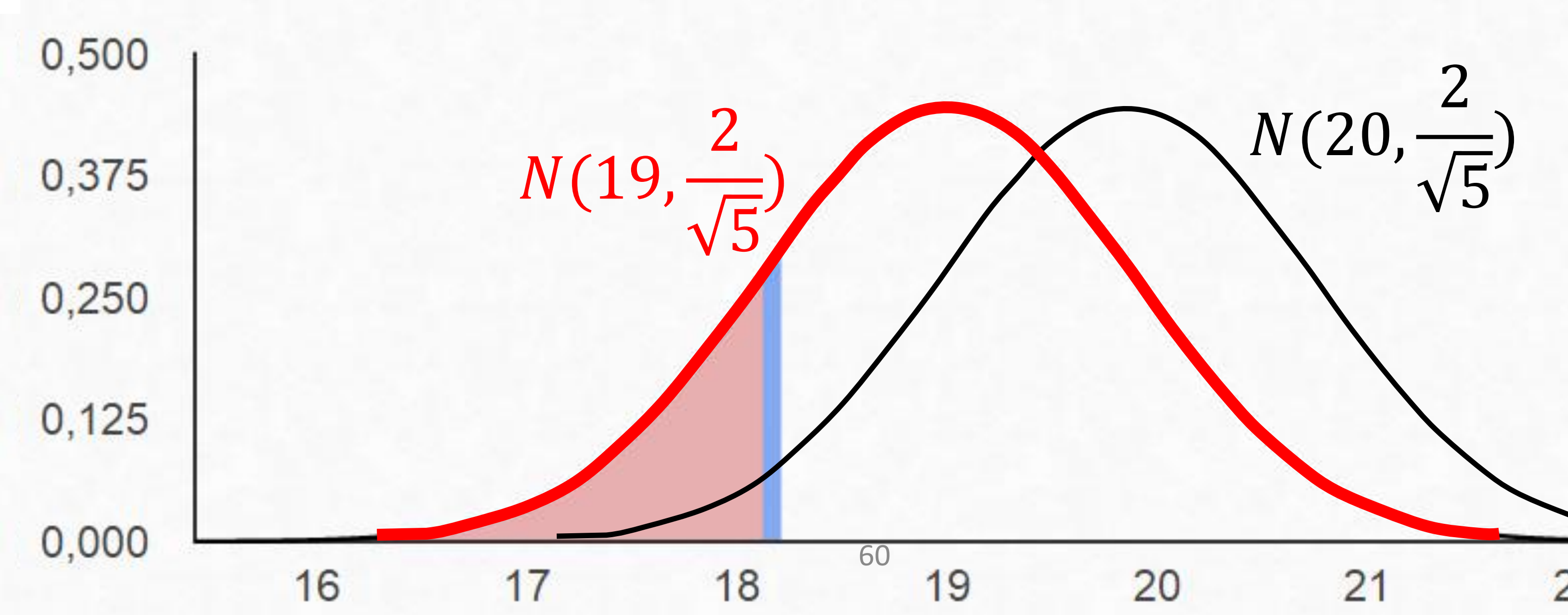
$$1-\beta = \mathrm{P}(H_0 \text{ verwerpen} | H_0 \text{ niet waar})$$
$$= P(H_0 \text{ verwerpen} \mid H_1 \text{ waar})$$

Dit is gelijk aan $1 - \mathrm{P}(\text{fout van de tweede soort})$.

$$1-\beta = P(\bar{X} < 18.1631 \mid \mu = 19.0)$$

We verwerpen H_0 als $\bar{x}$ in
het verwerpingsgebied ligt

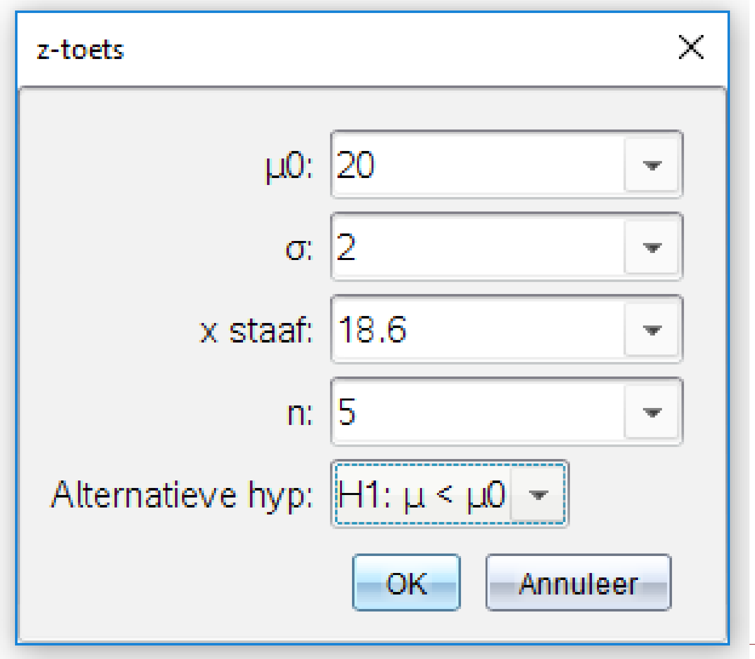
$$= \mathrm{normCdf}\left(-\infty, 18.1631, 19, \frac{2}{\sqrt{5}}\right)$$
$$= 0.174718$$



z-toets met de TI-Nspire

Menu – 6: Statistieken – 7: Statistiektoetsen – 1: z-toets...

- Kies ‘Gegevens’ als je alle uitkomsten van een steekproef hebt gekregen. Vul bij ‘Lijst’ de kolomnaam van de uitkomsten in.
- Voor het voorbeeld in deze les kies je ‘Stats’ en vul je in:



zTest 20,2,18.6,5,⁻1: *stat.results*

"Titel"	"z−toets"
"Alternatieve hypothese"	"μ < μ0"
"z"	⁻1.56525
"PVal"	0.058762
"$\bar{x}$"	18.6
"n"	5.
"σ"	2.

Het rekentoestel gebruikt altijd de p-waarde

Onderstaande grafieken horen bij een hypothesetoets waar de nulhypothese H_0: $\mu = 123$ getest wordt tegen de alternatieve hypothese H_1: $\mu < 123$. De kritische grens van de hypothesetoets ligt op 120. Veronderstel dat het werkelijke populatiegemiddelde hier gelijk is aan $\mu = 115$. Bij welke grafiek is voor deze hypothesetoets het onderscheidingsvermogen aangeduid (= donkerder van kleur)?