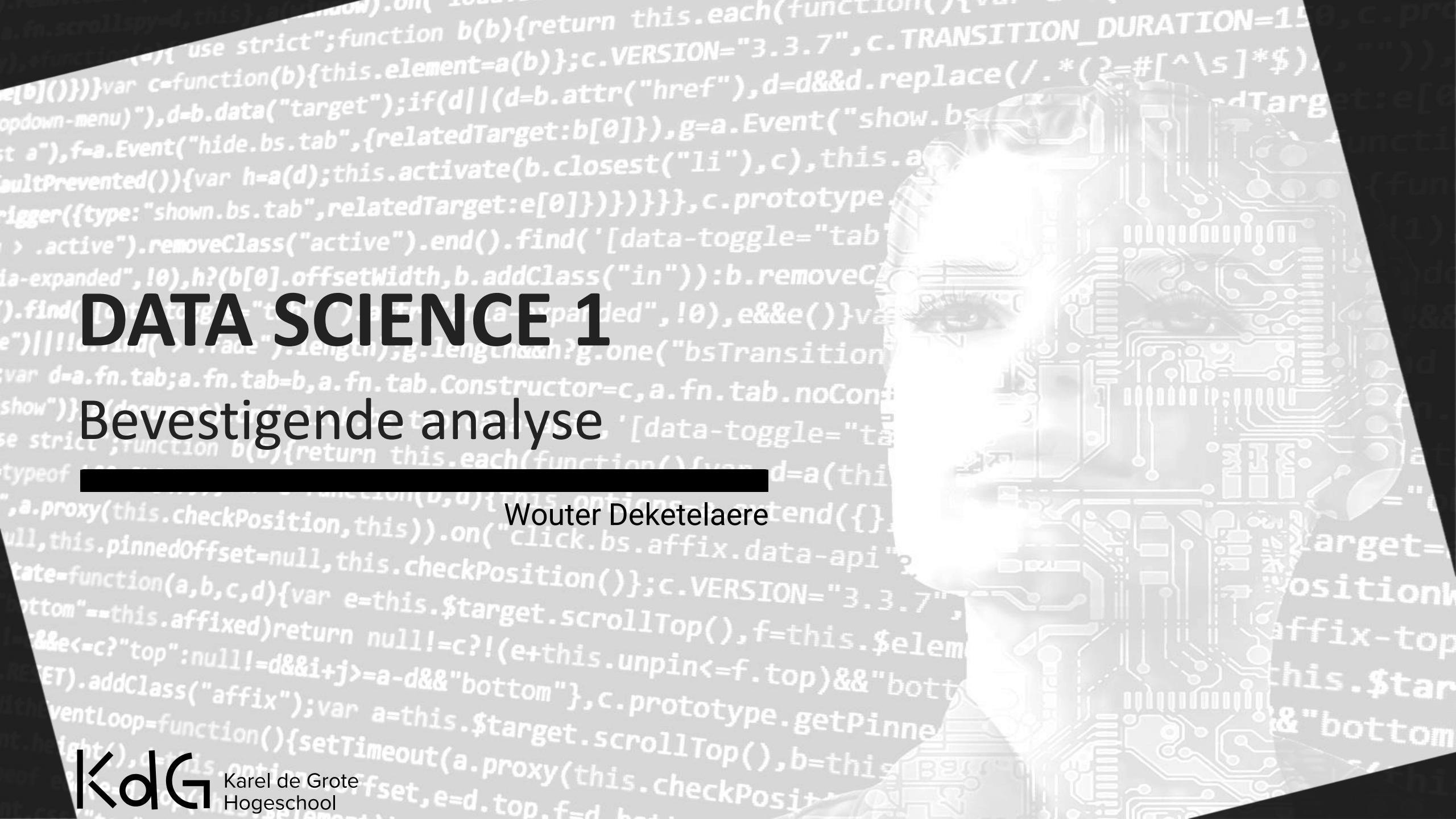
DATA SCIENCE 1

Bevestigende analyse

Wouter Deketelaere



KdG Karel de Grote Hogeschool

WAAROM

Waarom doen we dit?

Wat willen we bereiken?

Wat willen we als resultaat?

Hoe doen we dit?

Welke stappen nemen we

Welke technieken gebruiken we

HOE

LEERDOELEN

Na deze les …

- ken je wat het verschil is tussen een **populatie** en een **steekproef**
- weet je wat bedoeld wordt met een **steekproef –of populatieparameter**
- weet je wat bedoeld wordt met de **centrale limietstelling**
- weet je wat het **gestandaardiseerde steekproefgemiddelde** is
- weet je wat een **betrouwbaarheidsinterval** is
- kan je uitleggen hoe je een **betrouwbaarheidsinterval opstelt**
- weet je wat een **hypothesetoets** is
- ken je verschillende **soorten hypothesetoetsen**
- weet je wat een **nul- en alternatieve hypothese** is
- ken je het verschil tussen een **eenzijdige** en **tweezijdige** toets
- weet je wat een **significantieniveau** α betekent
- weet je wat de p**-waarde** betekent
- weet je wanneer iets **statistisch significant** is

LEERDOELEN

Na deze les kan je …

- **handmatig** een betrouwbaarheidsinterval berekenen
- een betrouwbaarheidsinterval berekenen in **GeoGebra**
- **handmatig** een bewering toetsen over een populatie m.b.v. 1 steekproef
- **handmatig** een bewering toetsen over het gemiddelde van 2 of meer steekproeven
- **handmatig** een bewering toetsen over de uitzonderlijkheid van een bepaalde verdeling
- een bewering toetsen over een populatie m.b.v. 1 steekproef in **GeoGebra** of **Orange**
- een bewering toetsen over het gemiddelde van 2 of meer steekproeven in **GeoGebra** of **Orange**
- een bewering toetsen over de uitzonderlijkheid van een bepaalde verdeling in **GeoGebra**

BEVESTIGENDE ANALYSE

I **WELK PROBLEEM WILLEN WE OPLOSSEN**
Kunnen we garanties geven?

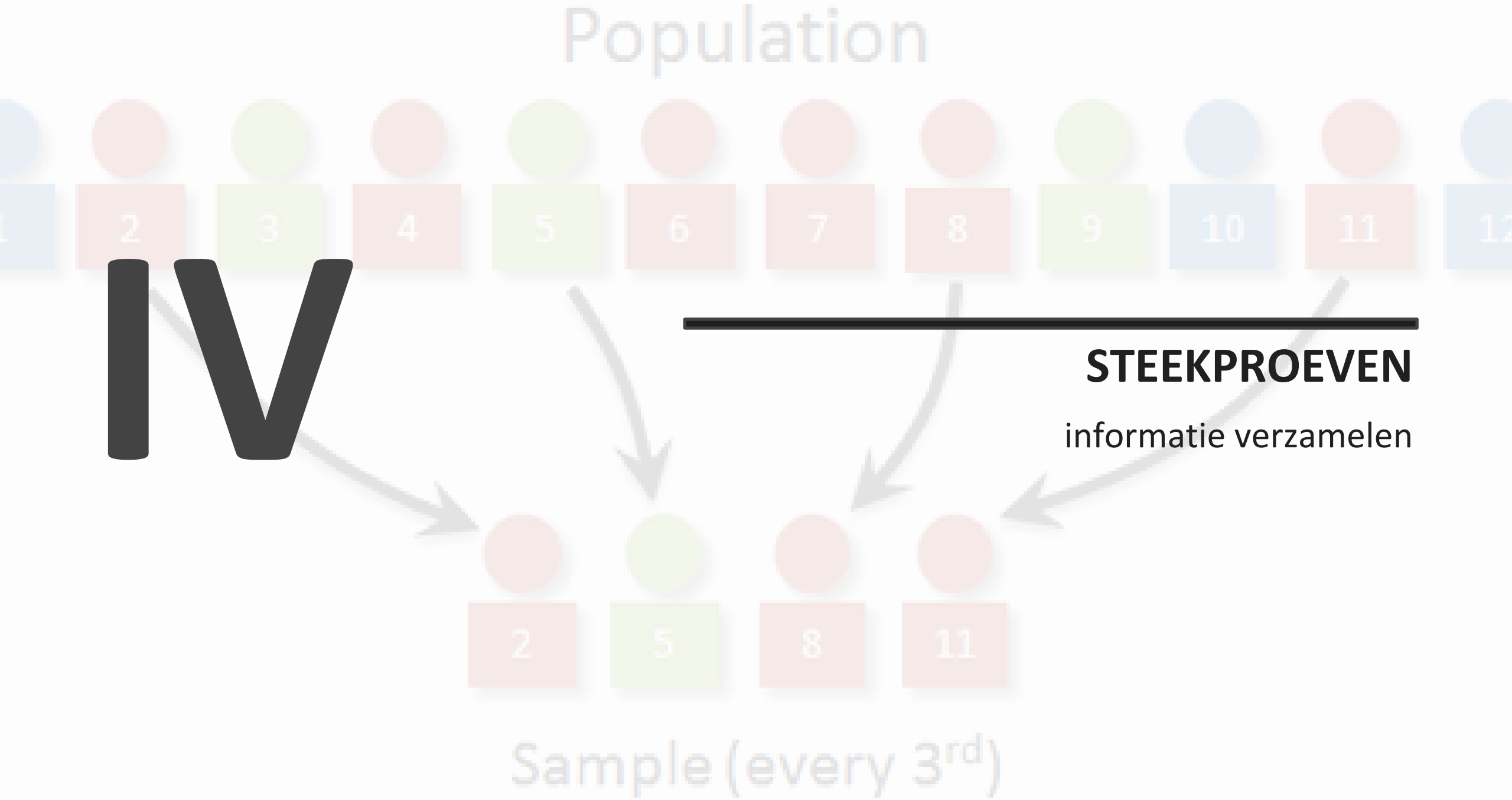
II **KANSREKENEN**
Wat is de kans?

III **KANSVERDELINGEN**
Het grotere plaatje

IV **STEEKPROEVEN**
Informatie verzamelen

V **BETROUWBAARHEID**
Grenzen stellen

VI **HYPOTHESES TOETSEN**
Beweringen controleren



IV

STEEKPROEVEN

informatie verzamelen

BEVESTIGENDE ANALYSE

TOEVAL OF NIET?

- Is dit patroon toeval?
- Wat is de kans dat we dit waarnemen?
- Welke echte conclusies kunnen we hieruit trekken?

KANSREKENEN

- Grondbeginselen van waarschijnlijkheidstheorie
- Enkele wetten van waarschijnlijkheid

KANSVERDELINGEN

- Normale verdeling
- Standaard normale verdeling
- Studentverdeling
- Chi-kwadraat verdeling

STEEKPROEVEN

- Populatie vs. steekproef
- Eigenschappen van steekproeven

BETROUWBAARHEIDSINTERVALLEN

- Binnen welke grenzen vallen nieuwe metingen?

HYPOTHESETESTS

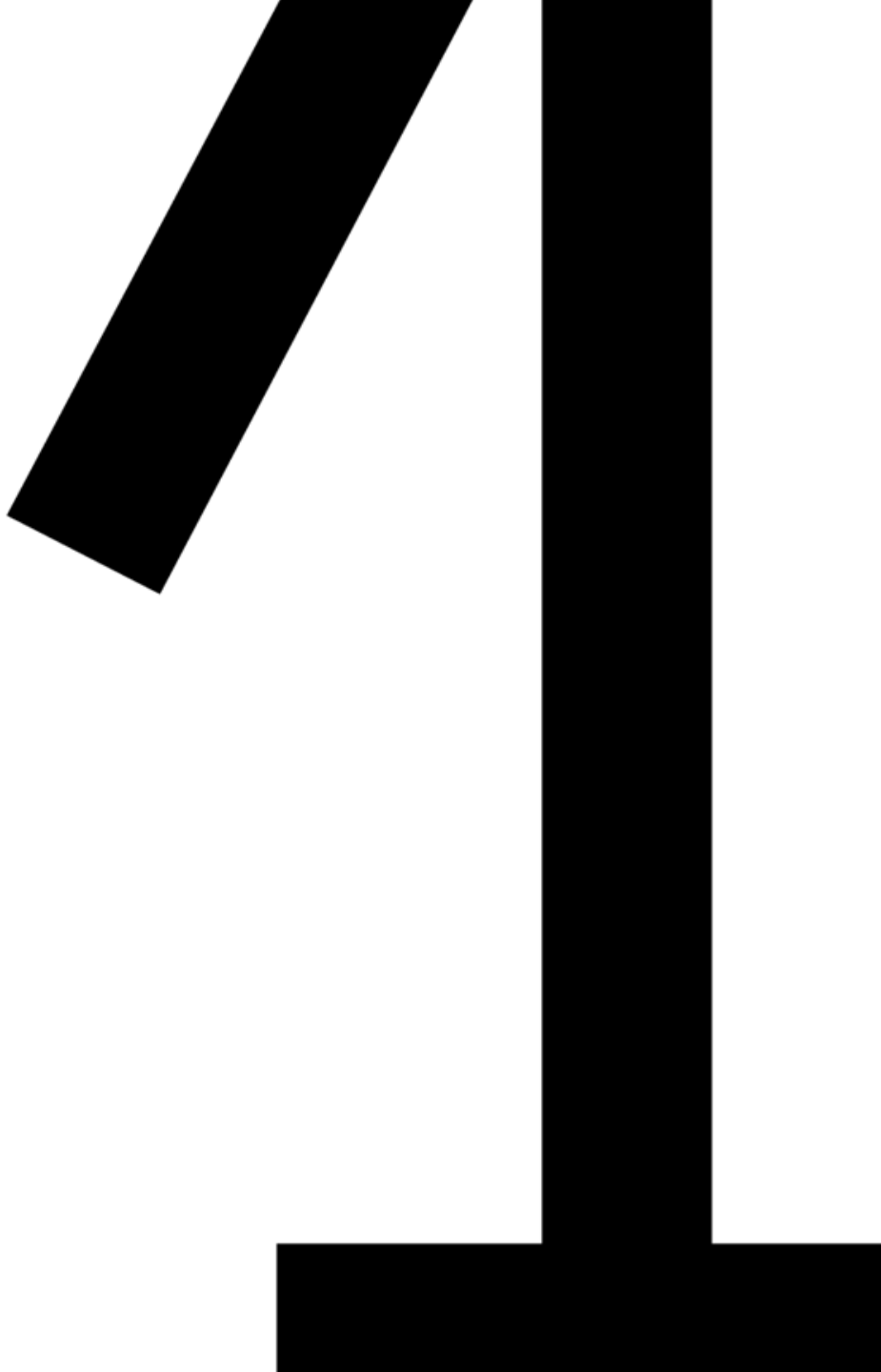
- Kunnen we beweringen verifiëren?
- Wanneer kunnen we beweringen weerleggen?

TECHNIEKEN

- Statistieken

WETEN WE ZEKER DAT HET GEBEURT?

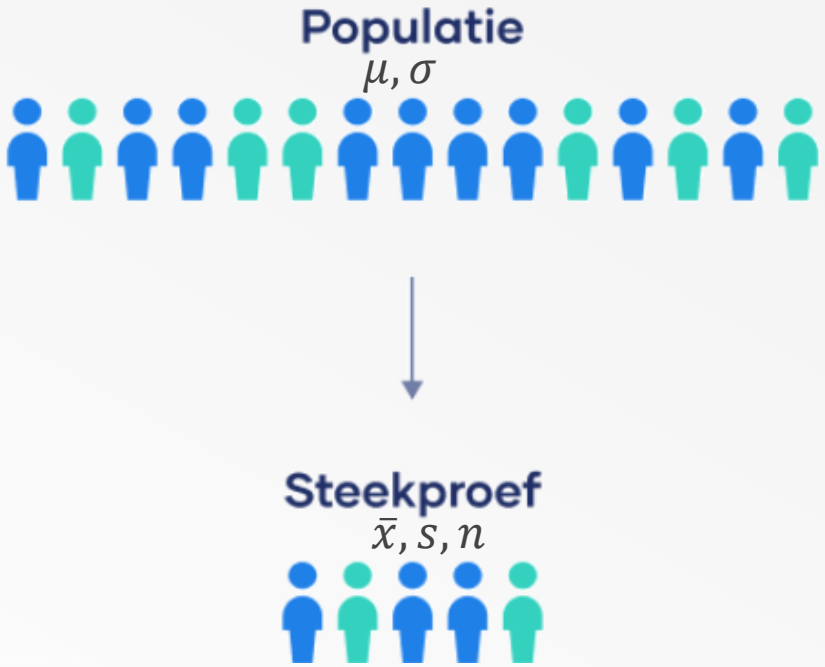
WAT IS EEN STEEKPROEF?



WAT IS EEN STEEKPROEF?

Populatie

- Verzameling van objecten waarop het onderzoek zich richt.
- Parameters
  - Populatiegemiddelde μ
  - Populatiestandaardafwijking σ



Steekproef

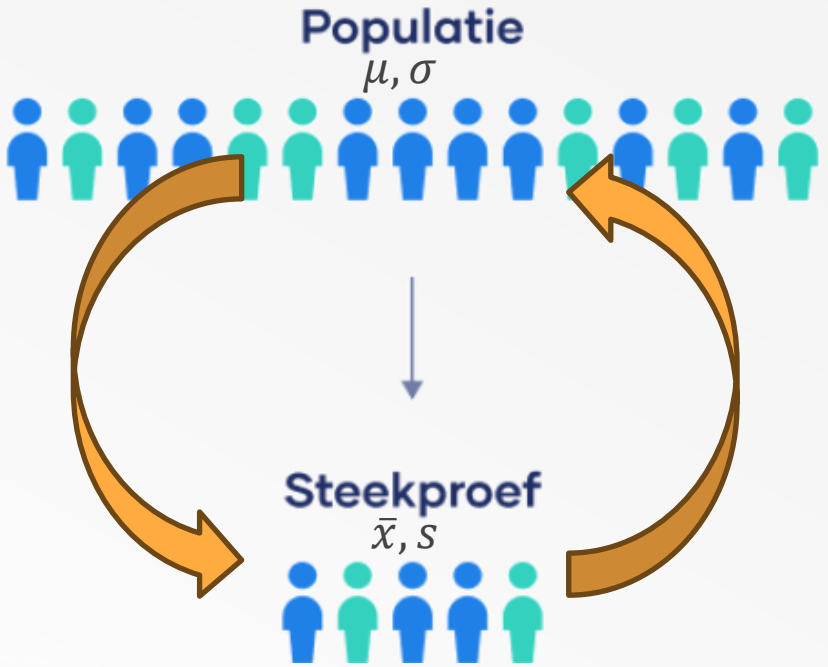
- Willekeurige selectie van objecten uit de populatie.
- Parameters
  - Steekproefgemiddelde $\bar{x}$,
  - Steekproefstandaardafwijking s
  - Steekproefgrootte n

WAAROM EEN STEEKPROEF?

Doel
We willen op basis van een steekproef uitspraken doen over de populatie.

Betrouwbaarheidsinterval
Grenzen bepalen voor μ en σ o.b.v. $\bar{x}$ en s.

Aanvaardingsinterval
Bewering over populatie controleren door zelf meting (steekproef) te doen.



2

EIGENSCHAPPEN VAN STEEKPROEVEN

WAT IS EEN STEEKPROEF?

Nieuwe definitie

Een steekproef nemen is eigenlijk een toevalsexperiment!

$\bar{x}$ en s zijn onderhevig aan toeval.

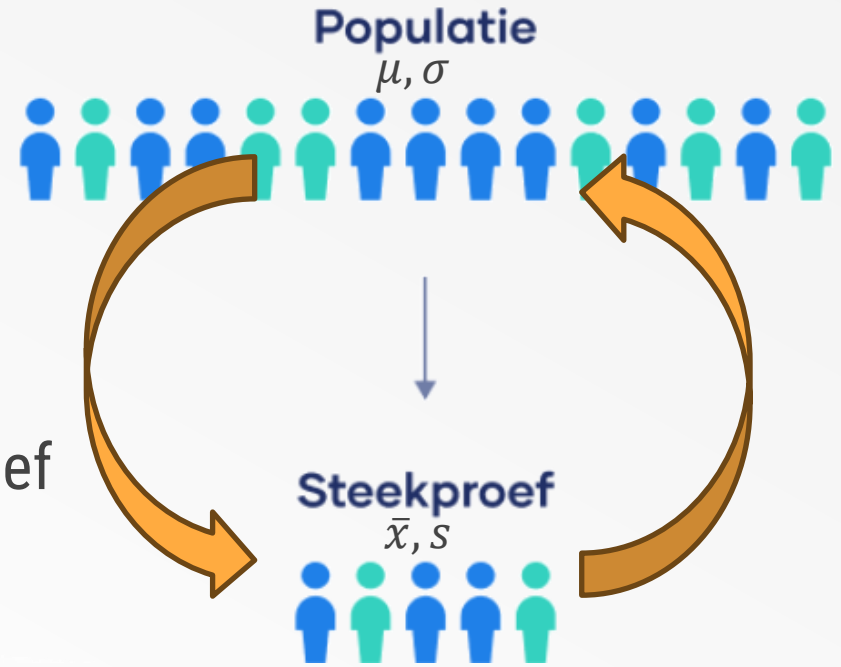
→ **continue toevalsvariabelen**

Toevalsvariabelen

$\bar{X}$ = Het steekproefgemiddelde $\bar{x}$ van een steekproef

S = Het steekproefstandaardafwijking s van een steekproef

$T = \frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}}$ = gestandaardiseerd steekproefgemiddelde



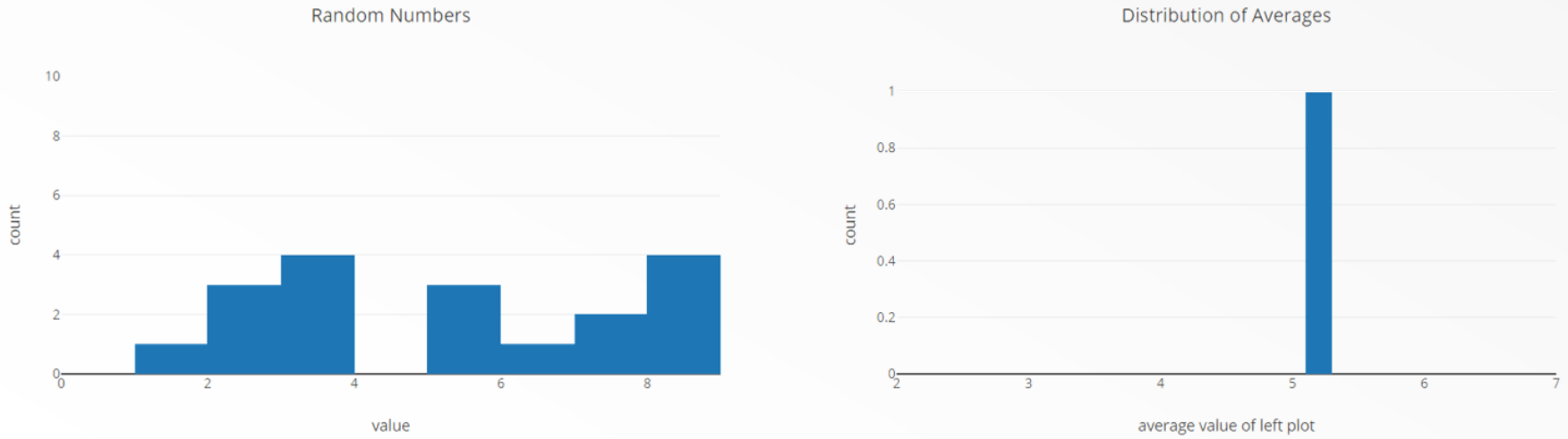
STEEKPROEFVERDELING

Centrale limietstelling

Geschaalde steekproefgemiddelden zijn normaal verdeeld als steekproefgrootte n groot is.

Verdeling $\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$ en dus $T \sim N(0,1)$

Gemiddelde van steekproefgemiddelden wordt gelijk aan populatiegemiddelde μ
Variantie van steekproeven wordt gelijk aan populatievariantie gedeeld door steekproefgrootte n



Generate 20 random numbers from 0 and 9. Find their average. Repeat 1000 times. The averages will approximate a normal distribution (bell curve) centered at 4.5.

STEEKPROEFVERDELING

Problemen

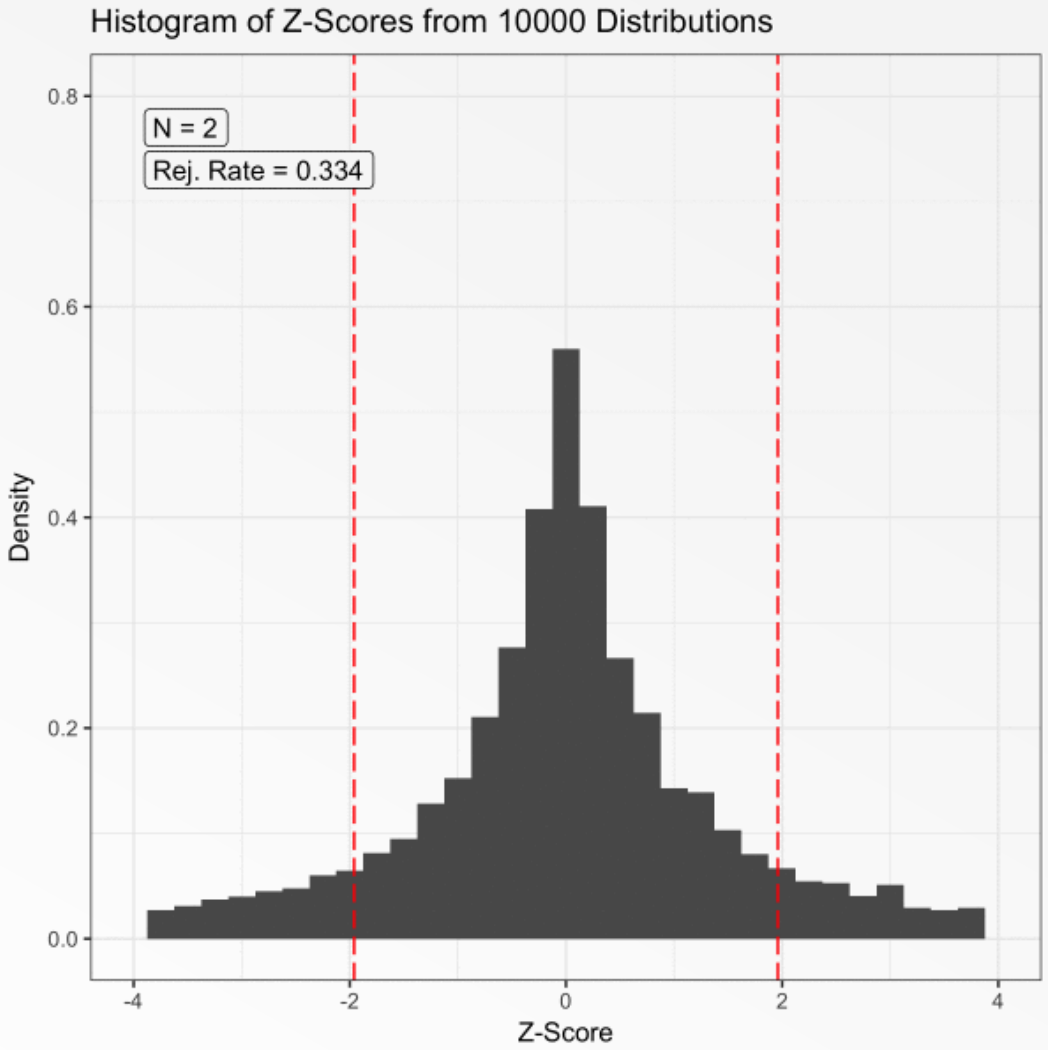
1. We kennen μ en σ niet!
   → **Benaderen** door $\bar{x}$ en s

2. Maar voor kleine steekproef ($n < 30$)
   → s slechte benadering voor σ
   → voor kleine n is de curve spitser

Oplossing

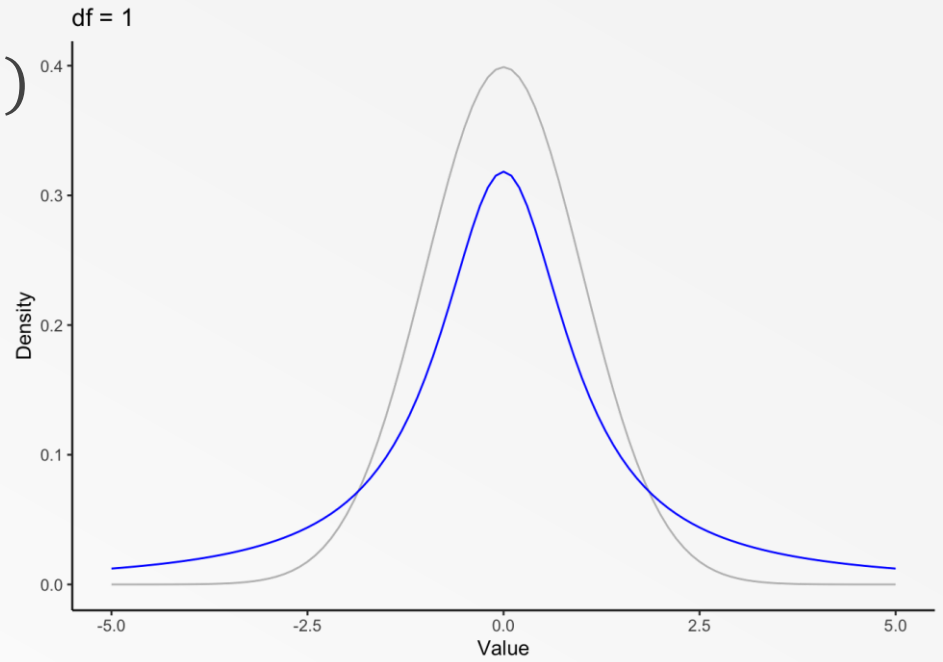
Andere verdeling gebruiken!

→ Studentverdeling $t(\nu)$



STUDENTVERDELING

Studentverdeling
Lijkt heel hard op standaardnormale verdeling $N(0,1)$



Notatie
$X \sim t(\nu)$

Definitie van $t(\nu)$

ν bepaalt uitzicht van verdeling

ν aantal vrijheidsgraden (degrees of freedom)

Hoe groter ν hoe meer de verdeling lijkt op normaalverdeling

STEEKPROEFVERDELING

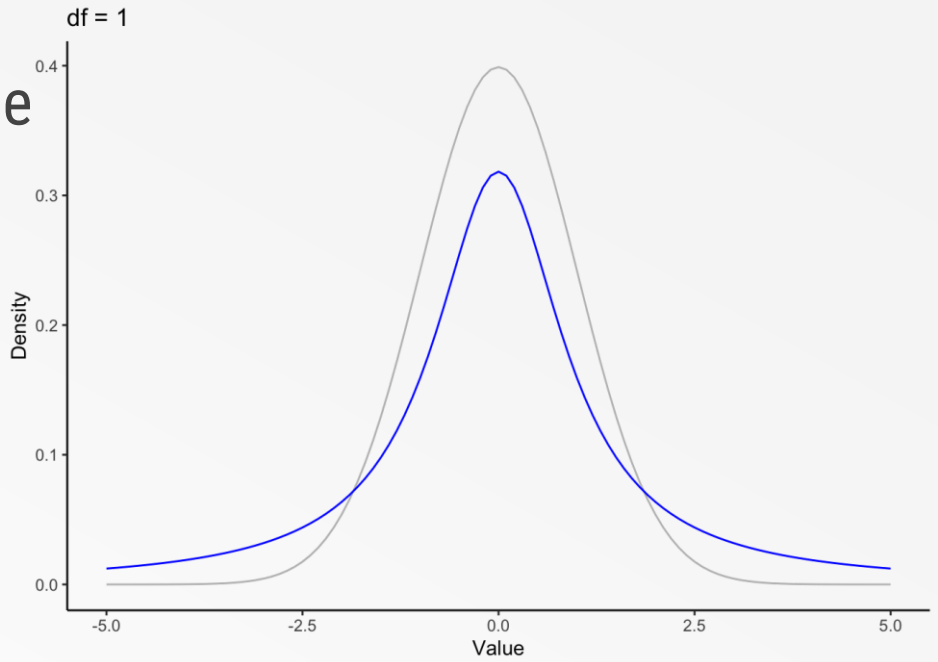
Toevalsvariabele

$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}}$ = gestandaardiseerde steekproefgemiddelde

Verdeling

$T \sim t(\nu)$ (= Studentverdeling)

$\nu = \mathrm{n} - 1$ altijd gelijk aan steekproefgrootte - 1



Correcter voor kleine steekproeven dan normale verdeling.
Even correct voor grote steekproeven!

→ Altijd deze verdeling gebruiken

BEVESTIGENDE ANALYSE

I **WELK PROBLEEM WILLEN WE OPLOSSEN**
Kunnen we garanties geven?

II **KANSREKENEN**
Wat is de kans?

III **KANSVERDELINGEN**
Het grotere plaatje

IV **STEEKPROEVEN**
Informatie verzamelen

V **BETROUWBAARHEID**
Grenzen stellen

VI **HYPOTHESES TOETSEN**
Beweringen controleren

V

BETROUWBAARHEID

GRENZEN STELLEN

BEVESTIGENDE ANALYSE

TOEVAL OF NIET?

- Is dit patroon toeval?
- Wat is de kans dat we dit waarnemen?
- Welke echte conclusies kunnen we hieruit trekken?

KANSREKENEN

- Grondbeginselen van waarschijnlijkheidstheorie
- Enkele wetten van waarschijnlijkheid

KANSVERDELINGEN

- Normale verdeling
- Standaard normale verdeling
- Studentverdeling
- Chi-kwadraat verdeling

STEEKPROEVEN

- Populatie vs. steekproef
- Eigenschappen van steekproeven

BETROUWBAARHEIDSINTERVALLEN

- Binnen welke grenzen vallen nieuwe metingen?

HYPOTHESETESTS

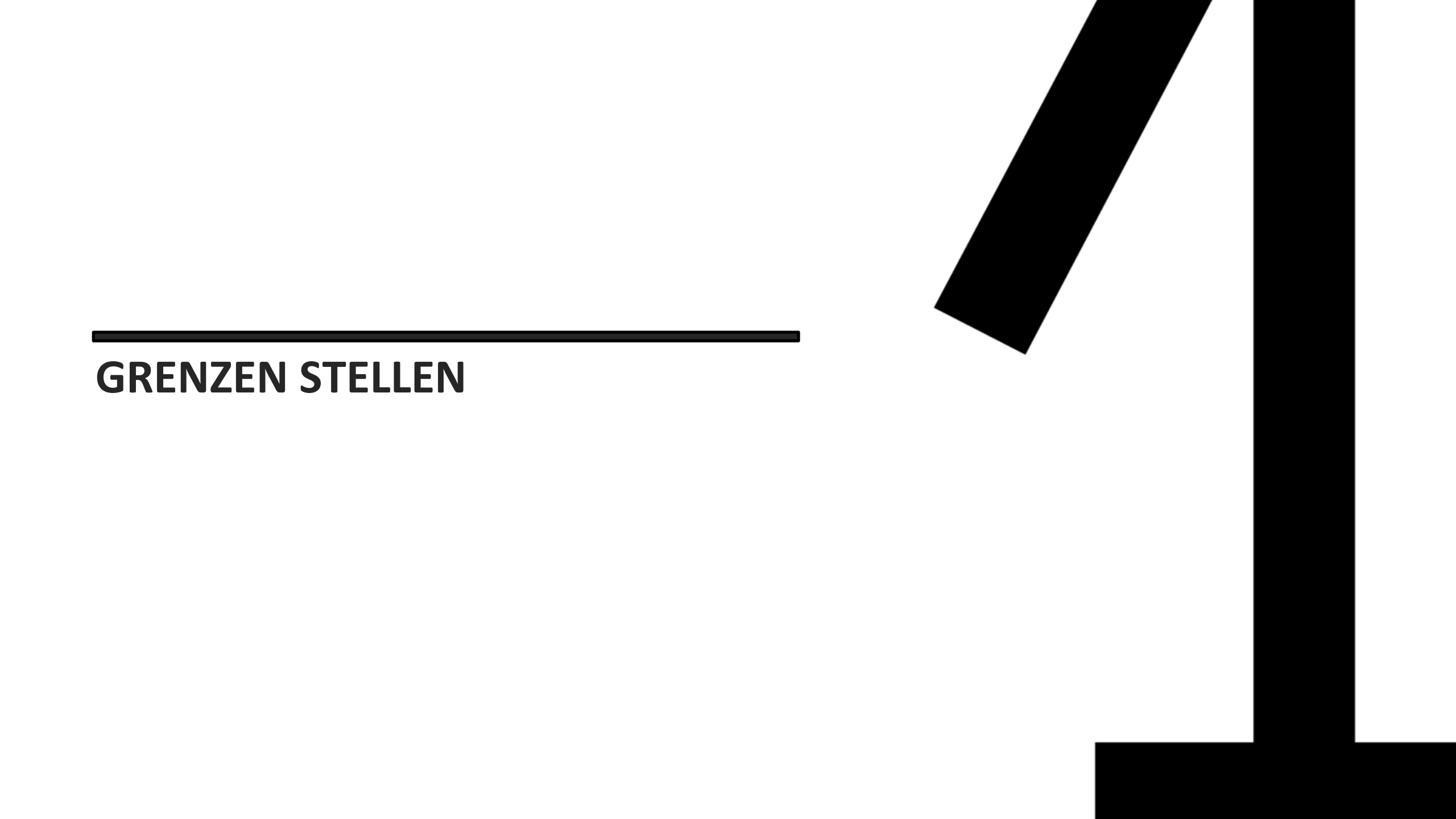
- Kunnen we beweringen verifiëren?
- Wanneer kunnen we beweringen weerleggen?

TECHNIEKEN

- Statistieken

WETEN WE ZEKER DAT HET GEBEURT?

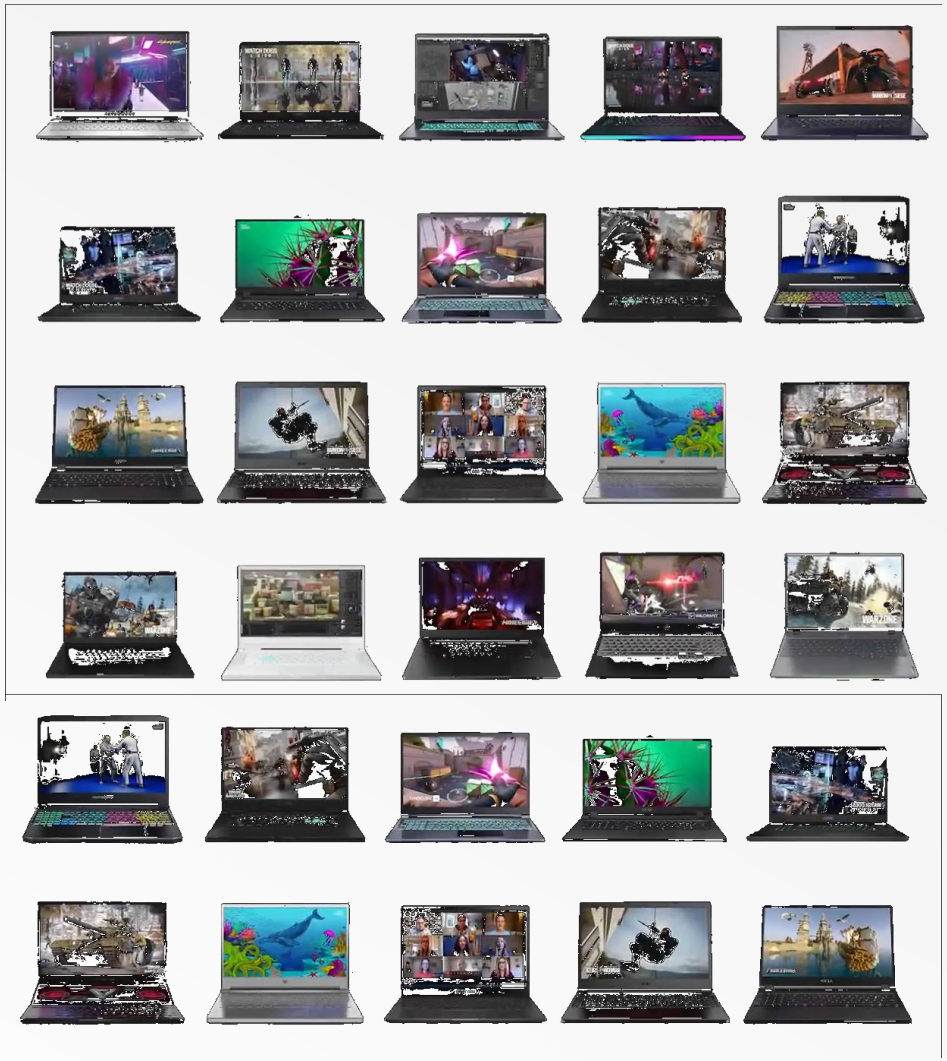
GRENZEN STELLEN



GRENZEN STELLEN

Voorbeeld
We meten het verbruik van een aantal laptops.

- aantal laptops getest:
  - $n = 30$
- steekproefgemiddelde:
  - $\bar{x} = 40\ W$
- steekproefstandaardafwijking:
  - $s = 20\ W$



Vraag

Wat zegt onze steekproef over het verbruik van alle laptops?

GRENZEN STELLEN

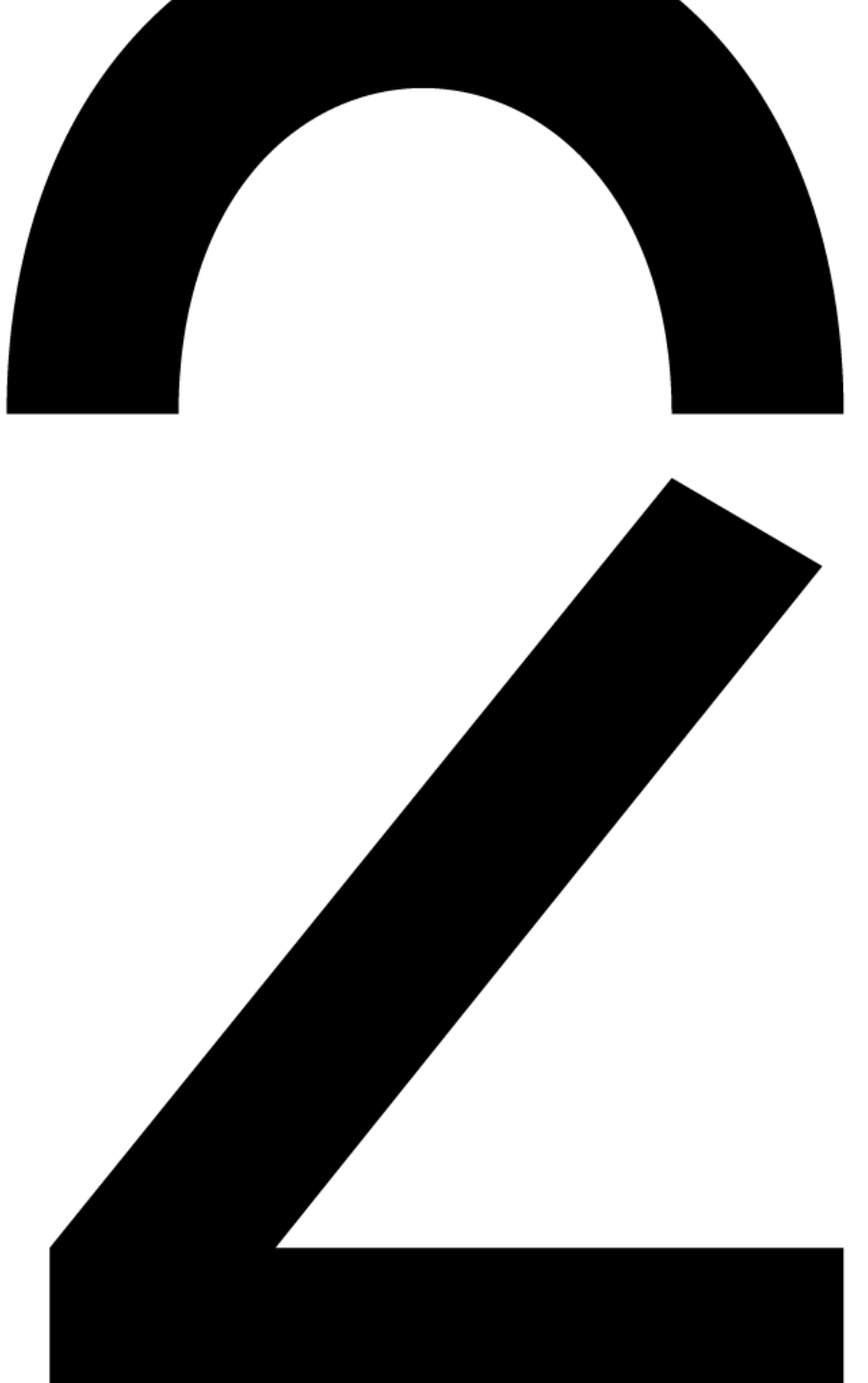
Vaststellingen

- Hoe meer laptops we testen, hoe zekerder we worden
- Hoe groter de standaardafwijking van onze steekproef, hoe onzekerder we worden
- We weten 100% zeker dat het gemiddelde verbruik van laptops tussen $-\infty$ en $+\infty$ W ligt
- We zijn minder zeker dat het verbruik tussen 30 en $50W$ ligt
- We zijn redelijk zeker dat het verbruik van laptops niet boven de $1000W$ ligt
- We zijn 100% zeker dat het gemiddelde verbruik niet exact $40{,}000000W$ is

Doel

Kunnen we dit niet formaliseren?

BETROUWBAARHEIDSINTERVAL



BETROUWBAARHEIDSINTERVAL

Doel

Interval vinden voor populatiegemiddelde μ

Oplossing

Vertrekken van $T = \frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}}$, het gestandaardiseerde steekproefgemiddelde.

Omvormen geeft $\mu = \bar{X} - \mathrm{T}\frac{S}{\sqrt{n}}$

Steekproef (met $\bar{x}, s, n$) geeft ons waarden voor $\bar{X}$, S en n: $\mu = \bar{x} - \mathrm{T}\frac{S}{\sqrt{n}}$

Maar welke waarde(n) heeft T? Het is een variabele.

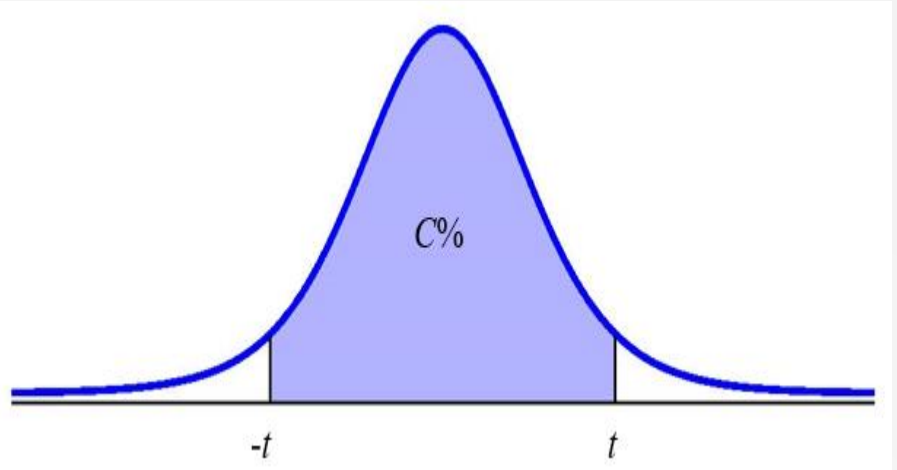
BETROUWBAARHEIDSINTERVAL

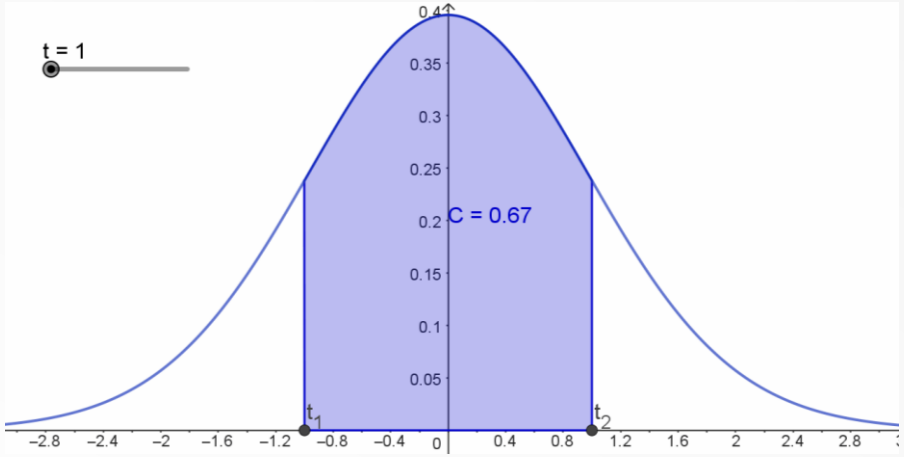
Waarde van de variabele T?

Hangt af van welke betrouwbaarheid $C\%$ we wensen.

$\mathrm{T} \sim t(n-1)$

C%	t1 en t2
90,0%	$\pm 1{,}65$
95,0%	$\pm 1{,}96$
95,5%	$\pm 2{,}00$
99,0%	$\pm 2{,}58$
99,7%	$\pm 3{,}00$
C%	$\pm t$





$$\mu = \bar{x} - \mathrm{T}\frac{s}{\sqrt{n}} \rightarrow \mu = \bar{x} \pm \mathrm{t}\frac{s}{\sqrt{n}}$$

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

BETROUWBAARHEIDSINTERVAL

Voorbeeld

$n = 30, \bar{x} = 40\,W, s = 20\,W$

Kies $C\% = 95{,}5\% \rightarrow t = 2$

Vraag

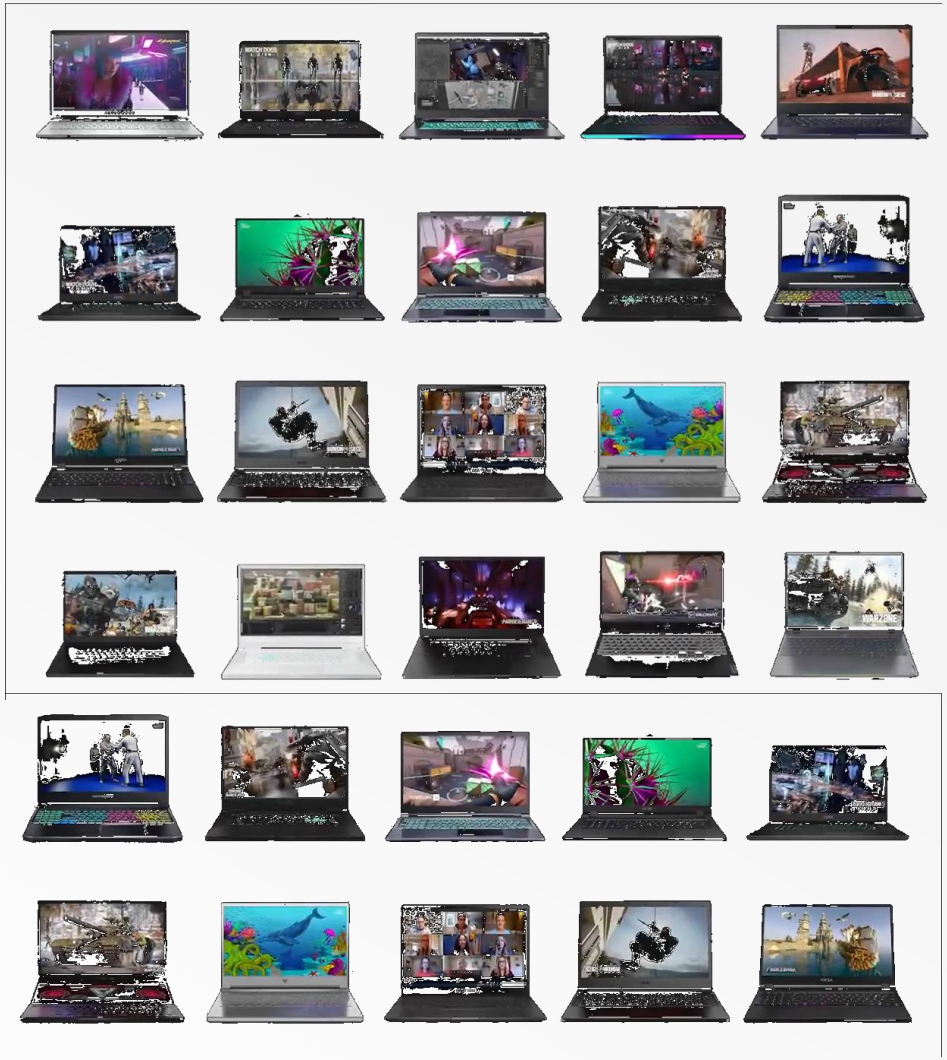
Wat zegt de steekproef over het verbruik van alle laptops?

Antwoord

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

$$40 - 2\frac{20}{\sqrt{30}} < \mu < 40 + 2\frac{20}{\sqrt{30}}$$

$$32.35 < \mu < 47.65$$



BETROUWBAARHEIDSINTERVAL

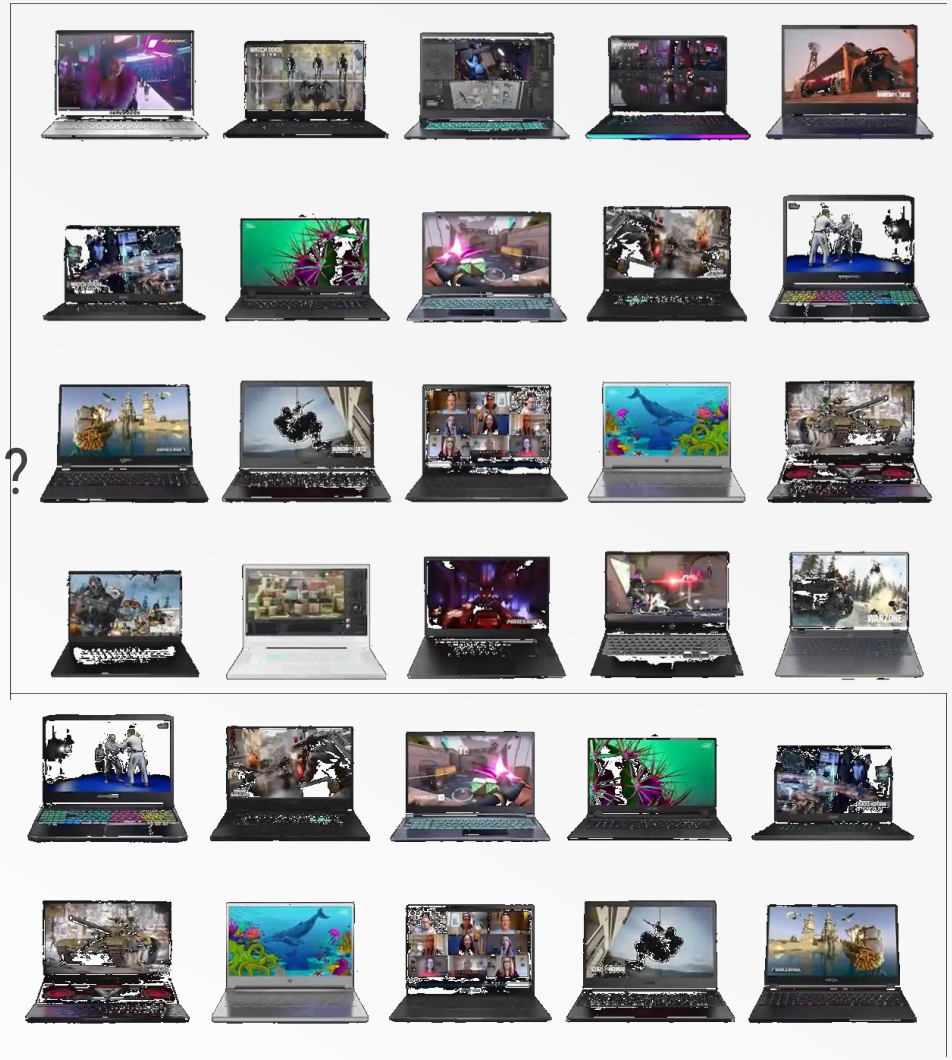
Voorbeeld

$n = 30, \bar{x} = 40\,W, s = 20\,W$

Kies $C\% = 90\% \rightarrow t = 1.65$

Vraag

Wat zegt steekproef over het verbruik van alle laptops?



Antwoord

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

$$40 - 1.65\frac{20}{\sqrt{30}} < \mu < 40 + 1.65\frac{20}{\sqrt{30}}$$

$$33.98 < \mu < 46.02$$

BETROUWBAARHEIDSINTERVAL

Voorbeeld

$n = 100, \bar{x} = 40\ W, s = 20\ W$

Kies $C\% = 90\% \rightarrow t = 1.65$

Vraag

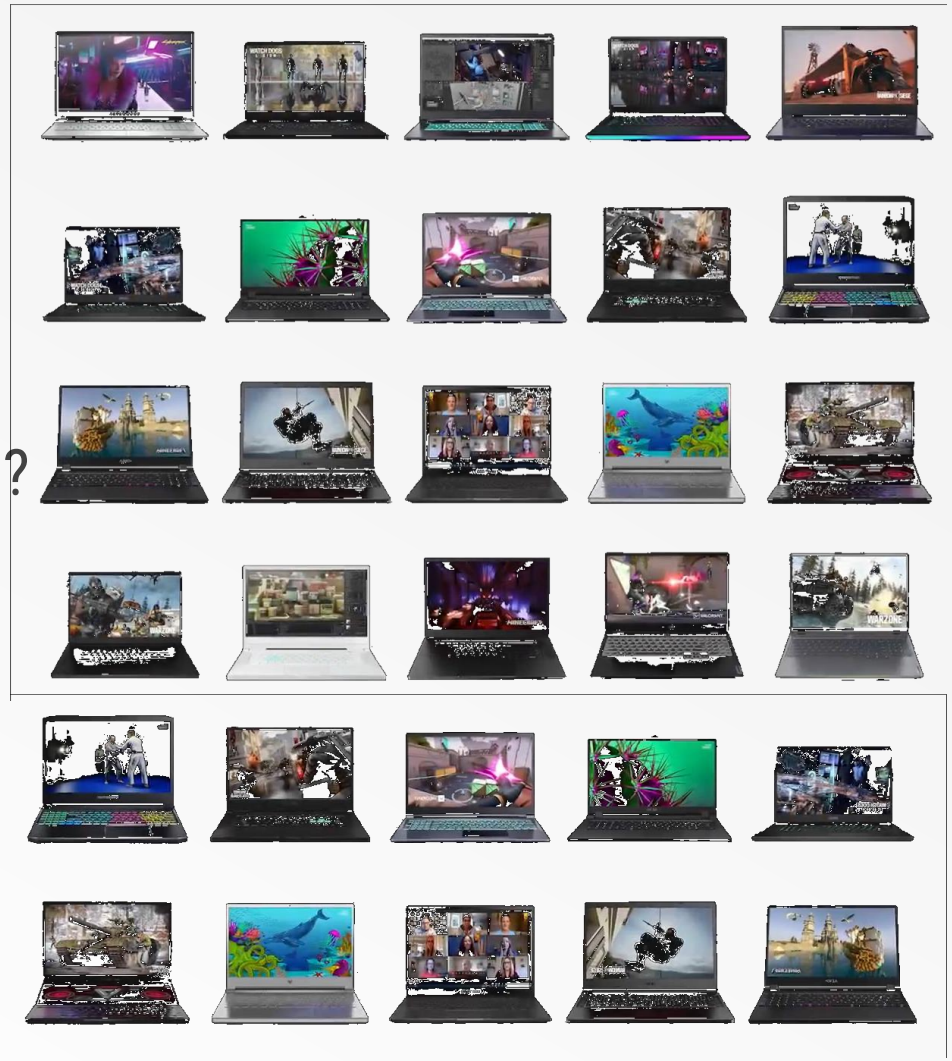
Wat zegt steekproef over het verbruik van alle laptops?

Antwoord

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

$$40 - 1.65\frac{20}{\sqrt{100}} < \mu < 40 + 1.65\frac{20}{\sqrt{100}}$$

$$36.70 < \mu < 43.30$$



BETROUWBAARHEIDSINTERVAL

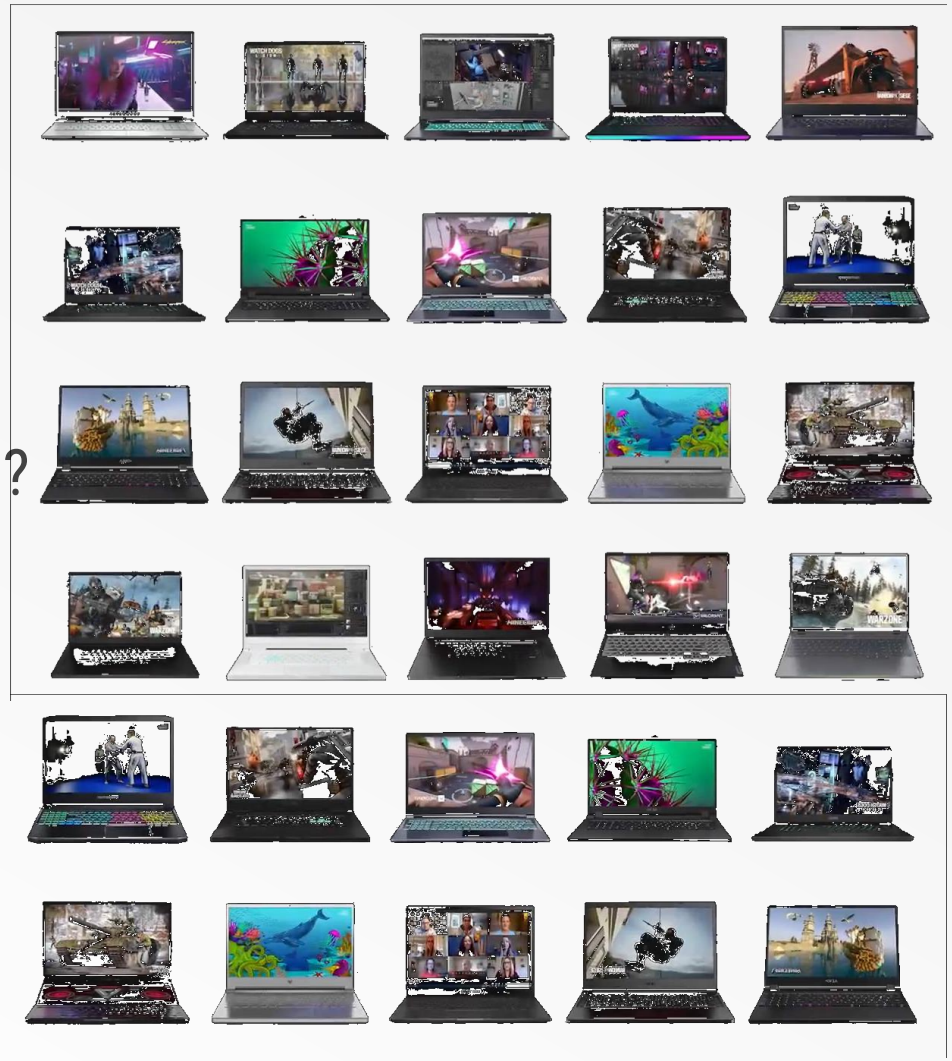
Voorbeeld

$n = 100, \bar{x} = 40\ W, s = 10\ W$

Kies $C\% = 90\% \rightarrow t = 1.65$

Vraag

Wat zegt steekproef over het verbruik van alle laptops?



Antwoord

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

$$40 - 1.65\frac{10}{\sqrt{100}} < \mu < 40 + 1.65\frac{10}{\sqrt{100}}$$

$$38.35 < \mu < 41.65$$

BETROUWBAARHEIDSINTERVAL

Conclusies

steekproef	grootte	gemiddelde	afwijking	$C\%$	t_1, t_2	ondergrens	bovengrens	breedte
1	$n = 30$	$\bar{x} = 40\ W$	$s = 20\ W$	95.5%	± 2.00	32.35	47.65	15.30
2	$n = 30$	$\bar{x} = 40\ W$	$s = 20\ W$	90.0%	± 1.65	33.98	46.02	12.04
3	$n = 100$	$\bar{x} = 40\ W$	$s = 20\ W$	90.0%	± 1.65	36.70	43.30	6.59
4	$n = 100$	$\bar{x} = 40\ W$	$s = 10\ W$	90.0%	± 1.65	38.35	41.65	3.30

Groter steekproef, kleine standardafwijking, lagere betrouwbaarheid → kleiner interval.

BETROUWBAARHEIDSINTERVAL

Conclusies

Betrouwbaarheidsinterval bepaalt de grenzen waarbinnen met $C\%$-zekerheid het echte populatiegemiddelde μ valt.

- 90% → kans van 90% dat μ ook effectief in het gevonden interval ligt
  $P(\mu \in [\bar{x} - \mathrm{t}\frac{s}{\sqrt{n}}, \bar{x} + \mathrm{t}\frac{s}{\sqrt{n}}]) = 0.9$

- 95% → kans van 95% dat μ ook effectief in het gevonden interval ligt
  $P(\mu \in [\bar{x} - \mathrm{t}\frac{s}{\sqrt{n}}, \bar{x} + \mathrm{t}\frac{s}{\sqrt{n}}]) = 0.95$
- …

Dat is de echte betekenis van een betrouwbaarheidsinterval.

PRAKTIJK

3

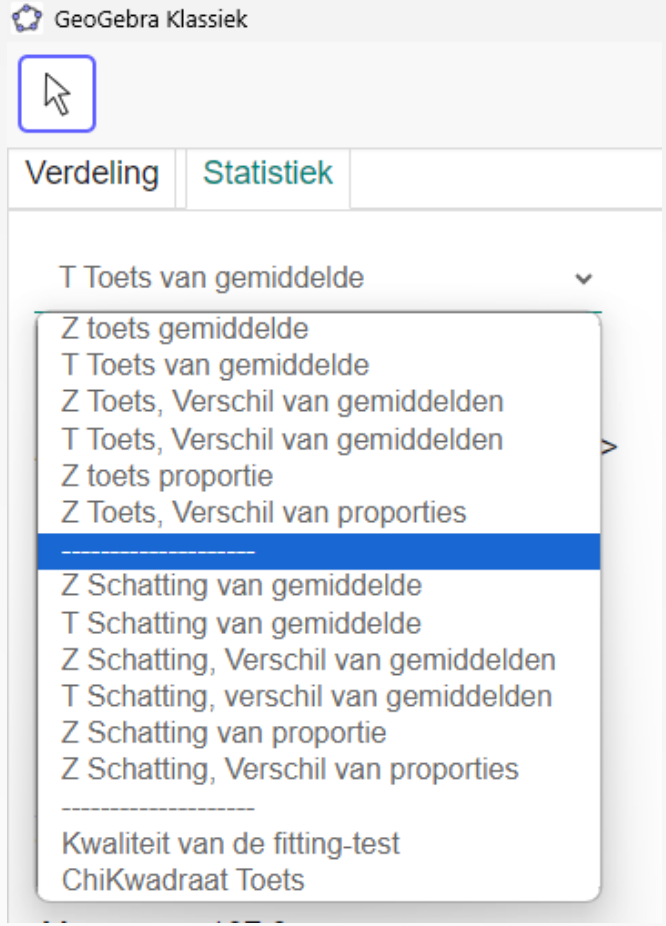
GEOGEBRA

Betrouwbaarheidsintervallen in GeoGebra

Makkelijkst via Schermindeling → Kansrekening → Statistiek en de juiste schatting kiezen:

- T Schatting van gemiddelde

Kan ook manueel via eigen berekeningen.



GEOGEBRA

Voorbeeld

$n = 30, \bar{x} = 40\ W, s = 20\ W$

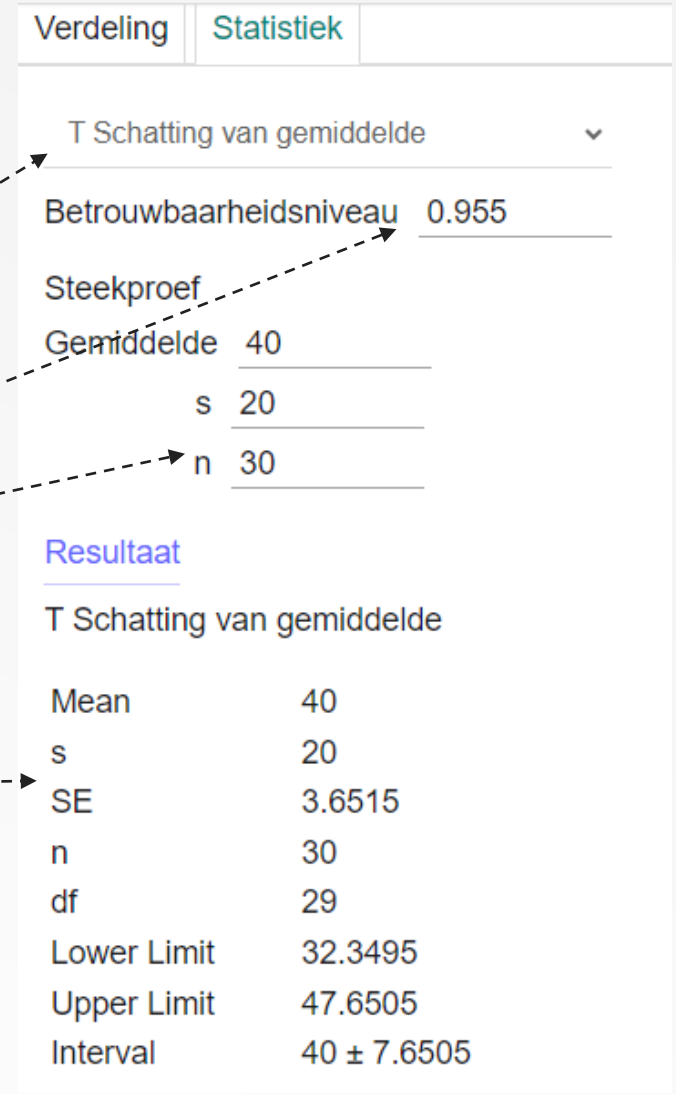
Kies $C\% = 95{,}5\% \rightarrow t = 2$

→ Kies voor T Schatting van gemiddelde

→ Stel gewenst betrouwbaarheidsinterval in

→ Gegevens invullen

→ Resultaten aflezen



BEVESTIGENDE ANALYSE

I
WELK PROBLEEM WILLEN WE OPLOSSEN
Kunnen we garanties geven?

II
KANSREKENEN
Wat is de kans?

III
KANSVERDELINGEN
Het grotere plaatje

IV
STEEKPROEVEN
Informatie verzamelen

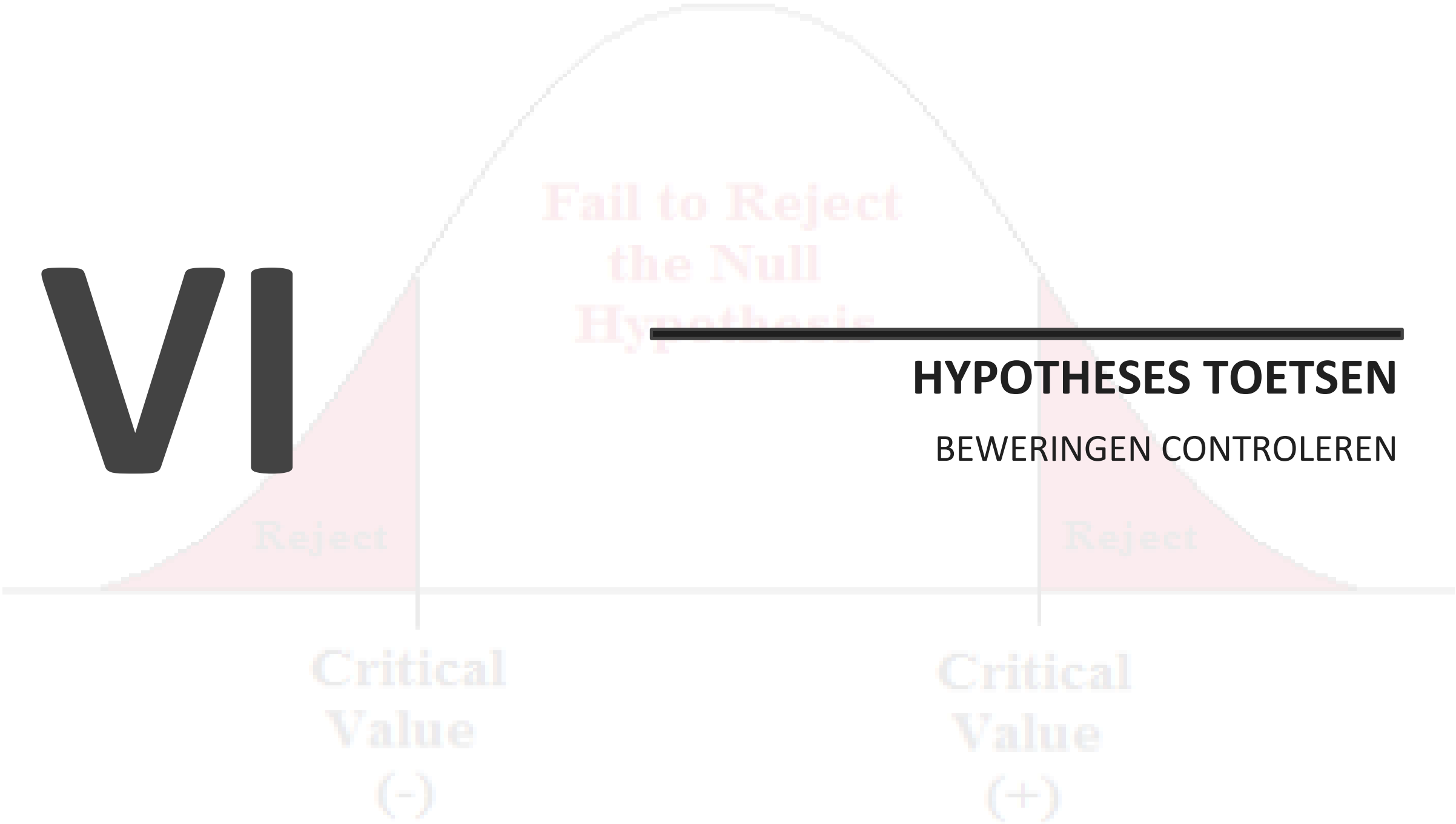
V
BETROUWBAARHEID
Grenzen stellen

VI
HYPOTHESES TOETSEN
Beweringen controleren

VI

HYPOTHESES TOETSEN

BEWERINGEN CONTROLEREN



BEVESTIGENDE ANALYSE

TOEVAL OF NIET?

- Is dit patroon toeval?
- Wat is de kans dat we dit waarnemen?
- Welke echte conclusies kunnen we hieruit trekken?

KANSREKENEN

- Grondbeginselen van waarschijnlijkheidstheorie
- Enkele wetten van waarschijnlijkheid

KANSVERDELINGEN

- Normale verdeling
- Standaard normale verdeling
- Studentverdeling
- Chi-kwadraat verdeling

STEEKPROEVEN

- Populatie vs. steekproef
- Eigenschappen van steekproeven

BETROUWBAARHEIDSINTERVALLEN

- Binnen welke grenzen vallen nieuwe metingen?

HYPOTHESETESTS

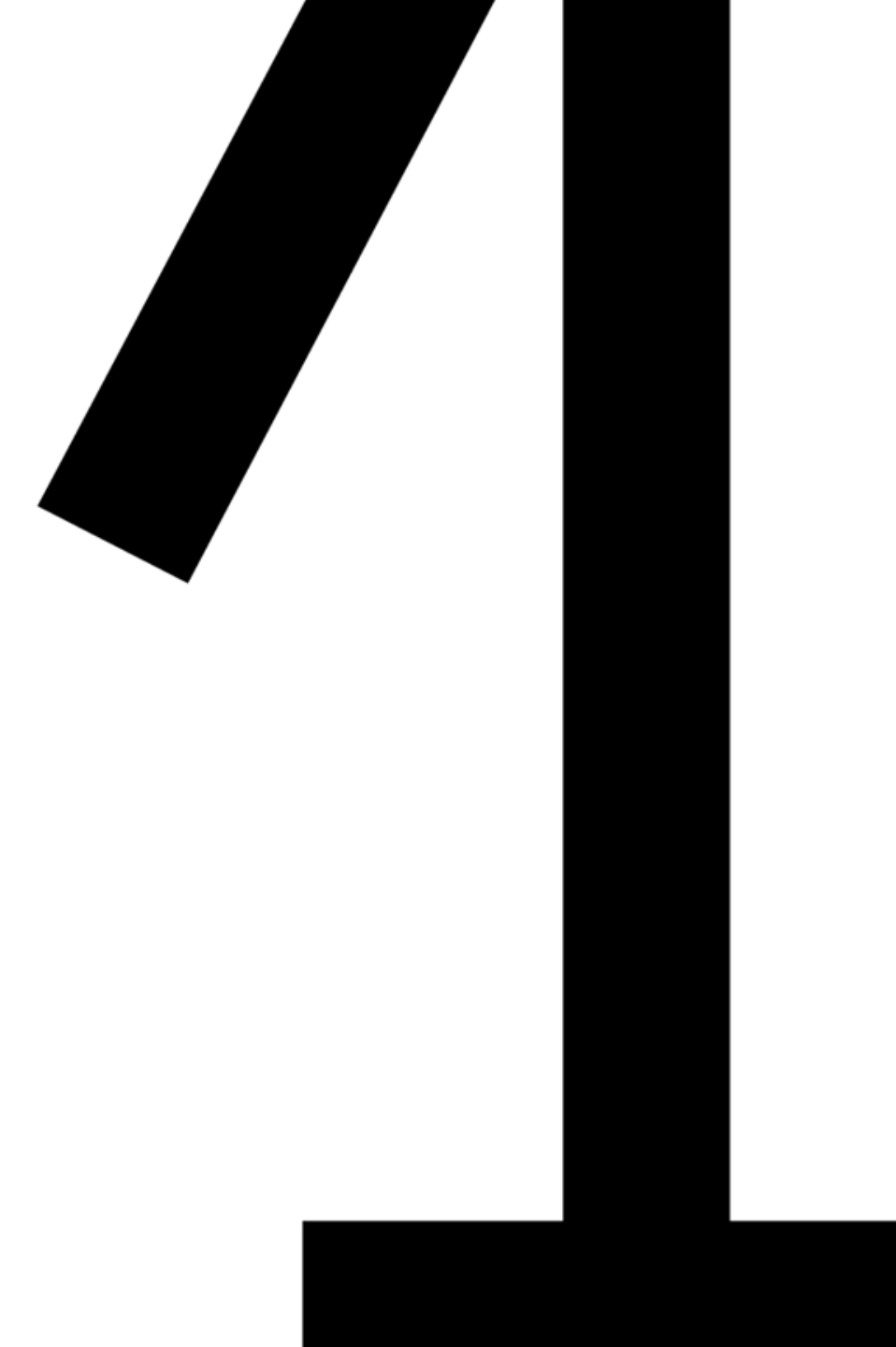
- Kunnen we beweringen verifiëren?
- Wanneer kunnen we beweringen weerleggen?

TECHNIEKEN

- Statistieken

WETEN WE ZEKER DAT HET GEBEURT?

EERST ZIEN EN DAN GELOVEN

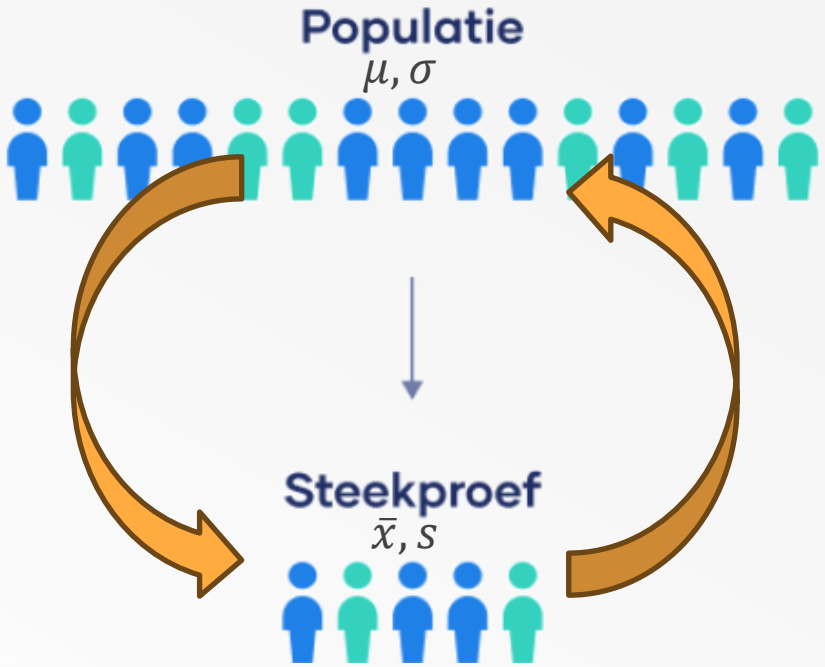


EERST ZIEN EN DAN GELOVEN

Ter herinnering
We willen op basis van een steekproef uitspraken doen over de populatie.

Betrouwbaarheidsinterval
Grenzen bepalen voor μ en σ o.b.v. $\bar{x}$ en s.

Aanvaardingsinterval
Bewering over populatie controleren door zelf meting (steekproef) te doen.



EERST ZIEN EN DAN GELOVEN

Voorbeeld

Een vriend stelt voor om een muntstuk op te gooien om te bepalen wie de rekening betaalt.

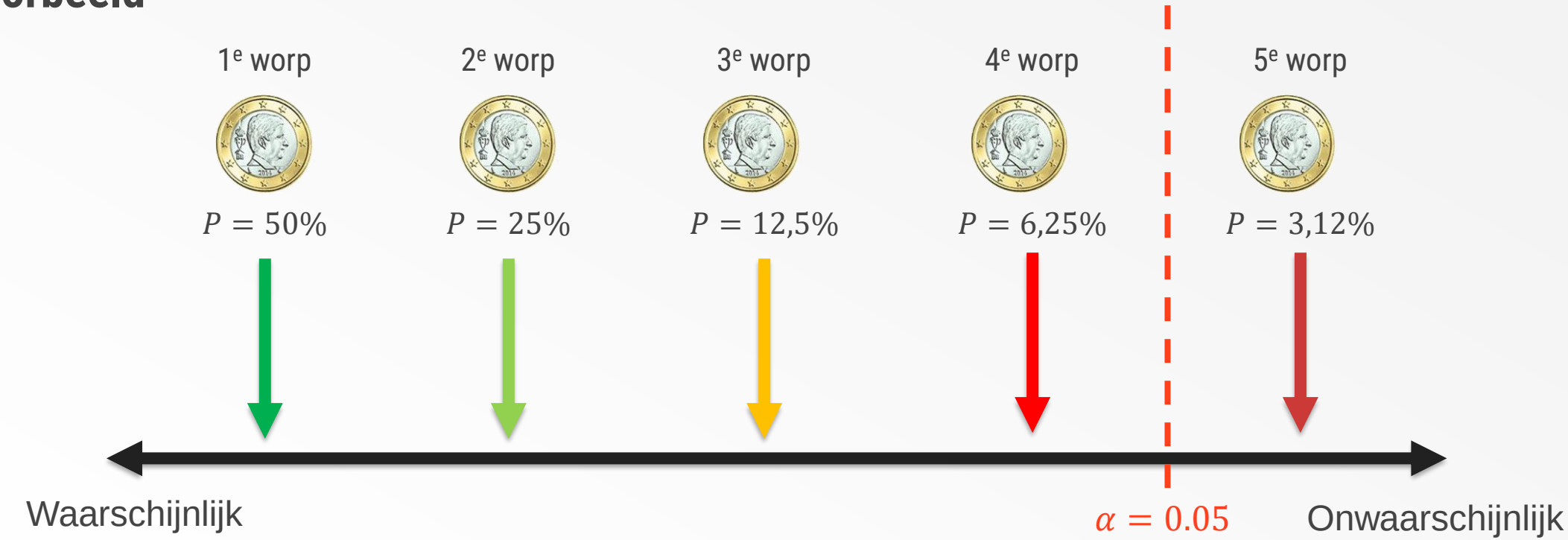
	jij betaalt
	vriend betaalt

Je aanvaardt het voorstel, maar houdt rekening met

Vriend gebruikt een eerlijk muntstuk	Nulhypothese H_0
Vriend gebruikt een vervalst muntstuk	Alternatieve hypothese H_1

IS HET MUNTSTUK VERVALST?

Voorbeeld



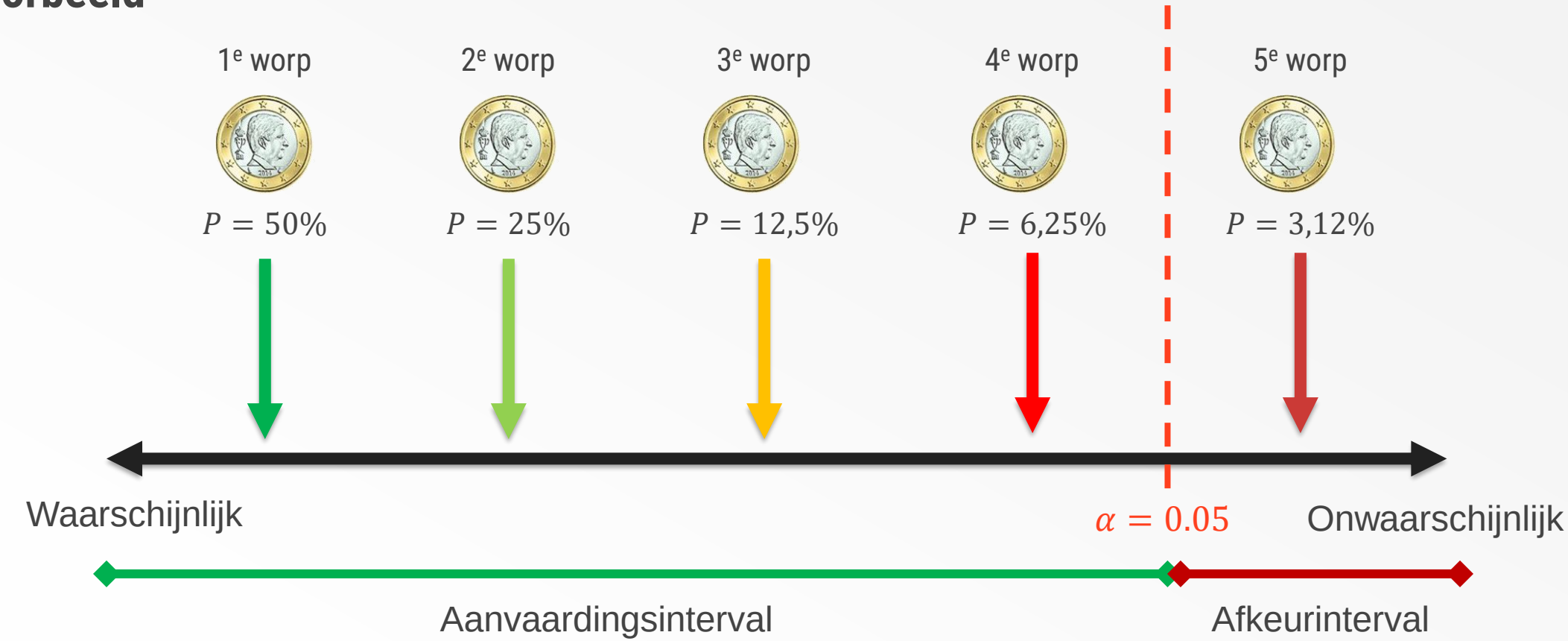
Vanaf welke kans worden we achterdochtig? Is dit door toeval? Of niet?

→ significantieniveau α

→ bv. $\alpha = 0.05$ (5%)

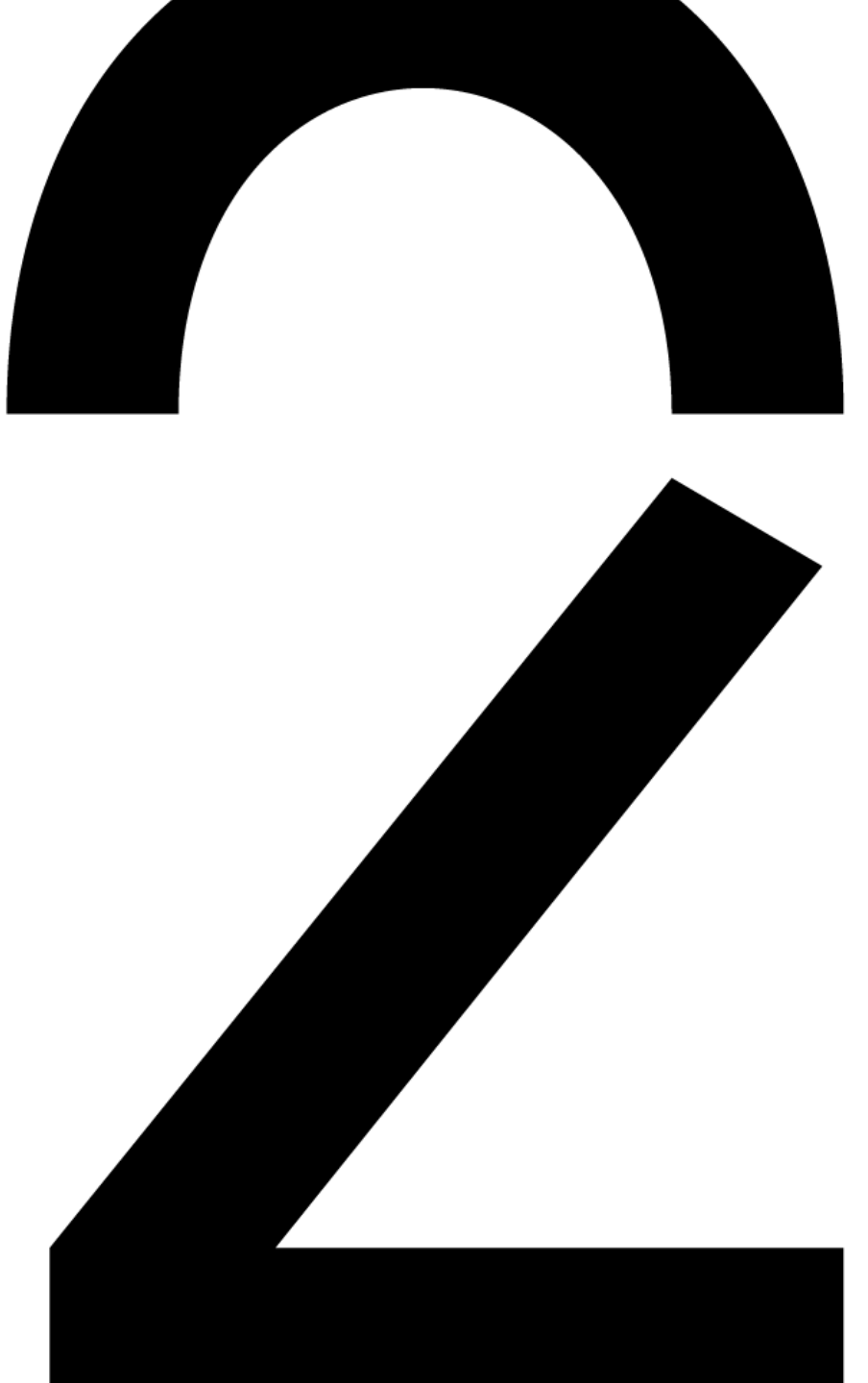
IS HET MUNTSTUK VERVALST?

Voorbeeld

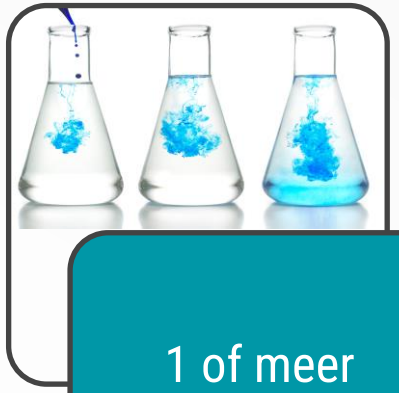


All kansen P gaan uit van de nulhypothese H_0, een eerlijk muntstuk

HYPOTHESES TOETSEN

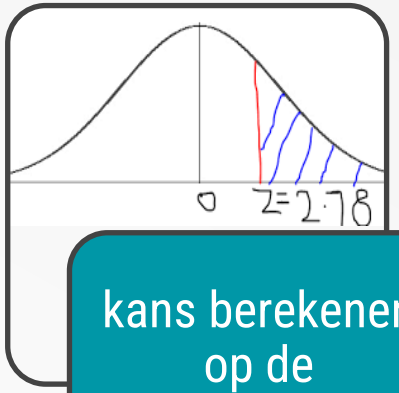


HOE WERKT EEN HYPOTHESETOETS

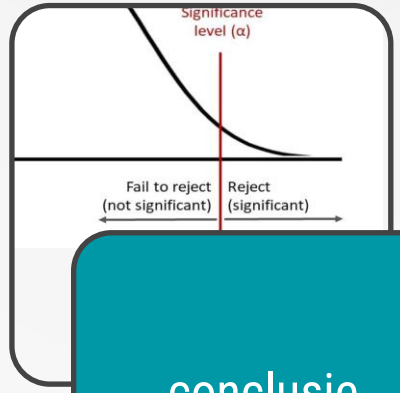


1 of meer steekproeven doen

teststatistiek berekenen o.b.v. steekproeven

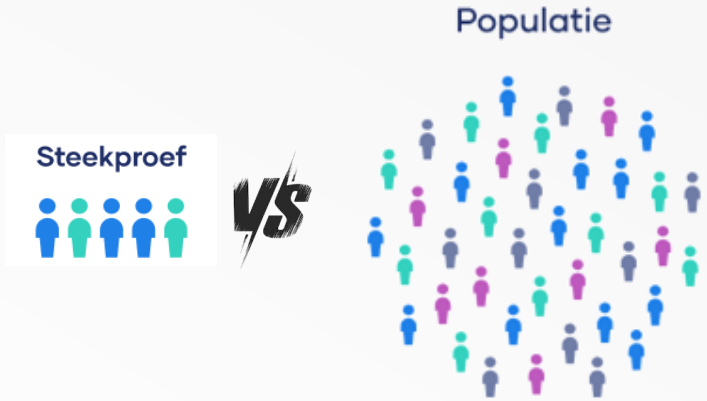


kans berekenen op de teststatistiek met een verdeling



conclusie trekken

WELKE TESTEN ZIJN ER?



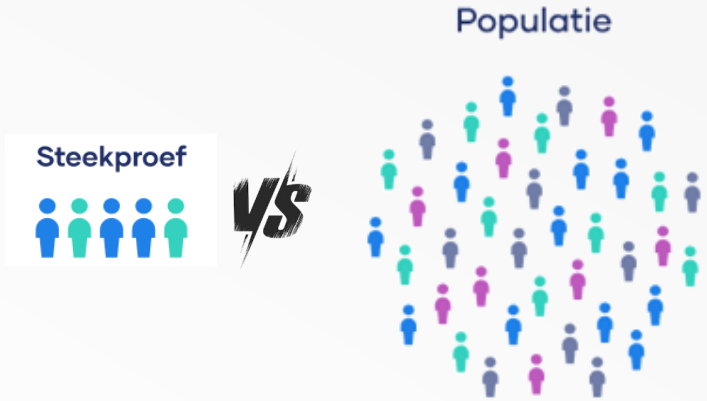
t-toets met **1** steekproef

t-toets met **2** steekproeven

ANOVA

Chi-kwadraat toets (χ^2)

WELKE TESTEN ZIJN ER?



t-toets met **1** steekproef

t-toets met **2** steekproeven

ANOVA

Chi-kwadraat toets (χ^2)

WAT HEB JE NODIG VOOR EEN TEST?

Hypotheses
Dit zijn veronderstellingen, genaamd H_0 en H_1.

Een significantieniveau α
Dit is een grenswaarde die je op voorhand kiest. Na de test is dit de kans dat je foute conclusie trekt.

Een steekproef
Een steekproef van met grootte n, gemiddelde $\bar{x}$, en standaardafwijking s

Een verdeling $T(\nu)$
Dit is de Student – verdeling met een aantal vrijheidsgraden ν, ν hangt af van de test.

Een toetsingsgrootheid t
Een maatstaf bepaald door de steekproef waarmee je kan testen, t hangt af van de test.

De p-waarde
De waarschijnlijkheid van de toetsingsgrootheid t, p hangt af van het soort toets.

WAT HEB JE NODIG VOOR DEZE TEST?

Hypotheses
Dit zijn veronderstellingen, genaamd H_0 en H_1.

H_0 - de nulhypothese	H_1 - de alternatieve hypothese
• onze steekproef komt overeen met de populatie • geen verschil in de parameters • $\bar{x} = \mu$	• onze steekproef toont heel andere resultaten • wel een verschil in parameters • $\bar{x} \neq \mu$ (tweezijdige toets) • $\bar{x} > \mu$, of $\bar{x} < \mu$ (eenzijdige toets)
We doen de test ervan uitgaande dat H_0 waar is.	Als het resultaat van de test onwaarschijnlijk blijkt, is H_0 vals en H_1 waar.

WAT HEB JE NODIG VOOR DEZE TEST?

Voorbeeld

We willen testen of trein tussen Mechelen en Antwerpen-Central er exact 17 minuten over doet.

H_0 - de nulhypothese	H_1 - de alternatieve hypothese
• onze steekproef komt overeen met de populatie • $H_0: \mu = 17$	• onze steekproef toont heel andere resultaten • $H_1: \mu \neq 17$ (tweezijdige toets)
We doen de test ervan uitgaande dat H_0 waar is.	Als het resultaat van de test onwaarschijnlijk blijkt, is H_0 vals en H_1 waar.

WAT HEB JE NODIG VOOR DEZE TEST?

Voorbeeld

We willen testen of de gemiddelde score voor Programmeren 1 van 1e jaarstudenten groter is dan of gelijk aan 10 op 20.

H_0 - de nulhypothese	H_1 - de alternatieve hypothese
• onze steekproef komt overeen met de populatie • $H_0: \mu \geq 10$	• onze steekproef toont heel andere resultaten • $H_1: \mu < 10$ (eenzijdige toets)
We doen de test ervan uitgaande dat H_0 waar is.	Als het resultaat van de test onwaarschijnlijk blijkt, is H_0 vals en H_1 waar.

WAT HEB JE NODIG VOOR DEZE TEST?

Voorbeeld

We willen testen of de hoeveelheid centiliter bier in een fles kleiner is dan 25cl.

H_0 - de nulhypothese	H_1 - de alternatieve hypothese
• onze steekproef komt overeen met de populatie • $H_0: \mu \leq 25$	• onze steekproef toont heel andere resultaten • $H_1: \mu > 25$ (eenzijdige toets)
We doen de test ervan uitgaande dat H_0 waar is.	Als het resultaat van de test onwaarschijnlijk blijkt, is H_0 vals en H_1 waar.

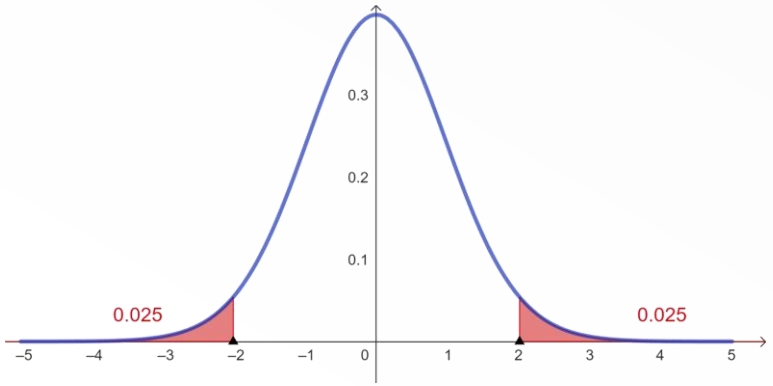
WAT HEB JE NODIG VOOR DEZE TEST?

Formuleren van de hypotheses

Hangt dus af van wat je juist wil toetsen.

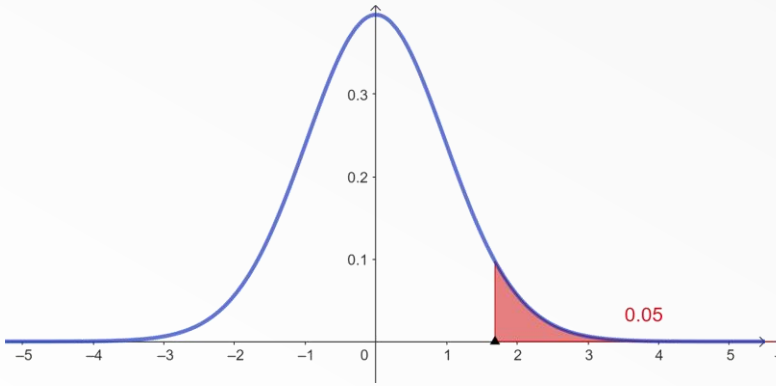
Tweezijdig toets

- $H_0: \mu = \mu_0$
- $H_1: \mu \neq \mu_0$

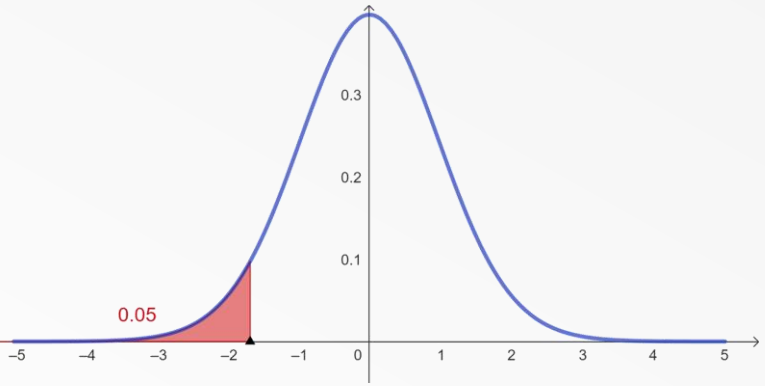


Eenzijdig toets

- $H_0: \mu = \mu_0$
- $H_1: \mu > \mu_0$



- $H_0: \mu = \mu_0$
- $H_1: \mu < \mu_0$



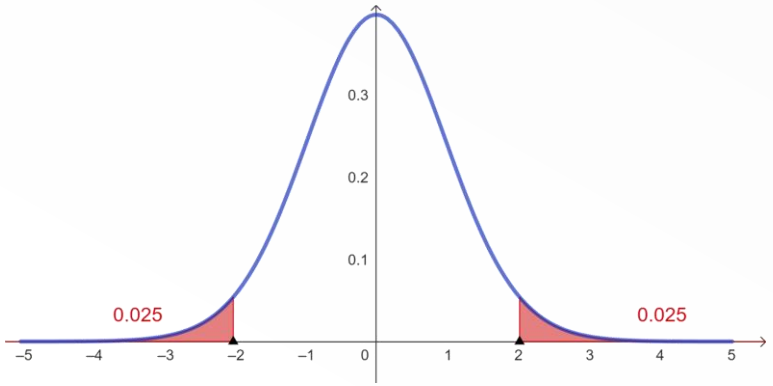
WAT HEB JE NODIG VOOR DEZE TEST?

Een verdeling $T(\nu)$

T is de Studentverdeling met ν het aantal vrijheidsgraden, meestal gelijk aan $n-1$

Tweezijdig toets

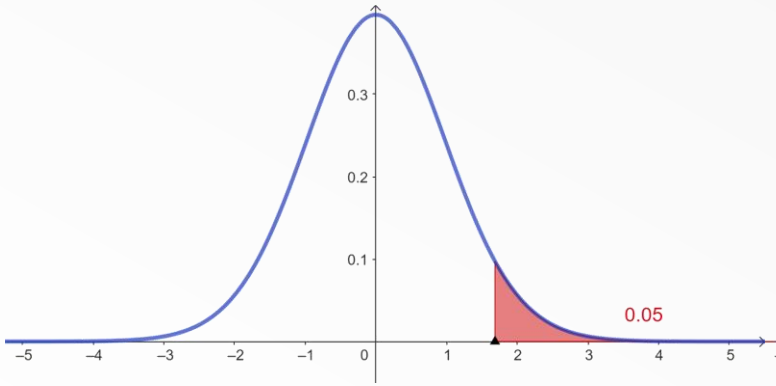
$T(n-1)$

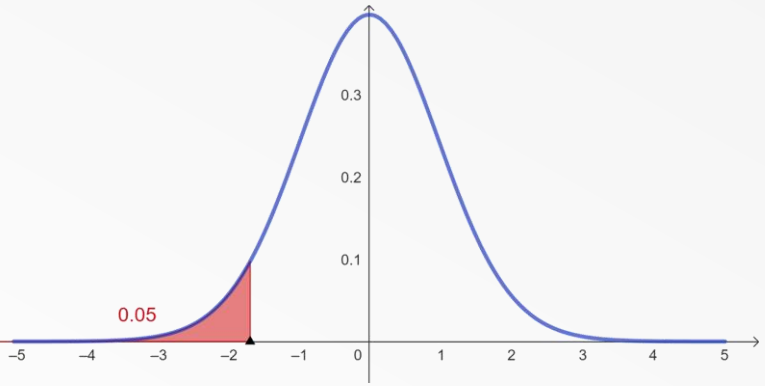


Eenzijdig toets

$T(n-1)$

$T(n-1)$





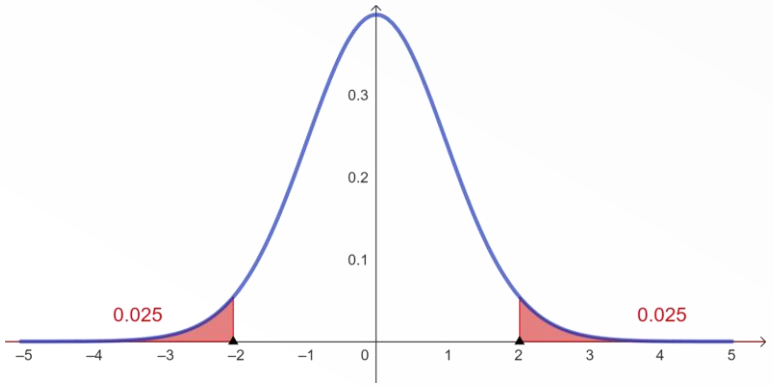
WAT HEB JE NODIG VOOR DEZE TEST?

Een toetsingsgrootheid t

Een moeilijk woord voor een getal dat uit de steekproef komt gerold. Hangt af van het soort test.

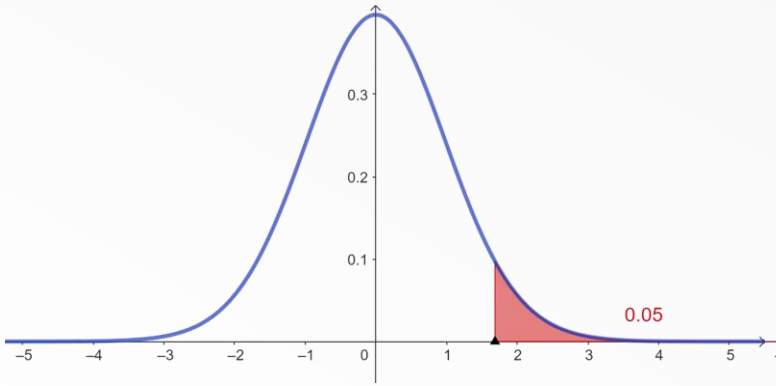
Tweezijdig toets

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

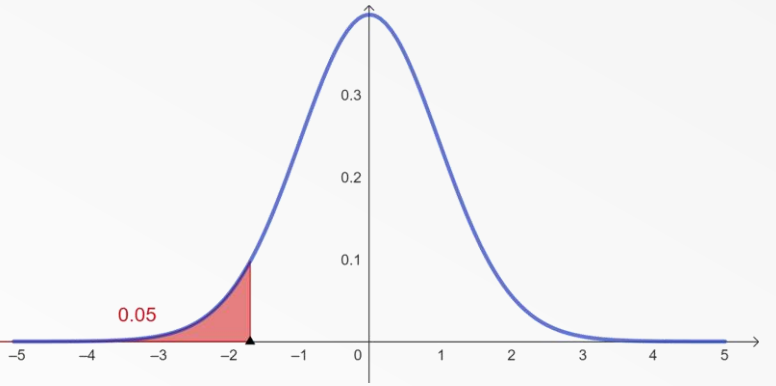


Eenzijdig toets

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$



$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$



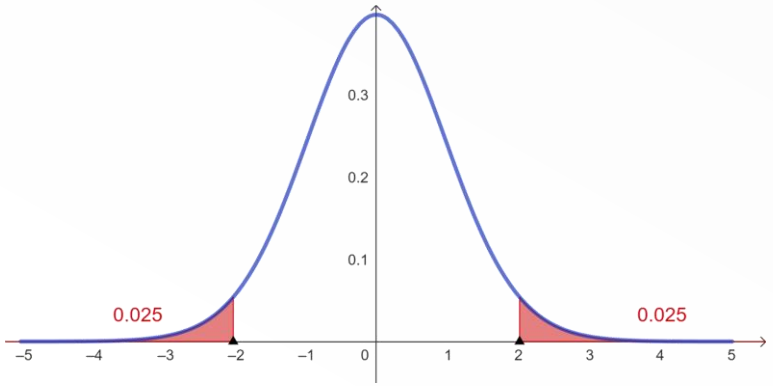
WAT HEB JE NODIG VOOR DEZE TEST?

De p-waarde

Als H_0 waar is kunnen we berekenen hoe waarschijnlijk de waarden uit onze steekproef zijn.

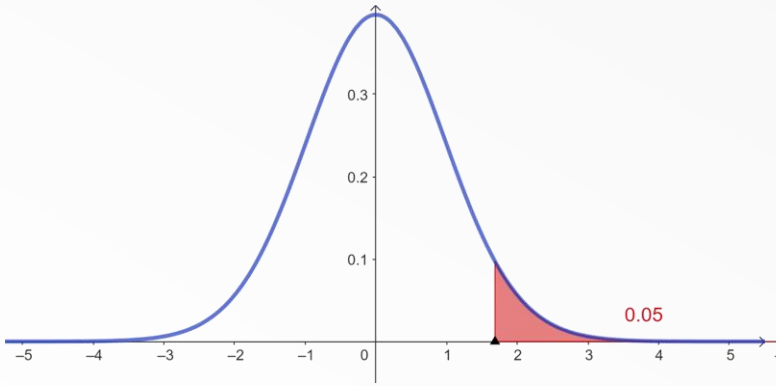
Tweezijdig toets

$p = 2 \cdot P(T < t)$
of
$p = 2 \cdot P(T > t)$

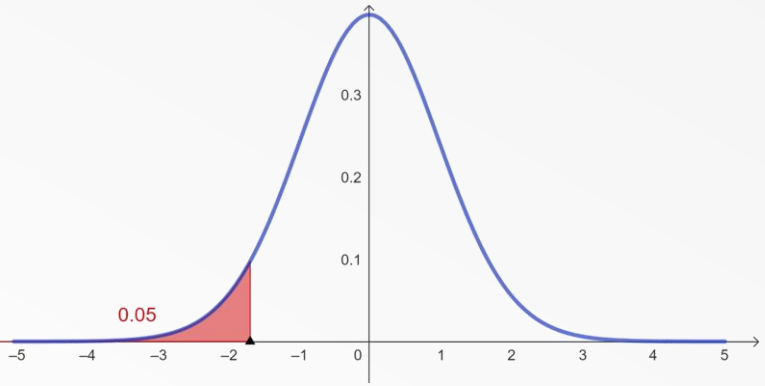


Eenzijdig toets

$p = P(T > t)$



$p = P(T < t)$



WAT HEB JE NODIG VOOR DEZE TEST?

Hypotheses

Dit zijn veronderstellingen, genaamd H_0 en H_1.

Significantieniveau

α

Een steekproef

n, $\bar{x}$ en s

Een verdeling $T(\nu)$

$\mathrm{T}(n-1)$

Een toetsingsgrootheid t

$t = \frac{\bar{x}-\mu}{s/\sqrt{n}}$

De p-waarde

tweezijdige toets: $p = 2 \cdot P(T < t)$ of $p = 2 \cdot P(T > t)$
eenzijdige toets: $p = P(T < t)$ of $p = P(T > t)$

t-toets met 1 steekproef

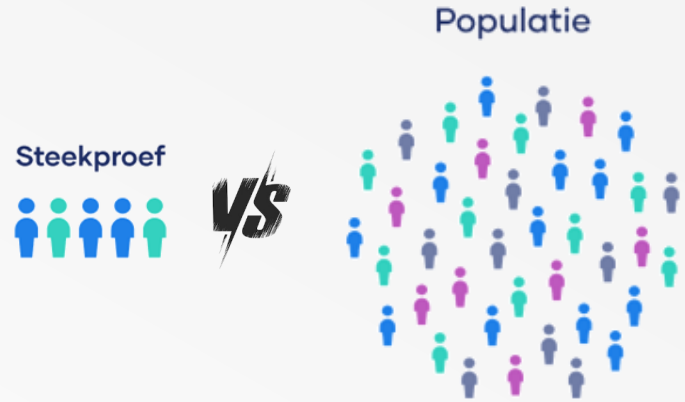
Voorbeeld

Iemand beweert dat de gemiddelde schermgrootte van alle verkochte televisies 40 inch is.

- $\mu = 40$ inch

We doen een steekproef van enkele verkochte televisies.

- $n = 50, \bar{x} = 43$ inch en $s = 10$ inch



t-toets met **1** steekproef

Stappenplan

1. Kies voor eenzijdige of tweezijdige test.
2. Formuleer H_0 en H_1
3. Kies significantieniveau α
4. Bepaal de p-waarde van toetsingsgrootheid $t = \frac{\bar{x}-\mu}{s/\sqrt{n}}$
5. Wat besluit je?

t-toets met 1 steekproef

Voorbeeld

Iemand beweert dat de gemiddelde schermgrootte van alle verkochte televisies 40 inch is.

1. Kies voor eenzijdige of tweezijdige test

   tweezijdige test

2. Formuleer H_0 en H_1

   $H_0: \mu = 40$

   $H_1: \mu \neq 40$

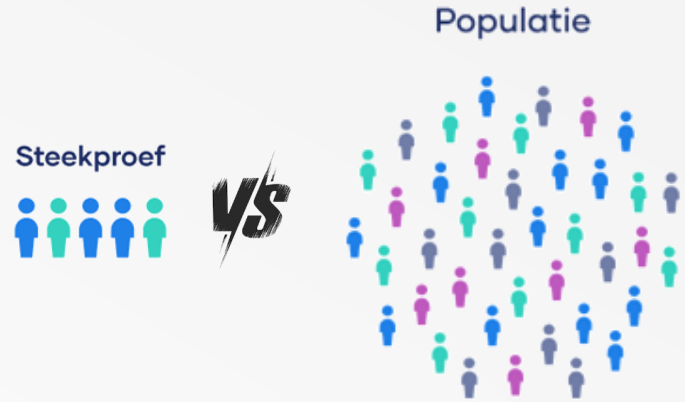
3. Kies significantieniveau α

   $\alpha = 0.05$

4. Bepaal de p-waarde van toetsingsgrootheid $t = \frac{\bar{x}-\mu}{s/\sqrt{n}}$

   $p = 2\,P(T < t)$ als t negatief is

   $p = 2\,P(T > t)$ als t positief is



t-toets met **1** steekproef

t-toets met 1 steekproef

Voorbeeld

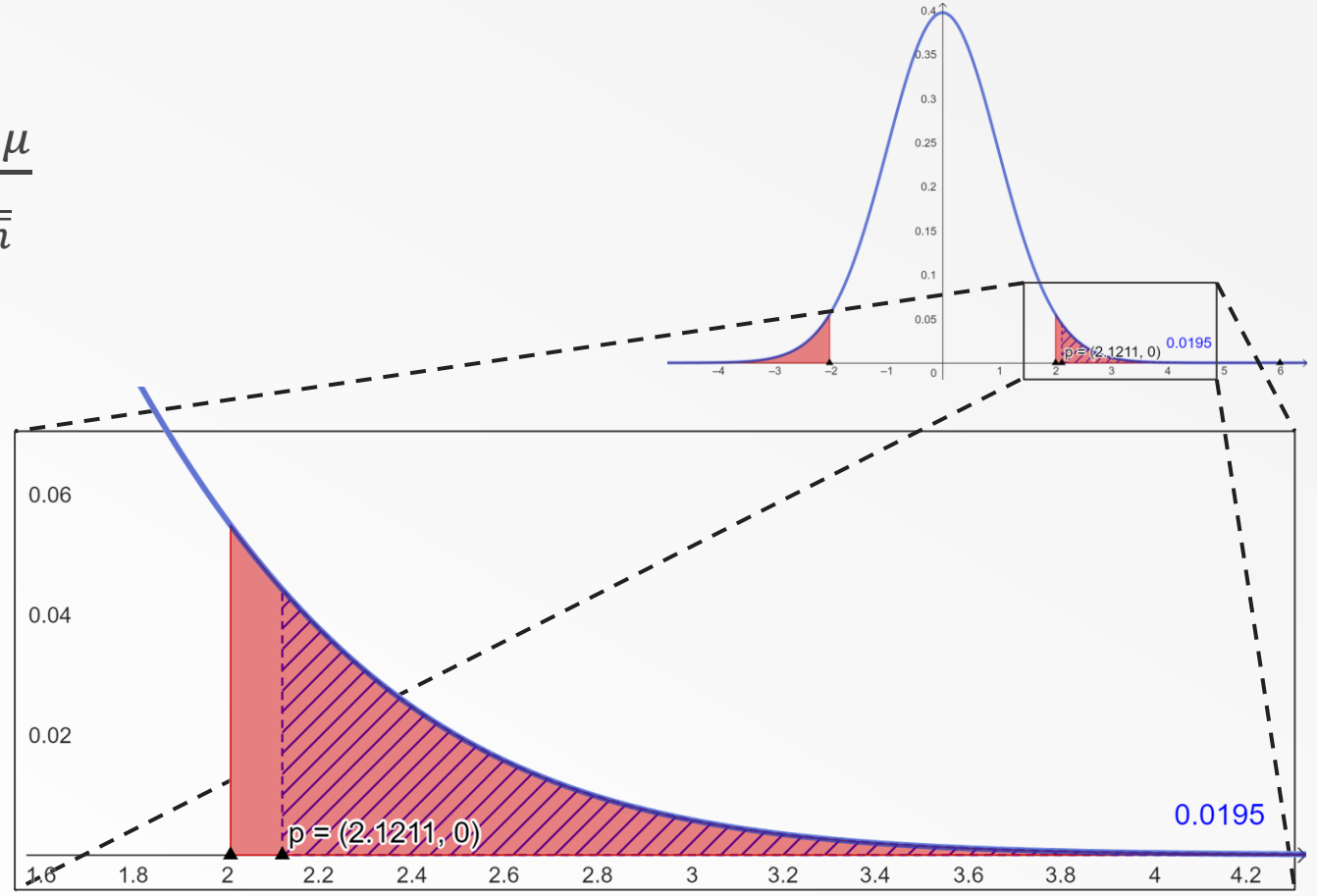
Iemand beweert dat de gemiddelde schermgrootte van alle verkochte televisies 40 inch is.

4. Bepaal de p-waarde met van $t = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}}$

→ $t = \frac{(43-40)}{\frac{10}{\sqrt{50}}} = 2.121$

→ $p = 2 \cdot 0.0195$

→ $p = 0.039$



t-toets met 1 steekproef

Voorbeeld

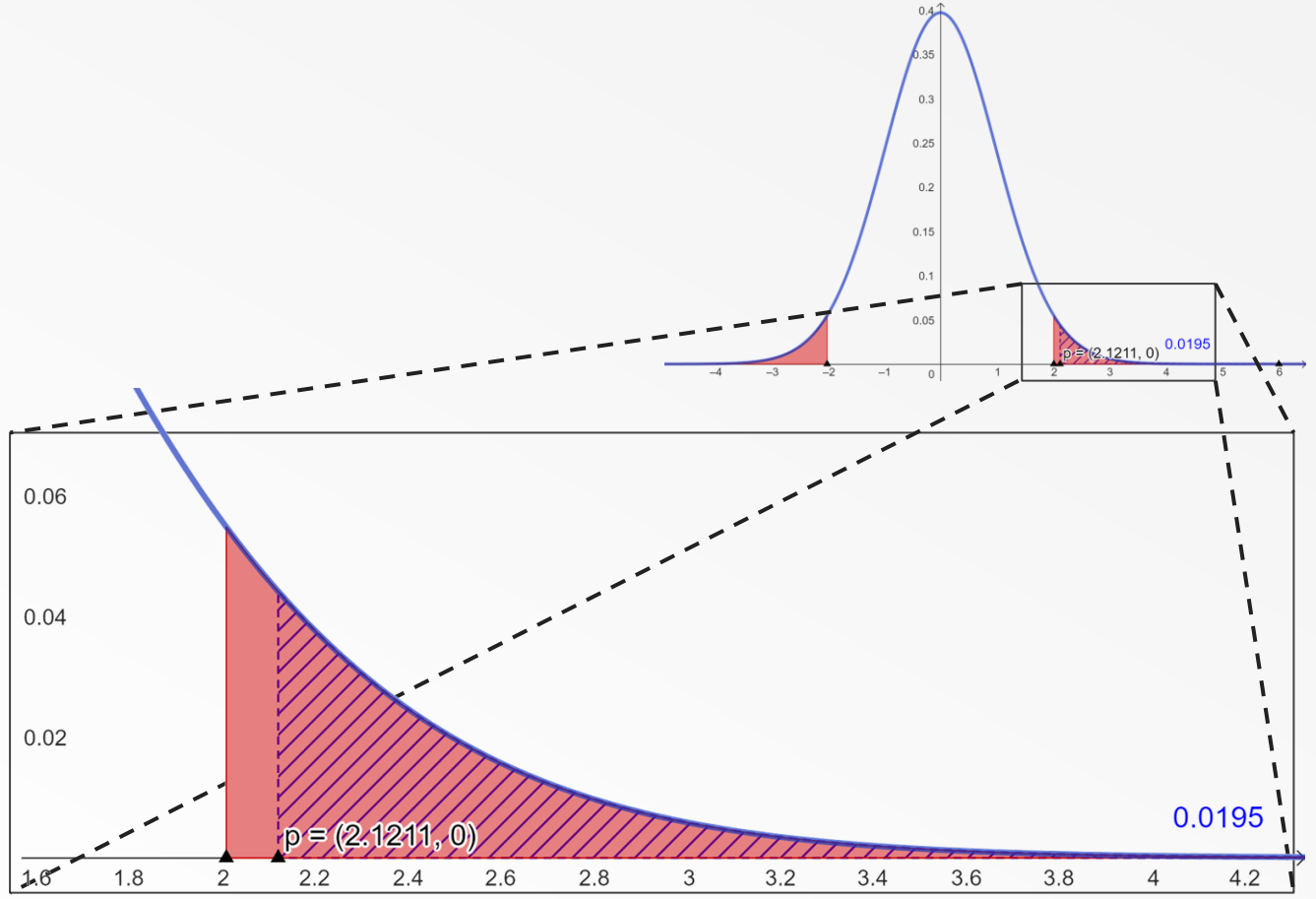
Iemand beweert dat de gemiddelde schermgrootte van alle verkochte televisies 40 inch is.

5. Wat besluit je ?

$p = 0.039$

$p < \alpha$

→ We verwerpen H_0!



t-toets met 1 steekproef

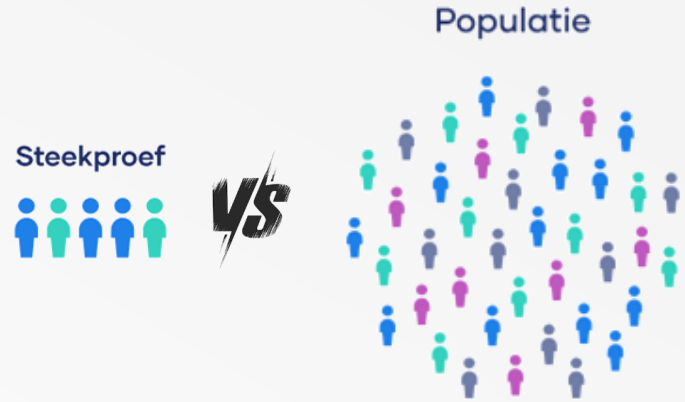
Voorbeeld

Iemand beweert dat de gemiddelde schermgrootte van alle verkochte televisies 40 inch is.

- $\mu = 40$ inch

We doen een steekproef van enkele verkochte televisies.

- $n = 50$, $\bar{x} = 42$ inch en $s = 10$ inch



t-toets met **1** steekproef

Stappenplan

1. Kies voor eenzijdige of tweezijdige test.
2. Formuleer H_0 en H_1
3. Kies significantieniveau α
4. Bepaal de p-waarde met van $t = \frac{\bar{x}-\mu}{s/\sqrt{n}}$
5. Wat besluit je?

t-toets met 1 steekproef

Voorbeeld

Iemand beweert dat de gemiddelde schermgrootte van alle verkochte televisies 40 inch is.

1. Kies voor eenzijdige of tweezijdige test

   tweezijdige test

2. Formuleer H_0 en H_1

   $H_0: \mu = 40$

   $H_1: \mu \neq 40$

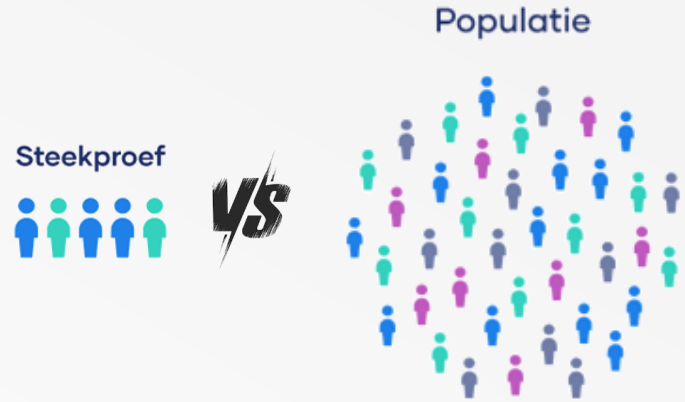
3. Kies significantieniveau α

   $\alpha = 0.05$

4. Bepaal de p-waarde van toetsingsgrootheid $t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$

   $p = 2\,P(T < t)$ als t negatief is

   $p = 2\,P(T > t)$ als t positief is



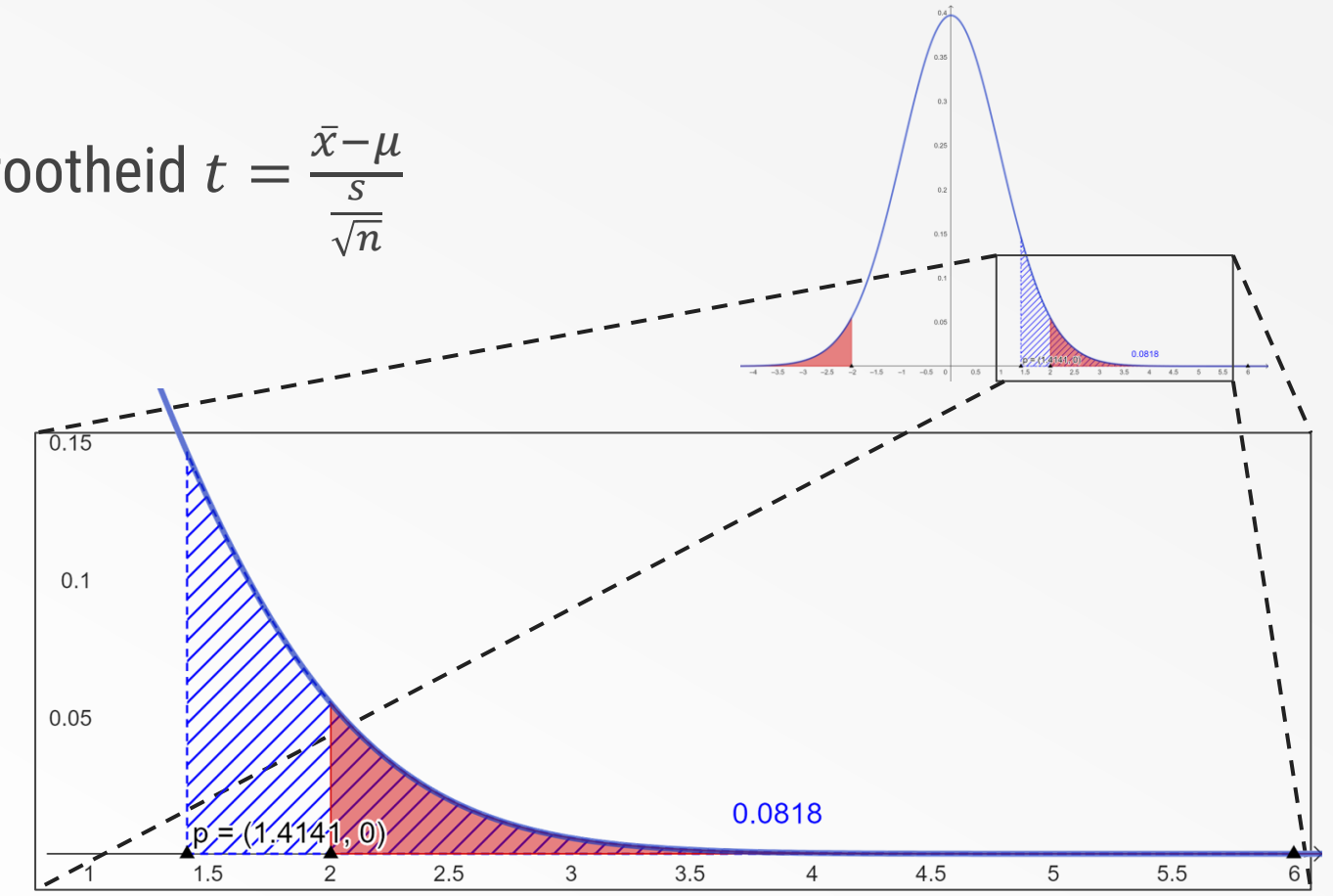
t-toets met **1** steekproef

t-toets met 1 steekproef

Voorbeeld

Iemand beweert dat de gemiddelde schermgrootte van alle verkochte televisies 40 inch is.

4. Bepaal de p-waarde van toetsingsgrootheid $t = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}}$



$\rightarrow t = \frac{(42-40)}{\frac{10}{\sqrt{50}}} = 1.414$

$\rightarrow p = 2 \cdot 0.0818$

$\rightarrow p = 0.1637$

t-toets met 1 steekproef

Voorbeeld

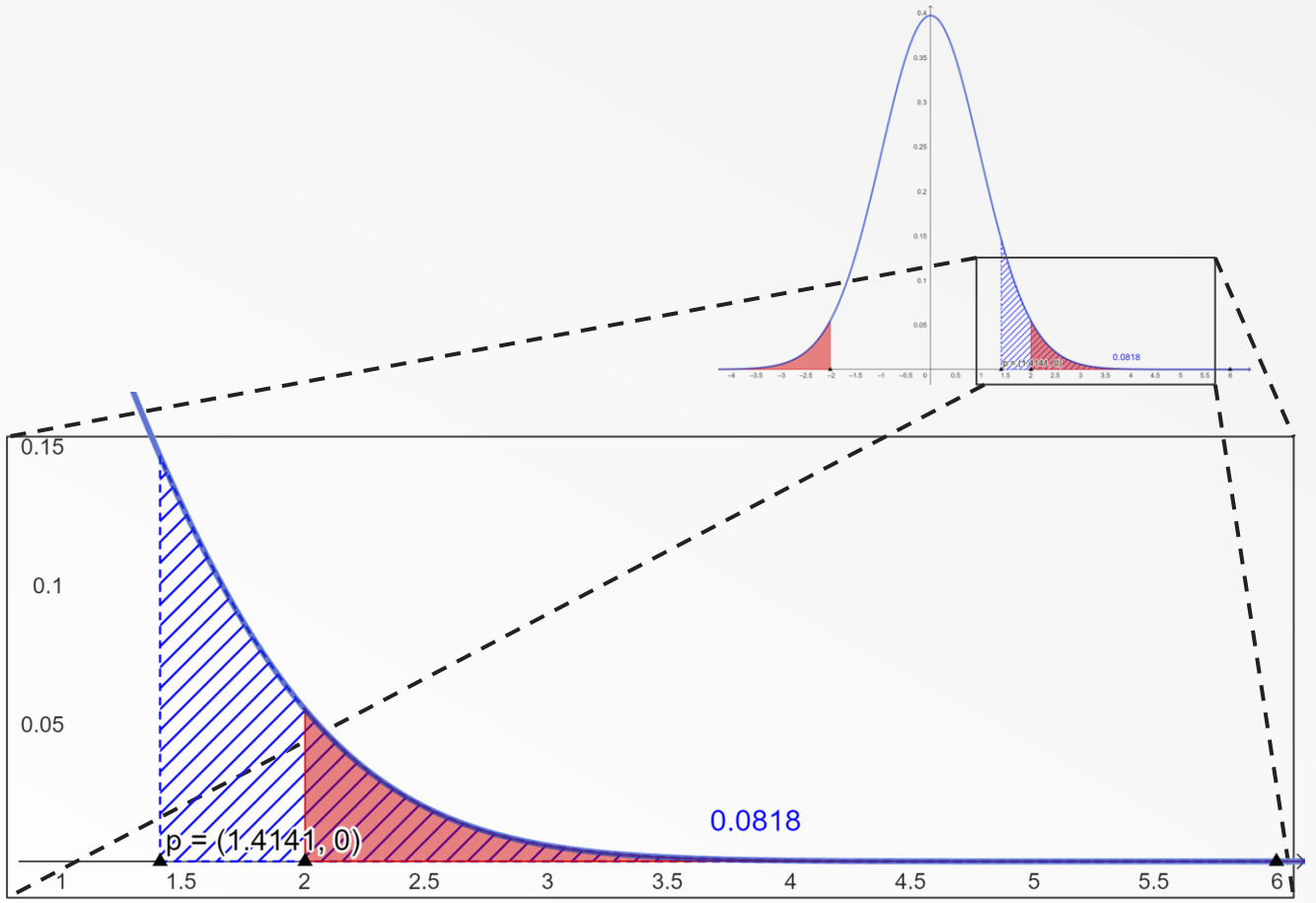
Iemand beweert dat de gemiddelde schermgrootte van alle verkochte televisies 40 inch is.

5. Wat besluit je ?

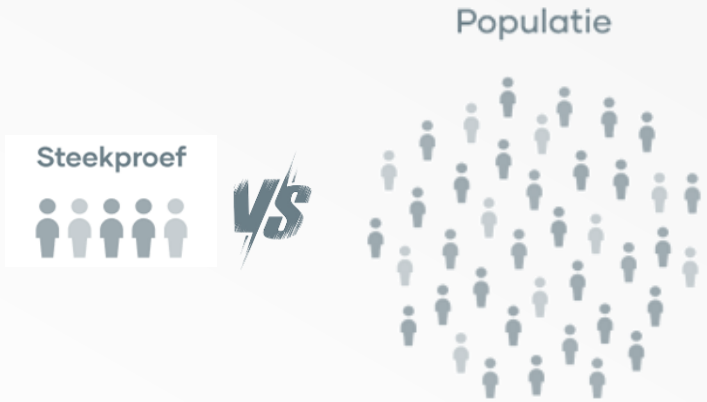
$p = 0.1637$

$p > \alpha$

→ We kunnen H_0 **niet** verwerpen!



WELKE TESTEN ZIJN ER?



t-toets met **1** steekproef

t-toets met **2** steekproeven

ANOVA

Chi-kwadraat toets (χ^2)

WAT HEB JE NODIG VOOR DEZE TEST?

Hypotheses

Dit zijn veronderstellingen, genaamd H_0 en H_1.

Twee steekproeven

$n_1, \bar{x}_1$ en s_1 en $n_2, \bar{x}_2$ en s_2

Een verdeling $T(\nu)$

$$\nu = \left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2 \Big/ \frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}$$

Een toetsingsgrootheid t

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

De p-waarde

tweezijdige toets: $p = 2 \cdot P(T < t)$ of $p = 2 \cdot P(T > t)$

eenzijdige toets: $p = P(T < t)$ of $p = P(T > t)$

t-toets met 2 steekproeven

Voorbeeld

Verkoopt de concurrentie grotere televisies?

We doen twee steekproeven: bij ons (1) en bij de concurrentie (2)

$n_1 = 50, \bar{x}_1 = 42$ inch en $s_1 = 10$ inch

$n_2 = 30, \bar{x}_2 = 43$ inch en $s_2 = 12$ inch

Stappenplan

1. Kies voor eenzijdige of tweezijdige test.
2. Formuleer H_0 en H_1
3. Kies significantieniveau α
4. Bepaal de p-waarde van toetsingsgrootheid $t = (\bar{x}_1 - \bar{x}_2) / \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$
5. Wat besluit je?

t-toets met 2 steekproeven

Stappenplan

1. Kies voor eenzijdige of tweezijdige test.

   eenzijdige toets

2. Formuleer H_0 en H_1

   H_0: $\bar{x}_1 = \bar{x}_2$

   H_1: $\bar{x}_1 < \bar{x}_2$

3. Kies significantieniveau α

   $\alpha = 0.01$

t-toets met **2** steekproeven

t-toets met 2 steekproeven

Stappenplan

4. Bepaal de p-waarde met van $t = (\bar{x}_1 - \bar{x}_2) / \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

$$t = (42 - 43) / \sqrt{\frac{10^2}{50} + \frac{12^2}{30}} = -0.3835$$

$$\nu = \left(\frac{10^2}{50} + \frac{12^2}{30}\right)^2 \Big/ \frac{\left(\frac{10^2}{50}\right)^2}{49} + \frac{\left(\frac{12^2}{30}\right)^2}{39} = 52.78$$

$$p = P(T < t) = 0.3515$$

t-toets met 2 steekproeven

Stappenplan

5. Wat besluit je?

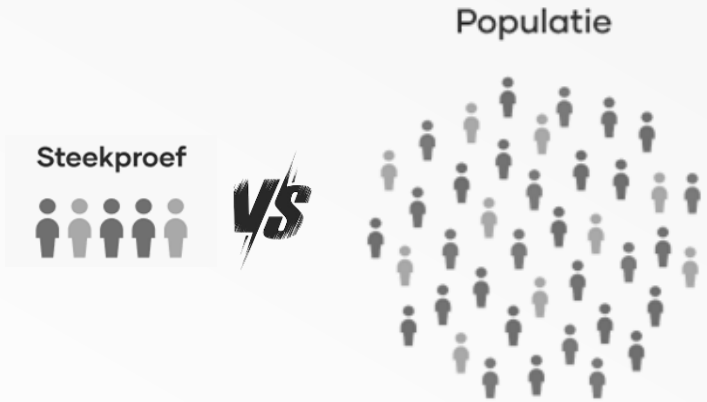
$p = 0.3515$

$p > \alpha$

→ We kunnen H_0 **niet** verwerpen!

→ De concurrentie verkoopt dus geen grotere televisies!

WELKE TESTEN ZIJN ER?



t-toets met **1** steekproef

t-toets met **2** steekproeven

ANOVA

Chi-kwadraat toets (χ^2)

WAT HEB JE NODIG VOOR DEZE TEST?

Hypotheses

Dit zijn veronderstellingen, genaamd H_0 en H_1.

Significantieniveau

α

k steekproeven met elk een andere waarde voor 1 categorische variabele

$n, n_i, \bar{x}_i$ steekproefgegevens voor de verschillende steekproeven.

Een verdeling $F(\alpha, \beta)$

$\alpha = k - 1, \beta = n - k$

Een toetsingsgrootheid f

$f = \frac{MSB}{MSW}$ (zie volgende slides)

De p-waarde

eenzijdige toets: $p = P(F > f)$

Variantieanalyse (ANOVA)

Voorbeeld

Is er een verschil tussen de scores op 3 verschillende vakken afgelegd door 5 studenten?

k groepen

Student	Programmeren 1	Data Science 1	User Interfaces 1
Bob	10	15	17
Nina	12	16	17
Tim	12	17	16
Kate	11	15	17
Alonzo	10	12	13
$\bar{x} = 14$	$\bar{x}_1 = 11$	$\bar{x}_2 = 15$	$\bar{x}_3 = 16$

gemiddelde van alle observaties $\bar{x} = 14$

steekproef 1 $n_1 = 5$ $\bar{x}_1 = 11$

steekproef 2 $n_2 = 5$ $\bar{x}_2 = 15$

steekproef 3 $n_3 = 5$ $\bar{x}_3 = 16$

ANOVA

Variantieanalyse (ANOVA)

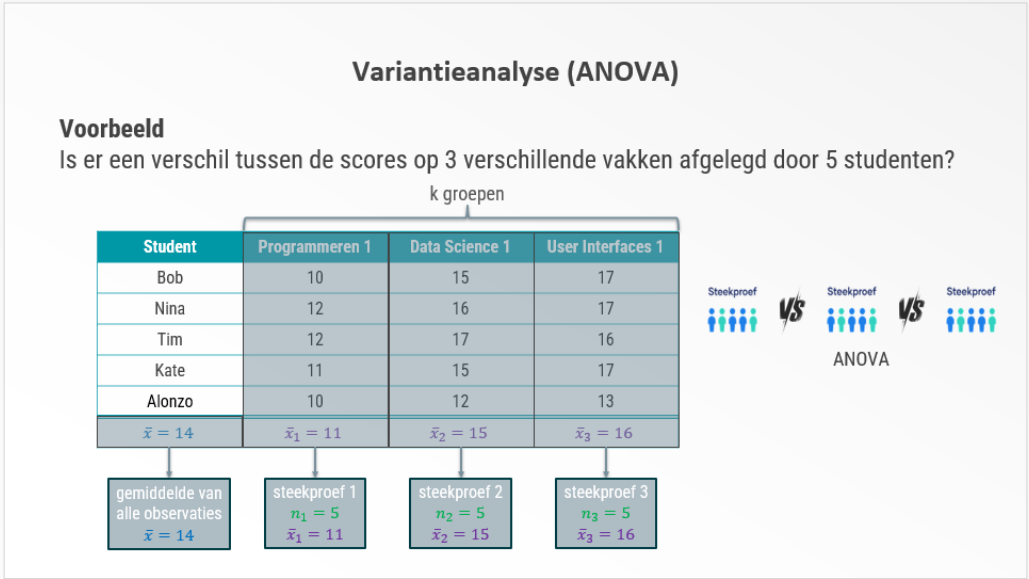
Sum of Squares Total = Sum of Squares Between + Sum of Squares Within

$$SST = SSB + SSW$$

- $SST = \sum_{i=1}^{n}(x_i - \bar{x})^2$
  $SST = (10 - 14)^2 + (12 - 14)^2 + \cdots + (13 - 14)^2$
  $SST = 100$

- $SSB = \sum_{i=1}^{k} n_i \cdot (\bar{x}_i - \bar{x})^2$
  $SSB = 5 \cdot (11 - 14)^2 + 5 \cdot (15 - 14)^2 + 5 \cdot (16 - 14)^2$
  $SSB = 70$

- $SSW = \sum_{j=1}^{k} \sum_{i=1}^{n_i} (x_i - \bar{x}_j)^2$
  $= (10 - 11)^2 + (12 - 11)^2 + \cdots + (10 - 11)^2 +$
  $(15 - 15)^2 + (16 - 15)^2 + \cdots + (12 - 15)^2 +$
  $(17 - 16)^2 + (17 - 16)^2 + \cdots + (13 - 16)^2$
  $= 30$



Variantieanalyse (ANOVA)

Mean SSB en Mean SSW

We delen de SSB – en SSW – waarden door hun respectievelijke degrees of freedom (df)

$$MSB = \frac{SS_B}{df_B} \text{ en } df_B = k - 1$$

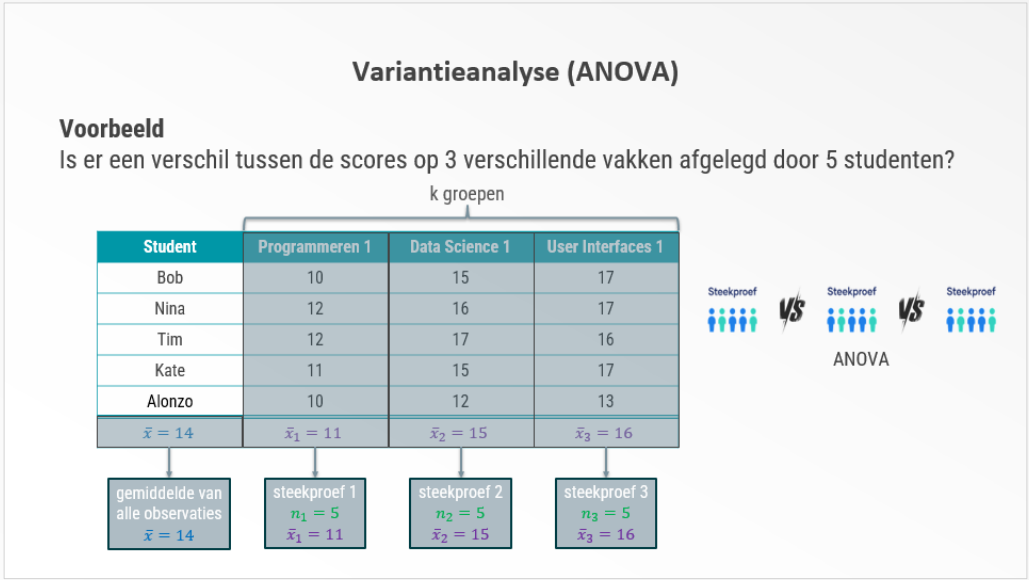
$$MSB = \frac{70}{3-1}$$

$$MSB = 35$$

$$MSW = \frac{SS_W}{df_W} \text{ en } df_W = n - k$$

$$MSW = \frac{30}{15-3}$$

$$= 2.5$$



Variantieanalyse (ANOVA)

Stappenplan

1. Kies voor eenzijdige of tweezijdige test.

   eenzijdige toets

2. Formuleer H_0 en H_1

   $H_0: \mu_1 = \mu_2 = \mu_3$

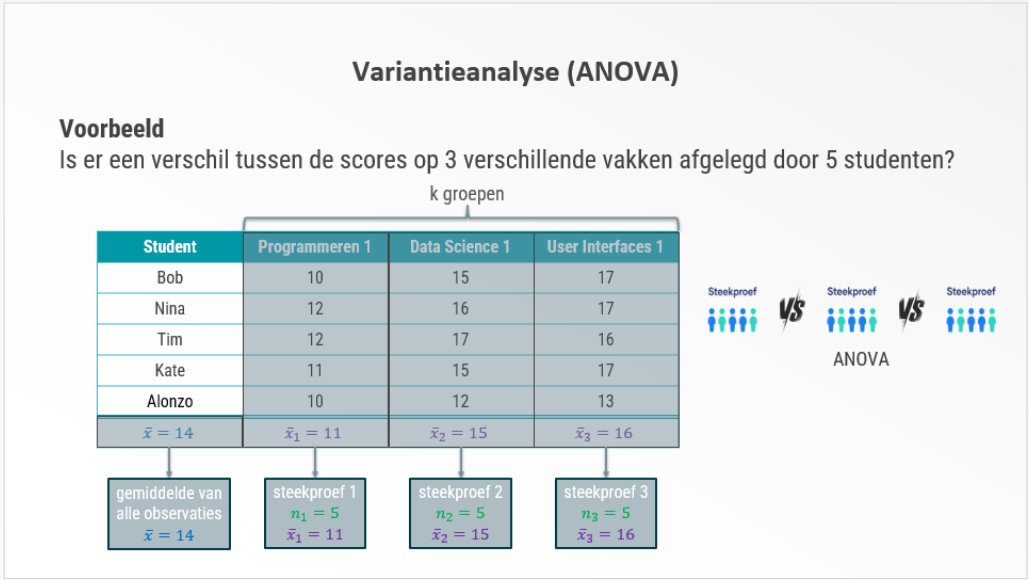
   $H_1: \mu_1 \neq \mu_2 = \mu_3$ of

   $\mu_1 = \mu_2 \neq \mu_3$ of

   $\mu_1 \neq \mu_2 \neq \mu_3$

3. Kies significantieniveau α

   $\alpha = 0.05$



Variantieanalyse (ANOVA)

Stappenplan

4. Bepaal de p-waarde van $f = \frac{MS_{between}}{MS_{within}}$

   $f = \frac{35}{2.5} = 14$

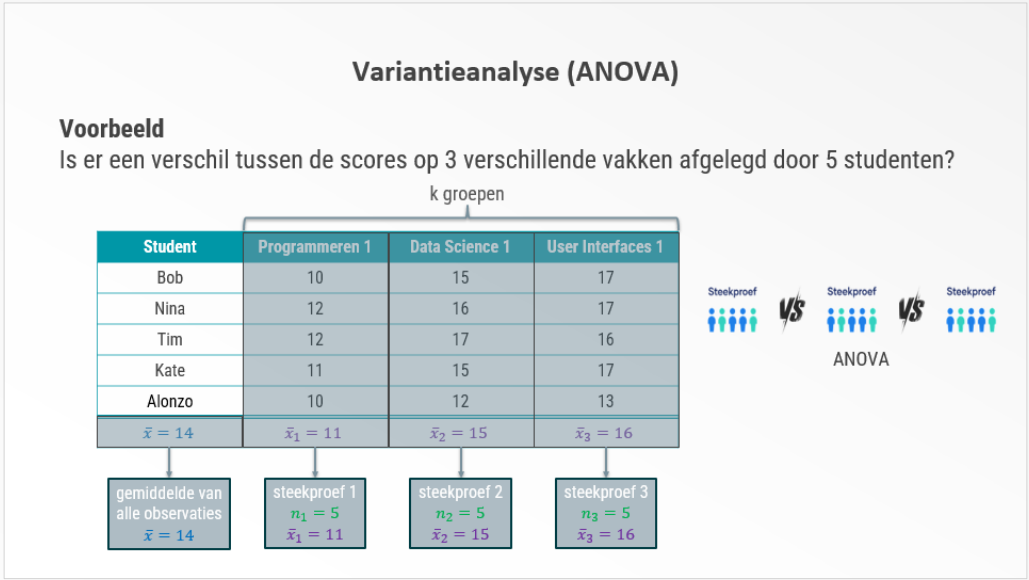
   $p = P(F > f) = 0.007$

5. Wat besluit je?

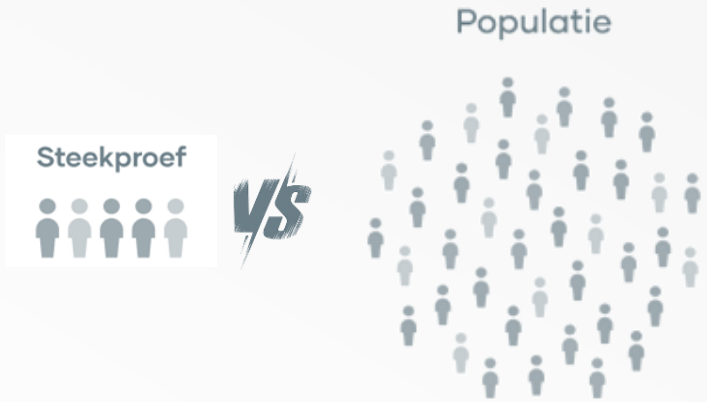
   $p = 0.007 < 0.05 = \alpha$

→ We verwerpen H_0!

→ Er is wel degelijk een verschil tussen de scores op de vakken!



WELKE TESTEN ZIJN ER?



t-toets met **1** steekproef

t-toets met **2** steekproeven

ANOVA

Chi-kwadraat toets (χ^2)

WAT HEB JE NODIG VOOR DEZE TEST?

Hypotheses

Dit zijn veronderstellingen, genaamd H_0 en H_1.

Significantieniveau

α

Een steekproef

Geobserveerde frequenties O en berekende verwachte frequenties E

Een verdeling $\chi^2(\nu)$

$\nu = (n-1) \cdot (m-1)$

Een toetsingsgrootheid Q

$$Q = \sum_i \frac{(o_i - e_i)^2}{e_i}$$

De p-waarde

eenzijdige toets: $p = P(\chi^2 > Q)$

χ^2-toets

Voorbeeld

Hebben witte producten een slechtere koeling?

Geobserveerde frequenties

	Wit merk	Geen wit merk	Totalen
Slechte koeling	1498	1513	**3011**
Goede koeling	504	6485	**6989**
Totalen	**2002**	**7998**	**10000**

Chi-kwadraat toets (χ^2)

Verwachte frequenties

	Wit merk	Geen wit merk	Totalen
Slechte koeling	$\frac{3011 \cdot 2002}{10000} = \mathbf{602.8022}$	$\frac{3011 \cdot 7998}{10000} = \mathbf{2408.1978}$	**3011**
Goede koeling	$\frac{6989 \cdot 2002}{10000} = \mathbf{1399.1978}$	$\frac{6989 \cdot 7998}{10000} = \mathbf{5589.8022}$	**6989**
Totalen	**2002**	**7998**	**10000**

χ^2-toets

Stappenplan

1. Kies voor eenzijdige of tweezijdige test.

   eenzijdige toets

2. Formuleer H_0 en H_1

   $H_0: Q = 0$

   $H_1: Q > 0$

3. Kies significantieniveau α

   $\alpha = 0.05$

Chi-kwadraat toets (χ^2)

χ^2-toets

Stappenplan

4. Bepaal de p-waarde met van $q = \sum_i \frac{(o_i - e_i)^2}{e_i}$

$$q = \frac{(1498-602.8)^2}{602.8} + \frac{(1513-2408.2)^2}{2408.2} + \frac{(504-1399.2)^2}{1399.2} + \frac{(6485-5589.8)^2}{5589.8} = 2378{,}3$$

$$\nu = (2-1)\cdot(2-1)$$

$$p = P(\chi^2 > Q) = 0$$

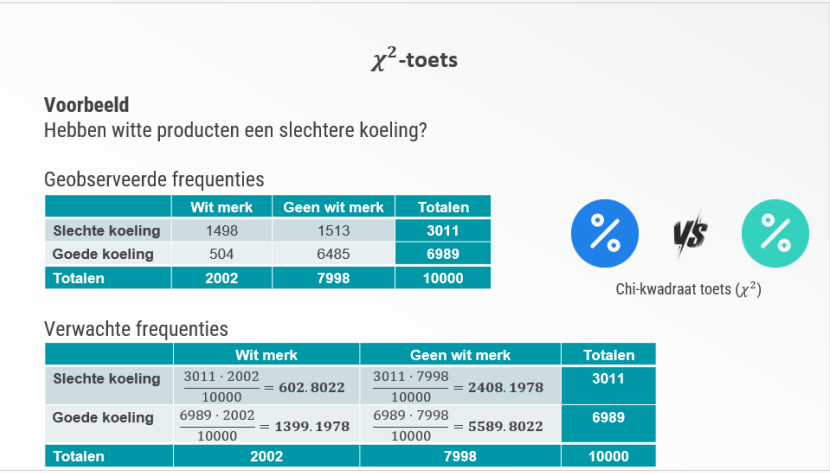
5. Wat besluit je?

$$p = 0 < 0.05 = \alpha$$

→ We verwerpen H_0!

→ Wit producten hebben een slechtere koeling.

Chi-kwadraat toets (χ^2)



χ^2-toets

Voorbeeld
Hebben witte producten een slechtere koeling?

Geobserveerde frequenties

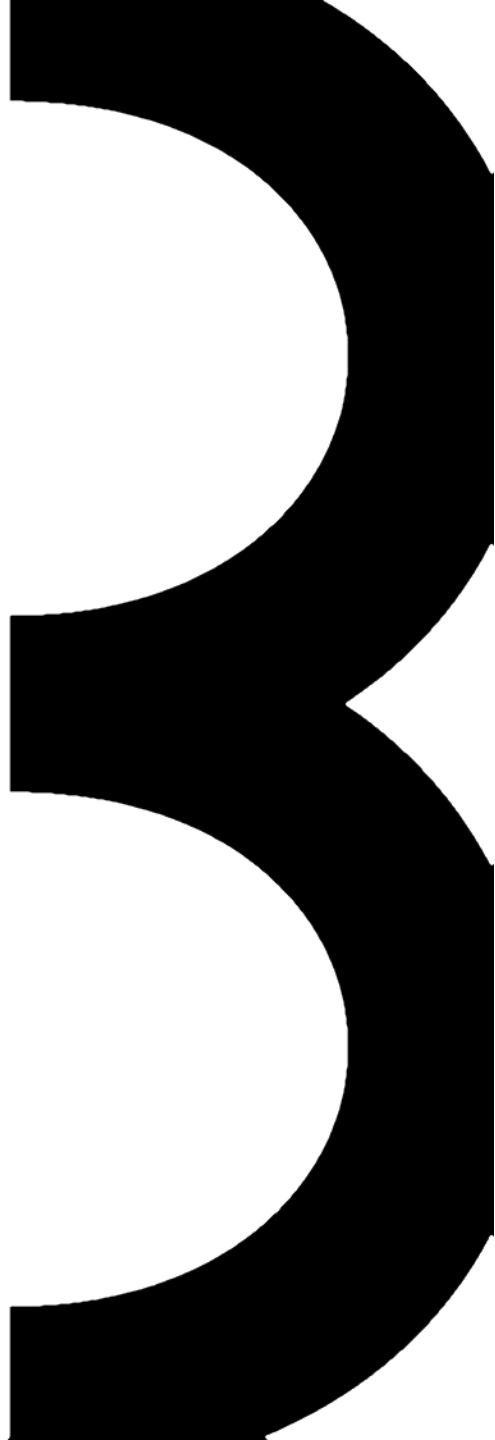
	Wit merk	Geen wit merk	Totalen
Slechte koeling	1498	1513	3011
Goede koeling	504	6485	6989
Totalen	2002	7998	10000

Chi-kwadraat toets (χ^2)

Verwachte frequenties

	Wit merk	Geen wit merk	Totalen
Slechte koeling	$\frac{3011 \cdot 2002}{10000} = 602.8022$	$\frac{3011 \cdot 7998}{10000} = 2408.1978$	3011
Goede koeling	$\frac{6989 \cdot 2002}{10000} = 1399.1978$	$\frac{6989 \cdot 7998}{10000} = 5589.8022$	6989
Totalen	2002	7998	10000

HOE ZEKER ZIJN WE NU?



HOE ZEKER ZIJN WE NU?

Fouten

Risico op het maken van twee soorten fouten bij de interpretatie van de resultaten:

Type I-fout

- de nulhypothese H_0 verwerpen, terwijl deze eigenlijk wel waar is.
- de kans hierop is α (significantie)

Type II-fout

- de nulhypothese H_0 niet verwerpen, terwijl deze eigenlijk onjuist is.
- de kans hierop is β

H_0 is …	Waar	Onwaar
Verworpen	Type I – fout α	Correcte beslissing
Niet verworpen	Correcte beslissing	Type II – fout β

HOE ZEKER ZIJN WE NU?

Voorbeeld

Je besluit je te laten testen op corona, omdat je milde symptomen hebt.

Er zijn twee fouten die mogelijk kunnen optreden:

Type I-fout

→ testresultaat beweert dat je corona hebt, maar dat heb je eigenlijk niet.

→ vals positief

Type II-fout

→ testresultaat beweert dat je geen corona hebt, maar dat heb je eigenlijk wel.

→ vals negatief

HOE ZEKER ZIJN WE NU?

Wat betekent dat?

p-waarde $>$ significantieniveau α

→ nulhypothese H_0 niet verworpen

→ resultaten niet statistisch significant

→ geen conclusie mogelijk (ook niet dat H_1 waar is)

→ mogelijk type II-fout gemaakt

p-waarde $<$ significantieniveau α

→ nulhypothese H_0 verworpen

→ resultaten statistisch significant

→ H_1 is (waarschijnlijk) waar

→ mogelijk type I-fout gemaakt

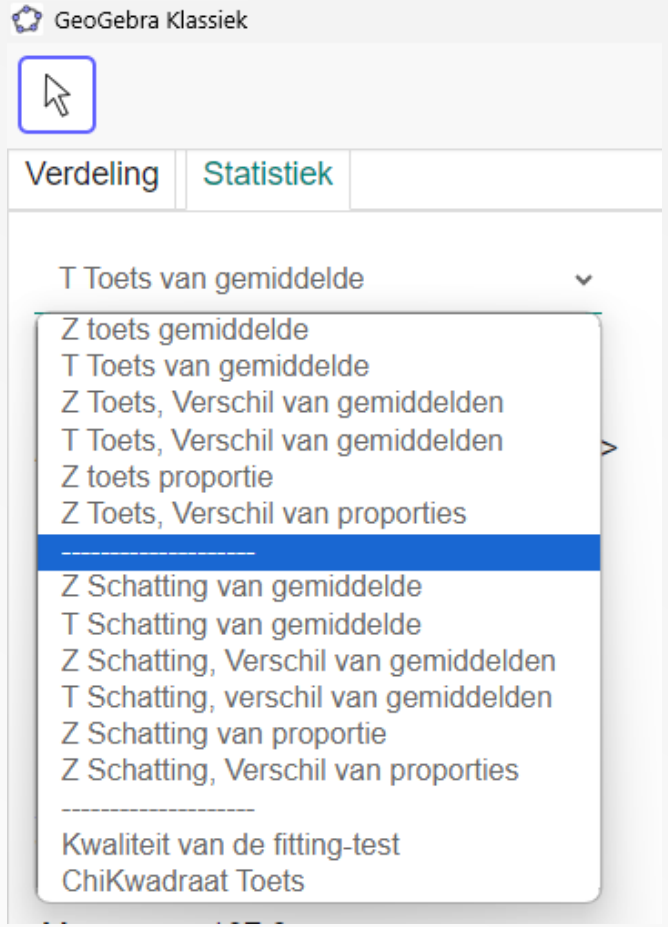
PRAKTIJK

4

GEOGEBRA

Toetsen in GeoGebra

Makkelijkst via Schermindeling → Kansrekening → Statistiek
en de juiste toets kiezen:



- T-toets van gemiddelde = t-toets met 1 steekproef
- T-toets, verschil van gemiddelden = t-toets met 2 steekproeven
- Kwaliteit van de fitting-test = χ^2 - toets
- ChiKwadraat Toets = χ^2 - toets
- F-test, variantieanalyse,ANalysis Of VAriance (ANOVA) (niet getoond)

Kan ook manueel via eigen berekeningen.

GEOGEBRA

Voorbeeld

Iemand beweert dat de gemiddelde schermgrootte van alle verkochte televisies 40 inch is.

- $\mu = 40$ inch

We doen een steekproef van enkele verkochte televisies.

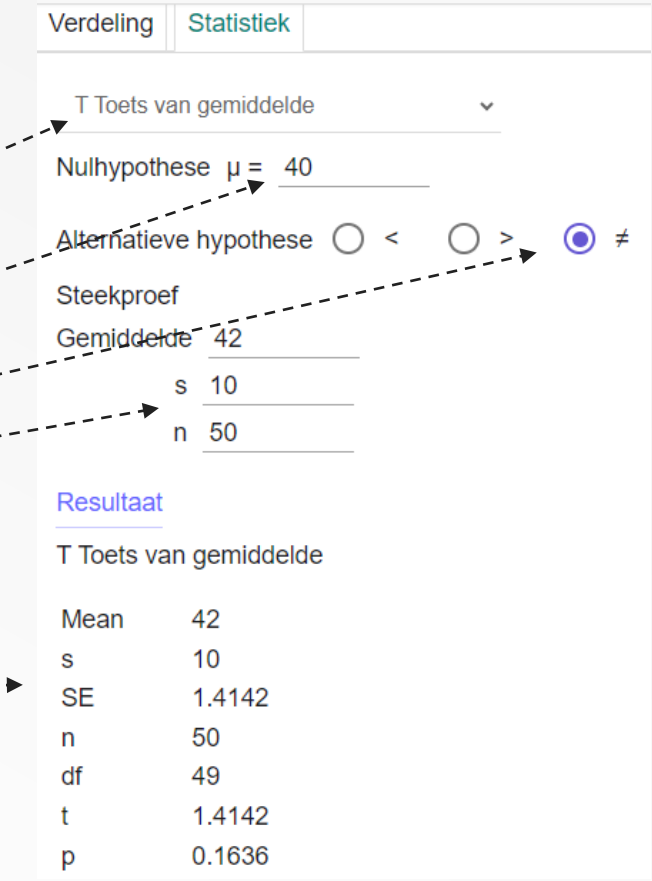
- $n = 50$, $\bar{x} = 42$ inch en $s = 10$ inch

→ Kies voor T-toets van gemiddelde

→ $\neq$ betekent tweezijdige toets

→ Gegevens invullen

→ Resultaten aflezen



GEOGEBRA

Voorbeeld

Verkoopt de concurrentie grotere televisies?

We doen twee steekproeven: bij ons (1) en bij de concurrentie (2)

$n_1 = 50, \bar{x}_1 = 42$ inch en $s_1 = 10$ inch

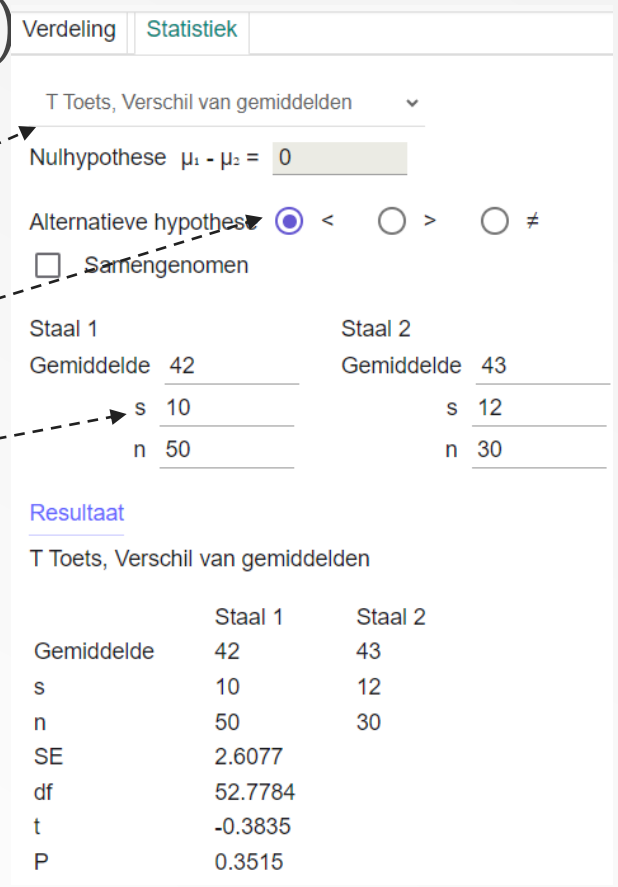
$n_2 = 30, \bar{x}_2 = 43$ inch en $s_2 = 12$ inch

→ Kies voor T-toets, verschil van gemiddelden

→ $<$ betekent eenzijdige toets

→ Gegevens invullen

→ Resultaten aflezen



GEOGEBRA

Voorbeeld

Is er een verschil tussen de scores op 3 verschillende vakken afgelegd door 5 studenten?

→ Schermindeling Rekenblad

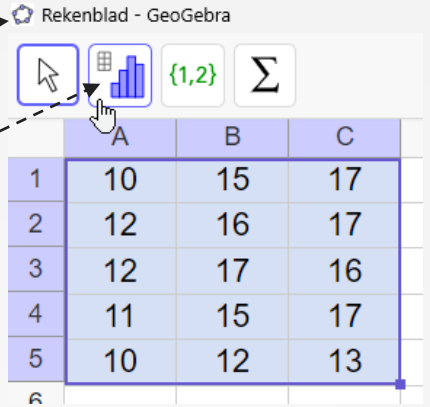
→ Observaties invullen en selecteren

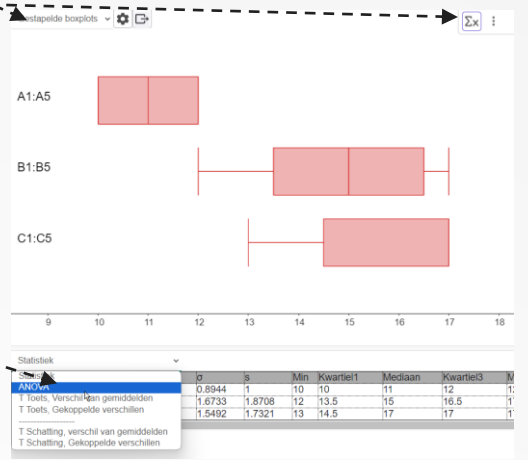
→ Knop induwen en Onderzoek meerdere variabelen kiezen

→ $\sum x$ kiezen klikken

→ ANOVA kiezen bij Statistiek

→ Resultaten aflezen





ANOVA

	df	SS	MS	F	P
Tussen groepen	2	70	35	14	0.0007
Binnen groepen	12	30	2.5		
Totaal	14	100			

	n	Gemiddelde	s
A1:A5	5	11	1
B1:B5	5	15	1.8708
C1:C5	5	16	1.7321

GEOGEBRA

Voorbeeld

Hebben witte producten een slechtere koeling?

Geobserveerde frequenties

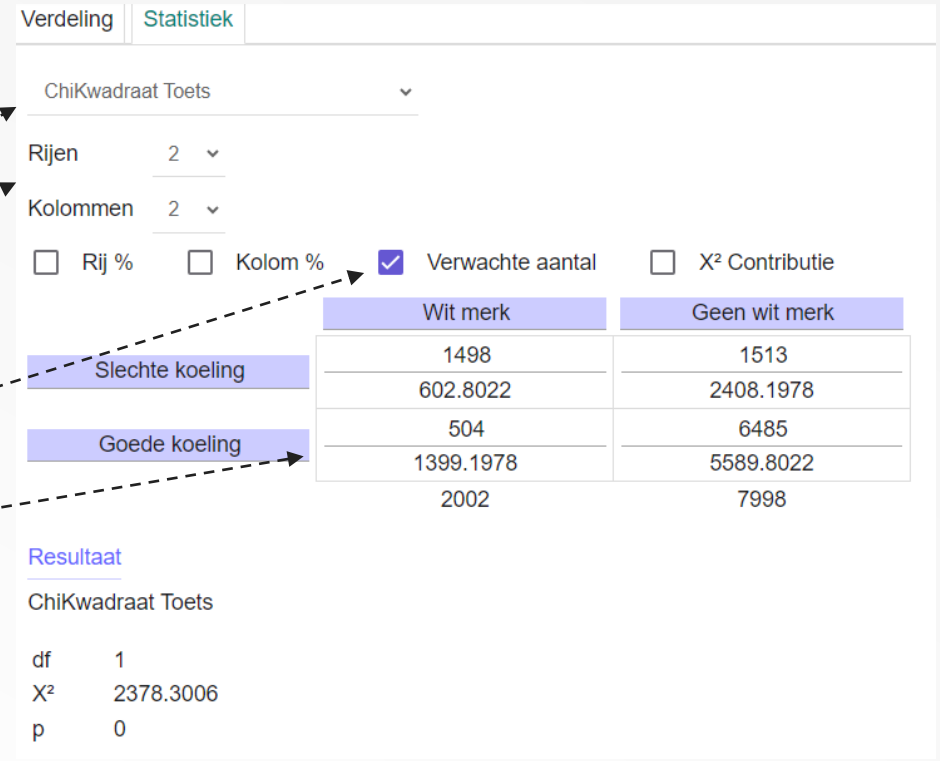
	Wit merk	Geen wit merk	Totalen
Slechte koeling	1498	1513	**3011**
Goede koeling	504	6485	**6989**
Totalen	**2002**	**7998**	**10000**

→ Kies voor ChiKwadraat Toets

→ Kies # rijen en kolommen

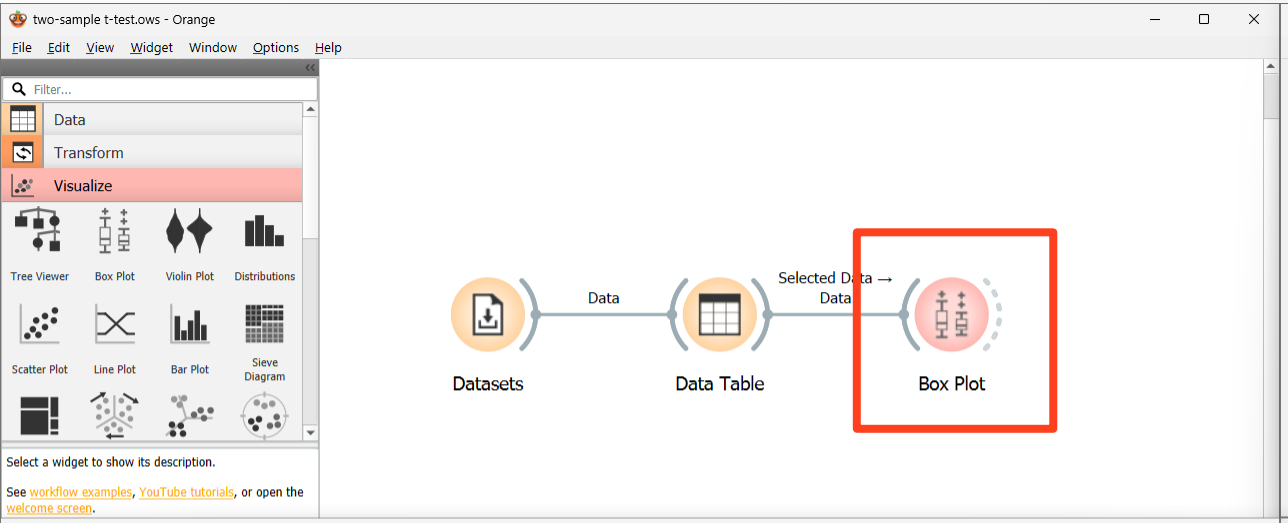
→ Gegevens invullen

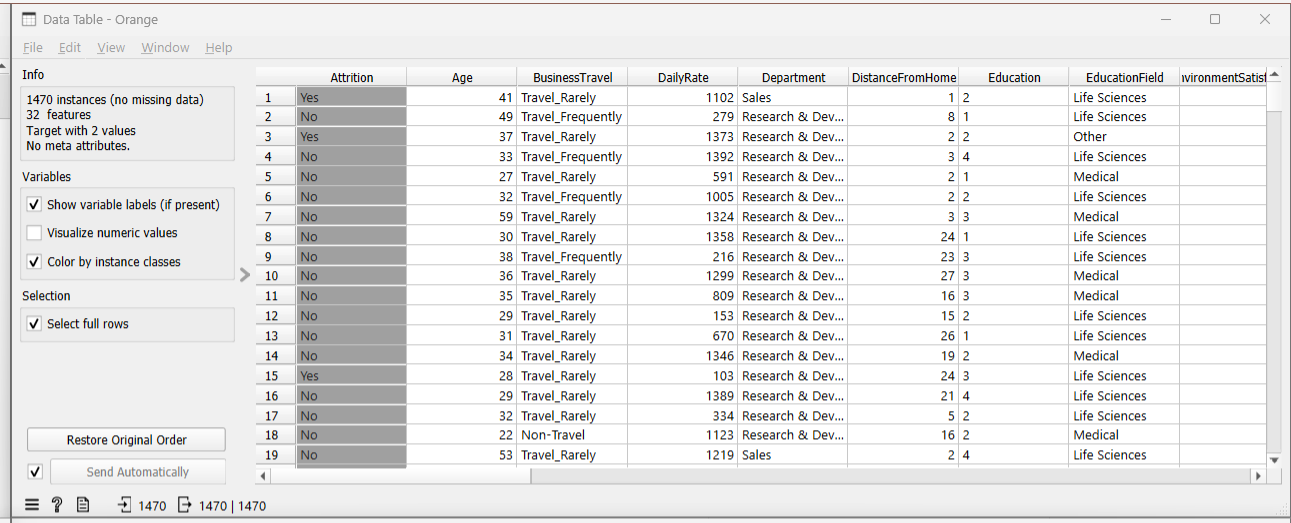
→ Resultaten aflezen



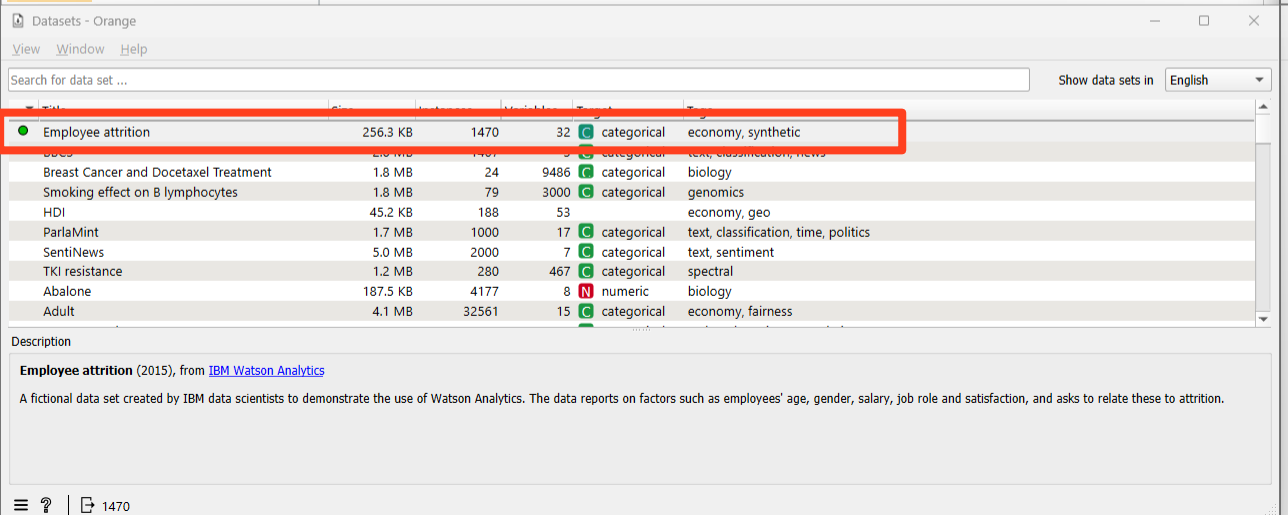
ORANGE

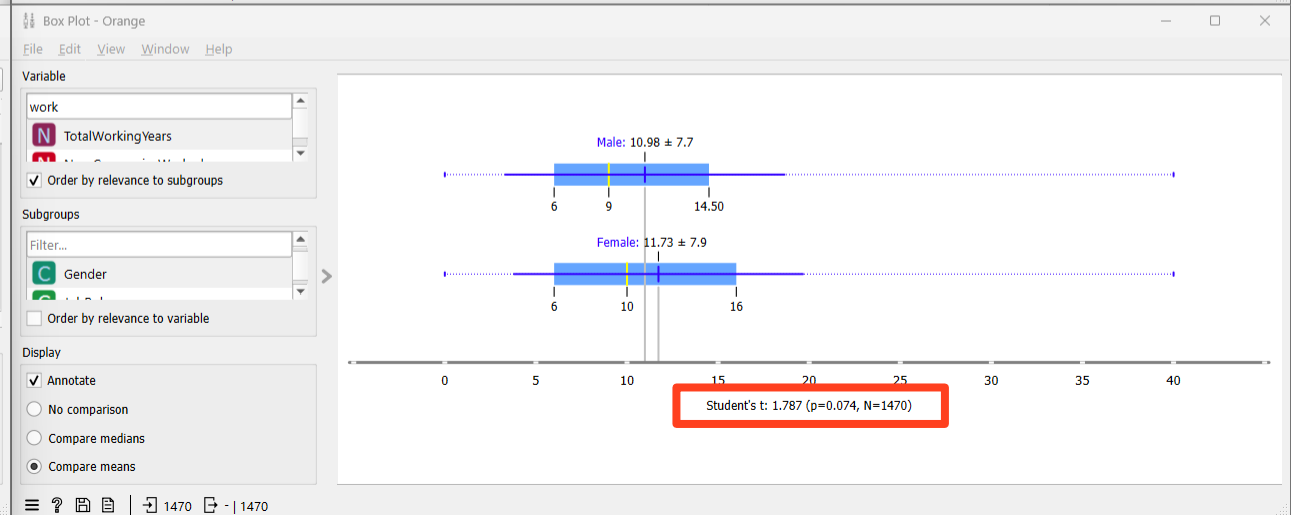
Werken mannen langer dan vrouwen? (significantieniveau α=0.10)





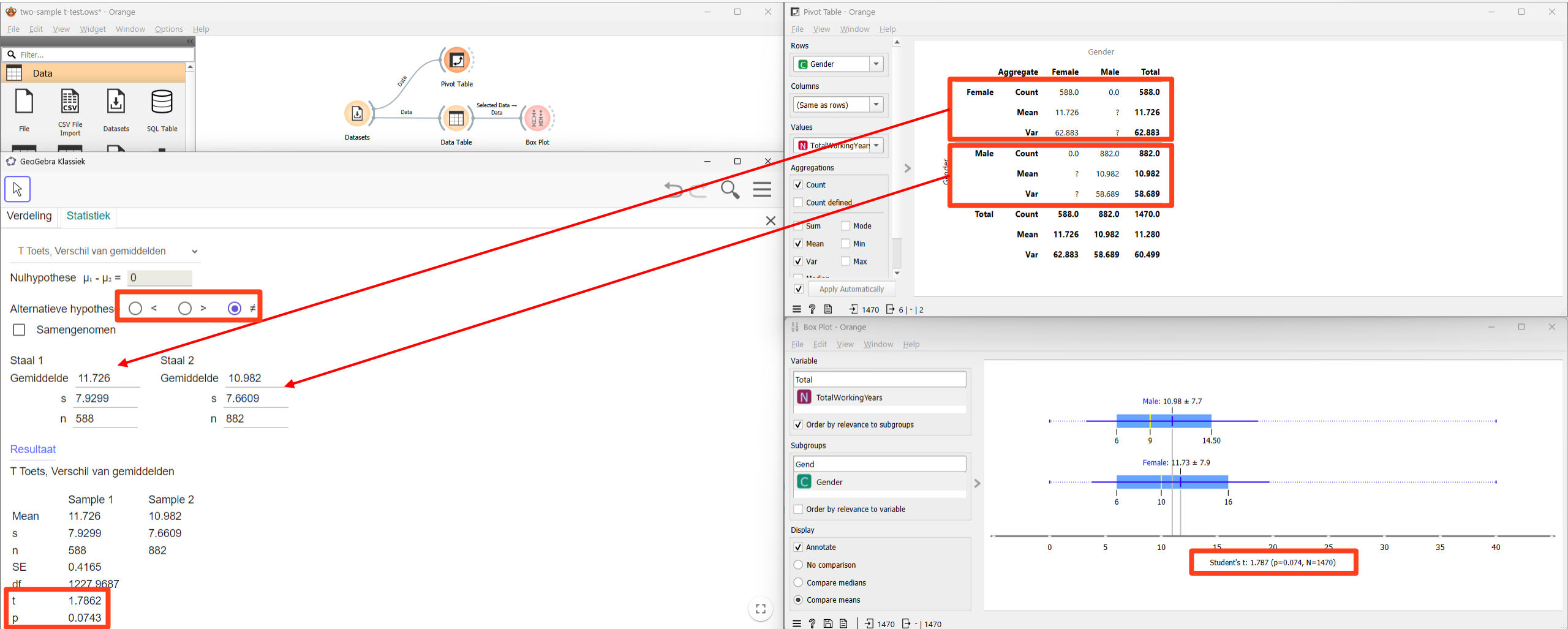
	Attrition	Age	BusinessTravel	DailyRate	Department	DistanceFromHome	Education	EducationField
1	Yes	41	Travel_Rarely	1102	Sales	1	2	Life Sciences
2	No	49	Travel_Frequently	279	Research & Dev...	8	1	Life Sciences
3	Yes	37	Travel_Rarely	1373	Research & Dev...	2	2	Other
4	No	33	Travel_Frequently	1392	Research & Dev...	3	4	Life Sciences
5	No	27	Travel_Rarely	591	Research & Dev...	2	1	Medical
6	No	32	Travel_Frequently	1005	Research & Dev...	2	2	Life Sciences
7	No	59	Travel_Rarely	1324	Research & Dev...	3	3	Medical
8	No	30	Travel_Rarely	1358	Research & Dev...	24	1	Life Sciences
9	No	38	Travel_Frequently	216	Research & Dev...	23	3	Life Sciences
10	No	36	Travel_Rarely	1299	Research & Dev...	27	3	Medical
11	No	35	Travel_Rarely	809	Research & Dev...	16	3	Medical
12	No	29	Travel_Rarely	153	Research & Dev...	15	2	Life Sciences
13	No	31	Travel_Rarely	670	Research & Dev...	26	1	Life Sciences
14	No	34	Travel_Rarely	1346	Research & Dev...	19	2	Medical
15	Yes	28	Travel_Rarely	103	Research & Dev...	24	3	Life Sciences
16	No	29	Travel_Rarely	1389	Research & Dev...	21	4	Life Sciences
17	No	32	Travel_Rarely	334	Research & Dev...	5	2	Life Sciences
18	No	22	Non-Travel	1123	Research & Dev...	16	2	Medical
19	No	53	Travel_Rarely	1219	Sales	2	4	Life Sciences





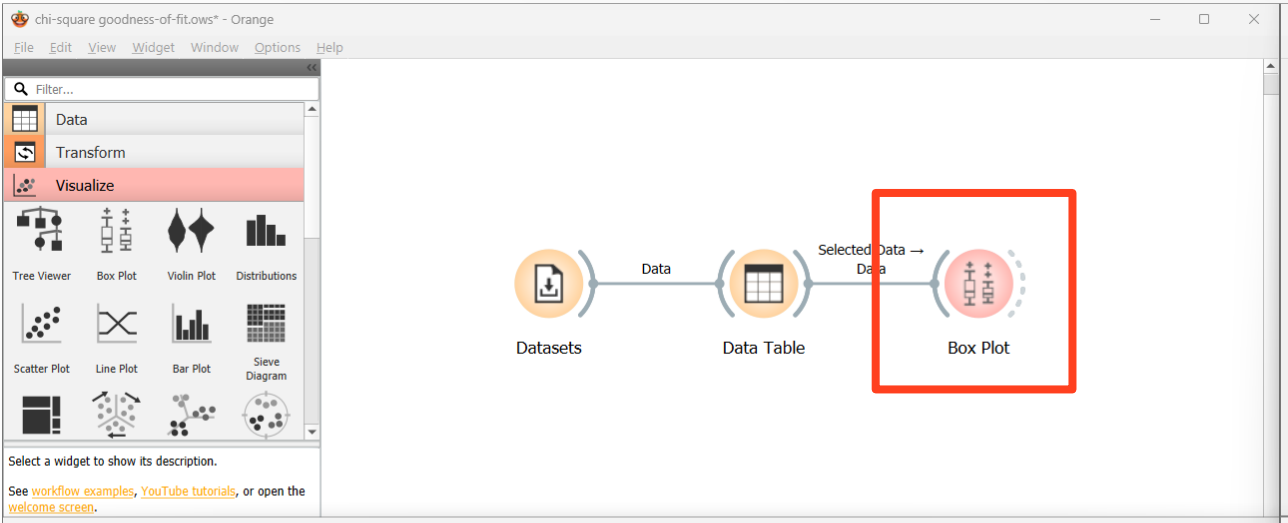
ORANGE

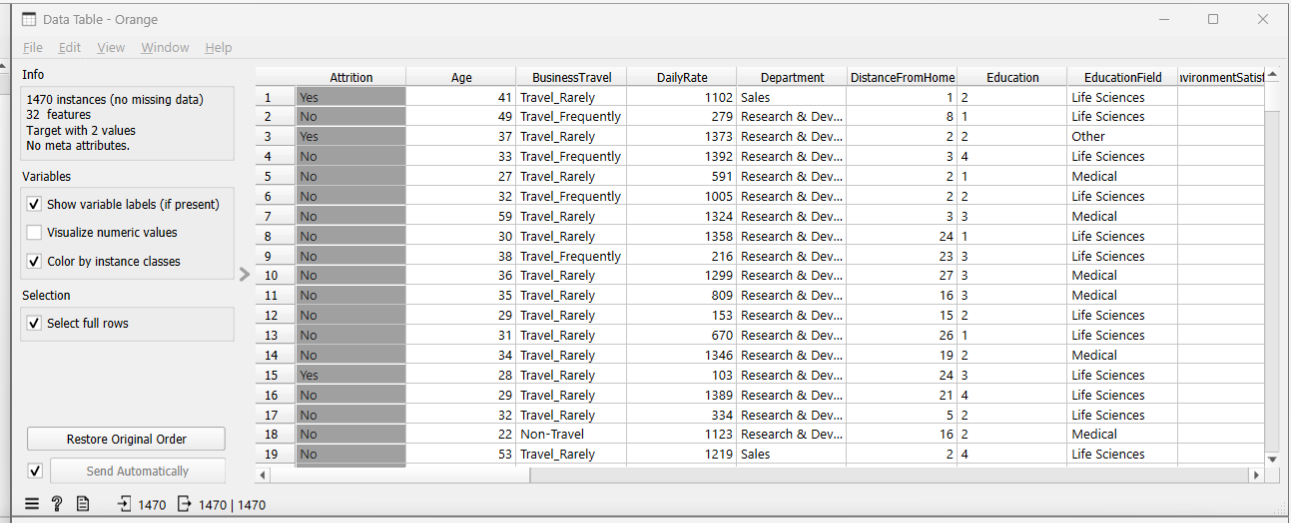
Werken mannen langer dan vrouwen? (significantieniveau α=0.10)

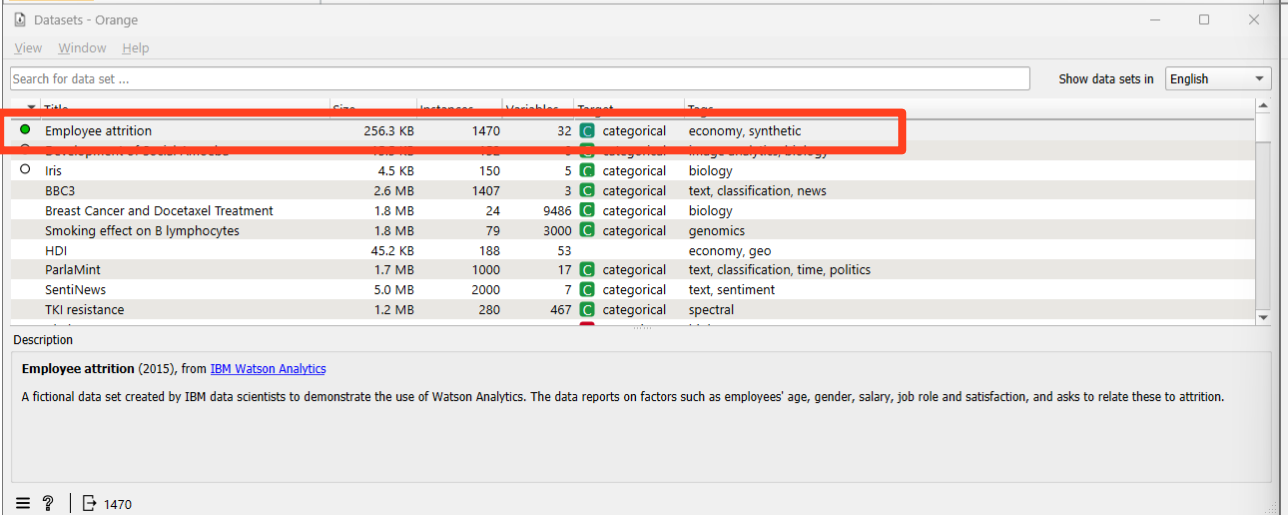


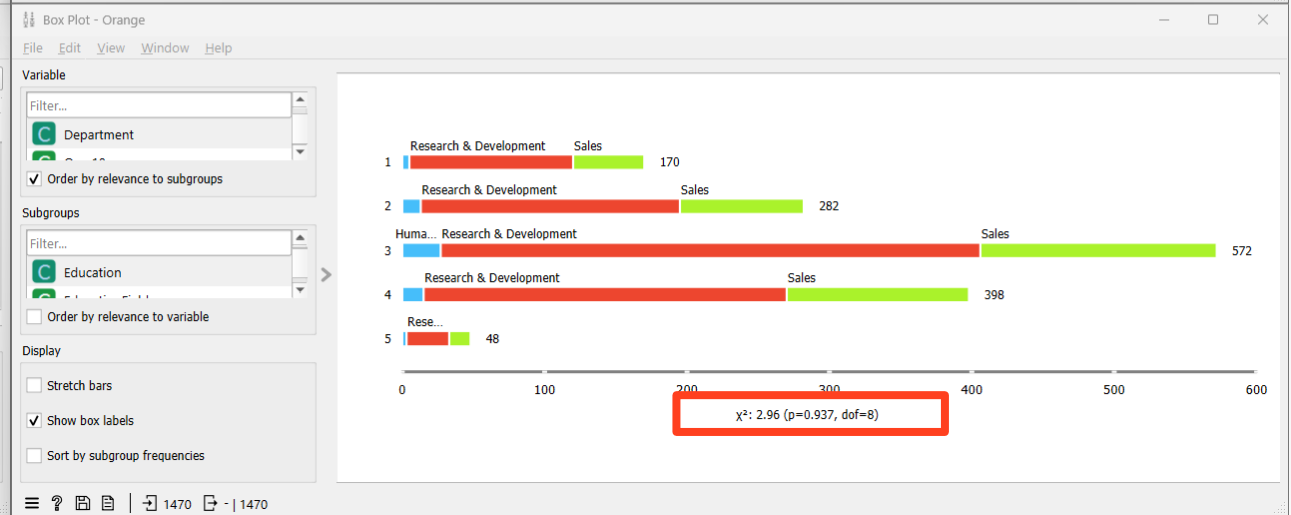
ORANGE

Heeft het departement waarvoor je werkt een invloed op je scholingsgraad (1-5)?



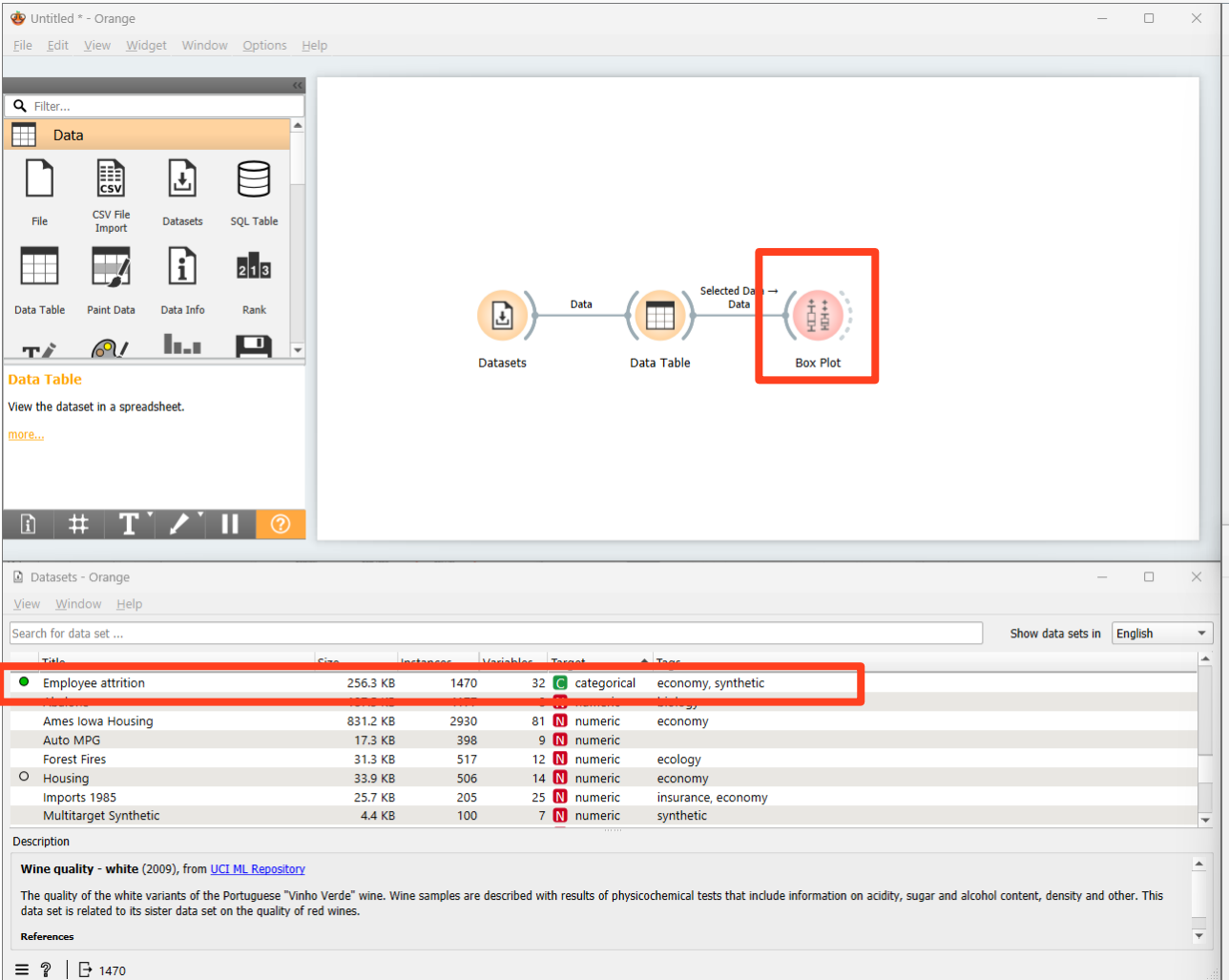


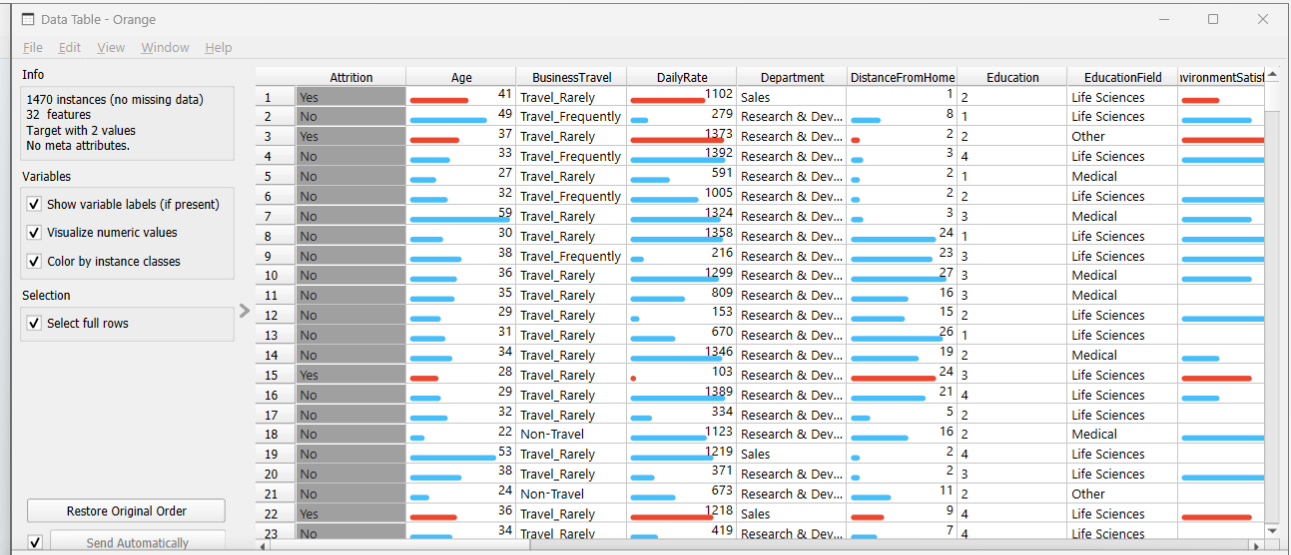


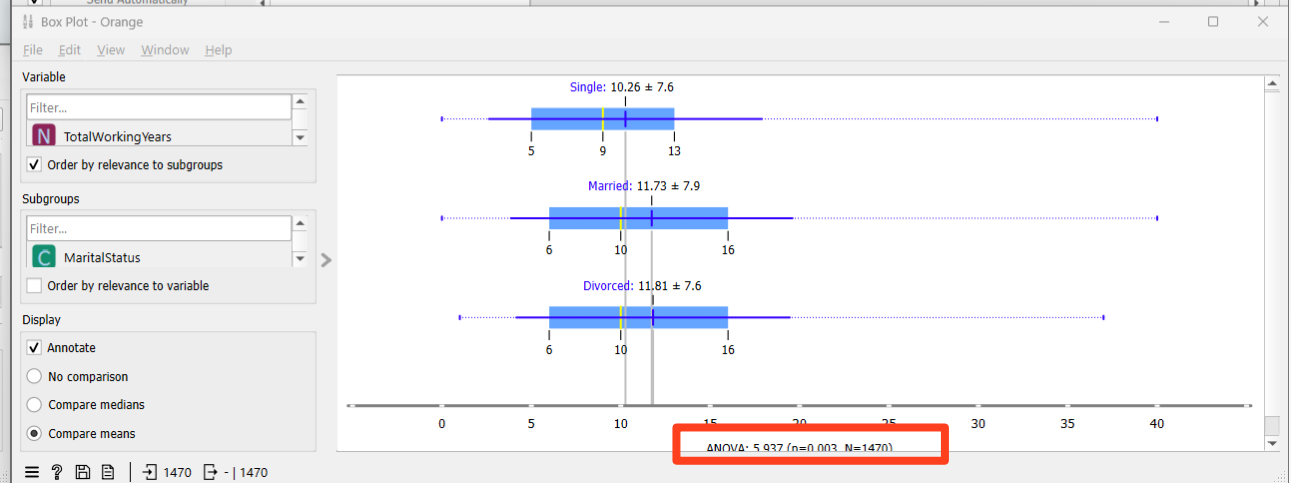


ORANGE

Is er een verschil in het totaal aan gewerkte uren naargelang je burgerlijke staat?

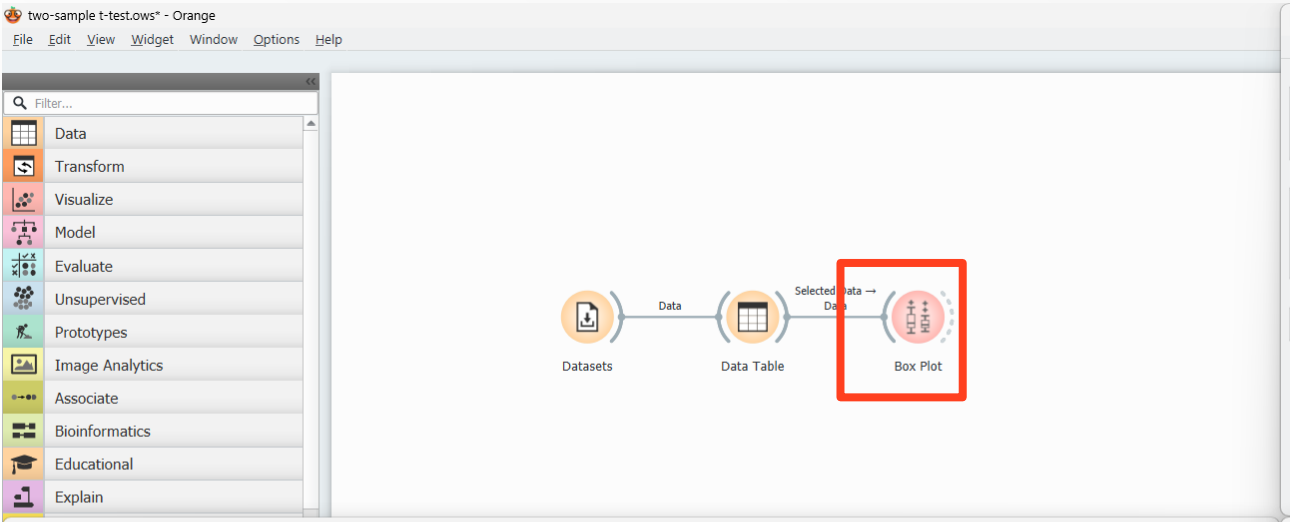


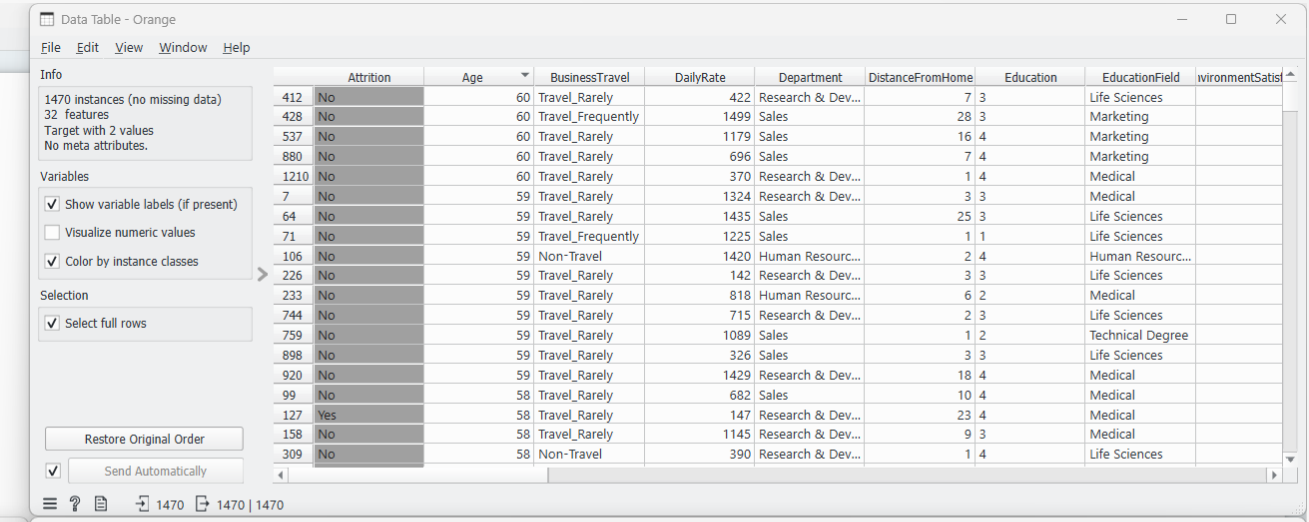


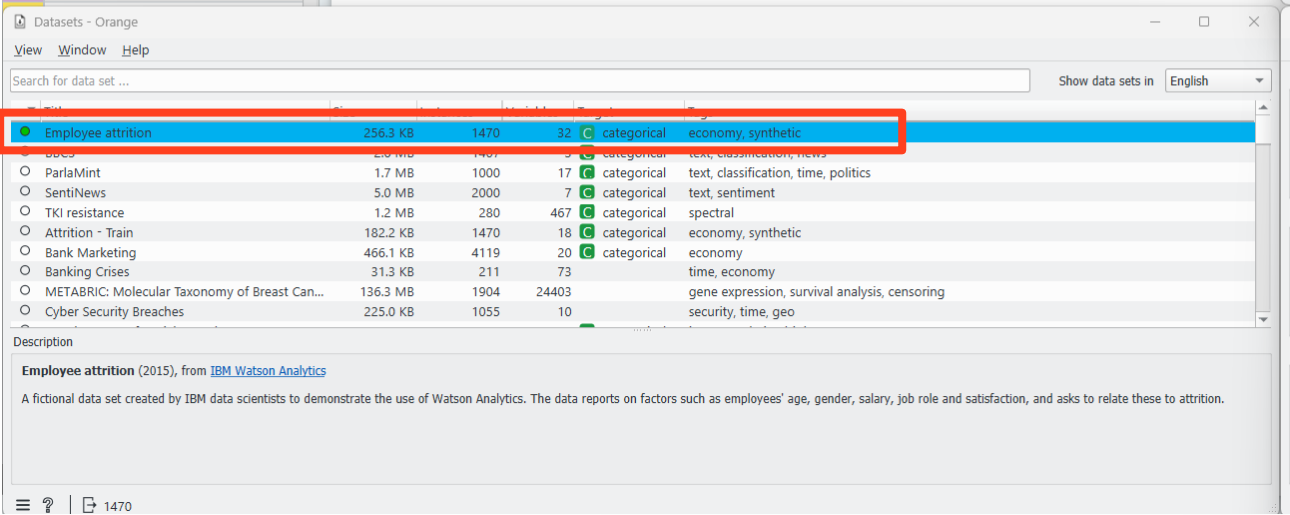


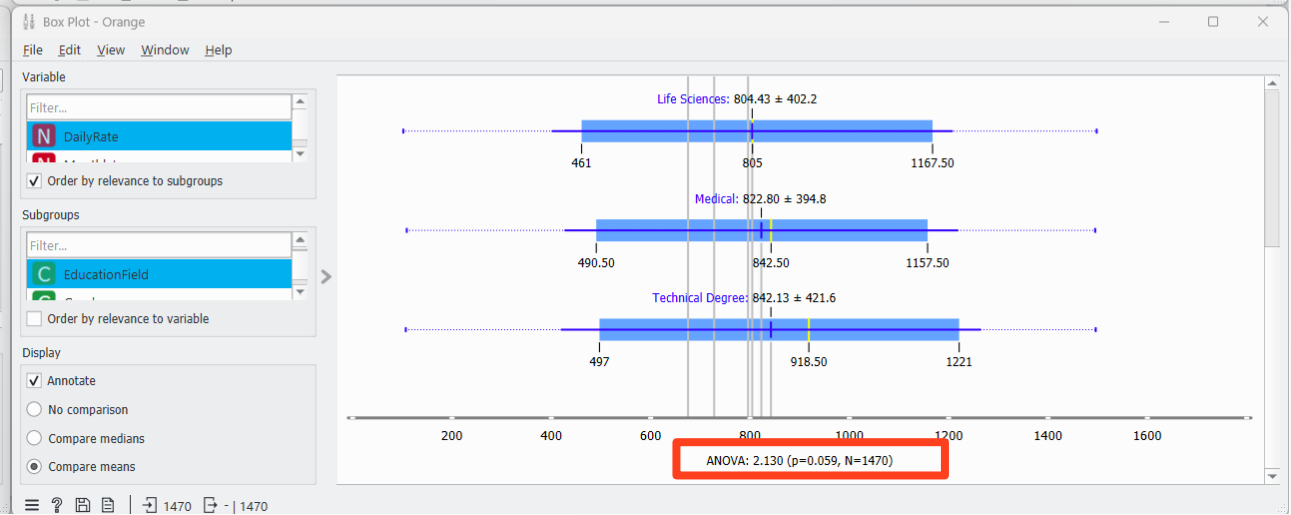
ORANGE

Wordt je meer betaald afhankelijk van je opleidingsniveau?









Wat hebben we geleerd?

- **populatie**
  de volledige verzameling van individuen, objecten of gebeurtenissen waarop een onderzoek gericht is.

- **steekproef**
  een deelverzameling van de populatie die wordt geanalyseerd om conclusies te trekken over de hele populatie. Deze deelverzameling komt tot stand door toeval.

- **populatieparameter**
  een kenmerk of eigenschap van de gehele populatie, bijvoorbeeld het gemiddelde of de standaardafwijking.

- **steekproefparameter**
  een kenmerk of eigenschap van de steekproef, gebruikt om uitspraken te doen over de populatieparameter.

- **centrale limietstelling**
  het gemiddelde van een toevalsvariabele zal, ongeacht de oorspronkelijke verdeling van de toevalsvariabele, normaal verdeeld zal zijn, en bij kleine steekproeven studentverdeeld.

- **steekproefverdeling**
  de verdeling van een steekproefparameter, zoals het gemiddelde of de variantie, gebaseerd op herhaalde trekkingen van steekproeven uit dezelfde populatie, vaak gelijk aan de Studentverdeling

Wat hebben we geleerd?

- **betrouwbaarheidsinterval**
  een interval voor een onbekende populatieparameter

- **hypothese**
  een bewering over de waarde van een populatieparameter

- **hypothesetest**
  op basis van bewijsmateriaal uit steekproeven, een procedure om te bepalen of de gestelde hypothese een redelijke of onredelijk verklaring is.

- **significantieniveau**
  kans op het maken van een type I fout

- **p-waarde**
  de kans dat een gebeurtenis puur toevallig plaatsvindt, aangenomen dat de nulhypothese waar is; hoe kleiner de p-waarde, hoe sterker het bewijs tegen de nulhypothese is

Wat hebben we geleerd?

- **type I-fout**
  fout die je maakt wanneer je nulhypothese foutief verwerpt

- **type II-fout**
  fout die je maakt wanneer je nalaat om de nulhypothese te verwerpen

- **statistisch significant**
  als p-waarde kleiner is dan α. Hoe kleiner p-waarde hoe meer statistisch significant de test is.