



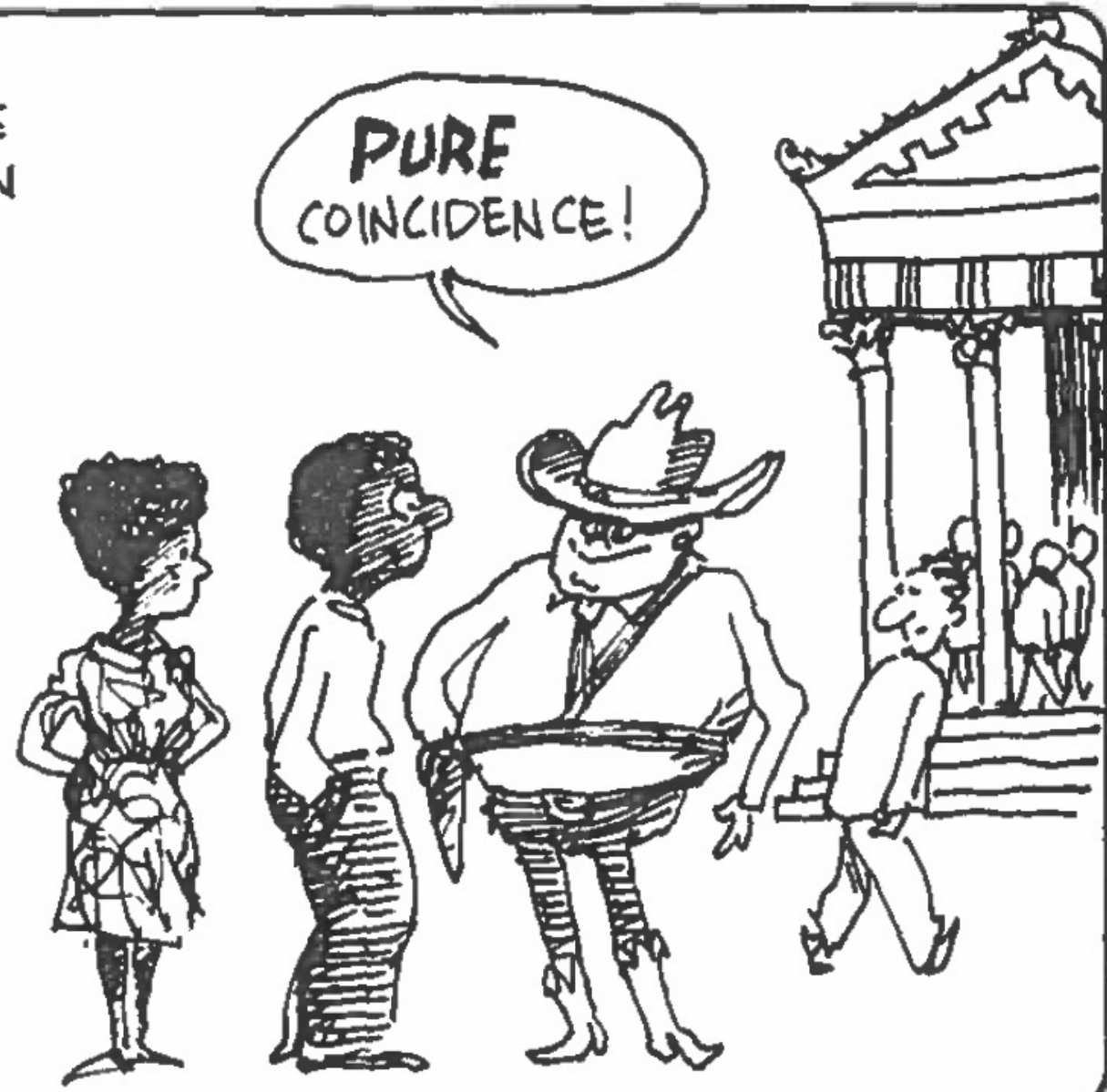
Hypothesetoetsen

Zie ook:
Cursusboek hoofdstuk 25

Inhoud

1. Introductie: Discriminatie bij juryselectie
2. Wat is een hypothesetoets?
3. John Arbuthnot en de eerste hypothesetoets
4. De P -waarde als maat voor het bewijs tegen de nulhypothese
5. De algemene vorm van een hypothesetoets
6. Hypothesetoetsen als besluitprocedures
7. Welke toetsen in deze cursus?

WE BEGIN WITH AN EXAMPLE
FROM THE LAW: A COMPOSITE
OF SEVERAL CASES ARGUED IN
THE SOUTH BETWEEN 1960
AND 1980, IN WHICH EXPERT
WITNESSES PRESENTED THE
CASE FOR RACIAL BIAS IN
JURY SELECTION.



1)

50% OF ELIGIBLE CITIZENS
WERE AFRICAN AMERICAN.



2)

ON AN 80-PERSON PANEL
OF POTENTIAL JURORS,
ONLY *FOUR* WERE
AFRICAN AMERICANS.

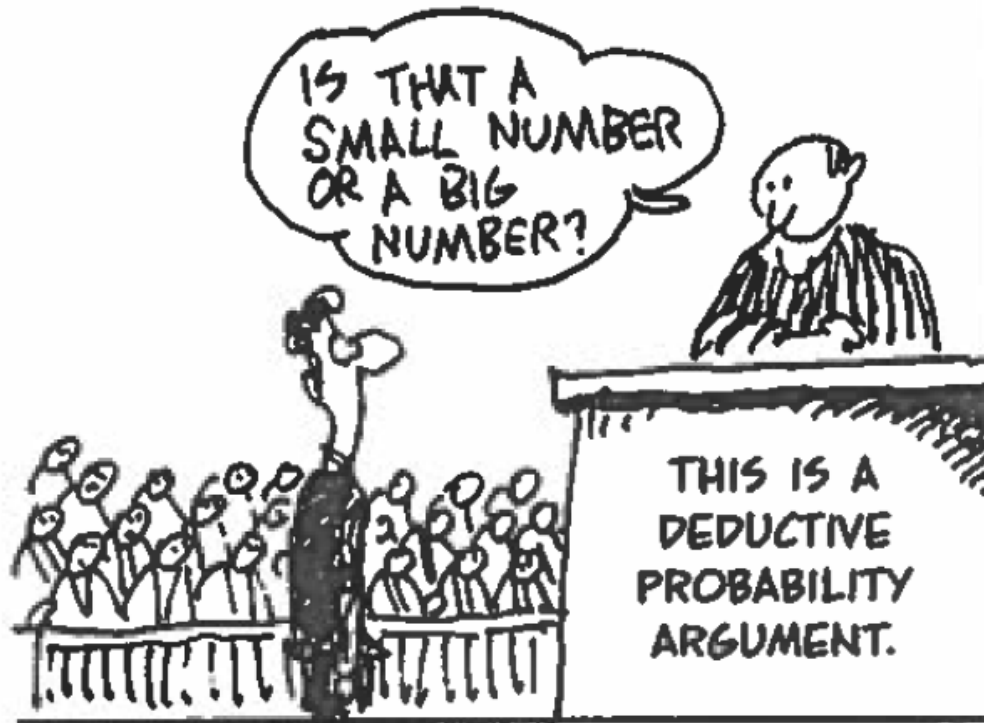


COULD THIS BE THE RESULT OF
PURE CHANCE?

FOR THE SAKE OF ARGUMENT,
SUPPOSE THAT THE SELECTION OF
POTENTIAL JURORS WAS *RANDOM*.
THEN THE NUMBER OF AFRICAN
AMERICANS ON THE 80-PERSON
PANEL WOULD BE THE *BINOMIAL*
RANDOM VARIABLE X WITH
 $n = 80$ TRIALS AND $p = .5$.



THUS, THE CHANCES OF GETTING A JURY WITH ONLY 4 AFRICAN AMERICANS IS $Pr(X \leq 4)$, WHICH WORKS OUT TO ABOUT .000000000000000000014 (!).



SINCE THE PROBABILITY IS SO SMALL, THE PARTICULAR PANEL WITH ONLY FOUR BLACK MEMBERS IS **STRONG EVIDENCE** AGAINST THE **HYPOTHESIS** OF RANDOM SELECTION.



De logische structuur van het argument

- Neem *for the sake of argument* aan dat er *geen* discriminatie heeft plaatsgevonden
- Bereken de kans op een uitkomst die minstens zo extreem is als de observatie $X = 4$.
- Die kans is *heel* erg klein.
- Dat is sterk bewijs dat er iets mis is met de aanname dat er geen discriminatie heeft plaatsgevonden.

Inhoud

1. Introductie: Discriminatie bij juryselectie
2. Wat is een hypothesetoets?
3. John Arbuthnot en de eerste hypothesetoets
4. De P -waarde als maat voor het bewijs tegen de nulhypothese
5. De algemene vorm van een hypothesetoets
6. Hypothesetoetsen als besluitprocedures
7. Welke toetsen in deze cursus?

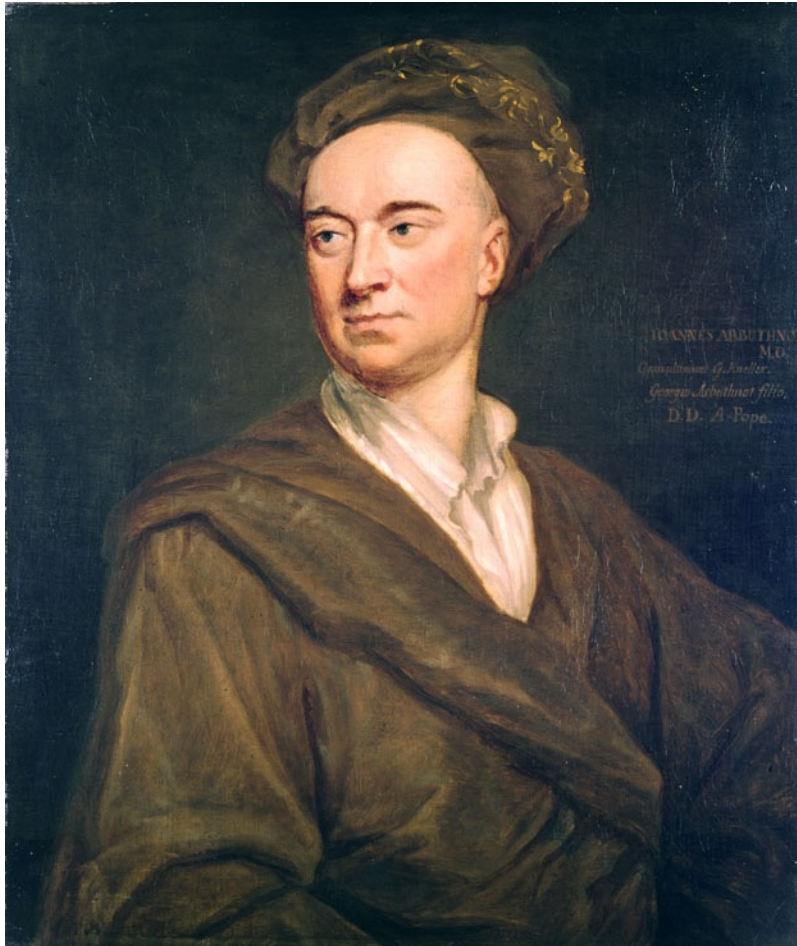
Een hypothesetoets is...

... een statistische procedure waarmee kan worden gekwantificeerd in hoeverre gegevens bewijs leveren tegen een bepaalde hypothese.

Inhoud

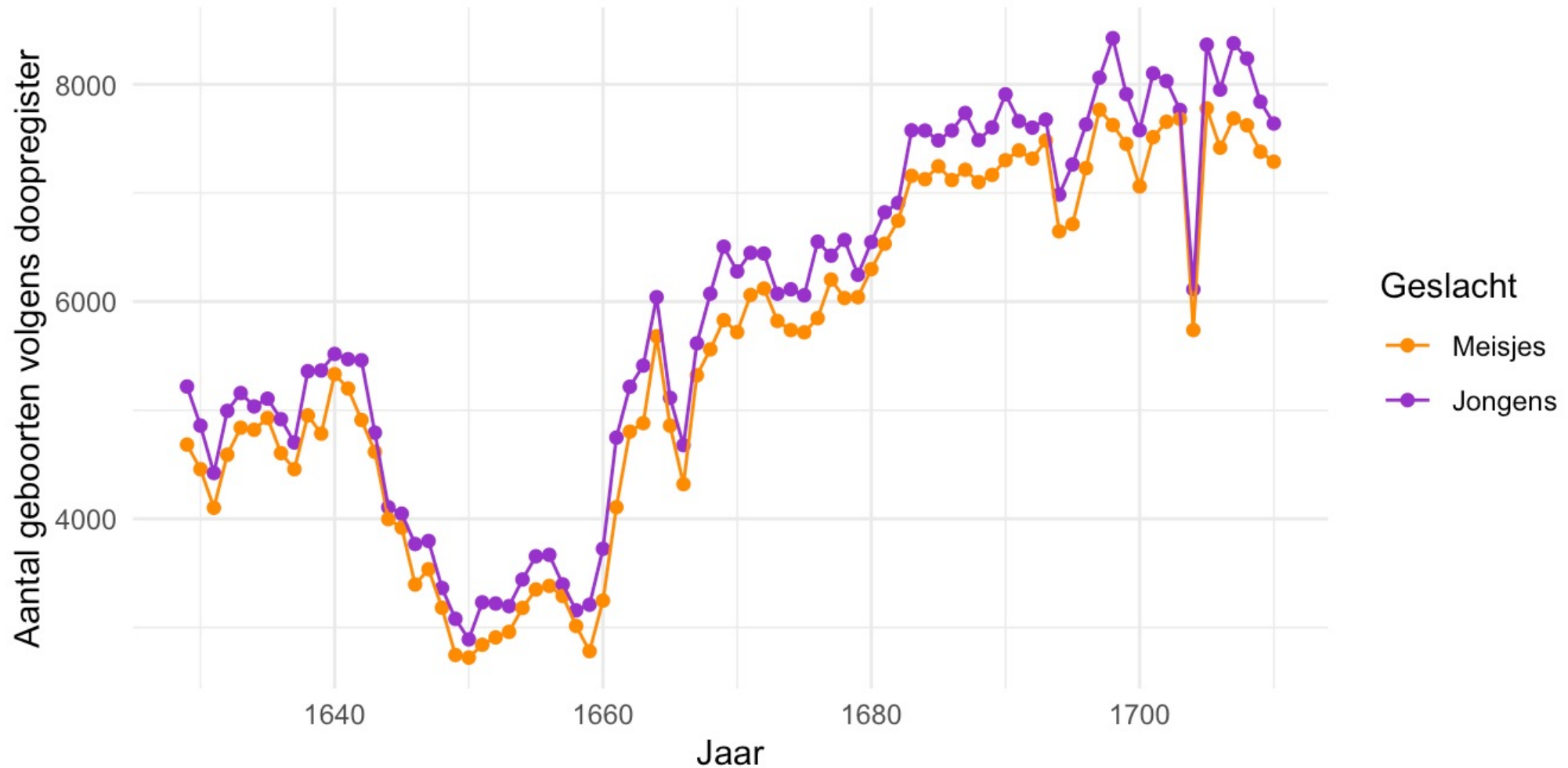
1. Introductie: Discriminatie bij juryselectie
2. Wat is een hypothesetoets?
3. John Arbuthnot en de eerste hypothesetoets
4. De P -waarde als maat voor het bewijs tegen de nulhypothese
5. De algemene vorm van een hypothesetoets
6. Hypothesetoetsen als besluitprocedures
7. Welke toetsen in deze cursus?

De eerste hypothesetoets: John Arbuthnot (1667–1735)

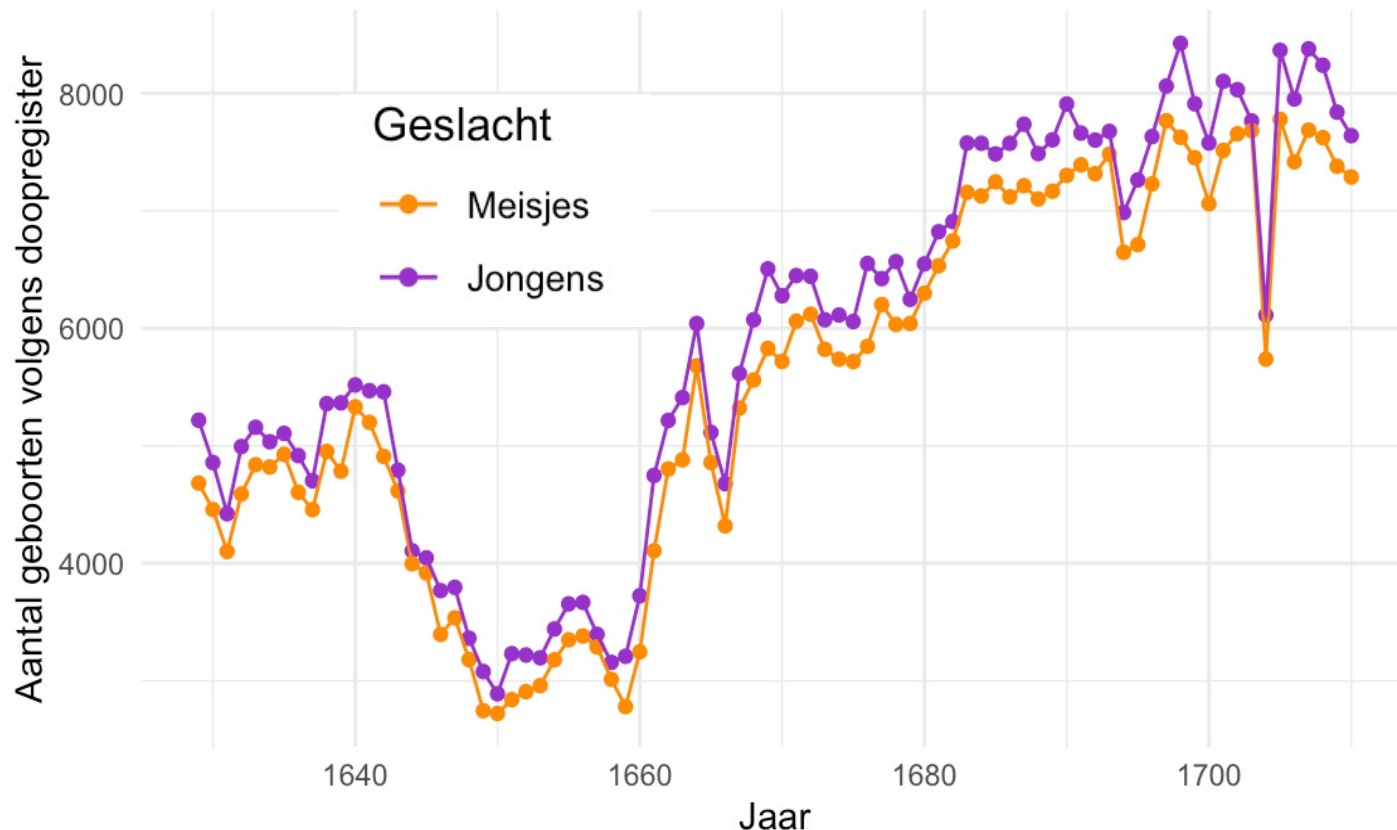


- Schotse arts, wiskundige, satiricus
- lijfarts van koningin Anne van Groot-Brittannië
- een van de oprichters van de Scriblerus Club
 - Jonathan Swift, Alexander Pope, ...
- hypothesetoets!

82 jaar meer jongetjes dan meisjes



Is it Art or Chance that governs?



Twee mogelijke hypotheses:

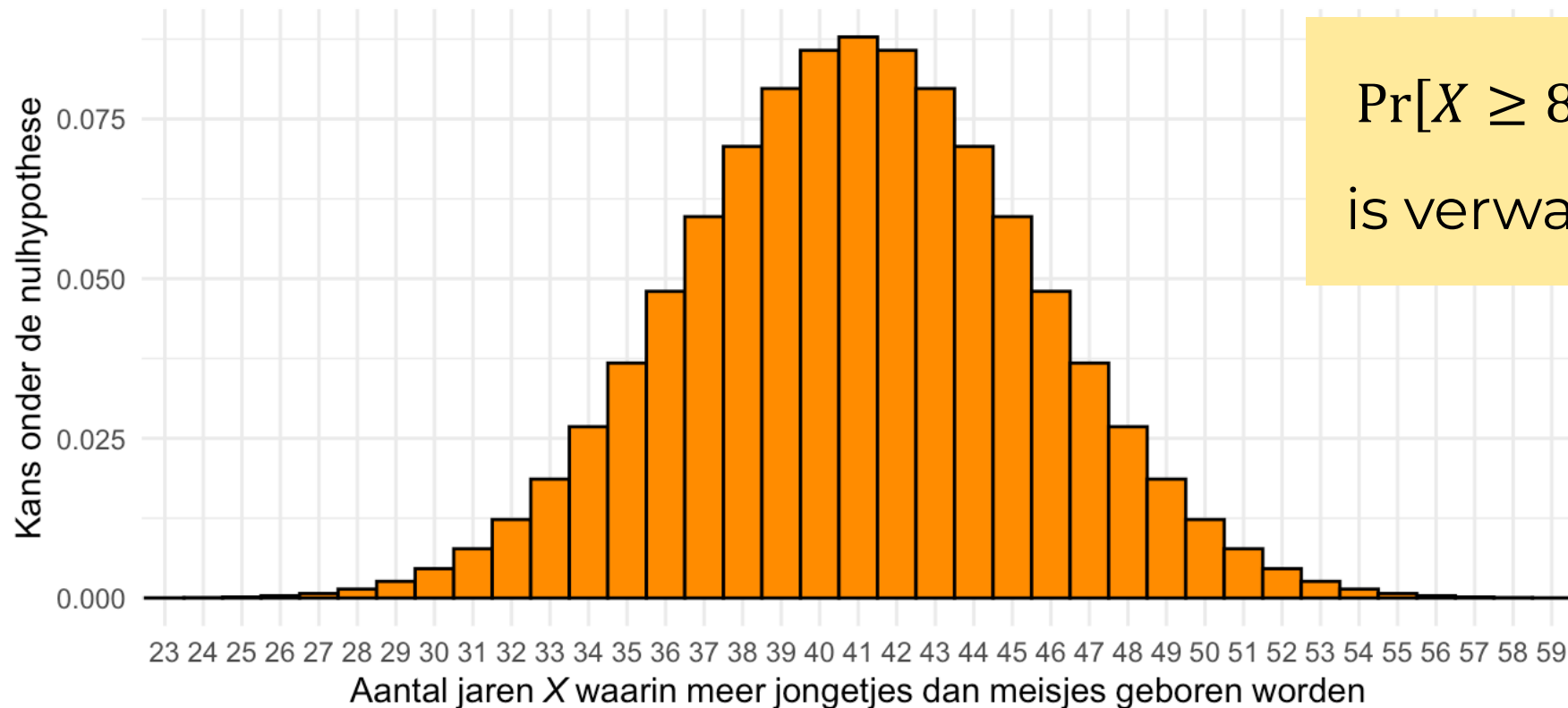
- Toeval. **Nulhypothese H_0**
- Kans op een jongetje daadwerkelijk groter dan de kans op een meisje.
Alternatieve hypothese H_A

Hoeveel toeval nodig?
Bereken kans!

For the sake of argument: een model

Model:

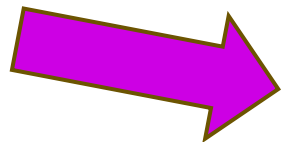
Ieder jaar onafhankelijk een kans van $p_J = \frac{1}{2}$ op meerderheid jongetjes



$$\Pr[X \geq 82] = \left(\frac{1}{2}\right)^{82} = 2 \times 10^{-25}$$

is verwaarloosbaar.

$X = 82$

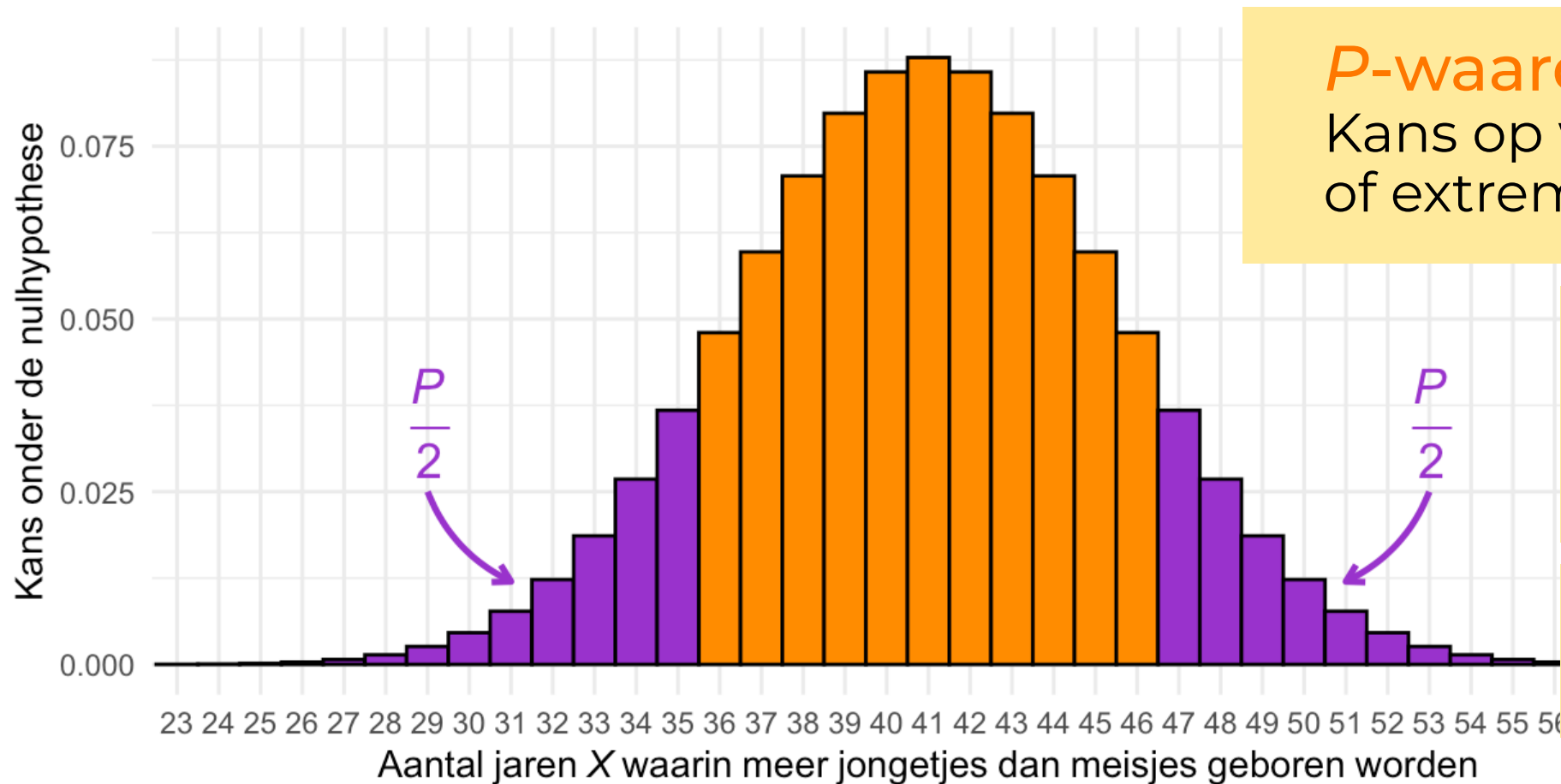


Inhoud

1. Introductie: Discriminatie bij juryselectie
2. Wat is een hypothesetoets?
3. John Arbuthnot en de eerste hypothesetoets
4. De P -waarde als maat voor het bewijs tegen de nulhypothese
5. De algemene vorm van een hypothesetoets
6. Hypothesetoetsen als besluitprocedures
7. Welke toetsen in deze cursus?

De P -waarde (of overschrijdingskans)

Stel dat Arbutnot $X = 47$ had gevonden:



P -waarde:

Kans op waargenomen X of extremer "onder H_0 "

Hoe verrassend is de waarde van X as H_0 waar is?

Kleine P is bewijs tegen H_0 .

“Significant” bewijs tegen de nulhypothese



Ronald Fisher (1890–1962)

$P < 0,05$: “significant” evidence

$P < 0,01$: “highly significant” evidence

Inhoud

1. Introductie: Discriminatie bij juryselectie
2. Wat is een hypothesetoets?
3. John Arbuthnot en de eerste hypothesetoets
4. De P -waarde als maat voor het bewijs tegen de nulhypothese
5. De algemene vorm van een hypothesetoets
6. Hypothesetoetsen als besluitprocedures
7. Welke toetsen in deze cursus?

Hypothesetoetsen in het algemeen

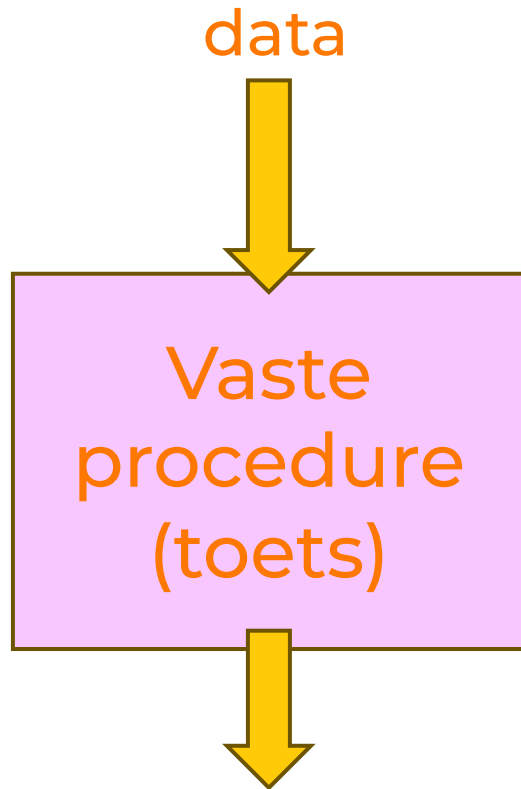
Structuur in 6 stappen

1. Kies een **statistisch model** dat past bij de casus.
2. Stel **hypotheses** H_0 en H_A op in termen van een parameter van het statistische model
3. Kies een **toetsingsgrootheid** en bepaal de **kansverdeling** van die grootheid “onder H_0 ”
4. **Bereken** de toetsingsgrootheid op basis van de waarnemingen.
5. Bereken de **P-waarde**.
6. Trek je **conclusies**: hoe sterk is het bewijs tegen H_0 ?

Inhoud

1. Introductie: Discriminatie bij juryselectie
2. Wat is een hypothesetoets?
3. John Arbuthnot en de eerste hypothesetoets
4. De P -waarde als maat voor het bewijs tegen de nulhypothese
5. De algemene vorm van een hypothesetoets
6. Hypothesetoetsen als besluitprocedures
7. Welke toetsen in deze cursus?

Hypothesetoetsen als besluitprocedures



Besluitprocedure:

een formele, "mechanische" procedure om tot een besluit te komen.

In het geval van hypothesetoetsen:

een formele, "mechanistische" procedure om te besluiten of H_0 moet worden "verworpen".

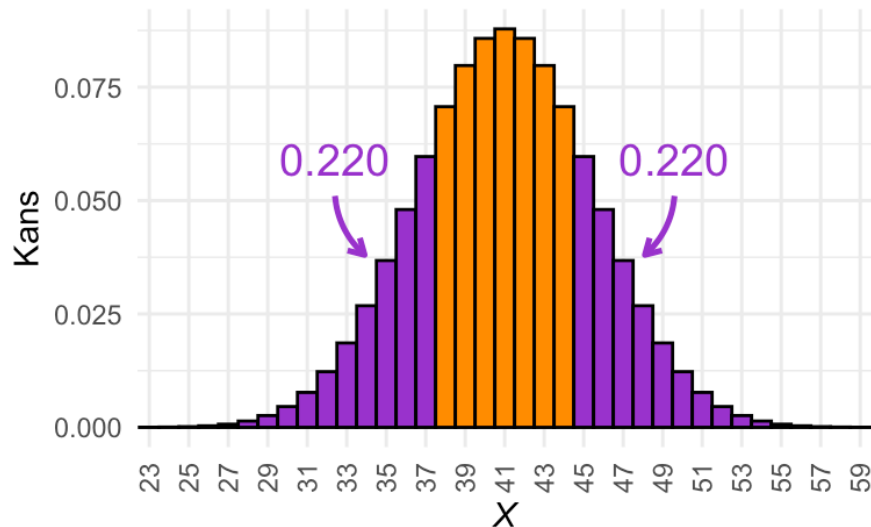
In dat geval wordt H_A "aangenomen".

H_0 verwerpen
 H_A aannemen

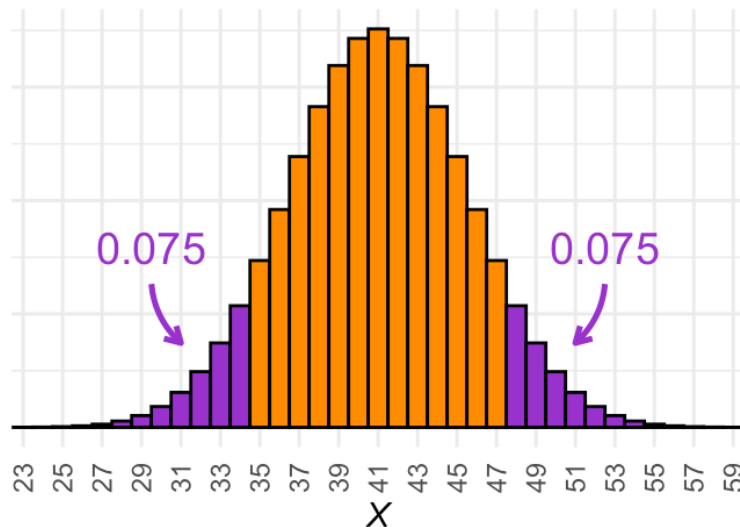
H_0 niet
werwerpen

Idee: verwerp H_0 als $P \leq$ een drempelwaarde α

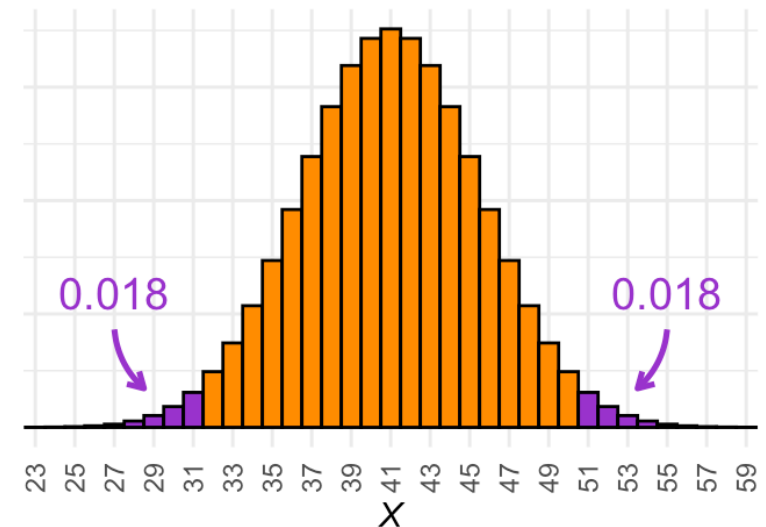
Bijna altijd wordt $\alpha = 0,05$ gekozen.



$P > 0,05$
dus H_0 **niet** verwerpen.



$P > 0,05$
dus H_0 **niet** verwerpen.



$P \leq 0,05$
dus H_0 **wel** verwerpen.

Type-1- en type-2-fouten: een dilemma

	H_0 waar	H_0 onwaar
Je verwierpt H_0 <i>niet</i>	Correcte beslissing (juist-negatief)	type-2-fout (onjuist negatief)
Je verwierpt H_0 <i>wel</i>	type-1-fout (onjuist-positief)	Correcte beslissing (juist-positief)

Kies je α klein, dan is de kans op een type-1-fout klein, maar de kans op een type-2-fout groot.

Kies je α groot, dan is de kans op een type-1-fout groot, maar de kans op een type-2-fout klein.

Nadelen van besluitprocedure

- Geeft de indruk dat het redelijk is om een besluit te nemen op basis van een procedure gebaseerd op één enkele dataset en toets.

Zijn voorkennis, ander onderzoek, coherentie van theorie irrelevant?

- Geeft de indruk dat het redelijk is om in zwart-wit-termen te denken.

Waarom niet denken in termen van *mate van bewijs*?

- Zet aan tot P -hacking.
- Versterkt publication bias.
- $P = 0,05$ is maar matig bewijs;
geen voldoende reden om een hypothesen uit te sluiten.

Inhoud

1. Introductie: Discriminatie bij juryselectie
2. Wat is een hypothesetoets?
3. John Arbuthnot en de eerste hypothesetoets
4. De P -waarde als maat voor het bewijs tegen de nulhypothese
5. De algemene vorm van een hypothesetoets
6. Hypothesetoetsen als besluitprocedures
7. Welke toetsen in deze cursus?

Volgende week:

- Toetsen van het gemiddelde: t-toetsen
- Toetsen van correlatie & regressie