

Toetsen van hypothesen

25 maart 2003

1 Het probleem

De busmaatschappij De Lijn heeft gemiddeld per dag 20000 reizigers in de stad Antwerpen. Tegenwoordig zijn er heel wat reizigers die proberen met de bus mee te rijden zonder te betalen. Deze mensen noemt men **zwartrijders**. De Lijn beweert dat in Antwerpen 10% van de reizigers zwartrijders zijn. Inspecteur Vrijsen beweert echter dat 20% van de reizigers zwartrijders zijn. Nu heeft De Lijn beslist dat als inspecteur Vrijsen gelijk heeft, men een extra controleur zal aanwerven om extra controle uit te oefenen op de bussen. Indien inspecteur Vrijsen ongelijk heeft, zal die extra controleur natuurlijk niet aangenomen worden.

Om dit te controleren heeft De Lijn beslist om een inspecteur op pad te sturen. Vermits het onmogelijk is om alle reizigers te controleren zal die inspecteur 200 reizigers controleren. Deze 200 reizigers worden willekeurig gekozen. De inspecteur kiest een willekeurig bus, kiest een willekeurige rit en controleert een willekeurig aantal reizigers. Dit doet hij de hele dag, tot hij 200 personen gecontroleerd heeft.

Oef 1 *Van de 200 gecontroleerde personen zijn er 36 die zwartrijden. Hoeveel procent van de gecontroleerde personen is zwartrijder?*

.....

Oef 2 *Zou jij op basis hiervan een extra controleur aannemen?*

.....

Oef 3 *Als je 200 andere personen controleert, kan je dan 14 zwartrijders vinden?*

.....

Oef 4 *Hoeveel procent van de reizigers rijdt zwart? Zou je op basis hiervan een extra controleur aannemen?*

.....

We zien dus dat de beslissing of we een extra controleur aannemen of niet afhangt van het **toeval**.



Oef 5 *Duid aan op het interval:*

- 10% van 200 = ...: *dit is het aantal zwartrijders dat we verwachten als we De Lijn geloven.*
- 20% van 200 = ...: *dit is het aantal zwartrijders dat we verwachten als we inspecteur Vrijsen geloven.*

Oef 6

1. *Hoeveel zwartrijders moet de controleur zeker vinden opdat er zo goed als zeker een extra controleur aangenomen wordt?*
2. *Bij hoeveel zwartrijders behouden we zo goed als zeker het huidige aantal controleurs in het controlesysteem?.....*

- A = het gebied waarvan we zeggen: als het aantal zwartrijders hierin ligt, dan nemen we geen extra controleur aan. Dit gebied noemen we het **aanvaardingsgebied**.
- V = het gebied waarvan we zeggen: als het aantal zwartrijders hierin ligt, dan nemen we zeker een extra controleur aan. Dit gebied noemen we het **verwerpingsgebied**.
- T = het gebied waarover we twijfelen.

Oef 7 *Duid deze 3 gebieden aan op het interval.*

2 De beslissingsregel

Nadat we de controle uitgevoerd hebben *moeten* we een beslissing nemen: ofwel nemen we een extra controleur aan ofwel niet. We moeten dus het twijfelgebied T opsplitsen in 2 delen, één deel sluit aan bij A, het ander bij V. Als dit gebeurd is, dan beschikken we over een **éénduidige regel** om te beslissen of we een extra controleur aannemen of niet:

REGEL:

Als het aantal zwartrijders in de steekproef in het aanvaardingsgebied A ligt, dan nemen we geen extra controleur aan.

Als het aantal zwartrijders in de steekproef in het verwerpingsgebied V ligt, dan nemen we een extra controleur aan.

Opmerking: elke waarde tussen 0 en 200 ligt ofwel in A ofwel in V, zodat geen enkele waarde in beide gebieden ligt.

Oef 8 *Als jij baas van De Lijn zou zijn, hoe zou je dan T verdelen over A en V? Waarom?*

.....

Oef 9 *Als jij inspecteur Vrijsen zou zijn, hoe zou je dan T verdelen en waarom?*

.....

Op deze manier zouden we eeuwig kunnen blijven discussiëren over hoe we het twijfelgebied zouden verdelen. Om dit te vermijden neemt men de volgende beslissing:

Algemeen wordt aangenomen dat er slechts 10% zwartrijders op de bussen zitten in Antwerpen. De Lijn heeft dus al min of meer beslist van geen extra controleur aan te nemen. Toch wil ze de mening van inspecteur Vrijsen niet negeren. De Lijn is bereid haar mening te veranderen, maar dan moet de test wel overtuigend zijn!!

Vermits de test uitgevoerd wordt op 200 willekeurige busreizigers, is het mogelijk dat er meer dan 40 zwartrijders gevonden worden (dus meer dan 20%), terwijl er eigenlijk toch slechts 10% zwartrijders zijn. In dat geval wordt er een extra controleur aangenomen terwijl dat niet echt nodig is. Men kan nooit 100% zeker zijn dat deze fout niet gemaakt zal worden.

REGEL:

De Lijn eist van zichzelf dat ze 95% zeker is dat ze die fout niet maakt: als het waar is dat er 10% zwartrijders zijn, dan wil De Lijn er 95% zeker van zijn dat er geen extra controleur aangenomen wordt.

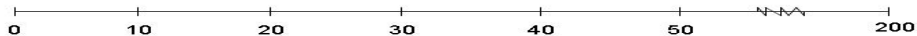
OF

Als het waar is dat 10% van de reizigers zwartrijders zijn, dan moet de kans dat er een extra controleur aangenomen wordt kleiner zijn dan 5%.

Wat betekent het nu om 95% zeker te zijn? Wat betekent het als iemand zegt: "Ik ben daar 99% zeker van" ? De waarde 99 op zich heeft geen enkele betekenis, deze waarde is niet berekend en is ook niet precies. Die persoon had evengoed kunnen zeggen dat hij 95% zeker was.

Als men in statistisch onderzoek voor 95% zeker wil zijn dat men zich niet vergist, dan kunnen we dat berekenen. Hoe we in de wiskunde de kans op een foutieve beslissing berekenen, zullen we later zien.

Stel nu dat De Lijn voor 95% zeker is als het aantal zwartrijders in de steekproef **kleiner is of gelijk aan** 30. Duidt nu op de figuur het aanvaardingsgebied (A) en het verwerpingsgebied (V) aan:



Oef 10 *De controleur vond 32 zwartrijders: ligt dit in het aanvaardings- of in het verwerpingsgebied?*

Oef 11 *Moeten we nu inspecteur Vrijsen geloven of De Lijn? En wat zal De Lijn beslissen, een extra controleur of niet?*

Alhoewel we niet helemaal zeker zijn dat we de juiste beslissing genomen hebben (we zijn voor 95% zeker), hebben we nu wel een sterk vermoeden van het aantal zwartrijders.

3 Verwerking

Vermits het praktisch onmogelijk is om al de busreizigers te controleren doen we het onderzoek op een deel van de busreizigers.

- Alle busreizigers in Antwerpen samen vormen de **populatie**.
- Het deel van de populatie dat gekozen wordt om model te staan voor de volledige populatie noemt men de **steekproef**.
- Als de steekproef effectief model staat voor de populatie, dan noemen we de steekproef **representatief**.
- Om de beslissing te nemen doet men een **test**. Hierin vergelijkt men de algemeen aanvaarde mening met de nieuwe, afwijkende mening.
- De algemeen aanvaarde mening noemen we de **nullhypothese**.
- De nieuwe mening noemen we het **alternatief**.

Oef 12 *Wat is in ons voorbeeld de nullhypothese, wat is het alternatief?*

.....
.....
.....
.....

We moeten een beslissing nemen, daarom splitsen we ons interval op in 2 delen: het aanvaardingsgebied en het verwerpingsgebied. Om op een objectieve manier het interval op te splitsen moeten er afspraken gemaakt worden. De afspraak is dat we *conservatief* zijn: we blijven zolang mogelijk de nullhypothese steunen, behalve als er voldoende bewijs is dat deze NIET juist is. Het grote probleem is: wanneer hebben voldoende bewijs? Aangezien we nooit helemaal zeker zijn dat we een juiste beslissing nemen, moeten we toelaten dat er soms foute beslissingen genomen worden. Er bestaan 2 mogelijkheden om een foute beslissing te nemen:

1. we kunnen beslissen dat de nullhypothese waar is, terwijl in de realiteit het alternatief waar is of,

2. we kunnen beslissen dat het alternatief waar is, terwijl in de realiteit de nulhypothese waar is.

Zie tabel:

	We besluiten dat de nulhypothese waar is	We besluiten dat het alternatief waar is
nulhypothese is waar	we nemen een juiste beslissing	we nemen een foute beslissing (Type-1-fout)
alternatief is waar	we nemen een foute beslissing (Type-2-fout)	we nemen een juiste beslissing

Het is bewezen dat het onmogelijk is om de kans op een type-1-fout en de kans op een type-2-fout tegelijkertijd klein te krijgen. Als we de kans op een type-1-fout klein willen krijgen, dan wordt de kans op een type-2-fout groot en omgekeerd. Aangezien we conservatief zijn, willen we zoveel mogelijk vermijden dat we een type-1-fout maken. We willen dat de kans op een type-1-fout klein is. Hoe klein we die kans willen hangt af van hoe streng we willen zijn. Meestal beslist men dat de kans op een type-1-fout kleiner moet zijn dan 5%.

- Deze 5%, de maat voor hoe streng we zijn, noemen we het **significantiëniveau**.

Oef 13 *Wat is in ons voorbeeld de type-1-fout?*

.....

Door de kans op een type-1-fout kleiner of gelijk aan 0.05 te stellen kunnen we de grens bepalen van het aanvaardingsgebied.

Eens we het aanvaardingsgebied en het verwerpingsgebied vastgelegd hebben, kunnen we gaan kijken of onze steekproef in het aanvaardingsgebied of in het verwerpingsgebied ligt.

Ligt de steekproef in het aanvaardingsgebied, dan aanvaarden we de nulhypothese, ligt de steekproef in het verwerpingsgebied, dan verwerpen we de nulhypothese.

Merk op dat we nooit zeggen dat we het alternatief aanvaarden!

Enkele voorbeelden:

Oef 14 *Bepaal van elk voorbeeld wat we als nulhypothese en wat als alternatief zouden nemen en verwoord ook wat de type-1-fout is.*

- De Amerikaanse krant, "The Washington Post", beweert dat 60% van de leden van de Europese Unie voor een aanval op Irak is. De EU daarentegen beweert dat er slechts 30% van hun leden voor een aanval zijn. Het is niet mogelijk om aan alle inwoners van de EU te vragen of ze voor of tegen een aanval op Irak zijn. Daarom neemt men een steekproef en beslist aan de hand daarvan of de Amerikaanse krant gelijk heeft of niet.

.....
.....
.....
.....

- De schepen van milieu van Kasterlee beweert dat slechts 5% van de bomen in de bossen van Kasterlee een bepaalde ziekte hebben. Volgens de inwoners van Kasterlee heeft 20% van de bomen de ziekte. Indien de inwoners gelijk hebben zal het gemeentebestuur extra geld vrij maken om die ziekte uit te roeien. Indien de bewoners ongelijk hebben zal er verder niets gebeuren aangezien de ziekte niet besmettelijk is voor de andere bomen. Het is niet mogelijk om alle bomen te controleren, daarom neemt men een steekproef en beslist men op basis daarvan of er extra geld moet vrijgemaakt worden of niet.

.....
.....
.....
.....

- De directeur van een suikerfabriek wil nagaan of de pakken suiker van 1 kg nog altijd 1 kg wegen. Als dit niet meer zo zou zijn dan moet hij de machines terug juist instellen. Het is onmogelijk om alle pakken suiker te wegen, daarom neemt hij een steekproef en beslist aan de hand daarvan of hij zijn machines moet aanpassen.

.....
.....
.....
.....

4 Het voorbeeld van De Lijn verder uitgewerkt

We keren nu terug naar het eerste voorbeeld. Daar hebben we 2 opinies: De Lijn zegt dat er 10% zwartrijders zijn. Dit is de *nulhypothese*. Inspecteur Vrijsen zegt dat er 20% zwartrijders zijn (het alternatief) en dat er een extra controleur aangenomen moet worden. Er wordt een controleur op pad gestuurd die een steekproef van 200 zal nemen, op basis hiervan zal een beslissing genomen worden. Om tot een beslissing te komen moeten we het interval $[0, 200]$ in 2 delen opsplitsen: het *aanvaardingsgebied* A en het *verwerpingsgebied* V:



Om tot een objectieve beslissingsregel te komen heeft De Lijn een bijkomende eis gesteld: ze vertrouwen erop dat ze zelf gelijk hebben, maar ze willen toch de mening van inspecteur Vrijsen niet negeren. Het zou echter wel een zware fout zijn om inspecteur Vrijsen te geloven, als er in werkelijkheid 10% zwartrijders zijn (dit hebben we een type-1-fout genoemd). De Lijn wil voor 95% zeker zijn dat ze geen type-1-fout maakt.

Dit wil zeggen dat de kans dat er een extra controleur aangenomen wordt, terwijl er toch 10% zwartrijders zijn, kleiner moet zijn dan 5%. Of meer wiskundig: we eisen dat de kans op een type-1-fout kleiner of gelijk is aan 5%.

Hoe kunnen we dat nu wiskundig beschrijven?

Als er 10% zwartrijders zijn, dan kunnen we zeggen dat de kans dat 1 reiziger een zwartrijder is, gelijk is aan $1/10$. We herhalen dit experiment 200 keer: we controleren 200 busreizigers, de kans dat ze zwartrijden is telkens gelijk aan $1/10$.

We hebben hier dus een experiment met 2 mogelijke uitkomsten: zwartrijder of geen zwartrijder, dit experiment wordt 200 keer herhaald. Het gaat hier dus eigenlijk om een *Binomiaalverdeling*. De kansverdeling vind je in de tabel op de volgende bladzijde. In de eerste kolom staan de mogelijke waarden voor het aantal successen (d.w.z. het aantal zwartrijders), in de tweede kolom staat de kans op dat aantal successen, in de derde kolom staat de kans dat het aantal successen kleiner of gelijk is dan dat aantal.

Oef 15 *We willen nu het interval $[0, 200]$ opsplitsen in twee delen zodat voldaan is aan de eis van De Lijn, namelijk dat de kans op een type-1-fout kleiner is dan 0.05. Wat zoeken we dan eigenlijk? Leg uit in woorden.*

.....
.....
.....

Oef 16 *Wat is de kans dat, als er 10% zwartrijders zijn, we minder dan 30 zwartrijders vinden bij de 200 proefpersonen?(Kijk in de tabel.)*

.....

Oef 17 *Hoeveel zwartrijders mogen er in de steekproef zitten opdat de kans op een type-1-fout kleiner is dan 0.05?*

.....

Oef 18 *Hoe moeten we het interval $[0, 200]$ opsplitsen in twee delen? Wat is dan het aanvaardingsgebied A en het verwerpingsgebied V ?*

Oef 19 *Als het aantal zwartrijders gelijk is aan 36, nemen we dan een extra controleur aan of niet?*

.....

aantal successen	kans	cumulatieve kans	aantal successen	kans	cumulatieve kans
0	7,05508E-10	7,05508E-10	60	2,76411E-15	1
1	1,5678E-08	1,63835E-08	61	7,04873E-16	1
2	1,73328E-07	1,89712E-07	62	1,75587E-16	1
3	1,27108E-06	1,46079E-06	63	4,27354E-17	1
4	6,95561E-06	8,4164E-06	64	1,01645E-17	1
5	3,02955E-05	3,87119E-05	65	2,36303E-18	1
6	0,000109401	0,000148112	66	5,37051E-19	1
7	0,000336884	0,000484997	67	1,19345E-19	1
8	0,000903037	0,001388034	68	2,5936E-20	1
9	0,002140532	0,003528566	69	5,51297E-21	1
10	0,004542684	0,00807125	70	1,14635E-21	1
11	0,008718283	0,016789533	71	2,33216E-22	1
12	0,015256996	0,032046529	72	4,64273E-23	1
13	0,024515514	0,056562043	73	9,04521E-24	1
14	0,036384136	0,092946179	74	1,72484E-24	1
15	0,050129255	0,143075434	75	3,2197E-25	1
16	0,064402167	0,207477601	76	5,88395E-26	1
17	0,077450972	0,284928574	77	1,05283E-26	1
18	0,087490913	0,372419487
19	0,093118984	0,465538471	160	3,0303E-120	1
20	0,093636311	0,559174782	161	8,3653E-122	1
21	0,089177439	0,648352222	162	2,2376E-123	1
22	0,080620008	0,72897223	163	5,7962E-125	1
23	0,069325418	0,798297648	164	1,453E-126	1
24	0,056808329	0,855105977	165	3,5223E-128	1
25	0,044436737	0,899542713	166	8,2518E-130	1
26	0,033232602	0,932775316	167	1,8667E-131	1
27	0,023796184	0,9565715	168	4,0741E-133	1
28	0,016336269	0,97290777	169	8,5714E-135	1
29	0,010765664	0,983673434	170	1,7367E-136	1
30	0,006818254	0,990491688	171	3,3854E-138	1
31	0,004154492	0,99464618	172	6,3421E-140	1
32	0,002437879	0,997084059	173	1,1405E-141	1
33	0,001379002	0,998463061	174	1,9664E-143	1
34	0,000752593	0,999215653	175	3,2461E-145	1
35	0,000396604	0,999612258	176	5,1233E-147	1
36	0,000201974	0,999814232	177	7,7187E-149	1
37	9,94709E-05	0,999913703	178	1,1082E-150	1
38	4,74086E-05	0,999961112	179	1,5133E-152	1
39	2,18809E-05	0,999982993	180	1,9617E-154	1
40	9,78563E-06	0,999992778	181	2,4085E-156	1
41	4,24309E-06	0,999997021	182	2,7938E-158	1
42	1,78479E-06	0,999998806	183	3,0533E-160	1
43	7,28675E-07	0,999999535	184	3,1344E-162	1
44	2,88894E-07	0,999999824	185	3,0121E-164	1
45	1,11278E-07	0,999999935	186	2,699E-166	1
46	4,16619E-08	0,999999977	187	2,2451E-168	1
47	1,51677E-08	0,999999992	188	1,725E-170	1
48	5,37189E-09	0,999999997	189	1,2169E-172	1
49	1,85154E-09	0,999999999	190	7,8282E-175	1
50	6,21293E-10	1	191	4,5539E-177	1
51	2,03037E-10	1	192	2,3718E-179	1
52	6,46421E-11	1	193	1,0924E-181	1
53	2,00567E-11	1	194	4,3795E-184	1
54	6,06653E-12	1	195	1,4973E-186	1
55	1,78932E-12	1	196	4,244E-189	1
56	5,14784E-13	1	197	9,5747E-192	1
57	1,44501E-13	1	198	1,6119E-194	1
58	3,95855E-14	1	199	1,8E-197	1
59	1,05859E-14	1	200	1E-200	1

Oef 20 *Stel dat De Lijn strenger zou zijn en eist dat ze 99% zeker kunnen zijn dat er geen type-1-fout gemaakt wordt. Hoeveel zwartrijders mogen er dan in de steekproef zitten opdat de kans op een type-1-fout kleiner is of gelijk aan 0.01?*

.....

Oef 21 *Als De Lijn voor 99% zeker wil zijn dat ze geen foute beslissing neemt i.p.v. 95%, zijn ze dan meer of minder overtuigd dat hun mening de juiste is?*

.....

Oef 22 *Als het aantal zwartrijders gelijk is aan 18, zou men dan een extra controleur aannemen of niet?*

.....

BESLUIT:

Als we een onderzoek uitvoeren waarvoor we een steekproef moeten nemen, kunnen we nooit helemaal zeker zijn dat we de juiste beslissing nemen. We weten wel dat als er 10% zwartrijders zijn, het zeer vreemd is om in een groep van 200 willekeurig gekozen reizigers meer dan 27 zwartrijders te vinden. De kans om meer dan 27 zwartrijders te vinden is immers kleiner dan 0.05. Als we strenger zijn en we voor 99% zeker willen zijn, dan moeten we ons aanvaardingsgebied groter maken!

Oef 23 *Stel nu dat inspecteur Vrijsen meent dat er 50% zwartrijders zijn, en dat er dus nog meer extra controleurs aangenomen moeten worden. Veronderstel dat men deze bewering wil onderzoeken met 95% zekerheid dat men geen foute beslissing neemt. Wat zal dan het aanvaardings- en verwerpingsgebied zijn?*