

## Antwoorden Onderzoek

- Opg. 1a klassengesprek  
1b klassengesprek  
1c omdat er onderzoek gedaan moet zijn en er resultaten moeten zijn vergeleken.
- Opg. 2a zie verder na deze opgave  
2b zie verder na deze opgave
- Opg. 3 24 van 100 000 = 0,024%  
10 000 van 61 miljoen = 0,016.. maar hier zitten heel veel kinderen bij.  
Beter is het om een grote groep Fransen van dezelfde leeftijdsamenstelling als bij France Telecom met die 24 per 10 000 te vergelijken.
- Opg. 4a zie verder na deze opgave  
4b het is een lichte stijging na een forse daling.
- Opg. 5 Ja, dat lijkt wel. 28 is veel minder dan 71 en ook beduidend minder dan 46.
- Opg. 6 van 5 naar 6
- Opg. 7a Ja, want bij mannen kiest 19% wisk.A/C en bij vrouwen kiest 35% wisk. A/C  
7b Nee, want bij mannen kiest 26% wisk. A/C en bij vrouwen kiest 29% wisk. A/C  
Dat verschilt niet zo veel.

Opg. 8a

	dood	niet dood
buikligging	127	199
geen buikligging	23	127
	150	326

- 8b bij dood had 85% buikligging, bij de niet dood had 61% buikligging  
8c ... groter dan wanneer er geen buikligging is.
- Opg. 9a ?, het percentage incidenten tijdens diensten van Lucia is wel erg hoog vergeleken met andere diensten.  
9b  $201 \times 27/1704 \approx 3$   
9c De verwachting is 3, bij Lucia 14 dus.... zie verder na deze opgave
- Opg. 10a  $51,26 / 48,74 = 1,05170.. \approx 1,0517$   
10b oddsN / oddsB = 1,0024  
dus oddsB = oddsN / 1,0024 =  $1,0517 / 1,0024 = 1,04918.. \approx 1,0492$   
 $J / M = J / (100 - J) = 1,0492$  dus  $J = (100 - J) \times 1,0492$  dit geeft  $J = 51,20\%$   
10c in China zijn dus 62,5% jongens en 37,5% meisjes.  
oddsC =  $62,5 / 37,5 = 1,66666.. \approx 1,6667$   
Oddsratio China-Nederland =  $1,6667 / 1,0517 = 1,58476.. \approx 1,5848$
- Opg. 11a dus  $W / V = 7$  dus  $W = 7V$  dus  $100 - V = 7V$ , dus  $8V = 100$  dus  $V = 12,5$  en  $W = 87,5$   
Kans op winnen is bij Ajax 87,5%  
11b  $600 / 400 = 1,5$   $500 / 400 = 1,25$  oddsratio =  $1,5 / 1,25 = 1,2$   
11c odds = winst / verlies oddsratio =  $(0,4 / 0,6) / (0,3 / 0,7) \approx 1,56$
- Opg. 12a ja, bij jongens rookt 64% en bij meisjes 52%  
12b  $\frac{a}{a+b} = \frac{c}{c+d}$  dus  $ac + ad = ac + bc$  dus  $ad = bc$  (kruisproduct)  
of  $\frac{a}{b} = \frac{c}{d}$  dus  $ad = bc$ .  
12c bij jongens  $\frac{a}{b}$  en bij meisjes  $\frac{c}{d}$   
12d  $\frac{a}{c}$  vermenigvuldig de teller en de noemer met  $bd$  en je krijgt het antwoord  
12e dan geldt  $ad = bc$  en de deling geeft dus 1

- Opg. 13a  $(30 \times 56) / (13 \times 55) \approx 2,35$   
 13b  $(14 \times 1490) / (187 \times 13) \approx 8,58$
- Opg. 14  $(127 \times 127) / (23 \times 199) \approx 3,52$
- Opg. 15a  $(37 \times 267) / (351 \times 14) \approx 2,01$   
 15b (enige) kaalheid bij  $37 / 388 \times 100\% \approx 9,5\%$  volle haardos bij  $267 / 281 \times 100\% \approx 95\%$   
 15c gezonde mannen (enige) kaalheid bij  $\approx 5\%$  en bij patiënten  $\approx 9,5\%$  en dat is bijna 2 maal zo veel. zie ook verder na deze opgave  
 15d  $(122 \times 200) / (266 \times 81) \approx 1,13$
- Opg. 16a  $(40 - 35) / 4 = 1,25$   
 16b  $\text{InvNormCD}(0.05, \mu=40, \sigma=4) \approx 33,4$  en  $\text{InvNormCD}(0.95, \mu=40, \sigma=4) \approx 46,6$   
 Dus tussen 33,4 en 46,6  
 16c Nee, 35 kan door toeval best wel voorkomen als  $\mu=40$
- Opg. 17a  $(2400 - 1850) / 150 \approx 3,67$   
 17b 550 meter meer, dat is 3,67 keer de sd.  
 17c ja, 3,67 sd's is uitzonderlijk veel
- Opg. 18 gemiddelde(j) = 6,4 sd(j)  $\approx 1,69$  gemiddelde(m)  $\approx 6,89$  sd(m)  $\approx 1,33$   
 Effectgrootte =  $\frac{6,89-6,4}{(1,69+1,33)/2} \approx 0,32$
- Opg. 19 gemiddelde(U) = 5,8 sd(U)  $\approx 1,25$  gemiddelde(J) = 7 sd(J)  $\approx 1,61$   
 Effectgrootte =  $\frac{7-5,8}{(1,61+1,25)/2} \approx 0,84$
- Opg. 20abc - - -
- Opg. 21a  $\text{BinCD}(X = 28, n = 100, p = 0,2) - \text{BinCD}(X = 11, n = 100, p = 0,2) \approx 0,9674$   
 $\text{BinCD}(X = 27, n = 100, p = 0,2) - \text{BinCD}(X = 12, n = 100, p = 0,2) \approx 0,9405 < 95\%$   
 21b 25 ligt in het 95% gebied, dus is dit resultaat niet significant  
 21c 65 t/m 95 geeft 0,947.. en 64 t/m 96 geeft 0,9611... , dus 64 t/m 96  
 21d 25% van 400 is 100 en dit ligt niet in het 95% gebied, dus is dit resultaat significant
- Opg. 22a 7 t/m 13 geeft 0,88.. en 6 t/m 14 geeft 0,958... , dus *tussen* 5 t/m 15  
 22b dit is significant, dus is de munt waarschijnlijk vals
- Opg. 23a  $\text{InvNormCD}(0.025, \mu=9,8, \sigma=1) \approx 7,84$  en  $\text{InvNormCD}(0.975, \mu=9,8, \sigma=1) \approx 11,76$  ,  
 dus tussen 7,84 en 11,76  
 23b nee, het resultaat ligt in bovenstaand gebied  
 23c 0,05 is de kans op TWEE staartgebieden samen,  
 dus middengebied wordt 95%, net als in opg.a
- Opg. 24a waarschijnlijk weinig jongeren, die slapen vaker uit  
 24b alleen lezers van die krant, een erg geselecteerde groep dus  
 24c alleen mensen met een telefoonnummer in deze boeken worden gebeld,  
 dus ook nu vallen jongeren met vaak alleen een mobiel buiten de boot.  
 Als er overdag gebeld wordt, krijg je de werkende mensen nooit aan de lijn
- Opg. 25 minimaal 10, maximaal heel veel, als je steeds al eerder gegooide getallen opnieuw gooit
- Opg. 26 29, 31, 09, 15, 12 en 23
- Opg. 27a 29, 31, 09, 15, 48 en (weer 09, die vervalt dus) 12  
 27b er zijn minder getallen nodig
- Opg. 28a  $21000 / (21000 + 117000) \times 150 \approx 23$  mannen en  
 $117000 / (21000 + 117000) \times 150 \approx 127$  vrouwen

28b leeftijd, les in groep 1, 2, .. of 8

- Opg. 29a dan kan de arts, door zijn gedrag, geen invloed hebben op de patiënt  
29b met medicijn geneest 61%, zonder geneest 67%, percentages liggen dicht bij elkaar of oddsratio  $\approx 1,26$  en dat is niet veel hoger dan 1  
29c tweede tabel is 61% tegen 12%, of oddsratio  $\approx 11,8$   
Geef ook eigen criterium .....

29d  $48 / 99 \times 59 = 28,6.. \approx 29$

29e afleestip: als het laatste getal 31 bij nr 937 staat komt 32 of meer 63 keer voor kans is in bovenstaand voorbeeld 63 van de 1000 en dat is 6,3%

29f de kans op 32 is  $\frac{\binom{32}{99} \binom{25}{57}}{\binom{127}{156}}$  of  $\frac{48}{99} \times \frac{47}{98} \dots \frac{17}{68} \times \frac{51}{67} \times \dots \frac{27}{43} \times \binom{57}{32}$

Zo ook de kansen op 33 t/m 48 en optellen  
Of op de casio GR 1 – kans (0 t/m 31) bij hypergeo

Opg. 30a

	polio
Salk	56
placebo	142
totaal	198

Neem populatieomvang 400.000, opgesplitst in twee groepen (polio en niet-polio), beide van omvang 200.000. Kies als proportie "blauwe" in VuStat dus 0,50. Neem daaruit een steekproef van omvang 198. Kijk hoeveel er polio ("blauw") zijn.

Als Salk niet helpt, verwacht je dat er ongeveer 99 polio bij zullen zijn. Als dat er beduidend minder zijn, concludeer je dat Salk helpt.

- 30b Ik heb de simulatie 40 keer uitgevoerd. Geen enkele keer was het aantal polio minder dan 88.
- 30c. Het gemeten aantal poliogeveallen onder de kinderen die Salk hadden gekregen lijkt dus uitzonderlijk laag te zijn. De conclusie dat Salk werkt is gerechtvaardigd.
- Opg. 31 18 en 9 graden Celsius
- Opg. 32 Bij 100 keer zo'n waarschuwing verwacht men dat het 65 keer ook echt gebeuren gaat
- Opg. 33  $5 \times 2 \times 40 = 400$  en 7% hiervan is 28. Hij verwacht dus 28 keer een grotere vertraging
- Opg. 34a  $1 / 17 \times 100\% \approx 6\%$   
34b tussen 5% en 15% lijkt het beste omdat de 6% uit vraag a daar tussen zit  
34c  $\text{BinCD}(X=1, n=17, p=0,15) \approx 0,2525$   
34d klassengesprek
- Opg. 35a  $400 \times 0,25 = 100$   
35b zoek g waarbij de kans op g of minder meer is dan 0,95  
 $Y = \text{BinCD}(X=X, n=400, p=0,25)$  met een tabel  
X = 113 geeft 0,938.. X = 114 geeft 0,951.. dus g = 114
- Opg. 36a  $1 - \text{BinCD}(X=5, n=16, p=0,2) \approx 0,082$   
36b  $Y = 1 - \text{BinCD}(X=X-11, n=X, p=0,2)$  tabel en zoek kleiner of gelijk aan 5%  
Je kunt ook letten op GEEN blessure, dan neem je  $Y = \text{BinCD}(X=10, n=X, p=0,8)$   
X = 16 geeft 0,081.. X = 17 geeft 0,037.. Neem een selectie van 17 of meer personen
- Opg. 37a  $1 - \text{BinCD}(X=1, n=20, p=0,06) \approx 0,340$   
37b tabel geeft afkeuren met kans 0,06 dus goedkeuren heeft kans 0,94  
tabel geeft afkeuren met kans 0,61 dus goedkeuren heeft kans 0,39  
37c nee, de kans dat een slechte partij (10% foute exemplaren) goedgekeurd wordt is wel 39%

- Opg. 38a 66 van de 200, dus 33%
- 38b VuStat/Simulaties/Steekproeven percentage blauw 0,33 / omvangpopulatie 3456  
omvangsteekproef 200/aantal steekproeven 100 (gebruik na afloop sorteren)
- 38c dus van 28% t/m 38%, dus tel hoe vaak de getallen 56 t/m 76 voorkomen
- Opg. 39a kies resultaten, kies foutenmarge zo klein mogelijk, maar zo dat "steekproeven binnenmarge"  $\leq 95\%$   
Als dit bijvoorbeeld 6% is, dan loopt het percentage van 33 - 6 t/m 33 + 6 en de aantallen van 54 t/m 78
- 39b kies foutenmarge enz. "steekproeven binnenmarge"  $\leq 90\%$
- 39c Bereken  $x$  zodat  $P(66 - x \leq X \leq 66 + x, n = 200, p = 0.33) \geq 0,95$   
Tabel met  $Y = \text{BinCD}(66 + x, n = 200, p = 0.33) - \text{BinCD}(66 - x - 1, n = 200, p = 0.33)$   
Dit moet  $\geq 0,95$  zijn.  $x = 12$  geeft 0,94... en  $x = 13$  geeft 0,957.. dus  $x = 13$   
Het interval loopt dus van 66 - 13 t/m 66 + 13, dus 53 t/m 79
- 39d Zoek nu  $Y \geq 0,90$   $x = 10$  geeft 0,88... en  $x = 11$  geeft 0,91.. dus  $x = 11$   
Het interval loopt dus van 66 - 11 t/m 66 + 11 dus 55 t/m 77
- Opg. 40a groter
- 40b nauwkeuriger
- Opg. 41 Kijk wat de hele klas heeft, misschien is het gemiddelde een 7  
Als een leerling een 4 een 5,5 en een 6,5 heeft gehaald staat hij onvoldoende.  
Maar weer een 6,5 is niet opzienbarend
- Opg. 42a Onderzoek doen bij mensen van 65 jaar en ouder die niet ingeënt zijn. Misschien is 3 sterfgevallen wel normaal.
- 42b mensen waarbij het niet hielp zijn allang geen klant meer.  
Er wordt niets gemeten, mensen geven aan wat ze voelen.  
Na iets gezonds eten voelen veel mensen zich gezonder.
- 42c Het kan zijn dat ontevreden vrouwen eerder aan zo'n onderzoek meedoen.  
De vrouwen hebben zichzelf geselecteerd.
- 42d vakantie in eigen land regelen veel mensen niet via een reisbureau.  
Gebruik een andere manier om aan een steekproef te komen
- Opg. 43a  $1 - \text{BinCD}(X = 11, n = 14, p = 0,5) \approx 0,00646..$
- 43b vast niet
- Opg. 44a Bij biologie leer je hoe kinderen verwekt worden, dat heeft niets met ooievaars en rode kool te maken
- 44b nee
- 44c bij veel oudere mensen gaat het geheugen achteruit.
- 44d ook mogelijk is: hogere studiekosten, hogere kamerhuur, maar doordat er sneller gestudeerd moet worden heb je minder tijd voor bijbaantjes
- 44e een koe krijgt een naam als er meer aandacht is voor de individuele koe, en die aandacht kan een hogere melkproductie geven
- 44f in de vakantie en met mooi weer wordt er altijd meer gefietst
- 44g zie verder na deze opgave
- 44h
- |            | bovenste helft | onderste helft | totaal |
|------------|----------------|----------------|--------|
| korte naam | 7              | 4              | 11     |
| lange naam | 2              | 5              | 7      |
| totaal     | 9              | 9              | 18     |
- 44i  $35 / 8 = 4,375$
- 44j de oddsratio is hoog, maar er geen logisch verband te bedenken, dus zal het toeval zijn
- Opg. 45a  $\left(\frac{1}{4000}\right)^2 \approx 6,25 \times 10^{-8}$
- 45b  $\left(\frac{1}{4000}\right)^3 \approx 1,56 \times 10^{-11}$
- 45c er kan een onbekende erfelijke oorzaak zijn,

misschien werd bij alle kinderen buikligging toegepast.

45d  $100 \times 100 \times \text{---} = 2,5$

Opg. 46a  $\text{---} \approx 0,00000019$

46b de kans is heel klein, maar het kan wel.

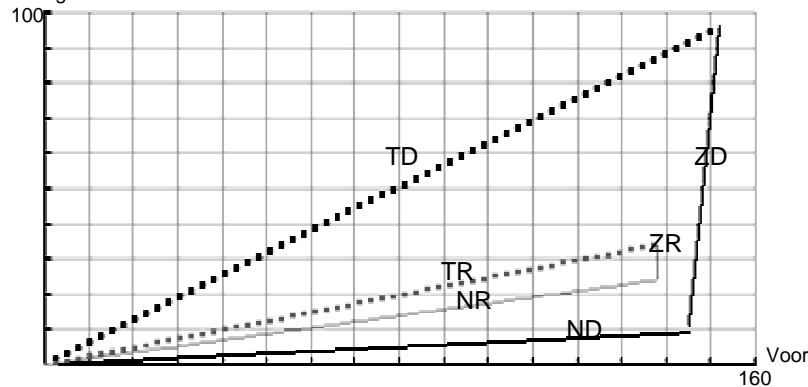
Opg. 47 zie verder na deze opgave

Opg. 48ab

aantal vóórstemmers	democraten	republikeinen	totaal
noordelijk	94%	85%	
zuidelijk	7%	0%	
totaal	61%	80%	

94 is meer dan 85 en 7 is meer dan 0, maar 61 is minder dan 80 !!!

Ter verduidelijking onderstaand plaatje, zwart zijn de Democraten, grijs de Republikeinen tegen



ND loopt van (0, 0) naar (145, 9) omdat er bij de Noordelijke Democraten 145 voor- en 9 tegenstemmers zijn, erachter is een lijn ZD geplakt die 7 naar rechts en 87 omhoog gaat omdat er bij de Zuidelijke Democraten 7 voor- en 87 tegenstemmers zijn.

De stippellijn TD geeft het totaal aantal aan, 152 voor- en 96 tegenstemmers

Net zo lopen de drie lijnen voor de Republikeinen

Als de helling van een lijn kleiner is, zijn er procentueel meer voorstemmers.

De hellingen van de N- en Z-lijnen zijn bij de Democraten kleiner dan bij de Republikeinen, maar de helling van de T-lijn is bij de Republikeinen kleiner.

Opg. 49a mannen met medicijn 60% geneest mannen met placebo 50% geneest  
vrouwen met medicijn 87,5% geneest vrouwen met placebo 80% geneest  
dus medicijn werkt bij mannen en vrouwen (fijn voor de fabrikant)

49b mensen met medicijn 72% geneest mensen met placebo 74% geneest  
dus medicijn werkt niet (jammer voor de fabrikant)

Opg. 50 zie verder na deze opgave

Opg. 51 er zijn ook hoogbegaafde kinderen met niet-hoogbegaafde ouders

Opg. 52 er wordt een effect geconstateerd, dat ergens aan toegeschreven kan worden, terwijl het effect er *altijd* is.

Opg. 53a de terugval naar het midden geeft aan dat van veel bedrijven, die erg goede jaarcijfers kunnen overleggen, het gemiddelde volgend jaar lager ligt. Dus te hooggespannen verwachtingen van de aandeelhouders.

53b de terugval naar het midden geeft aan dat van provincies waar 4 keer in een jaar de jackpot valt, het volgende jaar het gemiddelde dichter bij 10/12 komt te liggen. Verhuizen heeft niet zo veel zin.

- 53c de terugval naar het midden geeft aan dat van winkels met een uitzonderlijk hoog aantal diefstallen, het gemiddelde aantal de volgende week lager ligt, ook als er geen opziener in uniform loopt
- 53d de terugval naar het midden geeft aan dat van alle patiënten die zich erg beroerd voelen, de gemiddelde patiënt zich een dag later beter voelt.

Opg. 54ab -

54c  $195 \times 2 / 3 = 130$  erbij moet dus meer dan 65 worden opgeteld.

De kans is  $\text{NormCD}(65, 10^{99}, \mu=60, \sigma=8) = 0,2659.. \approx 0,266$

54d  $160 \times 2 / 3 = 106\frac{2}{3}$  erbij moet dus minder dan  $53\frac{1}{3}$  worden opgeteld.

De kans is  $\text{NormCD}(-10^{99}, 53\frac{1}{3}, \mu=60, \sigma=8) = 0,2023.. \approx 0,202$