

GRAFISCHE TOETSEN VOOR HET VERGELIJKEN VAN TWEE PERCENTAGES

J.M. Buhrman \*)

*SAMENVATTING Dit artikel beschrijft methoden om grafisch een toets uit te voeren op de gelijkheid van twee percentages. Zowel voor de situatie dat de percentages betrekking hebben op één steekproef, als voor het geval dat percentages uit verschillende steekproeven moeten worden vergeleken, is een nomogram beschikbaar. De gegeven methoden zijn vooral bedoeld voor onderzoekers uit andere gebieden dan de statistiek.*

Zeer veel mensen hebben te maken met gegevens die uit steekproeven zijn verkregen. Vaak hebben zij slechts een geringe kennis van wiskunde en statistiek. Maar ook zij willen antwoord op vragen als:

- Hoe betrouwbaar is een percentage van 18% in een steekproef van 500 respondenten?
- Als ik in een steekproef van 300 snurkers 23% rugslapers en 17% buikslapers vind, is dit verschil dan significant?

Zij weten dat gegevens uit steekproeven een beperkte nauwkeurigheid hebben, en dat verschillen in de steekproef m.b.v. een toets moeten worden onderzocht. Daarbij maken ze zich niet druk over de exacte betekenis van het begrip significant, maar hebben daarvan intuïtief wel een idee. Met wat formules en een zakrekenmachine kunnen de noodzakelijke toetsen wel worden uitgevoerd, maar aan wie hiermee niet voldoende vertrouwd is, kunnen goed toegankelijke tabellen en overzichtelijke grafieken meer uitkomst bieden.

Twee problemen zijn hier bekeken: het vergelijken van twee percentages behorende bij gescheiden categorieën in één steekproef (wat neerkomt op de tekentoets) en het vergelijken van twee percentages in twee steekproeven (de toets van Fisher). Van beide geven we een voorbeeld. Het gebruik van de figuren wordt aan de hand van deze voorbeelden uitgelegd.

---

\*) Inter/View B.V., Zocherstraat 30-36, 1054 LZ Amsterdam

## Twee percentages in één steekproef

In een aselechte steekproef van 500 personen in een zeker dorp bleek 19% werkzaam in de agrarische sector en 26% in de industrie. De vraag of in dit dorp ècht meer mensen in de industrie werkzaam zijn dan in de landbouw kan met figuur 1 als volgt bevestigend worden beantwoord. Men trekt een lijn door het punt 7 (= 26-19) op de bovenste schaal en het punt 45 (= 26+19) op de onderste schaal. Deze lijn snijdt de middelste schaal ongeveer bij 370. Ligt de grootte van de gebruikte steekproef op de middelste schaal rechts van dit snijpunt, dan (en alleen dan) is het verschil significant (op 95%-niveau), wat hier dus het geval blijkt te zijn.

## Twee percentages in verschillende steekproeven

In de dorpen A en B is aan 300 resp. 350 aselekt gekozen personen naar de religie gevraagd. In dorp A bleek 31% ongelovig, in dorp B 39%. Heeft het atheïsme in dorp B nu ècht harder toegeslagen dan in dorp A? Het antwoord is ja, wat blijkt door de toets van figuur 2 toe te passen. Men trekt een lijn door het punt 8 (= 39-31) op de bovenste schaal en het punt 70 (= 39+31) op de onderste schaal. Deze lijn snijdt de middelste schaal (ongeveer) bij 275. Omdat beide steekproeven groter waren dan 275, is het verschil significant (op 95%-niveau). Als ze beide links liggen van het snijpunt op de middelste lijn, dan is het verschil niet significant, en als de ène steekproef links van het snijpunt ligt en de andere rechts, dan kan met op deze manier geen uitspraak doen. Deze laatste situatie kan zich niet voordoen als de steekproeven even groot zijn, en zal vaker voorkomen naarmate de steekproeven meer in omvang verschillen.

De statisticus kan in een enkel geval ook wel gemak hebben van deze methoden: als de uitkomsten niet op de grens liggen tussen wel en niet significant, en tevens de nomogrammen binnen hand-

bereik zijn.

Concluderend kunnen we stellen dat deze nomogrammen vaak, maar niet altijd, behulpzaam kunnen zijn bij het toetsen van de gelijkheid van twee percentages.

Deze en enkele andere eenvoudige methoden staan beschreven in het rapport "Statistische technieken in het marktonderzoek" (Inter/View, 1981).

#### Meer nomogrammen

In oude jaargangen van *Statistica Neerlandica* staan o.a. de volgende interessante nomogrammen.

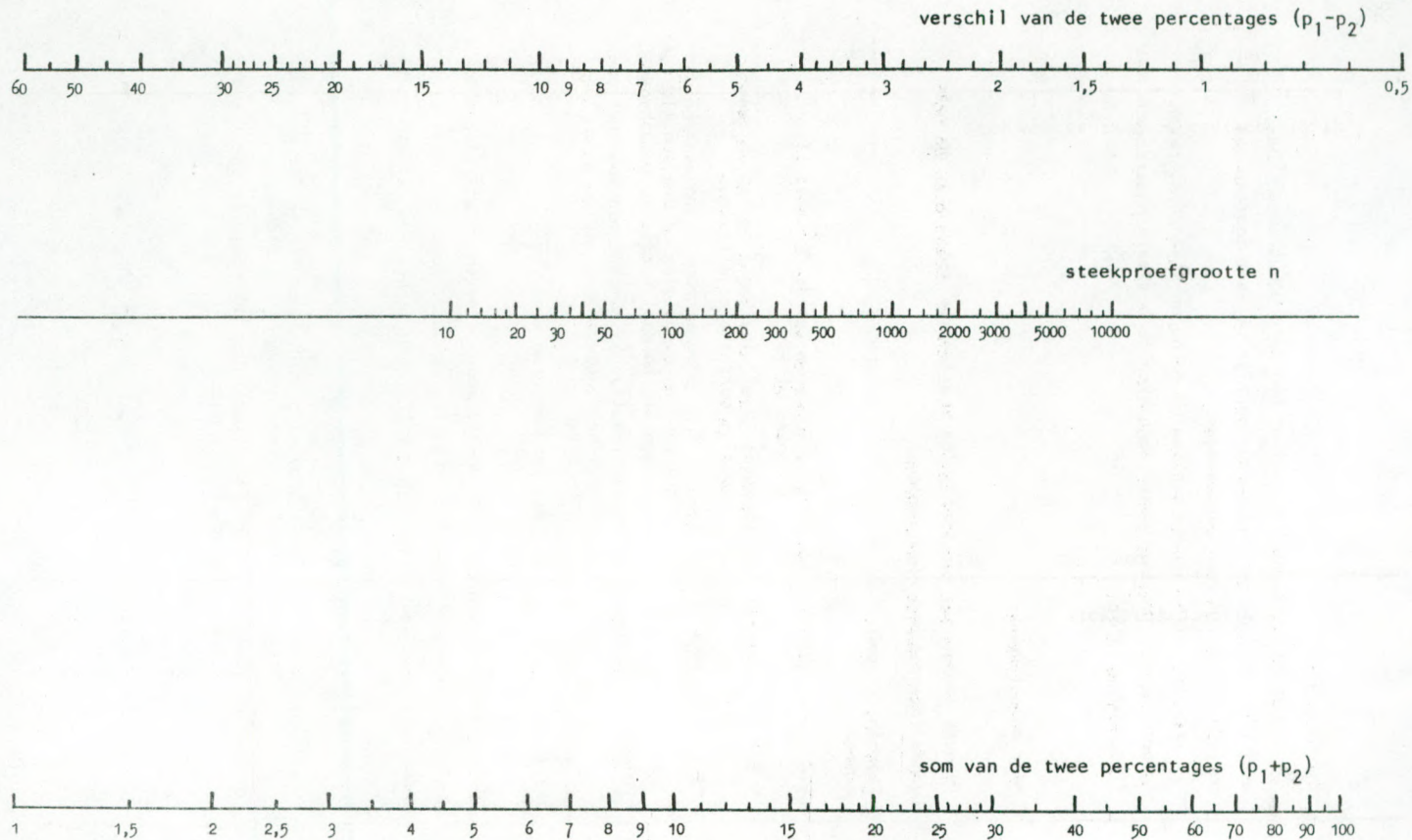
jaargang. nummer	pag.	
1.2	107-111	J. Sittig, Nomogram van de $\chi^2$ -test of goodness of fit
5.3/4	145-147	V. Varangot, Grafische bepaling van de tweevoudige correlatiecoëfficiënt
13.1	3-14	C. Levert, Een nomogram voor betrouwbaarheids-grenzen van quantielen en overschrijdings-kansen bij normaal verdeelde grootheden
14.2	127-150	M. Keuls, Tabellen en nomogrammen voor het onderscheidingsvermogen van de 5% en 10%-F-toets
14.2	155	L. Tas, Een nomogram voor $\sqrt{pq/n}$

Verder heeft de VvS ooit een boekje doen uitgeven: *Statistische Tabellen en nomogrammen*, onder redactie van M.J. Hogendijk, dat niet meer leverbaar is. In sommige bibliotheken is het nog wel aanwezig.

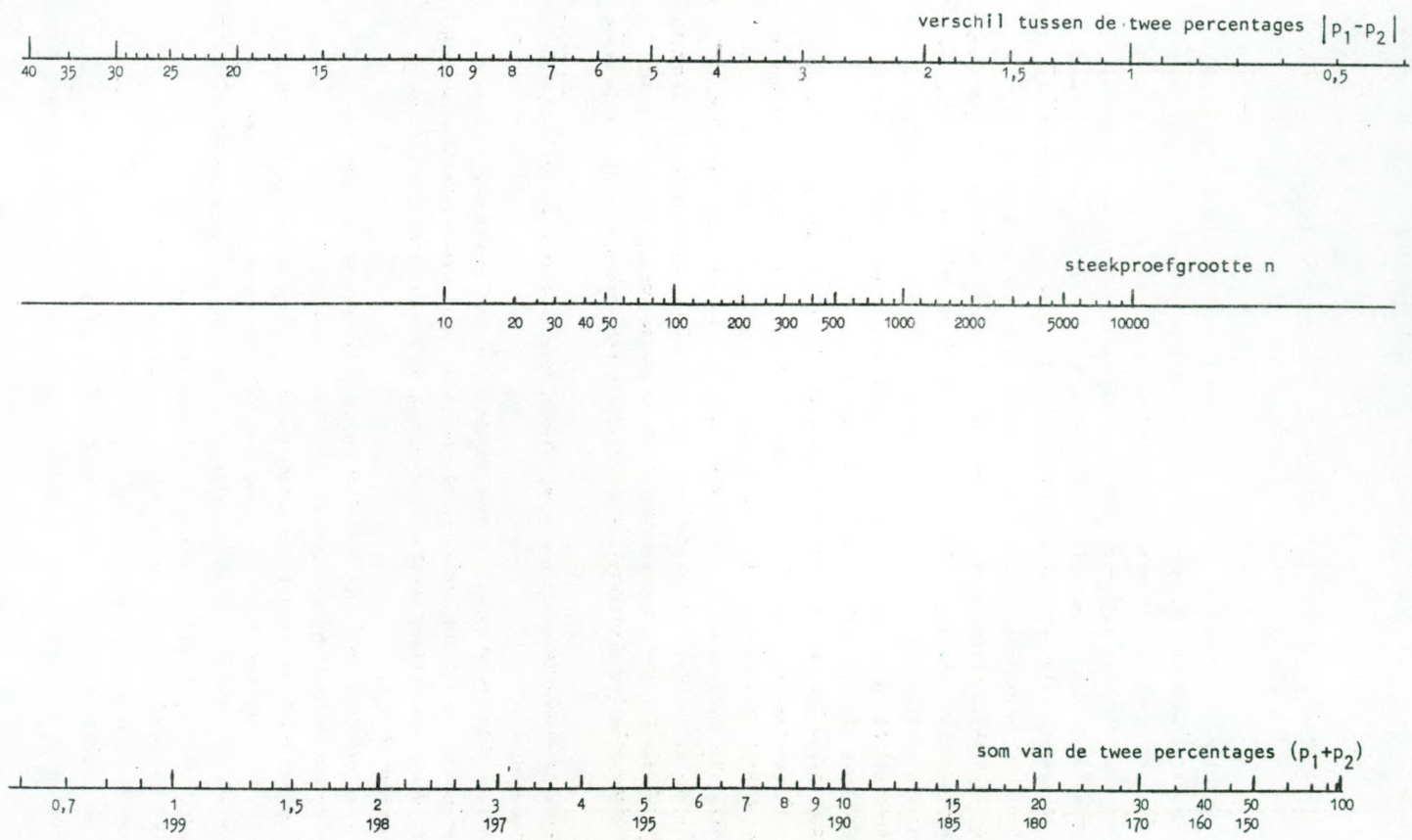
In *Biometrika Tables* vindt men

in deel 1, table 41, Chart providing confidence limits for  $p$  in binomial sampling, given a sample fraction  $c/n$

in deel 2, table 30, Chart for determining the power of the  $t$  and  $F$  tests: fixed effects model



FIGUUR 1 Nomogram voor uitvoering van de tekentoets met significantie-niveau 95%



FIGUUR 2 Nomogram voor het vergelijken van twee percentages uit twee groepen of steekproeven