

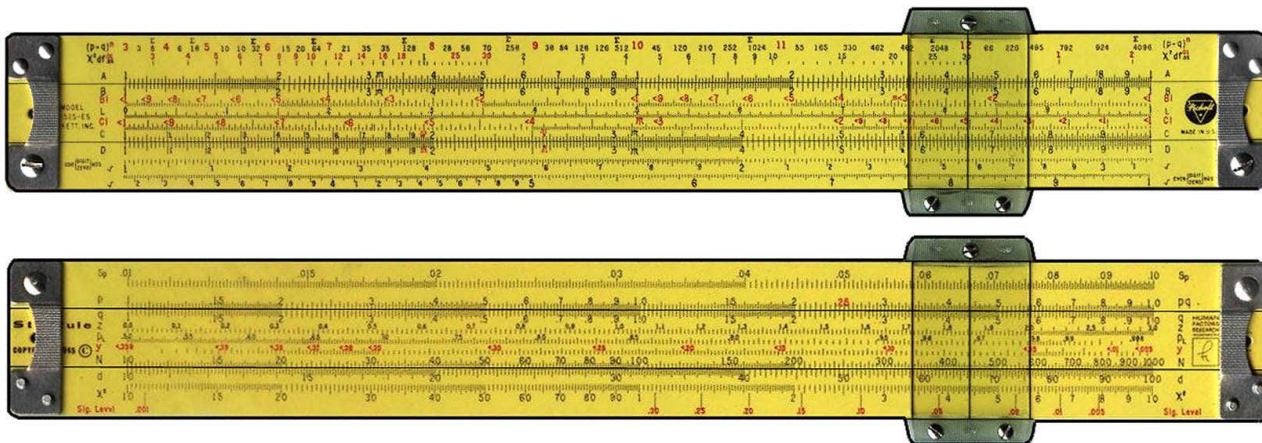
17. Verkiezingen en de Pickett N-525-T

Simon van der Salm

juni 2002

Abstract: A number of statistical scales of the Pickett N-525 statistical slide rule is explained by means of the calculation of a confidence interval of proportions and percentages that we could applicate for political election polls.

Keywords: statistics, statistical slide rule, Pickett N-525-T



De voorspellingen kwamen niet uit?

De dag na de geruchtmakende verkiezingen in mei, was er op TV een interview met een statisticus van een verkiezingsonderzoeksbureau. Het interview gaf de indruk dat hij door de ondervragende journalist op het matje was geroepen om eens te komen uitleggen waarom de “voorspellingen” betreffende de verkiezingsuitslagen niet waren uitgekomen. De statisticus begon met uit te leggen dat op steekproeven gebaseerd onderzoek geen “voorspellingen” opleveren, maar *verwachtingen* en dat noodzakelijkerwijs rond een verwacht percentage, dat een politieke partij zal krijgen tijdens de verkiezingen, een *onzekerheidsmarge* in rekening moet worden gebracht. De statisticus kon nog net de opmerking maken dat de grootte van die marge afhankelijk is van het in de steekproef waargenomen percentage. De ongeduldige

journalist viel hem in de rede, hij moest opschieten met de “verklaring” van de mislukte voorspellingen. Het gevolg was dat de statisticus niet veel meer te berde kon brengen, en dat de vooringenomen mening versterkt werd dat de verkiezingsonderzoeksbureau’s hun werk niet goed doen.

Wat bedoelde de statisticus te zeggen?

Statistische rekenlinialen

Statistische rekenlinialen vormden tot 1975, naast allerlei tabellen, handige hulpmiddelen bij het oplossen van statistische vraagstukken, maar hadden de beperking dat er geen optellingen (die het vaakst in de statistiek voorkomen) mee te verrichten zijn. Dat betekende dat de resultaten van veel tussenbewerkingen van de rekenliniaal op papier moesten worden overgenomen, op papier moesten worden verwerkt, van papier weer moesten worden overgebracht op de

rekenliniaal, waarmee vervolgens een nieuwe reeks bewerkingen moest worden uitgevoerd, enzovoorts. Een werkje dat veel geduld en nauwkeurigheid vereiste. Werk dat tegenwoordig door computers en elektronische rekenmachines wordt gedaan.

Ik bespreek hier een statistische rekenliniaal die, onder andere, gebruikt kan worden voor het bepalen van het *betrouwbaarheidsinterval* rond het in een steekproef waargenomen percentage stemmers dat vóór de verkiezingen zegt op een bepaalde politieke partij te gaan stemmen.

De betreffende liniaal is de

Pickett N-525-T StatRule,
copyright 1965, Human Factors
Research Incorporated.

De schalen op de Pickett N-525-T Statistical Rule

Op de CD-ROM van Herman van Herwijnen wordt de Pickett N-525-ES Statistical Slide Rule (ES voor Eye Saver) vermeld onder matchnummer 3650. Hiervan komen uitvoering en schaalindelingen volkomen overeen met die van de Pickett N-525-T (T voor Tradional White) die in mijn bezit is. Omdat er geen foto beschikbaar is van de T-versie, ziet u in figuur 1 de foto van de ES-versie.

De voorkant heeft de volgende schaalindeling:

$(p + q)^n$ $n(\text{rood})$;
 $\chi^2 df_{.05}^{.01}$ (spreek uit : Chi – kwadraat ,
 deg rees of freedom (.01 rood); A
 =====
 B; BI; L; CI; C;
 =====
 D; $\sqrt{\quad}$; $\sqrt{\quad}$

De achterkant laat de volgende schalen zien:

S_p ; $p(\text{links})$ en $pq(\text{rechts})$;
 =====
 q ; z ; p_L ; $y(\text{rood})$; N
 =====
 d ; χ^2 ; $\text{Sig.Level}(\text{rood})$

Voor het onderhavige artikel zijn de volgende schalen van belang:

N: steekproefgrootte;
 p: steekproefproportie;
 q: complementaire steekproefproportie ($p + q = 1$);
 S_p : standaardfout voor de proporties in steekproeven;

z: onafhankelijke variabele van de standaardnormale verdeling (standaardscore);
 y: afhankelijke variabele van de standaardnormale verdeling;
 p_L : probability left , de kans dat een standaardnormale score een waarde heeft links van een z-waarde, deze kans is identiek aan de oppervlakte onder de standaardnormale kromme links van z.

Proporties en percentages

De schalen S_p , p , q en N laten zien dat de N-525 speciaal bedoeld is voor het rekenen met proporties en percentages bij steekproeven.

De schaal S_p is de schaal voor de *standaarddeviatie voor proporties in steekproeven*, dus de schaal voor de *standaardfout voor proporties* (Eng: standard error). De schaal S_p is een 1-decade-schaal, waarvan de getallen niet van 1 tot 10 lopen, maar van 0.01 tot 0.1. Het is dus een C- of D-schaal, waarvan de waarden door 10 gedeeld zijn. Het gevolg van deze keuze wordt duidelijk als je een aantal keer de standaardfout, bij gegeven steekproefgrootte N en *steekproefproportie* p , probeert uit te rekenen. Je vindt voor praktische gevallen bijna altijd op de S_p -schaal een waarde die geen nadere vermenigvuldiging met een macht van 10 behoeft. Je kunt dus direct aflezen dat de standaardfout bijvoorbeeld 0.0234 is. Het is, behalve in een aantal uitzonderlijke gevallen, niet nodig de plaats van de decimale punt apart te bepalen.

De schalen p en q ($p =$ *steekproefproportie* en q is *complementaire steekproefproportie*) zijn 2-decadeschalen, zoals de gebruikelijke A- of B-schaal. De getallen op de p -schaal lopen echter niet van 1, via 10 tot 100, maar van 0.1 via 1.0 naar 1.0 (sic!). De linkerdecade is de p -decade; de rechterdecade de pq -decade, althans bij het werken met de liniaal lijkt deze keuze het handigst te zijn. De rechterkant van de p -schaal wordt ook niet voor niets met pq aangeduid. De p -schaal bevindt zich op het bovenste vaste deel van de rekenliniaal.

De q -schaal, op de tong van de liniaal, is identiek met de p -schaal, maar bevat aan de rechterkant gewoon q , in plaats van pq .

De N -schaal, eveneens op de tong van de liniaal, is ook een 2-decadenschaal. De waarden van N lopen daarbij van 10 via 100, halverwege de schaal, naar 1000.

De plaatsing van de schalen op de tong en op het vaste deel van de liniaal is zodanig dat met zo min mogelijk handelingen de standaardfout voor proporties is af te lezen.

Bijna voor alle in de praktijk voorkomende gevallen gaan we als volgt te werk:

1. Zoek in de linkerdecade van schaal p de waarde van p , bijvoorbeeld 0.75. Zet de looper op deze waarde;
2. Zet de waarde van N , bijvoorbeeld 20, op schaal N onder de looperstreep;
3. Verplaats de looper naar de waarde van $q = 0.25$ op de q -schaal;
4. Boven de waarde van q lezen we op de S_p -schaal de waarde van $S_p = 0.097$ direct in de juiste grootteorde.

De standaardfout van de *percentages* in de N waarnemingen ($100 \cdot S_p$) vinden we op de D-schaal. We hoeven alleen maar de liniaal om te draaien om op de voorkant het percentage te kunnen aflezen.

De formule die via de vier genoemde schalen S_p , p , q en N wordt toegepast is:

$$S_p = \sqrt{\frac{pq}{N}} = \sqrt{\frac{p(1-p)}{N}},$$

waarbij $p + q = 1$

Het betrouwbaarheidsinterval

De werkelijke proportie van aanhangers van politieke partij X in de Nederlandse bevolking is vlak voor de verkiezingen onbekend. Daarom neemt men een steekproef van N personen uit de bevolking, bepaalt men het aantal aanhangers van partij X in de steekproef en doet men, op grond van het steekproefresultaat, een uitspraak over de werkelijke proportie (of het percentage) in de populatie, in de volgende vorm: “uitgaande van de steekproef ligt het percentage aanhangers van partij X met 95% zekerheid tussen de percentages p_1 en p_2 ”.

We kennen dus niet de proportie van aanhangers van partij X in de bevolking, maar beschikken alleen over een in een steekproef waargenomen proportie p , die slechts als een benadering opgevat kan worden van de werkelijke proportie. Uit die waargenomen waarde van p kunnen we grenzen bepalen waartussen de waarde van de werkelijke proportie vermoedelijk zal liggen. Dat “vermoedelijk” wordt uitgedrukt in een *betrouwbaarheidspercentage*, dat weer gekoppeld is aan een waarde p_L .

De statisticus doet dus geen voorspelling over de werkelijke waarde van de proportie aanhangers! Hij accepteert zelfs dat hij er in 1 op de 20 gevallen naast zal zitten. In de statistiek zegt men: “hij is zeker over de onzekerheid waarmee hij zijn uitspraak doet”.

De vraag is: hoe bepalen we de grenswaarden p_1 en p_2 ?

De schalen z , p_L en y .

Voor het bepalen van een betrouwbaarheidsinterval hebben we, naast S_p een z -waarde nodig

De z -waarde volgt uit de klokvormige kromme die in de statistiek zoveel wordt gebruikt. Voor de z -waarden en bijbehorende y -waarden heeft de Pickett N-525 schalen.

De kans links van een gegeven waarde z kunnen we eveneens op de rekenliniaal vinden. Het is namelijk de waarde p_L , waarvan de schaal tussen de z -schaal en de y -schaal is getekend.

Zo vinden we voor bijvoorbeeld $z = 1.5$ de waarde $y = 0.129$ en voor de bijbehorende kans $p_L = 0.933$.

Uit het symmetrische betrouwbaarheidspercentage 95% = 0.95 volgt het asymmetrische percentage $p_L = 97,5\% = 0.975$ en op de z -schaal vinden we de bijbehorende z -waarde 1.96.

Voorbeeld:

We nemen een steekproef van 250 personen uit de populatie kiesgerechtigde Nederlanders en bepalen hieruit het aantal personen dat zegt op partij X te gaan stemmen.

Stel we vinden in de steekproef dat de proportie kiezers, dat zegt op partij X te gaan stemmen, gelijk is aan $p = 0,23$. Welk 95% betrouwbaarheidsinterval volgt er uit dit percentage?

We berekenen op de eerder beschreven wijze de standaardfout voor proporties:

$$S_p = \sqrt{\frac{pq}{N}} = 0,027 \text{ waarbij } p = 0,23 \text{ en } p + q = 1, \text{ dus } q = 0,77$$

We bepalen vervolgens, via de p_L - en z -schalen, de waarde van z bij $p_L = 0.975$ en vinden wederom $z = 1.96$.

Op grond van de steekproef kunnen we nu stellen dat met een kans van 95% de werkelijke proportie aanhangers van partij X ligt tussen $0.23 - 1,96 * 0.027 = 0.177$ en $0.23 + 1,96 * 0.027 = 0.282$.

Als we dit vertalen in percentages: als 23% van de ondervraagden zegt aanhanger te zijn van partij X, dan ligt met 95% zekerheid het werkelijke percentage tussen 17,7% en 28,2%. Vertaald in Tweede Kamerzetels betekent dat, op grond van hetgeen we in de steekproef hebben waargenomen, met 95% zekerheid geldt, dat het aantal zetels voor X zal liggen tussen 26,55 en 42,3, althans, als er op het moment van de steekproef verkiezingen zouden zijn gehouden.

Nemen we een steekproef van 1000 personen (4 x zogroot), dan wordt $S_p = 0.013$ (half zo groot).

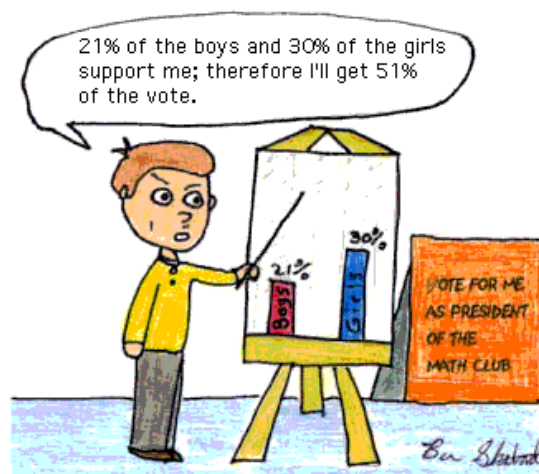
Met 95% zekerheid ligt het werkelijke percentage aanhangers van partij X dan tussen 20.4% en 25.6% en het te verwachten aantal zetels voor partij X tussen 30.6 en 38.4.

Ik meen dat dit laatste resultaat het meest realistisch is. Voor de grotere politieke partijen wordt in de polls inderdaad een onzekerheidsmarge van circa 3 zetels gehanteerd.

Tot slot

In het bovenstaande zijn de statistische schalen van de Pickett N-525-T liniaal, die in relatie met proporties staan, besproken. We hebben gezien hoe de besproken schalen kunnen worden gebruikt voor het bepalen van een betrouwbaarheidsinterval voor verkiezingspercentages.

De liniaal heeft echter nog meer statistische schalen. Deze schalen zal ik bespreken in een vervolgartikel.



[Cartoon]: courtesy of Ben Shabad: <http://Davidmlane.com/ben/cartoons.html>

Op [internet] wemelt het van de statistieksites, vaak met kleine interactieve programma's waarmee statistische situaties gesimuleerd kunnen worden. Een mij bekende site is: http://davidmlane.com/hyperstat/confidence_intervals.html

Hier kun je, zonder dat je zelf het rekenwerk hoeft te verrichten, experimenteren met betrouwbaarheidsintervallen.