14. Chi-kwadraat en de Pickett N-525-T

Simon van der Salm **november 2002**

Abstract: The Pickett N-525-T chi-square scale $\chi^2 df$ is explained by means of an example from the field of slide rule collecting.

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

De statistiek van het verzamelen

In MIR 31 verscheen het eerste artikel uit een kleine serie over één van de statistische rekenlinialen van Pickett, de N-525-T, StatRule ten behoeve van Human Factors Research. Dat artikel bevat ook een foto van de liniaal. In MIR 31 beschreef ik een aantal statistische schalen op deze liniaal. In dit vervolg bespreek ik de Chi-kwadraat-schaal aan de hand van het verzamelen van rekenlinialen.

De Chi-kwadraatverdeling, tezamen met de bijbehorende toetsen, behoort tot de meest gebruikte instrumenten in de statistiek. Zie voor de wiskundige en statistische achtergronden leerboeken over statistiek. Voor het begrijpen van dit artikel zijn die achtergronden niet nodig. Het is de bedoeling om te laten zien waarvoor een statistische rekenliniaal gebruikt kan worden. Om iets te weten over de toepassing van gauge-points op linialen voor vaten alcohol hoeft u tenslotte ook niet zelf alcohol te kunnen stoken.

Is het verzamelen van rekenlinialen voor iedereen mogelijk?

U heeft ongetwijfeld wel eens de verzameling van een medeverzamelaar bekeken. U zag prachtige en zeldzame exemplaren, rekenlinialen die uzelf niet in bezit heeft. Geeft u uw goede euro's aan iets anders uit? Zijn die dure linialen tegenwoordig bijna niet meer te verkrijgen?

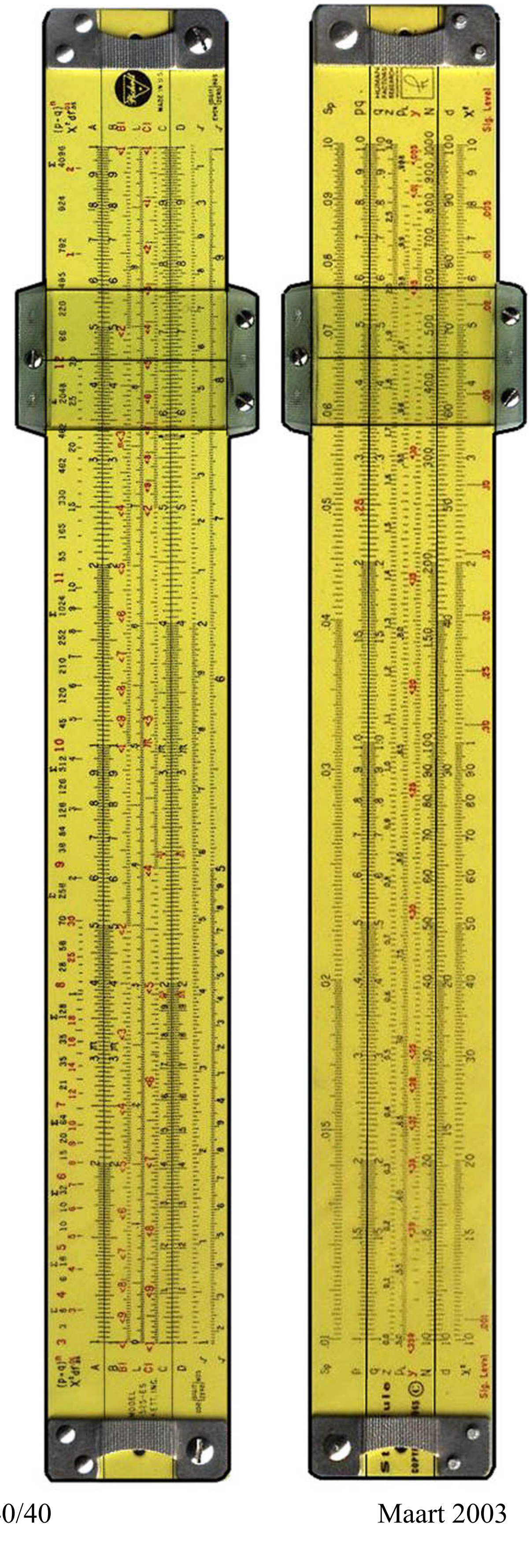
Iedere verzamelaar wordt wel eens geconfronteerd met de vraag of hij wel bereid is een flink hoog bedrag neer te tellen voor een object dat hem wordt aangeboden. Zijn overwegingen om al dan niet tot aankoop over te gaan, hebben te maken met zijn financiële mogelijkheden, maar ook met de verhouding tussen het bedrag dat het object waard is en de prijs die hij er voor wil betalen.

Bij het verzamelen van rekenlinialen komt de vraag op, of het inkomen van de verzamelaar een grote rol speelt bij zijn inkoopbeleid. Vergeleken met andere te verzamelen objecten lijken rekenlinialen in het algemeen gesproken niet zo heel erg duur. Wat de verzamelaar duur vindt, is deels bepaald door zijn subjectieve opvatting over de gevraagde prijs. Bij objecten die iedere verzamelaar zou kunnen betalen, als hij daarvoor zou kiezen, spelen andere zaken dan zijn inkomen een grote rol bij het al dan niet aankopen van een object.

De prijs van de objecten in zijn verzameling en de inkomensklasse van de verzamelaar lijken onafhankelijk van elkaar te zijn.
Daarom bestaat het vermoeden dat de prijs (= financiële waarde) van de rekenlinialen die de verzamelaar in zijn verzameling heeft, niet sterk correleert met zijn inkomen, maar met andere zaken die zijn keuze beïnvloeden. Bijvoorbeeld met de keuze of hij zijn geld wel of niet aan iets anders wil besteden. Een beetje ongenuanceerd gezegd: iedere verzamelaar zou alle rekenlinialen kunnen verzamelen, als hij dat zou willen, onafhankelijk van zijn inkomen.

De vraag of de gemiddelde prijs van rekenlinialen in verzamelingen afhankelijk of juist onafhankelijk is van de inkomensklasse van de verzamelaar, is een statistische vraag, die een groot aantal verzamelaars betreft. Deze vraag betreft dus niet de individuele verzamelaar. Er zijn arme verzamelaars die de prachtigste stukken in hun verzameling hebben. Daarnaast zijn er schatrijke verzamelaars met een fraaie, doch niet opvallende collectie, als we althans kijken naar de financiële waarde van de objecten. Het omgekeerde komt ook voor.

We stellen ons dus de vraag of



“gemiddeld genomen” de kwaliteit van verzamelingen rekenlinialen al of niet samenhangt met de inkomensklasse van de verzamelaar.

Als we deze vraag willen beantwoorden, dan moeten we de collectie van een groot aantal verzamelaars bestuderen en op statistische overwegingen een beslissing nemen over de afhankelijkheid/onafhankelijkheid van prijs en inkomen.

Het onderzoek onder rekenlinialenverzamelaars

De schrijver heeft geen authentieke gegevens uit veldonderzoek onder verzamelaars van rekenlinialen. Daarom werken we met gefingeerde getallen.
De gemiddelde prijzen van rekenlinialen verdelen we in drie klassen: laag, midden en hoog. De inkomens van verzamelaars verdelen we ook in drie klassen: lager dan modaal, modaal en bovenmodaal.

Er bleken 525 verzamelaars van rekenlinialen mee te willen doen in het onderzoek. Van iedere verzamelaar werd bepaald tot welke prijsklasse zijn verzameling hoort en in welke klasse zijn inkomen valt.

De onderstaande tabel bevat drie matrices. De bovenste geeft een overzicht van de gevonden waarden. Deze waarden noemt men *waargenomen frequenties* W_{ij}.
Langs de horizontale as staat de prijsklasse van de verzameling vermeld; de verticale as geeft de inkomensklassen.

Het lezen van de matrix is niet moeilijk. Zo leest u in de middelste cel van de bovenste matrix dat 85 van de 525 verzamelaars een middenklasverzameling hebben en tegelijkertijd een modaal inkomen hebben.

In de bovenste matrix zien we dat 240 van de 525 verzamelaars een modaal inkomen hebben en dat maar liefst 225 verzamelaars een collectie van hoge kwaliteit in huis hebben. U ziet: de cijfers lijken in elk geval realistisch.

Relatie inkomensklasse en prijsklasse van rekenlinialen

Waarnem.	laag	midden	hoog	(Prijs)
< modaal	35	40	75	150
modaal	65	85	90	240
>modaal	30	45	60	135
(Inkomen)	130	170	225	525

Verwacht	laag	midden	hoog	(Prijs)
< modaal	37,1	48,6	64,3	150,0
modaal	59,4	77,7	102,9	240,0
>modaal	33,4	43,7	57,9	135,0
(Inkomen)	130,0	170,0	225,0	525,0

Toets	laag	midden	hoog	(Prijs)	
< modaal	0,1	1,5	1,8	3,4	
modaal	0,5	0,7	1,6	2,8	
>modaal	0,4	0,0	0,1	0,5	
(Inkomen)	1,0	2,2	3,5	**6,7**	**(Toets-grooth)**

p-waarde =	**0,15**		
alfa =	**0,05**	**Krit.grs =**	**9,49**
betrouwb =	**95%**		

Onafhankelijkheid

We vermoeden dat de prijsklasse van de rekenlinialenverzameling "gemiddeld genomen" onafhankelijk is van de inkomensklasse van de verzamelaar.
Bij volkomen onafhankelijkheid zou de matrix van waargenomen getallen eruit zien als de middelste matrix in de tabel. De getallen in deze matrix vind je door het product van bijbehorend rijgetal en kolomgetal met elkaar te vermenigvuldigen en te delen door 525. Bijvoorbeeld, voor het getal in de middelste cel: 77,7 = (170 * 240)/525. Op overeenkomstige wijze kunnen we de andere getallen vinden. De getallen in de tweede matrix heten *verwachte frequenties* V_{ij} .

Tegenwoordig gebruiken we een computerprogramma of uitgebreide rekenmachine voor het opstellen van de tabel met verwachte frequenties. Vroeger was het een tijdrovende en foutgevoelige zaak om deze tabel op te stellen, zeker als het aantal klassen groot werd.

De Chi-kwadraatscore

Uit de twee bovenste matrices wordt de derde afgeleid. Met een computer een

koud kunstje, maar tot halverwege de jaren zeventig bewerkelijk.
Voor iedere cel wordt het kwadraat van het verschil tussen verwachte en waargenomen frequentie berekend en gedeeld door de verwachte frequentie:

$$\frac{\left(V_{ij} - W_{ij}\right)^2}{V_{ij}}$$

Met deze formule bereken je het gekwadrateerde relatieve verschil tussen de waarnemingsresultaten W_{ij} en de resultaten V_{ij} die je bij onafhankelijkheid zou verwachten, gemeten ten opzichte van hetgeen je zou verwachten. .

Vervolgens worden de aldus berekende waarden bij elkaar opgeteld:

$$\chi^2 = \sum_i \sum_j \frac{\left(V_{ij} - W_{ij}\right)^2}{V_{ij}}$$

Deze waarde is de *Chi-kwadraatscore*. Hier vinden we de score 6,7.

Dit is de toetsgrootheid op grond waarvan we uiteindelijk beslissen of er al dan niet sprake is van onafhankelijkheid tussen de prijsklasse van de verzameling en de inkomensklasse van de verzamelaar.

De beslissing door middel van de rekenliniaal

Bij volkomen onafhankelijkheid is de Chi-kwadraatscore gelijk aan 0. Hoe meer er sprake is van afhankelijkheid tussen inkomen en kwaliteit van de verzameling, des te groter is de score. Hoe beslissen we nu of er al dan niet sprake is van afhankelijkheid?

Daarvoor hebben we de statistische rekenliniaal nodig, in het bijzonder de schaal $\chi^2 df$
De letters df staan voor *degrees of freedom*. De twee $\chi^2 df$-schalen van de Pickett 525 lopen van 1 tot en met 30 vrijheidsgraden.

Voor problemen zoals hierboven geldt dat het aantal df gelijk is aan:

df =
(aantal rijen –1) * (aantal kolommen –1)

Voor ons onderzoek geldt dus:
df = (3 –1) * (3-1) = 4

Naast het vaststellen van het aantal vrijheidsgraden, dienen we een betrouwbaarheidsniveau voor onze toets af te spreken. Veel voorkomende waarden zijn 95% en 99%.
De Pickett 525 bevat daarom voor beide percentages bijbehorende df-schalen. De .01-schaal (in rode cijfers) hoort bij het betrouwbaarheidspercentage 99%. De schaal wordt met het complement (alfa) van het betrouwbaarheidpercentage aangeduid. De .05-schaal (in zwarte cijfers) hoort bij het betrouwbaarheidspercentage 95%.

In de tabel van het onderzoek is van dat laatste percentage (en dus van een alfa van 5%) uitgegaan.

Het getal df = 4, op de zwarte schaalverdeling, komt op de A-schaal overeen met de zogenaamde *kritische grenswaarde* 9,49, dezelfde waarde die het spread sheet programma leverde voor de bovenstaande tabel.

Omdat de gevonden waarde van de toetsgrootheid (= 6,7) kleiner is dan de kritische grenswaarde (= 9,49), beslissen we dat er onafhankelijkheid bestaat tussen de prijsklasse van de rekenlinialenverzameling en de inkomensklasse van de verzamelaar.

Wat opvalt

Opvallend is dat heel veel werk niet met de rekenliniaal gedaan kan worden, omdat er sommen en verschillen moeten worden berekend. Voor de quotiënten kan men natuurlijk wel de rekenliniaal hanteren.

De statistische schalen van de liniaal komen pas in het laatste stadium van de berekeningen aan de orde. Juist hiervoor gebruikte men vroeger meestal statistische tabellen. Deze tabellen geven gewoonlijk voor veel andere betrouwbaarheidspercentages en aantallen vrijheidsgraden, de kritische grenzen voor de Chi-kwadraatscores. Kennelijk waren de schalen op de Pickett 525 voldoende voor de specifieke toepassingen (Human Factors Research) waarvoor hij ontworpen was.

De belastingdienst: nog een toepassing

De Chi-kwadraattoets kent vele toepassingen. De belastingdienst onderzoekt tegenwoordig volkomen geautomatiseerd of er aanleiding is om uw belastingaangifte onder de loep te nemen. Het is namelijk bekend dat bijvoorbeeld de laatste cijfers van de getallen in de aangiftes een specifieke frequentieverdeling dienen te volgen. Met de Chi-kwadraattoets wordt bepaald of de frequentieverdeling van de bedragen in uw aangifte niet te veel afwijkt van deze specifieke verdeling. Het is bekend dat fraudeurs de getallen zodanig kiezen dat al snel van die specifieke verdeling wordt afgeweken. De fraudeur kan met een statistische rekenliniaal zoals de Pickett 525 nagaan hoever hij kan gaan in het verzinnen van bedragen, terwijl de inspecteur kan beoordelen of er al dan niet sprake is van mogelijke fraude.

Tot slot

Het artikel beschrijft één kenmerkende toepassing van de Chi-kwadraattoets. Aan de hand van een gefingeerd onderzoek onder rekenlinialenverzamelaars wordt getoond wat de betekenis is van de twee schalen $\chi^2 df$ op de Pickett N-525-T Statistical Slide Rule.

De Pickett N-525-T is een duplex rekenliniaal, dus het is wat lastig om over voor- en achterkant te spreken. De voorkant noem ik die zijde van de liniaal waarop de A, B, C en D-schalen zich bevinden. De achterkant bevat een viertal schalen, N, d, χ^2 en Sig.Level genaamd, die kunnen worden gebruikt bij toetsen die gebaseerd zijn op de $\chi^2 distributie$
In een vervolgartikel zal ik hierop ingaan.