Examenuitslag: de score-zeshoek

Simon van der Salm

Scores in plaats van cijfers

Als u deze maand weer op veel plaatsen de schooltassen van geslaagde scholieren ziet hangen aan de vlaggenstok, dan realiseert u zich waarschijnlijk niet wat daaraan is voorafgegaan. “Docenten geven cijfers en dat is het wel” zult u denken, maar in deze gecompliceerde tijd is dat vanzelfsprekend veel te simpel gedacht. Docenten geven geen cijfers meer, maar scores, een groot verschil!

Elk vraagstuk krijgt een aantal scorepunten; het examen voor elk vak kent een maximale score, de zogenaamde scorelengte L. Voor elk vak en van elke leerling in het voortgezet onderwijs wordt het gemaakte examenwerk door twee correctoren nagekeken, die het samen eens moeten zien te worden over de scores, die bij elkaar opgeteld de score $S \leq L$ voor de betreffende leerling oplevert. De correctoren, dus ook de docenten, weten niet welk cijfer C overeenkomt met welke score S. Het cijfer kan wel 2 punten hoger of lager uitvallen dan gedacht.

Deze maand, op 12 juni, zaten ’s morgens vroeg op alle scholen voor voortgezet onderwijs de examensecretarissen met het zweet in de handen te wachten op de normeringstermen van het CvTE, het College voor Toetsing en Examens. Pas als deze zogenaamde N-termen, voor elk vak, bekend zijn, kunnen de cijfers worden bepaald. Daarvoor is een ingenieus computerprogramma nodig. Het werk van de examensecretarissen is het inkloppen van de door het CvTE bepaalde N-termen en…. vervolgens rolt er voor elke leerling, voor elk vak, een cijfer uit de computer.

De N-termen worden door het CvTE bepaald op grond van een statistische analyse van de examenscores. Vallen die scores lager uit dan op grond van de statistiek kan worden verwacht, dan wordt een hoge N-term bepaald; vallen ze daarentegen hoger uit dan verwacht, dan wordt een lage N-term berekend. Zo zorgt het CvTE ervoor dat het niveau van bijvoorbeeld het HAVO-examen Geschiedenis van jaar op jaar een gelijke becijfering krijgt, onafhankelijk van de toevallige fluctuaties in moeilijkheidsgraad.
In tabel 1 ziet u een voorbeeld van door het CvTE toegekende N-termen.

Van examens krijg je uitslag

Op scholen kunnen gewoonlijk alleen de wiskundeleraren uitleggen wat er in dat computerprogramma gebeurt. Hoewel, de gevolgde methode is weleens op het eindexamen VWO in een vraagstuk aan de orde geweest. Dus de examenkandidaten in dat jaar wisten precies wat hun te wachten stond. Wie zegt dat wiskundigen geen gevoel voor humor hebben?

Enfin, de secretaris, die in het dagelijks leven docente Duits is, drukt op de knop van de computer en accepteert maar op goed vertrouwen (en op goed geluk) het resultaat dat er uitkomt. “Eerlijker dan dit kan het niet gaan”, mompelt zij.

HAVO

jaar	2018		2017		2016		2015		2014		2013	
tijdvak	1e[10]	2e	1e	2e	1e	2e	1e	2e[6]	1e	2e[7]	1e[8]	2e[9]
Nederlands	0,7	1,1	1,6	1,6	0,9	1,0	1,1	1,1	1,2	1,3	1,1	1,1
Fries	0,2	-	0,0	-	0,0	-	0,0	-	0,0	-	0,5	-
Frans	0,4	0,5	0,0	0,0	0,0	0,0	0,0	0,2	0,4	0,4	1,0	1,0
Duits	0,6	0,8	0,2	1,0	0,5	0,8	1,0	1,0	0,4	0,4	0,2	0,3
Engels	1,4	1,4	1,4	1,4	1,1	1,1	0,5	0,5	0,0	0,0	0,3	0,3
Spaans	1,9	1,9	1,3	1,3	1,2	1,2	1,0	1,0	1,5	1,5	1,3	1,3
Russisch	0,7	-	0,7	-	0,7	-	0,7	-	0,7	-	0,7	-
Turks	0,7	-	0,7	-	1,0	-	0,7	-	0,7	-	0,7	-
Arabisch	0,7	-	0,7	-	0,7	-	0,7	-	0,7	-	0,7	-
Geschiedenis	1,0	1,0	0,7	1,2	1,3	1,3	1,1	1,5	1,7	1,8	1,6	1,6
Aardrijkskunde	1,6	1,6	1,1	1,1	1,0	1,0	1,0	1,1	1,3	1,4	1,2	1,2
Wiskunde A	1,2	1,2	1,5	1,5	1,0	1,0	1,0	1,0	1,1	1,1	0,5	0,5
Wiskunde A (bezem)	-	-	1,1	1,2	-	-	-	-	-	-	-	-
Wiskunde B	1,2	1,7	1,8	1,8	1,5	1,6	1,2	1,6	1,4	1,4	0,8	1,1
Wiskunde B (bezem)	-	-	2,0	-	-	-	-	-	-	-	-	-
Natuurkunde	0,6	0,5	1,2	1,1	1,2	1,5	1,2	1,2	1,3	1,3	1,3	1,4
Natuurkunde (bezem)	-	-	-	-	-	-	1,4	-	-	-	-	-
Scheikunde	1,0	1,4	1,2	1,4	1,7	1,7	1,8	1,8	1,4	1,4	1,5	1,8

Tabel. 1. Een voorbeeld van N-termen, zoals gepubliceerd door het CvTE.

Toch is met een beetje wiskunde van rechte lijnen (al bedacht door Descartes) best te begrijpen welke relatie er bestaat tussen het uiteindelijk toegekende cijfer C bij de behaalde score S, nadat de waarde van N is bekend gemaakt.

Ik reken het u voor op mijn TI-NSpire CX-CAS, grafische rekenmachine, die geeft niet zulke mooie plaatjes als de modernere rekenmachines, maar in elk geval voldoende duidelijk om te begrijpen hoe die relatie in elkaar steekt. Zie figuur 1 voor het voorbeeld met $L = 77$.

Het meest linker hoekpunt (0;1,0) van de score-zeshoek ligt op de roze, horizontale stippellijn door het cijfer 1,0; het meest rechter hoekpunt (77,10) ligt op de groene, horizontale stippellijn door het cijfer 10. In blauw is de hoge cijfergrens aangeduid; in rood de lage cijfergrens. De zwarte lijn is de bij L = 77 behorende normaalrelatie. Elke combinatie van score S en cijfer C, dus het punt (S,C), ligt binnen of op deze zeshoek.

We gaan uit van N = 1,5. Omdat er sprake is van een N-term groter dan 1, telt de hoge cijfergrens (blauw). We zien de twee snijpunten (4,28;2,0) en (68,4;9,5) van de normaalrelatie en de hoge cijfer-grens. Scores S van 0 tot en met 4 geven daarom een cijfer C volgens de hoge cijfergrens (blauw); scores van 5 tot en met 68 geven een cijfer C volgens de normaalrelatie (zwart); scores van 69 tot en met 77 volgen weer de hoge cijfergrens (blauw).

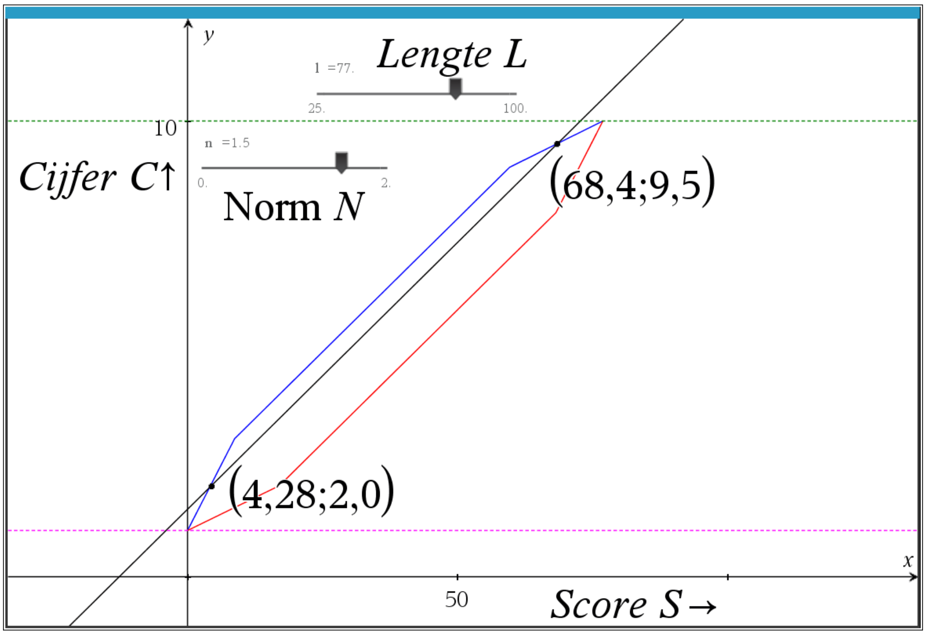

Fig. 1. Cijfertoekenning bij een scorelengte van L = 77 en normeringsterm N = 1,5. De lage cijfergrens is rood gekleurd; de hoge cijfergrens blauw; de normaalrelatie zwart. Omdat de N-term groter is dan 1, zijn er zijn twee snijpunten van de normaalrelatie en de hoge cijfergrens.

Op deze wijze komt er altijd eenduidig een cijfer C van 1,0 tot en met 9,9 (1 decimaal!) en vervolgens 10 uit, wat wettelijk is voorgeschreven. Op overeenkomstige wijze gaan we te werk voor N-termen kleiner dan 1, waarbij dan de lage cijfergrens van de zeshoek (rood) een rol speelt.

Normeringsterm N

De N-term is dus een variabele die Nederland gebruikt wordt bij het berekenen van het cijfer bij een examen. De definitieve N-term wordt door het CvTE afgeleid uit het aanvankelijke gemiddelde cijfer bij $N = 1$ en nog enkele andere statistische parameters. Als het gemiddelde cijfer lager is dan verwacht, wordt $N = 1$ naar $N > 1$ aangepast. Als dit gemiddelde cijfer hoger is dan verwacht, wordt $N = 1$ naar $N < 1$ aangepast. Als de N-term bekend is, gaat de computer als volgt te werk:

Normaalrelatie (zwart in figuur 1)

$$C_N(S) := \frac{9}{L} \cdot S + N; \quad 0 \leq S \leq L \tag{1}$$

- C het cijfer, in twee significante cijfers
- S de score, een geheel getal
- L een geheel getal, de lengte van de scoreschaal, zoals vastgelegd in het correctievoorschrift
- N kan, op uitzonderingen na, iedere waarde van alle tienden van 0,0 tot en met 2,0 aannemen

Zeshoekzijden

Lage cijfergrens (rood): $$C_l(S) := \begin{cases} \frac{9}{L} \cdot \frac{S}{2} + 1; & 0 \leq S < \frac{2}{9}L \\ \frac{9}{L} \cdot S; & \frac{2}{9}L \leq S < \frac{8}{9}L \\ \frac{9}{L} \cdot 2S - 8; & \frac{8}{9}L \leq S \leq L \end{cases} \tag{2}$$

Hoge cijfergrens (blauw): $$C_h(S) := \begin{cases} \frac{9}{L}\cdot 2S + 1; & 0 \le S < \frac{1}{9}L \\ \frac{9}{L}\cdot S + 2; & \frac{1}{9}L \le S < \frac{7}{9}L \\ \frac{9}{L}\cdot \frac{S}{2} + \frac{11}{2}; & \frac{7}{9}L \le S \le L \end{cases} \qquad (3)$$

Snijpunten normaalrelatie met zeshoekzijden

$$\begin{cases} N < 1 \to C_N(S) = C_l(S) \leftrightarrow S_l = \frac{2(1-N)}{9}L \vee S_h = \frac{N+8}{9}L \\ N \ge 1 \to C_N(S) = C_h(S) \leftrightarrow S_l = \frac{(N-1)}{9}L \vee S_h = \frac{11-2N}{9}L \end{cases} \qquad (4)$$

Cijfertoekenning

$$\begin{cases} N < 1 \to \begin{cases} 0 \le S < S_l \to C = C_l(S) \\ S_l \le S < S_h \to C = C_N(S) \\ S_h \le S \le L \to C = C_l(S) \end{cases} \\ N \ge 1 \to \begin{cases} 0 \le S < S_l \to C = C_h(S) \\ S_l \le S < S_h \to C = C_N(S) \\ S_h \le S \le L \to C = C_h(S) \end{cases} \end{cases} \qquad (5)$$

Bij een herexamen in een vak geldt gewoonlijk minimaal dezelfde *N*-term als bij het examen in dat vak. Een enkele keer zijn, bij extreem moeilijke examens, *N*-termen boven de 2,0 toegepast.

Waarom zo moeilijk?
Dit is natuurlijk een hoop gedoe, maar geeft wel de mogelijkheid tot het eerlijker becijferen van examens. En de computer doet het vreselijke rekenwerk, dus zoveel inspanning wordt er nu ook weer niet van de examensecretaris verwacht.

Deze methodiek is in eind jaren negentig bedacht na een rechtszaak gevoerd door de ouders van een briljante leerlinge. Die had, ik noem maar wat, voor het vak VWO wiskunde-B een 10 behaald. Dat jaar was het examen nogal moeilijk, werden er veel onvoldoendes behaald en werd besloten het cijfer van iedere leerling met zoiets als 0,8 punten op te hogen. De betreffende leerling zou dus 10,8 hebben moeten krijgen, maar op haar cijferlijst stond een 10. Genoeg reden voor een rechtszaak. De rechter besloot dat ze alsnog 10,8 op haar lijst moest krijgen. De minister zei “dat nooit meer”, en vroeg een wiskundeleraar een oplossing te bedenken.

Enfin, … na lezing van dit artikel kent u de gevolgen.