

Prototypes rond statistiek en datasets

Schoolexamens statistiek met grote datasets. Moeilijk om zelf te maken, maar gelukkig worden we op prototypes getrakteerd door het jarige Cito.

Inleiding

Vanwege het 50-jarig bestaan heeft Cito getrakteerd op gratis schoolexamens statistiek met grote datasets en ict voor havo wiskunde A. De waardering van het veld voor deze traktatie heeft geleid tot schoolexamenopgaven voor vwo wiskunde A een jaar later. Na deze ervaringen deelt Cito haar deskundigheid met wiskundesecties. In de loop van 2020 zullen wederom grote datasets met bijbehorende vragen en correctiemodellen ter beschikking worden gesteld.

Aanleiding

Sinds 2015 is statistiek met grote datasets en ict onderdeel van de examenprogramma's voor havo wiskunde A en voor vwo wiskunde A/C en een verplicht onderdeel van het schoolexamen. Vanuit het veld kregen we signalen dat het vinden van grote datasets en het construeren van bijbehorende vragen nog niet meevalt. Het 50-jarig bestaan van Cito bood de toetsdeskundigen wiskunde de gelegenheid een prototype van schoolexamens hierover te maken en gratis ter beschikking te stellen. Het maken van >

schoolexamens is geen kerntaak voor Cito maar voor deze gelegenheid kon gelukkig een uitzondering worden gemaakt.

2018: Havo wiskunde A

Voor de constructie van de opgaven heeft Cito twee ervaren constructeurs aangetrokken die zeer bekend zijn met statistiek in het nieuwe examenprogramma. Deze constructeurs zijn onder leiding van de toetsdeskundigen aan de slag gegaan. Dit heeft uiteindelijk geleid tot twee sets schoolexamenopgaven, zie tabel 1.

Excel en VuStat

Om docenten de vrijheid te geven, is besloten om de opgaven en correctiemodellen in Word te maken. Op deze manier kunnen opgaven naar believen worden aangepast, aangevuld of verwijderd. Omdat op scholen verschillende software wordt gebruikt voor het omgaan met grote datasets, is ervoor gekozen om de databestanden in het formaat van twee gangbare softwarepakketten mee te leveren: als Excel-bestand en als VuStat-bestand.^[1] De opgaven kunnen met beide programma's gemaakt worden. De correctiemodellen zijn gebaseerd op het gebruik van VuStat omdat dit naar onze mening leerlingvriendelijker is dan Excel voor wat betreft statistische analyses.

Classificering

In een Excelbestand zijn kenmerken van opgaven meegeleverd. Hiermee kunnen docenten keuzes van opgaven maken. De kenmerken hebben betrekking op het aantal punten per vraag, het subdomein uit de syllabus waarover de vraag gaat en een classificering van de vraag. De vragen zijn ingedeeld in de volgende categorieën: *procedureel*, *conceptueel* en *computervaardigheid*. Een vraag is procedureel genoemd als de vraag het uitvoeren van een standaardprocedure betreft, bijvoorbeeld het berekenen van het 95%-betrouwbaarheidsinterval aan de hand van de formulekaart. Bij vragen die conceptueel zijn

genoemd, gaat het om meer dan alleen het uitvoeren van een standaardprocedure. Denk hierbij bijvoorbeeld aan vragen waarbij de leerling moet aangeven waarom een bepaalde procedure niet werkt, of onjuist is, of waarbij de leerling het verschil tussen twee groepen moet bepalen, zonder dat daarbij wordt verteld welke maat hij hiervoor moet gebruiken. We merken op dat de mate waarin een vraag procedureel of conceptueel is, afhangt van de voorkennis van een leerling. Vragen die leerlingen al eens hebben gezien of waarmee specifiek geoefend is in de les, kun je al gauw als minder conceptueel zien. Onze indeling moet daarom met enige voorzichtigheid worden geïnterpreteerd. De vragen worden beantwoord met behulp van Excel of VuStat. Beide vragensets vereisen dat leerlingen overweg kunnen met het betreffende programma. Vragen die hier sterk een beroep op doen zijn bijvoorbeeld het maken van een cirkel- of staafdiagram, frequentietabel, kruistabel of het splitsen van variabelen. Een enkele vraag is geclassificeerd als alleen een computervaardigheid. De meeste vragen zijn geclassificeerd als computervaardigheid in combinatie met procedureel of conceptueel. Tot slot zijn er twee vragen die alleen betrekking hebben op het toepassen van een definitie. In deze vragen wordt niet naar een berekening gevraagd of naar het uitvoeren van een procedure maar bijvoorbeeld om het benoemen van het type variabele als kwalitatief of kwantitatief.

Enquête

In de WiskundeE-brief is een bericht geplaatst waarin docenten werden uitgenodigd om zich aan te melden. Om enigszins te voorkomen dat leerlingen zich zouden aanmelden, is een check gedaan op de BRIN-code van de school en het e-mailadres van de docent. Het materiaal is als zipbestand gestuurd naar ruim 300 wiskundedocenten. Na een maand is onder deze docenten een enquête uitgezet. Iets minder dan 25% van de aangeschreven docenten heeft gereageerd. Uit deze reacties bleek dat de opgaven en het correctiemodel zeer worden gewaardeerd.

Titel	Staandehoudingen	COOL
Omschrijving dataset	bevat gegevens van staandehoudingen in het verkeer in de staat Florida in de Verenigde Staten	bevat gegevens van havo-leerlingen die eind-examen hebben gedaan
Aantal records	1740	6976
Aantal vragen	16	28
Duur toets	90 minuten	afhankelijk van de keuze van de vragen
Aard van de opgavenset	compleet schoolexamen	vragen om zelf een schoolexamen uit samen te stellen
Niveau	havo wiskunde A	havo wiskunde A

tabel 1 Kenmerken van de datasets en bijbehorende opgavensets voor havo wiskunde A

Titel	Speeddating	Schoolverlaters
Omschrijving dataset	bevat gegevens van deelnemers aan een speeddating-experiment in de VS	bevat gegevens van schoolverlaters en hun vervolgoopleidingen in Nederland
Aantal records	8388	3690
Aantal vragen	18	23
Duur toets	90 minuten	afhankelijk van de keuze van de vragen
Beschrijving opgavenset	compleet schoolexamen	vragen om zelf een schoolexamen uit samen te stellen
Niveau	vwo wiskunde A; in Excelbestand is aangegeven welke vragen geschikt zijn voor havo wiskunde A/ vwo wiskunde C	vwo wiskunde A; in Excelbestand is aangegeven welke vragen geschikt zijn voor havo wiskunde A/ vwo wiskunde C

tabel 2 Kenmerken van de datasets en bijbehorende opgavensets voor vwo wiskunde A

Men vond de opgaven goed op niveau en divers genoeg. De meeste respondenten laten leerlingen Excel gebruiken en hadden interesse in een vergelijkbare set voor vwo.

2019: vwo wiskunde A/C

In december 2019 zijn wederom twee datasets opgeleverd, nu met schoolexamens die geschikt zijn voor vwo wiskunde A, zie tabel 2. In de meegeleverde analyse van de opgaven is per vraag aangegeven of deze past binnen de examenprogramma's van vwo wiskunde C en van havo wiskunde A. Ook dit materiaal is verspreid onder meer dan 300 docenten die zich als geïnteresseerden hadden aangemeld. Op basis van de constructie van de vier sets binnen Cito is een beknopte handleiding ontwikkeld. Hierin staat onder meer beschreven waar geschikte datasets gevonden kunnen worden, welke eigenschappen een dataset geschikt maken, welke voor- en nadelen er zijn wat betreft software en hoe gezorgd kan worden voor een diversiteit in vragen. Wil je deze handleiding ontvangen? Stuur dan een mail naar se-statistiek@cito.nl dan sturen we je deze toe.

2020 en verder ...

De ervaringen die we hebben opgedaan met de constructie van deze prototypes van schoolexamens willen we delen met het veld. Daarom hebben we in het najaar van 2019 een oproep gedaan aan scholen in de WiskundE-brief en in de nieuwsbrief van de NVvW om zich bij ons te melden met als doel om samen materiaal te ontwikkelen. Sinds december 2019 zijn een wiskundesectie en een docentenontwikkelteam aan de slag om datasets te vinden en opgaven met bijbehorende correctiemodellen te construeren. De bedoeling is dat al het ontwikkelde materiaal beschikbaar wordt gesteld via WisBase.^[2] Deze site, gehost door de NVvW, biedt toetsmateriaal voor docenten. Op de site kun je gratis een account aanmaken. Voor het inzien van toetsen is het niet nodig om zelf materiaal in te sturen en ook niet om

lid te zijn van de NVvW. De prototypes van schoolexamens statistiek die door Cito zijn ontwikkeld, zijn in WisBase opgenomen. We willen je uitnodigen om jouw eigen toetsen statistiek met grote datasets en ict aan WisBase toe te voegen. Op die manier kun je elkaar werk uit handen nemen en ontstaat er een ruime keuze.

Heb je schoolexamenmateriaal ontwikkeld en wil je feedback van Cito? Dan nodigen we je uit om de dataset, opgaven en het correctiemodel naar se-statistiek@cito.nl te mailen.^[3]

Tot slot

Met het materiaal dat we hebben ontwikkeld hopen we dat we een bijdrage hebben geleverd aan de kwaliteit van schoolexamens statistiek met grote datasets en ict. Door de handleiding en ons materiaal beschikbaar te maken, hopen we handvatten te hebben aangereikt aan docenten om zelf schoolexamens statistiek met grote datasets en ict te kunnen maken. Wij zien jouw materiaal met plezier tegemoet!

Noten

- [1] Voor meer informatie over VUSTAT, zie <http://www.vusoft.be/index.html>
- [2] Zie <https://wisbase.nl/>
- [3] We houden hierbij een slag om de arm: mocht de belangstelling te groot worden, dan zijn we genooddaakt een keuze te maken.

Over de auteurs

Ivo Claus, Ger Limpens en Irene van Stiphout zijn toetsdeskundigen bij Cito.
E-mailadres: se-statistiek@cito.nl.