

Cesurbepaling houdt veel meer in dan afspreken dat 55% een voldoende is

Door: Jolanda Soeting en Esther Haykens

Stel: je hebt een volledige toets of opdracht samengesteld. De toets bestaat bijvoorbeeld uit een aantal gesloten vragen en een aantal open vragen. Je telt alle punten op die de studenten maximaal kunnen behalen. Daarna sta je voor een lastig vraagstuk: hoeveel punten moet een student minimaal halen voor een voldoende?

Het minimaal aantal punten dat studenten moeten behalen voor een voldoende, wordt de *cesuur* of de *zak/slaaggrens* genoemd. Er zijn twee methoden om de cesuur te bepalen:

1. De **absolute cesuur**, waarbij het aantal punten voor een voldoende **vooraf** wordt bepaald.
2. De **relatieve cesuur**, waarbij het aantal punten voor een voldoende **achteraf** wordt bepaald, op basis van de scores van alle studenten.

Beide methodes kun je combineren. Hierna lees je meer over de beide methodes en het combineren ervan.

De absolute cesuur

Deze cesuurmethode wordt het meest gebruikt in de praktijk en is in de meeste gevallen ook aan te raden. Bij deze methode bepaal je vóór afname van het examen hoeveel punten de studenten moeten halen voor een voldoende. Hiervoor kun je kijken naar een aantal zaken:

1. Bekijk welke **leerdoelen** in het examen worden getoetst. Hoeveel procent van deze leerdoelen zouden de studenten minimaal moeten beheersen om een voldoende te halen?
2. Bestaat het examen uit veel relatief gemakkelijke vragen, ten opzichte van de **moeilijkheidsgraad** van de leerdoelen? Kies er dan voor om de cesuur iets hoger te leggen. De studenten moeten dan meer van die gemakkelijke vragen goed beantwoorden voor een voldoende.
Bestaat het examen juist uit veel vragen die relatief moeilijk zijn? Kies er dan voor om de cesuur iets lager te leggen.
3. Zijn er veel gesloten vragen opgenomen waarbij de student door **gokken** een gedeelte van de punten al kan halen? Indien ja: leg de cesuur hoger dan je bij een toets met vooral open vragen zou doen. De gokkans is vaak eenvoudig te berekenen en als je werkt met TestVision wordt de gokkans automatisch berekend.

De absolute cesuur heeft een aantal voordelen:

- Je kunt goed toetsen of de studenten het vereiste minimumniveau beheersen.
- Studenten hebben vooraf duidelijkheid over het aantal punten dat ze moeten behalen voor een voldoende.

Maar, denk wel elke keer opnieuw na over welke absolute cesuur bij je toets past! Te vaak zien we dus nog dat 55% van de punten als standaard zak/slaag-grens wordt gehanteerd.

Er zijn verschillende methodes om de absolute cesuur te berekenen. Zo is er de **methode Ebel**, waarbij je de cesuur bepaalt aan de hand van de relevantie en de moeilijkheid van elke toetsvraag.

Aan de hand van deze kenmerken kun je berekenen hoeveel vragen de studenten goed zouden moeten beantwoorden voor een voldoende. Deze methode kost relatief veel tijd.

Een andere methode is de **methode Nedelsky**. Deze methode is alleen geschikt bij toetsen met één-uit-meer vragen (de standaard multiple-choice toetsen). In deze methode moeten experts bij iedere vraag bedenken tussen welke alternatieven een student die de stof nét voldoende beheerst zou gokken (en welke alternatieven de student meteen zou afstrepen). Het aantal punten dat de gokkende student op deze manier gemiddeld zal halen (de gokkans voor deze student), wordt als cesuur genomen.

Tenslotte is er de **methode Angoff**. Hierbij schatten experts hoeveel procent van de studenten die de stof nét voldoende beheersen, een vraag goed zal beantwoorden. Het gemiddelde van deze percentages bij alle vragen, geeft de cesuur. Een voorbeeld hiervan zie je in de tabel hieronder.

Item	Expert			Gemiddelde
	1	2	3	
1	70%	60%	70%	67%
2	90%	50%	80%	73%
3	80%	80%	70%	77%
4	70%	80%	60%	70%
Cesuur				71,75

Afbeelding: Voorbeeld van de Angoff-methode (uit Teelen Kennismanagement, 2015)

De relatieve cesuur

In de praktijk wordt de relatieve cesuur minder vaak gebruikt dan de absolute cesuur. In sommige gevallen is het echter juist goed om wél met een relatieve cesuur te werken. Soms is het verstandig om pas achteraf te bepalen bij welke score studenten zakken of slagen. De moeilijkheidsgraad van de toets en de kwaliteit van het onderwijs zijn bijvoorbeeld niet goed in te schatten, waardoor deskundigen het moeilijk vinden een absolute cesuur te bepalen. Als een opleiding bijvoorbeeld net gestart is (met nieuwe vakken en leermiddelen, nieuwe lessen en nieuwe toetsen) kan het lastig zijn om voor elke toets vooraf te bepalen hoeveel punten studenten moeten halen om een voldoende te krijgen. In zo'n geval helpt het om achteraf de scores van de studenten te bekijken en dan pas te bepalen waar de cesuur ligt. Een andere reden om te kiezen voor een relatieve cesuur, is de selectieve functie van een toets. Dit is bijvoorbeeld het geval als de beste dertig studenten geselecteerd moeten worden voor een opleiding met een beperkt aantal beschikbare plaatsen of als een groep zwakke studenten geselecteerd moeten worden voor een bijspijkerkursus.

Er zijn verschillende manieren om de relatieve cesuur te berekenen.

Op basis van slagingspercentages

Bij deze cesuurmethode wordt voor afname van het examen een slagingspercentage bepaald. Dat is het percentage van de studenten dat een voldoende moet krijgen. Dit kan bijvoorbeeld 75 % van de studenten zijn. De zak/slaag-grens komt te liggen bij de student die als laatste binnen de groep van 75% valt. De afbeelding maakt de methode duidelijk:

Aan een examen hebben 12 kandidaten deelgenomen. Ze konden in het examen maximaal 30 punten halen.



Volgens de relatieve cesuur moet 75% van de kandidaten slagen. Dat zijn er 9 van de 12. De score van de vierde persoon is dan de zak-/slaaggrens: 8 punten.

De methode Wijnen

Bij deze methode wordt de gemiddelde score van de studenten als uitgangspunt genomen. Wijnen veronderstelt dat de gemiddelde student de leerdoelen voldoende beheerst en dus zou moeten slagen. Van de gemiddelde score wordt tweemaal de standaardmeetfout van het groepsgemiddelde afgetrokken om de cesuur te bepalen. Bijvoorbeeld: de gemiddelde score op de hiervoor beschreven toets is 14,1 en de standaardmeetfout is 1,8. Volgens de methode Wijnen wordt de cesuur dan vastgesteld op 10,5 ($14 - 2 * 1,8$). Een belangrijk voordeel van de methode Wijnen is dat de betrouwbaarheid van de toets meeweegt in de cesuur. Als een toets erg onbetrouwbaar is, valt de cesuur lager uit dan bij een betrouwbare toets.

Aan de relatieve cesuur kleven een aantal nadelen:

- Als er veel zwakke studenten zijn, verhoogt dit de slaagkans van andere studenten (die wellicht ook de leerdoelen maar matig beheersen).
- De kwaliteit van de toets en het onderwijs worden niet zichtbaar in de toetsuitslag.
- De relatieve cesuur mag niet zonder meer gehanteerd worden bij kleine groepen en bij herkansingsgroepen.
- Bij groepen waarbij de toetsscores niet normaal verdeeld zijn (bijvoorbeeld omdat er veel studenten het maximaal aantal te behalen punten haalde), is het statistisch gezien minder verantwoord om met een relatieve cesuur te werken.

Beide methodes combineren

Elke methoden heeft dus voor- en nadelen. Zo geeft een absolute cesuur al van te voren duidelijkheid, maar is ze relatief bewerkelijk vast te stellen. Bij een relatieve cesuur kun je de moeilijkheid van de toets meewegen, maar loop je het risico dat de cesuur erg laag uitvalt. Hanteren van een combinatiemethode heft dit laatste nadeel op.

Een scoregebied hanteren voor beide cesuren met een absolute ondergrens

Een eerste manier is om vooraf twee gebieden voor beide cesuren te bepalen. Je stelt dan bijvoorbeeld dat de cesuur (absoluut) tussen 55% en 65% van de punten moet liggen en het slaagpercentage tussen 60% en 70%. Aan de hand van de scores van de studenten bepaal je achteraf de daadwerkelijke cesuur. Bij een makkelijke toets, zal de cesuur dichterbij de 65% komen te liggen, bij een moeilijke toets dichterbij de 55%. 55% blijft de absolute ondergrens. Het is dus mogelijk dat toch minder dan 60% van de studenten slaagt voor de toets.

De absolute cesuur combineren met een relatieve onder- en/of bovengrens

Een tweede manier is om een absolute cesuur te hanteren, maar een bovengrens en/of een ondergrens te hanteren wat betreft de slagingspercentages. De absolute cesuur is bijvoorbeeld 65% van de punten. Het slagingspercentage mag echter niet lager zijn dan 55% en/of niet hoger dan 90%.

Je weet dan zeker dat minimaal 55% en maximaal 90% van de studenten zal slagen voor deze toets. De absolute cesuur zal daarvoor naar boven of beneden bijgesteld kunnen worden.

De cesuurmethode Cohen-Schotanus

Bij de methode Cohen-Schotanus bepaal je vooraf een absolute cesuur. Normaal gesproken krijgt een student de maximale score (een 10) als hij de (theoretisch) hoogst haalbare score heeft behaald (100% van de punten). De methode van Cohen-Schotanus gaat er echter vanuit dat de score van de beste (vijf of tien) student(en) de hoogst haalbare score is. De cesuur wordt vervolgens berekend aan de hand van die hoogste haalbare score.

Voorbeeld

Een toets bestaat uit 50 vragen. De absolute cesuur is 60%. Elke vraag kan één punt opleveren. Het is theoretisch dus mogelijk om 50 punten te halen. Bij een normale cesuur krijgt een student pas een 10 als hij/zij alle 50 punten haalt. Bij 30 punten krijgt hij/zij een voldoende (5,5).

Na afname van het examen bij 300 studenten blijkt dat de beste vijf studenten gemiddeld 43 punten haalden. Volgens de methode Cohen-Schotanus moet de cesuur zo aangepast worden, dat 43 punten de maximale score oplevert. Als gevolg hoeft een student niet 60% van 50 punten te halen voor een voldoende, maar 60% van 43 punten. De daadwerkelijke cesuur ligt dan op 26 punten (in plaats van de oorspronkelijke 30).

Zo wordt bij deze cesuurmethode rekening gehouden met de moeilijkheid van de toets. Als de groep echter niet homogeen is en de beste studenten bijvoorbeeld meer voorkennis hebben dan de rest van de groep, dan zal de cesuur vaak alsnog hoger liggen dan wanneer de methode Wijnen gebruikt wordt.

Tot slot

Bij veel examens wordt achteraf nog een controle gedaan op de cesuur, waarbij deze nog kan worden bijgesteld. Dit is bijvoorbeeld het geval bij de centrale schoolexamens. Als blijkt dat een examen moeilijker was dan van tevoren verwacht, wordt de cesuur met de zogenaamde n-term gecorrigeerd. Deze variabele wordt na afloop van de examens voor elk vak apart berekend. Op die manier wordt de absolute cesuur dus achteraf, relatief, bijgesteld.

Literatuur

Angoff, W. H. (1971). Scales, norms and equivalent scores. In: R.L. Thorndike (red.), Educational Measurement. Washington D.C.: American Council on Education.

Cohen-Schotanus, J., Vleuten, C.P.M. van der, & Bender, W. (1996). *Een betere cesuur bij tentamens. Onderzoek van onderwijs*, 25, 54-55.

Ebel, R.L. (1972). *Essentials of Educational Measurement*. Englewood Cliffs, N.J.: Prentice Hall.

Teelen Kennismanagement (2015). *Toetskwaliteit in de praktijk. Hoe maak ik goede toetsen met gesloten en open vragen?* Wilp: Teelen Kennismanagement

Wijnen, W.H.F.W. (1971). *Onder of boven de maat; een methode voor het bepalen van de grens voldoende/onvoldoende bij studietoetsen*. Lisse: Swets & Zeitlinger.