

Zijn meisjes beter in taal?

Vaak wordt beweerd dat meisjes beter in taal zijn dan jongens. Maar hoe kun je dat onderzoeken? Een bekende methode is om na een toets het gemiddelde van de jongens te vergelijken met dat van de meisjes. Maar er zijn ook andere methoden. In deze opgave bekijken we een manier van vergelijken die is opgesteld door Frank Wilcoxon. Hij kijkt vooral naar de rangschikking van de jongens en de meisjes.

Vier meisjes en drie jongens

Bij een taaltoets die door vier meisjes en drie jongens wordt gemaakt, zijn maximaal 250 punten te behalen. De meisjes halen de scores 162, 217, 231 en 240 en de jongens halen de scores 119, 145 en 179.

Je kunt nu deze scores rangschikken van laag naar hoog en daarbij noteren of het om een jongen (J) of een meisje (M) gaat. Er ontstaat zo een rij met driemaal een J en viermaal een M. In dit geval krijg je de rij J J M J M M M.

- 3p **16** Bereken hoeveel verschillende rijen er mogelijk zijn met driemaal een J en viermaal een M.

Van elke rij kunnen we de zogenoemde *U*-waarde berekenen. Dit gaat als volgt: tel voor elke M het aantal J'tjes dat ervoor staat en tel daarna alle uitkomsten bij elkaar op.

De *U*-waarde van het rijtje J J M J M M M is dus $2+3+3+3=11$. Deze *U*-waarde is hoog want op één na scoren alle meisjes hoger dan alle jongens.

De *U*-waarde geeft aan in hoeverre de scores van jongens en meisjes verschillen. Dat betekent:

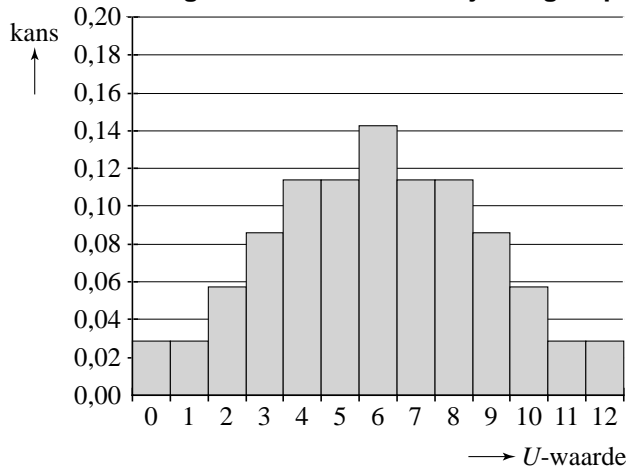
- *U*-waarde hoog: de meisjes scoren hoog en de jongens laag;
- *U*-waarde laag: de jongens scoren hoog en de meisjes laag;
- *U*-waarde middelmatig: er is weinig verschil tussen meisjes en jongens.

- 3p **17** Geef een voorbeeld van een rij bestaande uit driemaal een J en viermaal een M met *U*-waarde 5.

Door bij alle mogelijke rijen de U -waarde te bepalen kun je de kansverdeling van U opstellen. Daarbij ga je er dus van uit dat alle rangschikkingen even waarschijnlijk zijn. Deze kansverdeling is weergegeven in de figuur.

figuur

kansverdeling van de U -waarde bij een groep van vier meisjes en drie jongens



Bij grotere groepen gaat dit kanshistogram op de grafiek van een normale verdeling lijken. De U -waarde is dan bij benadering normaal verdeeld met:

$$\text{gemiddelde} = 0,5 \cdot n_m \cdot n_j$$

$$\text{standaardafwijking} = \sqrt{\frac{n_m \cdot n_j \cdot (n_m + n_j + 1)}{12}}$$

Hierin is n_m het aantal meisjes en n_j het aantal jongens.

Een havo 5-groep met 70 meisjes en 75 jongens maakt de taaltoets. De U -waarde die je mag verwachten als jongens en meisjes ongeveer even goed scoren, zal rond het gemiddelde liggen.

- 4p **18** Bereken de kans dat de U -waarde tussen 2400 en 2800 ligt, uitgaande van de hierboven genoemde formules.

Als de U -waarde van een groep niet in de buurt van het gemiddelde ligt maar veel hoger is, kun je zeggen dat de meisjes in die groep beter in taal zijn dan de jongens in die groep. Maar wat is precies 'veel hoger'?

Wilcoxon trekt de grens als volgt: als de U -waarde bij de hoogste 5% zit, uitgaande van bovenstaande normale verdeling, dan wordt besloten dat de meisjes beter in taal zijn dan de jongens.

Na verwerking van de scores van genoemde havo 5-groep blijkt de U -waarde gelijk te zijn aan 2984.

- 4p **19** Onderzoek of voor deze havo 5-groep besloten wordt dat de meisjes beter in taal zijn dan de jongens.