

Spiekbriefjes bij Statistiek

Centrummaten

Er zijn drie getallen die een hoeveelheid gegevens (waarnemingen) kunnen samenvatten:

- de **modus** (de modale waarde) is de waarneming die het vaakst voorkomt en dus de grootste frequentie heeft. Er is geen modus als twee waarnemingen het meest voorkomen.
- de **mediaan** is het midden van alle waarnemingen als die op volgorde staan.
- het **gemiddelde** van de waarnemingen vind je door alle waarden bij elkaar op te tellen en dat getal te delen door het totaal aantal cijfers. Daarbij moet je rekening houden met de frequenties (de wegingen) van de waarnemingen.

De modus, de mediaan en het gemiddelde zijn **centrummaten**. Deze drie getallen kunnen verschillend zijn.

Je kunt de mediaan en het gemiddelde alleen bepalen als je getallen als waarneming hebt.



meer info

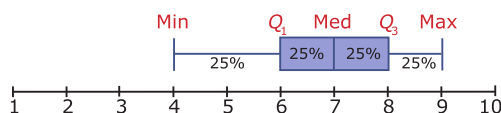
Spreidingsmaten

Je kunt van een hoeveelheid gegevens een samenvatting maken van vijf getallen:

- de kleinste waarneming;
- het **eerste kwartiel** Q_1 , dit is de mediaan van de eerste helft waarnemingen;
- het **tweede kwartiel** is de mediaan;
- het **derde kwartiel** Q_3 , dit is de mediaan van de tweede helft waarnemingen;
- de grootste waarneming.

Met deze vijf getallen maak je een **boxplot** met de **spreidingsmaten**:

- de **spreidingsbreedte** is de grootste min de kleinste waarneming;
- de **(inter)kwartielfstand** is het derde min het eerste kwartiel.



meer info

Klassenindeling

Een groep gegevens (liefst getallen) noem je **ruwe data**, ruwe gegevens. Die ruwe data verdeel je vaak in **klassen**, groepjes getallen die dicht bij elkaar liggen. Kies eerst de **klassengrenzen**. Neem alle klassen even breed.

Het verschil van de klassengrenzen van een klasse heet de **klassenbreedte**.

Het **klassenmidden** is meestal het gemiddelde van de twee klassengrenzen.

Het aantal getallen in een bepaalde klasse is de **absolute frequentie**.

Vaak zijn **relatieve frequenties** handiger. Die bepaal je door de frequentie te delen door het totaal. Ook kun je het **percentage** berekenen, je vermenigvuldigt dan de relatieve frequentie met 100.

Door zo'n **klassenindeling** te maken, krijg je **geordende data**.

Het gemiddelde kun je nu alleen nog maar schatten met behulp van de klassenmiddens.



meer info

Schatten

Vanuit een klassenindeling kun je van de verdeling van de cijfers de centrummaten (modus, mediaan, gemiddelde) en de spreidingsmaten (spreidingsbreedte, kwartielafstand) niet meer precies berekenen.

In plaats daarvan gebruik je:

- de **modale klasse**, dat is de klasse met de hoogste frequentie;
- de klassenmiddens om het **gemiddelde te schatten**;
- het aantal klassen maal (keer) de klassenbreedte voor de **spreidingsbreedte of variatiebreedte**.

Om een **boxplot** te maken moet je de hoogste en de laagste waarde, kwartielen en de mediaan schatten.



meer info

Statistische uitspraken

Een **statistische uitspraak** over een groep is een uitspraak die is gebaseerd op gegevens van een deel van de groep. Soms is een groep zo groot dat je niet even ieder lid van de groep kunt bekijken/bevragen.

Een vraag over zo'n grote groep beantwoord je met behulp van **statistisch onderzoek**.

Je bekijkt/bevraagt een beperkt aantal leden van de groep en verzamelt daarvan gegevens. Je noemt dat een **steekproef** trekken uit de totale **populatie** (bijvoorbeeld alle Nederlanders).

Zo'n steekproef moet een goede vertegenwoordiging van de totale populatie zijn, **representatief** zijn voor de populatie. Pas dan kun je met een enige betrouwbaarheid zeggen dat het gevonden gegevens ook iets zeggen over de hele groep. Meer kun je met statistiek niet doen...



meer info