

1. Introductie

Introductie

Statistiek wordt in de huidige maatschappij overal gebruikt. Om statistische uitspraken te kunnen doen, moet er eerst data verzameld worden. Data zijn verzamelde observaties van een kenmerk van belang. Bijvoorbeeld de mening van 1000 mensen over marihuana gebruik. Zulke data kan gegenereerd worden aan de hand van o.a. enquêtes, experimenten en observaties.

De term statistiek verwijst vervolgens naar methoden voor het verkrijgen en analyseren van data. Aan de hand van een 'design' wordt een planning gemaakt hoe data wordt verzameld. Vervolgens zijn er twee vormen van statistische analyse, te weten beschrijvende statistiek (samenvatten) en inferentiële statistiek (voorspellen). Afhankelijk van het doel wordt bepaald van welke vorm van statistiek er sprake is: als er wordt samengevat is het beschrijvend, als er voorspellingen worden gedaan is het inferentieel.

Beschrijvende en inferentiële statistiek

Eerst zijn een aantal basisbegrippen belangrijk, namelijk de populatie en de steekproef. De populatie (N) is de totale set van subjecten die je wilt onderzoeken (bijv. de Nederlandse bevolking). De steekproef (n) is een subset die je hieruit hebt getrokken waarover je daadwerkelijk data gaat verzamelen (bijv. 1000 Nederlanders uit verschillende provincies). Het uiteindelijke doel is iets te leren over de populatie, maar omdat het onmogelijk is om de gehele populatie te onderzoeken, trekt men een steekproef. Aan de hand van beschrijvende statistiek kan de data van de steekproef worden samengevat (bijv. gemiddelden, mediaan, modus). Aan de hand van inferentiële statistiek kan worden onderzocht in hoeverre de informatie uit de dataset de gehele populatie weerspiegelt.

Wanneer er wordt gekeken naar de statistieken van een steekproef en de populatie maakt men onderscheid tussen parameters en statistieken. Een statistiek is feitelijk een numerieke samenvatting van de steekproefdata. Een parameter is de numerieke samenvatting van de populatie. Wanneer men data verzamelt, worden er dus statistieken verzameld (steekproef). Hiermee wil men uiteindelijk iets gaan zeggen over de parameters (populatie).