

Hoofdstuk 7

BS HC 8B – 4 november

Bij een enkelvoudige regressieanalyse moet als een dichotome variabele in een regressieanalyse wordt opgenomen deze als 0 en 1 zijn gecodeerd. De 0-categorie is de vergelijkingscategorie. De intercept (a) heeft hier betekenis: Als iemand in de 0-categorie valt, scoort diegene *gemiddeld*... De intercept heeft dus alleen betekenis als je een dichotome variabele hebt gebruikt. De regressiecoëfficiënt (b) is het *gemiddelde verschil* ten opzichte van de 0-categorie.

Standaardiseren is een variabele ontdoen van zijn meeteenheid. Om dit te doen gebruik je z-scores. De **gestandaardiseerde regressiecoëfficiënt** noemen we Beta en duiden we aan met b^* . De ongestandaardiseerde regressiecoëfficiënt (b) is afhankelijk van hoe we x meten. We kunnen leeftijd bijvoorbeeld in jaren en in maanden meten. Dit is onhandig als je meerdere variabelen wilt vergelijken. Je kunt de ongestandaardiseerde regressiecoëfficiënt zo omzetten dat deze gestandaardiseerd wordt. Omdat b dan gestandaardiseerd is tot Beta, neemt Beta bijna altijd een waarde aan tussen de -1 en de 1, zodat we hem kunnen interpreteren als een associatiemaat. Beta is het zuivere effect van x op y.

Bij een enkelvoudige regressieanalyse zijn de Beta en de R aan elkaar gelijk. Je kunt dus gewoon de wortel trekken van R^2 . Er wordt alleen geen rekening gehouden met + en -, want als je de wortel van iets trekt krijg je altijd een positief getal, terwijl het ook een negatief getal had kunnen zijn. Je kunt aan R alleen de sterkte van het verband aflezen, niet de richting. Dus: bij een enkelvoudige regressieanalyse (= één onafhankelijke variabele) is $|r| = R = |Beta|$. Bij een enkelvoudige regressie kan de richting alleen worden afgelezen door naar Beta te kijken.

Een meervoudige regressieanalyse is als je niet één maar meerdere onafhankelijke variabelen hebt. Er is dan nog steeds maar één intercept, maar meerdere lijnen met eigen ongestandaardiseerde regressiecoëfficiënten.

$$\hat{y} = a + b_1x_1 + b_2x_2 + b_3x_3$$

b_1 geeft het effect van x_1 op y weer onder constant houding van x_2 en x_3 . Beta geeft het zuivere effect op y aan. Aan Beta kun je zien welke variabele het sterkste effect heeft op de afhankelijke variabele. Het maakt dan niet uit of het effect positief of negatief is, het gaat om de sterkte. Bij beschrijving van een dichotome variabele gebruik je de woorden 'gemiddeld'. Bij R^2 zie je hoe veel de variantie in de afhankelijke variabele verklaard wordt door de variantie in de onafhankelijke variabelen samen.

Bij een practicumvraag zijn de eerste vier stappen. Stap 1: hoeveel variabelen zijn er? Stap 2: wat is het meetniveau van de variabelen? Stap 3: bewerken van je gegevens. Je haalt bijvoorbeeld 'missing' of 'niet van toepassing' als antwoordmogelijkheden uit de gegevens. Stap 4: Bedenk welke associatiemaat het meest geschikt is. Dan ga je pas verder met de vragen. Aan het eind van de vraag lever je je Syntax in. Bij alles wat je doet, schrijf je in Syntax een toelichting. Begin de tekst met een * en eindig met een punt! Als je dit goed doet, wordt de toelichting grijs.