

12. ANCOVA

Inhoud

In dit hoofdstuk wordt de *analyse van covariantie* (ANCOVA) besproken. Het principe is hetzelfde als ANOVA, alleen wordt er hier nog een variabele bij betrokken die ook invloed heeft op de uitkomstvariabele.

ANCOVA

De regressievergelijking van een ANOVA kan worden uitgebreid met continue variabelen die de uitkomstvariabele ook voorspellen. Deze variabelen, die geen deel zijn van de experimentele manipulatie, worden *covariaten* genoemd. Zodra we de covariaten in de ANOVA analyse betrekken, wordt het een ANCOVA. Door covariaten mee te nemen kun je corrigeren voor de invloed die ze op de afhankelijke variabele hebben.

De covariaten worden als eerst ingevoerd zodat je kunt kijken wat voor effect de onafhankelijke variabele heeft, nadat het effect van de covariaat al bekend is. Er zijn twee redenen om covariaten in de ANOVA mee te nemen:

- Het verminderen errorvariantie binnen de groepen. Door covariaten kunnen we de onverklaarde variantie verminderen en daardoor nauwkeuriger bepalen wat het effect van de onafhankelijke variabele is.
- Het elimineren van vertekeningen. Bij elk experiment kunnen er variabelen zijn die de resultaten vertekenen doordat ze ook de uitkomstvariabele beïnvloeden. Door deze variabele als covariaat mee te nemen, kan deze vertekening worden weggehaald.

Door covariaten toe te voegen komt de regressievergelijking in het voorbeeld van de invloed van Viagra op libido er zo uit te zien:

Afhankelijke variabele $i = b_0 + b_3 \text{ covariaat}_i + b_2 \text{ Hoge dosis} + b_1 \text{ Lage dosis} + \text{error}$

Dit model test het verschil tussen groepsgemiddelden, nadat er gecorrigeerd is voor de covariaat.

ANCOVA en het lineaire model

ANCOVA kan ook gedaan worden als hiërarchische multiple regressie. De output bij deze regressie laat eerst de goodness of fit zien en vervolgens hoe de ANOVA eruit ziet zonder en met covariaat. De ANOVA tabel is in twee delen gesplitst. Het bovenste deel is het effect van alleen het covariaat en het onderste deel laat het hele model zien (zie blz. 482). ANCOVA hoeft niet op deze manier gedaan te worden en beide analyses uitvoeren is ook niet nodig. Maar als je graag de contrasten wilt weten is het regressiemenu handiger. Contrasten kunnen dan gevonden worden door een ANOVA in de regressie te doen en de variabelen als dummy's te coderen.

Assumpties

ANCOVA heeft dezelfde assumpties als ANOVA met nog twee aandachtspunten. Ten eerste moet de covariaat onafhankelijk zijn van het experimentele effect. De ideale situatie is wanneer het covariaat volledig in het gedeelte van de onverklaarde variantie zit. Dit is niet altijd het geval en soms zit het tussen de verklaarde en onverklaarde variantie in. Daardoor wordt het experimentele effect kleiner omdat een deel van het effect verklaard wordt door de covariaat. In dit geval moet geen ANCOVA gebruikt worden.

Wanneer de experimentele groepen van elkaar verschillen op een covariaat, zorgt het meenemen van deze covariaat in de analyse er niet voor dat er gecontroleerd wordt voor deze verschillen. kan je niet altijd een covariaat toepassen. De covariaat en de onafhankelijke variabele hebben dan gedeelde variantie, die je niet kunt opdelen in een covariaat deel en een deel van de onafhankelijke variabele. Je kan dit probleem voorkomen door personen random in groepen in te delen.

Een ander aandachtspunt bij de ANCOVA is de *homogeniteit van de regressiehellingen*. Bij een ANCOVA kijk je naar de hele relatie tussen de afhankelijke variabele en de covariaat, waarbij je negeert uit welke groep iemand komt. Daarom moet deze relatie tussen de uitkomstvariabele en de covariaat hetzelfde zijn in alle groepen. Deze assumptie kan het best bekeken worden door een scatterplot te maken van elke groep met de covariaat op de ene as en de uitkomstvariabele op de andere as. Als de regressielijnen bij deze groepen allemaal ongeveer hetzelfde zijn (oftewel, dezelfde b-waardes), is voldaan aan deze assumptie.

Wanneer niet aan deze assumptie voldaan wordt, kan je met SPSS gebruik maken van bootstrap voor de model parameters en de post hoc tests, maar niet bij de belangrijkste analyse, de F-statistiek. Dit kan wel met het programma R.

ANCOVA in SPSS

Eerst moet gekeken worden of de covariaat onafhankelijk is van de experimentele manipulatie. Dit kan door eerst een ANOVA uit te voeren, met de covariaat als uitkomstvariabele en de manipulatie als onafhankelijke variabele. Wanneer het hoofdeffect niet significant is, kan de covariaat in de analyses toegepast worden.

Voor ANCOVA ga je naar Analyze – General Linear Model – Univariate. Zodra een covariaat geselecteerd is kunnen er geen post hoc testen meer worden gedaan. Met contrasten kunnen nog wel een aantal standaard vergelijkingen worden gedaan (zie blz. 491).

Bij opties kan je toch een aantal post hoc tests kiezen. Dit doe je door de onafhankelijke variabele naar het veld Display Means for te slepen, en dan te klikken op Compare main effects. Bij confidence interval adjustment kan je kiezen voor LSD (none), Bonferroni (aan te raden) en Sidak. De *Sidak correctie* kan je kiezen als minder verlies van power wilt dan bij de bonferroni. Bij opties kan je nog meer aanvinken:

- Descriptive statistics. Geeft een tabel met gemiddelden en standaardafwijkingen.
- Estimates of effect size: Dit geeft de gedeelde η^2 weer.
- Observed power. Dit geeft een powermeting, maar het is niet echt nuttig.
- Parameter estimates: Deze optie laat de regressiecoëfficiënten zien en de significantie van de variabelen in het model.

- Contrast coefficient matrix: Dit is bruikbaar wanneer groepen vergeleken worden in een contrast.
- Homogeneity test. Produceert Levene's test.
- Spread vs. Level plot. Geeft een grafiek waarbij elk groepsgemiddelde van de factor tegen de standaardafwijking van die groep wordt geplott.
- Residual plot. Dit plot kan gebruikt worden om de assumptie van homoscedasticiteit te controleren.

Output

In de output is eerst een tabel van Levene's test te zien en daarna een ANOVA tabel. De tabel is niet heel anders met een covariaat dan bij een gewone ANOVA. De covariaat zorgt voor een extra regel in de tabel, waarbij je kunt kijken of de covariaat een significante voorspeller is.

Het meenemen van een covariaat in de analyses betekent dat de gemiddelden van de groepen worden aangepast, gecorrigeerd voor de covariaat. Daarom moet je kijken naar de aangepaste gemiddelden, een belangrijke reden om de Display Means for optie te gebruiken.

Een positieve b coëfficiënt van de covariaat betekent een positieve relatie tussen de covariaat en de uitkomst. De grootte geeft aan met hoeveel eenheden de uitkomstvariabele stijgt of daalt bij een verandering van de covariaat. Dit verband kun je ook zien als je een scatterplot maakt, met de covariaat op de ene as en de afhankelijke variabele op de andere as.

De assumptie van homogeniteit tussen de hellingshoeken

Bij deze assumptie nemen we aan dat de relatie tussen de covariaat en de uitkomstvariabele voor elke experimentele groep hetzelfde is. In een scatterplot zou dat betekenen dat de regressielijnen parallel lopen.

Om de assumptie te testen doen we een ANCOVA maar dan met een aangepast model. Eerst ga je naar de gewone ANCOVA zoals in eerder in dit hoofdstuk is uitgelegd. Daarna klik je op model en krijg je het scherm zoals op blz. 501. Normaal gesproken test de ANCOVA alleen het hoofdeffect van de onafhankelijke variabele en de covariaat. Nu wil je echter de interactie tussen de covariaat en de onafhankelijke variabele testen. De hoofdeffecten moeten nog wel allebei worden meegenomen.

Voor een aangepast model vink je Custom aan en plaats je de covariaat en de onafhankelijke variabele in het vak Model, terwijl je Main effects aangevinkt hebt bij Type. Daarna selecteer je de covariaat en de onafhankelijke variabele tegelijkertijd, klik je bij Type op Interaction, en plaats je zo de interactieterm in het vak Model. Wanneer de interactieterm significant is, is de assumptie van homogeniteit van de slopes geschonden.

De effectgrootte

Bij ANCOVA kunnen we voor elk effect een η^2 uitrekenen. De partiële η^2 kan ook uitgerekend worden, die kijkt naar de proportie variantie die een variabele verklaart, dat niet verklaard wordt door de andere variabelen in de analyse.

$$\eta^2 = \frac{SS_{\text{Effect}}}{SS_{\text{Totaal}}}$$

$$\text{Gedeelde } \eta^2 = \frac{SS_{\text{Effect}}}{SS_{\text{Effect}} + SS_{\text{Residu}}}$$

ω^2 kan ook gebruikt worden, maar alleen bij gelijke groepsgroottes. Met een t-toets kan de effectgrootte van de contrasten berekend worden:

$$r_{\text{contrast}} = \sqrt{\frac{t^2}{t^2 + df}}$$

Het rapporteren van een ANCOVA is hetzelfde als een ANOVA. Bij ANCOVA moet alleen het effect van het covariaat er ook bij vermeldt worden.