

Practicum 2: Antwoorden



Dataset

We bouwen voort op de dataset van vorige week. Download hiervoor de dataset van deze week op Canvas:
CWIDM_practicum_2_dataset.csv

Variabelen

SuWB, Prestatietevredenheid, Lichaamstevredenheid, Extraversie, Neuroticism.

7-punts Likertschaal

Een hogere score staat voor een hogere mate van subjectief welzijn, prestatietevredenheid, lichaamstevredenheid, extraversie en neuroticism.

Zelfver

11-punts Likertschaal

Een hogere score staat voor een hogere mate van zelfvertrouwen.

Dummy typeblootstelling

Experimentele conditie waarbij de helft van de participanten zijn blootgesteld aan conditie (0) en de andere helft van de participanten aan conditie (1)

0 = 'Body positivity Instagrampost'

1 = 'Ideaal lichaamstype Instagrampost'

Doel practicum: lineaire regressie met mediatie

Nadat we vorige week factoranalyses en betrouwbaarheidsanalyses hebben uitgevoerd gaan we in dit practicum beginnen met de daadwerkelijke analyses. Hierbij komen we dezelfde variabelen tegen die we zelf vorige week hebben aangemaakt met daarbovenop een aantal nieuwe variabelen. In dit practicum zullen we terugblikken op het uitvoeren van een lineaire regressieanalyse (zoals jullie dit ook bij BIS hebben geleerd) en zullen we een stap verder gaan door een lineaire regressieanalyse met mediatie uit te voeren. Hiervoor gebruiken wij zowel de Baron en Kenny methode als bootstrapping middels Process.

Opdracht 1: Voor beschrijvende en verkennende analyses uit bij deze steekproef.

Voordat we de daadwerkelijke data-analyse beginnen is het goed om verkennende analyses uit te voeren om een beter idee te krijgen van onze dataset. Voor een frequentie-analyse uit en maak een correlatiematrix. Zorg dat je hiermee de volgende vragen kan beantwoorden. Rond af op 1 decimaal.¹

A) Wat is de gemiddelde extraversiescore?

Antwoord: $M = 3.8$

B) Wat is de standaarddeviatie van lichaamstevredenheid?

Antwoord: $SD = 1.1$

C) Hoeveel participanten hebben niet deelgenomen aan het experimentele gedeelte van dit onderzoek?

Antwoord: $N = 151$

D) Wat is de correlatie tussen neuroticisme en lichaamstevredenheid?

Antwoord: $r = -.56$

E) Wat is de geobserveerde p-waarde van de correlatie tussen prestatietevredenheid en subjectief welzijn?

Antwoord: $p < .001$

F) Hoe sterk is de correlatie tussen extraversie en zelfvertrouwen volgens de vuistregel van Cohen?

Antwoord: Sterk. $r = .57$ (0.1 = zwak, 0.3 = middelmatig, 0.5 = sterk)

¹ Bij een p-waarde rond je op 3 decimalen af. Anders vragen we je altijd op 2 decimalen af te ronden bij dit vak tenzij anders aangegeven.

Opdracht 2: Lineaire regressieanalyse.

We gaan nu onze eerste daadwerkelijke hypothesetoetsing uitvoeren. In dit geval willen we weten of iemands prestatietevredenheid en lichaamstevredenheid invloed hebben op het subjectieve welzijn dit individu. Daarbij willen we ook weten welk effect sterker is. Voer hiervoor zowel een ongestandaardiseerde als een gestandaardiseerde lineaire regressieanalyse uit².

- A) Wat is de ongestandaardiseerde regressiecoëfficiënt (B) van lichaamstevredenheid op het subjectieve welzijn?

Antwoord: $B = .34$

- a. Wat is de bijbehorende standaardfout?

Antwoord: $SE = .05$

- B) Wat is de gestandaardiseerde regressiecoëfficiënt (Beta) van prestatietevredenheid op het subjectieve welzijn?

Antwoord: $Beta = .43$

- a. Wat is de bijbehorende standaardfout?

Antwoord: $SE = .06$

- C) Welk van de twee onafhankelijke variabelen heeft het sterkste effect op het subjectieve welzijn van een individu?

Antwoord: **Levenstevredenheid**

- D) Verwacht je dat lichaamstevredenheid in de populatie een positief effect heeft op iemands subjectieve welzijn? Noteer in je antwoord in ieder geval de geobserveerde p-waarde en het geobserveerde betrouwbaarheidsinterval.

Antwoord: Ja, $t(.05) = 7.48$, $p < .001$, $CI[.30:.52]$. $B = .34$. De geobserveerde p-waarde ($p < .001$) is kleiner dan de kritieke p-waarde ($p = .05$). Het betrouwbaarheidsinterval bevat geen 0 en enkel positieve waarden ($CI[.30:.52]$). De B (en Beta) zijn positief ($B = .34$, $Beta = .41$).

² Bij het uitvoeren van een lineaire regressieanalyse zijn er allemaal assumpties waaraan voldaan moeten worden. Deze leren jullie in de master en hier hoeven jullie voor dit vak nog geen rekening mee te houden. Yay!

Opdracht 3: Multipele lineaire regressieanalyse.

Omdat betere uitspraken te kunnen doen over causale relaties willen we bij een vragenlijstonderzoek zo veel mogelijk variabelen toevoegen waarvan we verwachten dat deze een effect hebben op het subjectieve welzijn van een individu. In dit geval verwachten wij, als onderzoekers, dat ook de persoonlijkheidskenmerken neuroticisme en extraversie een effect hebben op het subjectieve welzijn van individuen. Voer nu een lineaire regressieanalyse³ uit met de 4 hiervoor genoemde onafhankelijke variabelen en subjectief welzijn als afhankelijke variabele. Zorg dat deze analyses overzichtelijk in één tabel worden weergegeven in R.

- A) Geef de uitkomst van het effect van neuroticisme op iemands subjectieve welzijn wanneer je hebt gecontroleerd voor de andere drie onafhankelijke variabelen. Noteer dit volgens de APA-richtlijnen. Noteer hierbij dus in ieder geval de geobserveerde t -waarde en p -waarde en noteer de ongestandaardiseerde regressiecoëfficiënt.

Antwoord: $t(.07) = -6.51, p < .001. B = -.47$

- B) Geef de resultaten van de algehele analyse aan volgens de APA-richtlijnen. Oftewel: geeft de testwaarden van het model waarin je alle onafhankelijke variabelen meeneemt met subjectief welzijn als afhankelijke variabele. Noteer hierbij dus in ieder geval de geobserveerde F -waarde en p -waarde.

Antwoord: $F(4, 200) = 62.16, p < .001$. Dit geeft aan dat het algehele model een goede voorspeller is van iemands subjectief welzijn in de populatie.

- C) Wat is de verklaarde variantie van model 2 (het model mét extraversie en neuroticisme)?

Antwoord: $R^2 = .55$ (55%)

- D) Is model 2 een significant betere voorspeller van iemands subjectieve welzijn dan model 1 (het model mét alleen lichaamstevredenheid en prestatietevredenheid)? Noteer in je antwoord ook de geobserveerde F -waarde en p -waarde waar je dat uit af kan lezen volgens de APA-richtlijnen.

Ja, $F(2, 200) = 36.68, p < .001$. Model 2 is een significant betere voorspeller van subjectief welzijn dan model 1.

³ Wanneer wij niet specificeren of je een gestandaardiseerde of ongestandaardiseerde regressieanalyse moet uitvoeren moet je in de basis kiezen voor een ongestandaardiseerde. Kies voor een gestandaardiseerde wanneer dit expliciet wordt gevraagd of wanneer wij vragen om de sterktes van verschillende variabelen met elkaar te vergelijken (en dus de Beta te rapporteren).

Opdracht 4: Lineaire regressie met een dichotome variabele.

Nu we weten op welke manier verschillende persoonskenmerken en verschillende vormen van tevredenheid samenhangen met iemands subjectieve welzijn willen we ook weten welke invloed externe factoren kunnen hebben op een onderdeel van iemands welbevinden. Daarom hebben de onderzoekers onder een deel van de steekproef een experiment uitgevoerd waarbij de participanten aselekt over de twee verschillende condities van Instagram blootstelling werden verdeeld. In één conditie kregen de participanten een post te zien met een model met een realistisch lichaamsbeeld (body positivity post) en in de andere conditie kregen de participanten een post te zien met een onrealistisch lichaamsbeeld (ideaal lichaamstype Instagram post). Omdat de onderzoekers verwachtten dat iemands subjectieve welzijn niet direct van een dergelijke post aangepast zal worden hebben ze een ietwat volatiel onderdeel van iemands subjectieve welbevinden geanalyseerd; iemands zelfvertrouwen. Voer nu een lineaire regressieanalyse uit waarbij deze experimentele conditie de onafhankelijke variabele is en zelfvertrouwen de afhankelijke variabele.

- A) Wat is het voorspelde zelfvertrouwen voor een participant in de ‘body positivity’ conditie?

Antwoord: Voorspeld zelfvertrouwen = 6.43

- B) Wat is het voorspelde subjectieve welzijn voor een participant in de ‘ideaal lichaamstype’ conditie.

Antwoord: Voorspeld zelfvertrouwen = 4.95

- C) Noteer de resultaten van deze analyse zo bondig maar volledig mogelijk. Volg hierbij de APA-richtlijnen.

Antwoord: Het type blootstelling heeft een significant effect op iemands zelfvertrouwen $t(.32) = -4.67, p < .001$. Wanneer iemand blootgesteld wordt aan een ‘body positivity’ Instagrampost wordt in de populatie verwacht dat deze een hoger zelfvertrouwen ($M = 6.43$) heeft dan wanneer deze wordt blootgesteld aan een ‘ideaal lichaamstype’ Instagrampost ($M = 4.95$)

Opdracht 5: Lineaire regressie met mediatie.

Wij, als onderzoekers, verwachten dat deze relatie tussen het type blootstelling en iemands zelfvertrouwen (deels) verklaard kan worden door verschillende vormen van iemands tevredenheid. Oftewel, lichaamstevredenheid en prestatietevredenheid zijn volgens onze verwachting mediërende variabelen in de relatie tussen type blootstelling en zelfvertrouwen. Middels verschillende lineaire regressies kunnen wij achterhalen of dit het geval is door gebruik te maken van de Baron en Kenny-methode⁴. Voer de verschillende lineaire regressieanalyses uit.

- A) Vul de ontbrekende waarden van de twee onderstaande conceptuele modellen in. Geef bij iedere regressie-coëfficiënt ook aan of deze significant is met behulp van een *.

Antwoord: Zie hieronder.

- B) Is er volgens de Baron en Kenny methode sprake van een mediatie bij lichaamstevredenheid als mediator? Leg uit!

Antwoord: Voor de Baron en Kenny methode moet c significant zijn, a en b moeten significant zijn en c' moet niet significant zijn (volledige mediatie) of kleiner zijn dan c (partiële mediatie). In dit geval is aan alle voorwaarden van een partiële mediatie voldaan.

- C) Is er volgens de Baron en Kenny methode sprake van een mediatie bij prestatietevredenheid als mediator? Leg uit!

Antwoord: Voor de Baron en Kenny methode moet c significant zijn, a en b moeten significant zijn en c' moet niet significant zijn (volledige mediatie) of kleiner zijn dan c (partiële mediatie). In dit geval is a niet significant en dus is niet aan alle voorwaarden van een mediatie voldaan.

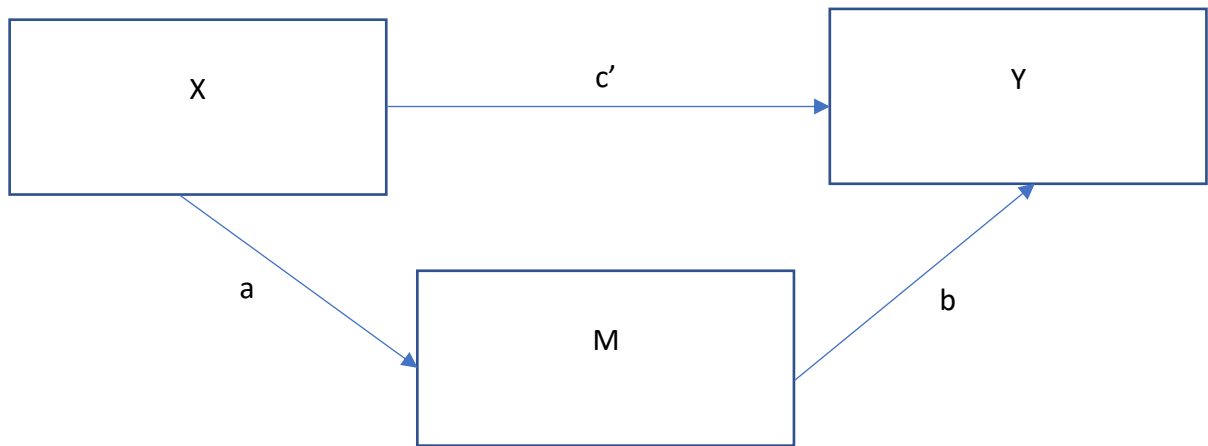
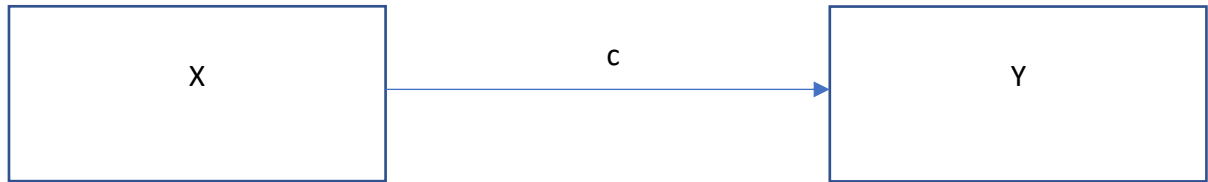
- D) Zou er ook sprake kunnen zijn van een mogelijk schijnverband in plaats van een mogelijk mediërend verband? Let uit!

Antwoord: Nee. Er is sprake van een schijnverband wanneer X geen invloed op M heeft, maar M invloed op X heeft. Bij een experimenteel design met aselechte verdeling over de condities kunnen we spreken van causaliteit en dus kan er geen sprake zijn van een achterliggend verband. Wanneer we een indirect effect vinden is dit dus ook daadwerkelijk een mediatie!

⁴ Bij de volgende vraag gaan we gebruik maken van Process maar bij deze vraag blijven we bij 'standaardregressies'.

Causaal model 1

Lichaamstevredenheid als mediërende variabele.



X = Type blootstelling

Y = Zelfvertrouwen

M = Lichaamstevredenheid

$c = -1.48^{***}$

$c' = -1.25^{***}$

$a = -.60^*$

$b = .39^{**}$

Indirecte effect ($a*b$) = $-.60 * .39 = -.23$

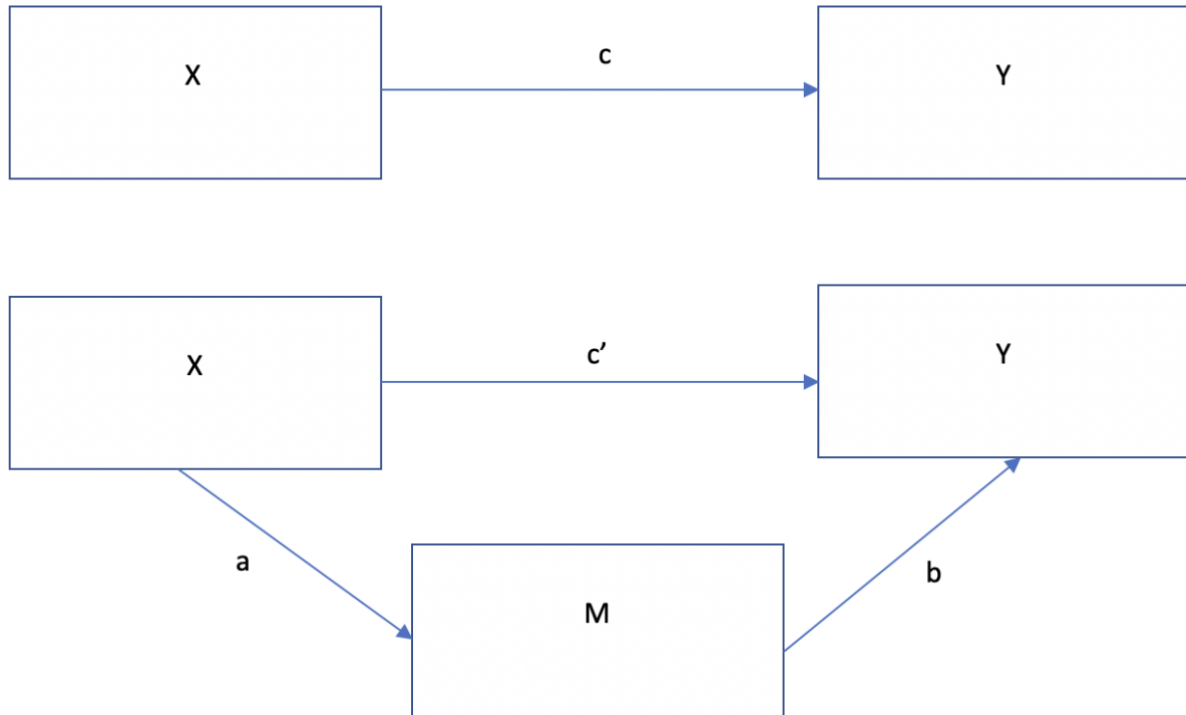
Indirecte effect ($c - c'$) = $-1.48 - 1.25 = -.23$

Directe effect = -1.25

Causale effect = -1.48

Causaal model 2

Prestatietevredenheid als mediërende variabele.



X = Type blootstelling

Y = Zelfvertrouwen

M = Prestatietevredenheid

$c = -1.56^{***}$

$c' = -1.41^{***}$

$a = -.25$

$b = .61^{**}$

Indirecte effect ($a*b$) = $-.25 * .61 = -.15$

Indirecte effect ($c - c'$) = $-1.56 - 1.41 = -.15$

Directe effect = -1.41

Causale effect = -1.56

Opdracht 6: Mediatie middels Process.

Hoewel de Baron en Kenny een mooi inzicht geeft in hoe een mediatie werkt is het niet de makkelijkste, meest complete en meest correcte methode om een mediatie uit te voeren. Zo weten we bijvoorbeeld niet of het algehele indirecte effect ($a*b$) significant is. Gelukkig is er een methode waarmee we op een makkelijke manier een mediatie kunnen uitvoeren en inzicht krijgen in de significantie van het indirecte effect: Process. Onderzoek nu middels Process of het effect van type blootstelling op zelfvertrouwen wordt verklaard door iemands levenstevredenheid.

A) Wat is de regressie coëfficiënt van type blootstelling op prestatietevredenheid?

Antwoord: $B = - .25$

B) Wat is de regressiecoëfficiënt van het directe effect van x op y? Is deze significant? Noteer hiervoor het geobserveerde betrouwbaarheidsinterval.

Antwoord: $B = - 1.41$. Dit effect is significant, $CI[-2.03 : -.81]$.

C) Wat is de regressiecoëfficiënt van x op m op y? Is deze significant? Noteer hiervoor de geobserveerde p-waarde.

Antwoord: $B = - .15$. Dit indirecte effect is niet significant, $p = .349$.

D) Wat is de regressiecoëfficiënt van het totale causale effect? Is deze significant? Noteer hiervoor de geobserveerde p-waarde en het betrouwbaarheidsinterval.

Antwoord: $B = - 1.56$. Dit totale effect is significant, $p < .001$, $CI[-2.30 : - 1.03]$.

E) Hoe groot is de verklaarde variantie van het model waarin levenstevredenheid als afhankelijke variabele fungeerde. Geef je antwoord in een percentage.

Antwoord: $R^2 = 2\%$

Opdracht 7: Los het volgende raadsel op ;)⁵

Als Suzans' dochter mijn dochters moeder is. Wat ben ik van Suzan?

- A. Grootmoeder
- B. Moeder
- C. Dochter
- D. Kleindochter
- E. Ik ben Suzan

⁵ Niet verplicht maar laat wel zien dat, hoe lastig statistiek ook is, er ergere dingen zijn om op te lossen. ☺