**Samenvatting TOE DT1**

**Carr H7 online**

7.1 – Surveys

Surveys kunnen op twee manieren wordne verwerkt:

1. **Primary data collections:** Onderzoekers nemen af/ontwerpen eigen surveys
2. **Secondary data source:** surveys die uitgevoerd zijn door anderen

De meeste surveys gaan over:

1. Leven van individuen
2. Gedachten
3. Gedrag

Respondenten antwoorden over zichzelf. Ze kunnen ook gebruikt worden om informatie over sociale instituties te vergaren. Hier wordt een **key informant** uitgezocht, die antwoorden geeft namens de instituties.

7.2 – Survey formats

Surveys komen voor in verschillende vormen, bijvoorbeeld telefonisch of via een **self-administered questionnaire (SAQ)**. Al deze **modes of administration** hebben hun eigen voor- en nadelen.

Er zijn twee soorten vragen:

* **Closed-ended:** respondenten moeten een antwoord kiezen van een lijst vooraf ingestelde **responscategorieën** → worden vaak gebruikt in surveys
* **Open-ended:** erg breed → worden vaak gebruikt in diepte-interviews.

7.3 – Timing van de survey

Er zijn **single-observatie surveys; cross-sectionele surveys.** Deze zijn goed geschikt voor het vergelijken van ervaringen van subgroepen op een bepaald moment.

Ook zijn er **longitudinale surveys**, die op meerdere momenten in de tijd worden afgenomen. Hiervan zijn er twee soorten:

1. **Repeated cross-sectional survey:** Onderzoekers interviewen verschillende samples van individuen op meerdere momenten; hierdoor erg geschikt voor documenteren van historische trends
2. **Panel survey:** Het interviewen van dezelfde sample individuen op meerdere momenten; geschikt voor onderzoeken hoe mensen veranderen over tijd.

|  |  |  |
| --- | --- | --- |
| Karakteristiek | Cross-sectional survey | Panel survey |
| Kosten | Vrij laag, omdat respondenten maar een keer gecontacteerd worden | Vrij hoog, omdat respondenten getraceerd, behouden en opnieuw geïnterviewd moeten worden |
| Gemak van uitvoering | Vrij makkelijk, omdat respondenten maar een keer geïnterviewd worden | Moeilijk, omdat het opvolgen en lokaliseren van subjecten omvat, lange wachttijden tussen interviews |
| Causale inferentie | Kan geen causale ordening vaststellen, omdat alle metingen op hetzelfde moment worden uitgevoerd | Excellent, metingen van een interview kunnen uitkomsten van volgende interviews voorspellen |
| Bronnen van bias | Coverage en samples errors, excluderen vaak degenen die moeilijk bereikbaar zijn | Zelfde bias als cross-sectional survey, additioneel met selectieve uitval |
| Mogelijkheid om verandering te documenteren | Kan interne verandering niet beoordelen | Interne verandering is goed beoordeelbaar |
| Aandacht voor sociale geschiedenis | Kan niet leeftijd, periode en cohort effecten losmaken | Panel studies over een cohort kunnen niet de leeftijd vs. Periode effect differentiëren. Multicohort longitudinale studies kunnen niet leeftijd versus cohort effect differentiëren. |

Het grootste nadeel van panel surveys is **uitval/attrition**; het verliezen van sample leden gedurende de tijd.

7.4 – Breedte van onderwerpen

Meestal wordt er allereerst demografische data verzameld, omdat sociale wetenschappers geïnteresseerd zijn in de ongelijkheden hierin. Daarnaast zijn surveys die door de staat gesponsord zijn verplicht om etnische data te verzamelen, omdat het een kritische factor is voor het maken van beleidsbeslissingen; en gebruiken staten ook etnische data om te voldoen aan wetgevende herograniserende eisen, evalueren van gelijke kansen op werk, en beoordelen van raciale verschillen in gezondheid.

Sommige surveys focussen op één onderwerp met een specifiek doel. Dit kan bijvoorbeeld door middel van een **poll**, een korte survey met vaak alleen een ja/nee antwoord.

Andere surveys focussen op een breed scala aan onderwerpen, en worden ook wel omnibus surveys genoemd. De ontwerpers van deze surveys willen vaak veel en diverse antwoorden, maar niet te veel tijd in beslag nemen van hun respondenten (max. 30-45 min), om geen foutieve antwoorden/onafgemaakte vragenlijsten te ontvangen. Een manier om de lengte van een survey te imiteren is door gebruik te maken van een **split-ballot ontwerp**, waarbij vaak (op basis van een 50%-50% verhouding) de survey vragen worden verdeeld over de respondenten.

7.5 – Vergelijkingen tussen groepen en gedurende de tijd

We gebruiken vaak **fixed response vragen** en **responscategorieën**, om ervoor te zorgen dat de data makkelijk te vergelijken is en om subjectieve fenomenen te kunnen vergelijken. Ze zijn zo ontworpen dat ze een bijna compleet scala aan responses dekken.

7.6 – Bronnen van errors/fouten in surveys

Er zijn 4 belangrijke soorten fouten:

1. **Nonresponse:** Respondenten doen niet mee aan de survey of vullen bepaalde vragen niet in.
2. **Measurement error:** De meting van een bepaalde variabele heeft invloed op de gegeven respons
3. **Coverage error:** Sampling frame dekt niet adequaat alle mensen van de targetpopulatie
4. **Sampling error:** Refereert aan verschillen tussen de karakteristieken van de survey sample en de populatie waar de sample van is genomen.

7.7 – Responspercentages

Een van de grootste challenges in surveys is om een zo hoog mogelijk **responspercentage** te hebben. Deze verschillen vaak van 20%-80%. Om ervoor te zorgen dat deze toch hoog is, kan een onderzoeker een alternatieve versie van het responspercentage aanbieden, om zo ook deels-afgemaakte surveys te includeren. Deze zijn echter niet zo informatief als complete surveys.

7.8 – Modes van administratie

De meest gebruikte manieren om surveys af te nemen zijn:

1. Face-to-face
2. Telefonisch
3. E-mail
4. Internet
5. Zelf-geadministreerd

De belangrijkste criteria om te kiezen voor een bepaalde modus zijn:

1. Kosten
2. Responspercentage
3. Controle van onderzoeker over data-collectieproces
4. Potentie voor **interviewer effecten**

7.9 – Face-to-face interview

De interview ontmoet de respondent persoonlijk en stelt vragen van het **interview schema**. Interviewer observaties zijn een type **paradata;** data over het proces hoe het interview is afgenomen. Er zijn verschillende manieren voor een face-to-face interview:

1. **PAPI:** Paper-and-pencil interview. De interviewer stelt vragen en schrijft de antwoorden op een voorgedrukte kopie van het survey boek.
2. **CAPI:** Computer-assisted personal interview. Interviewer leest vragen voor vanaf een voorgeprogrammeerde computer met alle vragen/antwoorden. Hierbij kan een **showcard** gebruikt worden voor de respondent om de antwoorden te kunnen onthouden. Deze manier helpt met **skip-pattern vragen**, die altijd beginnen met een **screener-vraag**.

Face-to-face is voordelig. Omdat de respondenten vragen kunnen stellen als ze een bepaalde vraag niet goed begrijpen. Echter, kan de aanwezigheid van een interviewer zorgen voor inaccurate antwoorden door bijvoorbeeld de **sociale wenselijkheid-bias**.

Om de setting van het face-to-face interview comfortabeler te maken, kan gebruik gemaakt worden van **ACASI:** Audio computer-assisted self-interview; waarbij de respondent een laptop of tablet gebruikt om te luisteren naar en te antwoorden op vooraf opgenomen vragen.

7.10 – Telefoon survey

Dit is een gestructureerd interview, telefonisch uitgevoerd. Respondenten worden vaak gegenereerd door RDD, random digit dialing. Hier kan gebruik gemaakt worden van **CATI**: Computer-assisted telephone interview. Deze manier helpt door ervoor te zorgen dat de skip-patterns correct gevolgd worden.

Extra voordelen van telefoon surveys zijn:

1. Lagere kost
2. Geconducteerd vanuit een centrale locatie
3. Weinig voorafgaande planning nodig
4. Geen reistijden/kosten etc.

Extra nadelen zijn:

1. Geen persoonlijke dingen telefonisch willen bespreken
2. Positief vooringenomen beeld over houding en gedrag (bias)
3. Interviewer effecten
4. Mensen willen thuis niet (telefonisch) lastig worden gevallen
5. Sample bias (bijv. door niet opnemen van telefoon)

Je kan vooraf een brief sturen om deze effecten tegen te gaan.

Wat ook voor kan komen is **respondent fatigue**; mensen hebben geen zin om lang te telefoneren. Hierdoor moeten vragen simpel en helder zijn. Ook is er tijdens telefoon surveys een nieuwe limiet ontstaan; telefoonnummers komen niet altijd overeen met woonlocatie. Vragen om locatie kost extra tijd (en eventueel geld).

7.11 – Mail surveys

SAQ’s kunnen ervoor zorgen dat mensen aan de ene kant eerlijkere antwoorden geven, maar aan de andere kant kunnen ze hierdoor belangrijke vragen niet begrijpen of zelfs overslaan.

SAQ’s worden vaak via e-mail verspreid. Ze zijn minder/niet gevoelig voor interviewer effecten. Een ander punt is lagere kosten. Hierdoor kan het onderzoeksteam makkelijker een grote & geografische diverse sample krijgen.

Toch zijn er ook nadelen. Er is vaker missende data en “weet niet”-antwoorden. Ook is er een laag responspercentage, vaak tussen de 20% en de 40%. Strategieën om dit tegen te gaan zijn dezelfde mensen opbellen, het sturen van een kaartje, of een gepersonaliseerde brief. Dit kost echter wel meer, met een laag succespercentage.

7.12 – Online internet surveys

Deze gebruiken gecomputeriseerde SAQ’s, en worden vaak voor marktonderzoek gebruikt met korte vragenlijsten. Tegenwoordig wordt het ook vaak gebruik door academische en regeringsonderzoekers.

De voordelen zijn:

* Lage kosten
* Hoeft niet geprint te worden
* Er hoeft geen extra persoon ingehuurd te worden om de data in te voeren
* Makkelijk te ontwerpen en uit te voeren
* Makkelijk navigeerbaar tussen vragen
* Minder missings dan mail surveys omdat de online survey zo geprogrammeerd is dat respondenten niet verder kunnen gaan totdat ze alle vragen hebben beantwoord op elk scherm.
* Respondenten kunnen over heel de wereld bereikt worden.

Nadelen:

* Niet iedereen/alle subgroepen hebben internettoegang of maken hier gebruik van
* Responspercentage blijft toch enigszins gelijk aan e-mail surveys.

7.13 – Mixed-models benadering

Er wordt ook vaak gebruik ggemaakt van meerdere modes, omdat elke modus zijn limieten en nadelen heeft. De redenen voor het gebruik van de mixed-modes benadering zijn als volgt:

* Nieuwe of verschillende typen informatie per respondent
* Beoordelen kwaliteit van de data & potentiële bias bronnen
* Je kunt mensen bereiken die bijvoorbeeld niet via e-mail willen participeren, maar wel via telefoon
* Verhoogd responspercentage

Het is een belangrijke strategie voor het minimaliseren van coverage & nonresponse errors, verhogen responspercentage en verminderen van bias. De data-analyse moet echter wel rekening houden met mode-effecten → de manieren waarop de manier van uitvoering per mode de antwoorden van respondenten kan beïnvloeden

De vragen hebben meestal dezelfde structuur. De **stem** stelt de vraag, en de responscategorieën zorgen voor de antwoorden.

7.14 – Closed-ended vragen

De meest basis variant hiervan heeft een **dichotomous outcome category**, wat wil zeggen dat er maar twee opties zijn (vaak ja/nee). De meeste vragen hebben echter wel meerdere responscategorieën. Er zijn twee voorwaarden:

* **Mutually exclusive:** Categorieën moeten niet overlappen
* **Exhaustive:** Er moet minstens één accurate respons zijn voor een respondent

Ze worden vaak gebruikt voor het meten van frequenties van bepaald gedrag. Vaak wordt gebruik gemaakt van een **rating scale**, met de **Likert scale** als meest voorkomende variant.

Vaak wordt ook gebruik gemaakt van een **forced option**, die een indicatie geeft van hun algemene richting van overeenstemming. Sommige experts vinden dat eens/niet eens-categorieën onderhevig zijn aan de **acquiescence bias;** sommige respondenten antwoorden continu “mee eens”. Het geven van twee contradictoire keuzes gaat deze bias tegen.

Een andere vorm van closed-ended vragen is het **rank ordening** van prioriteiten of preferenties. Rating vragen stellen dat mensen elk item apart moeten evalueren, terwijl ranking items vragen om directe vergelijkingen tussen items/onderwerpen.

Voor moeilijk te beantwoorden vragen (zoals bijvoorbeeld inkomen) kunnen **unfolding questions/unfolding brackets** gebruikt worden, waarbij ze moeten zeggen of hun inkomen bijv. hoger is dan 50.000 euro, met als follow-up vraag “hoger dan 60.000 euro”, etc.

7.15 – Open-ended en semi-gestructureerde vragen

Vragen meestal om een kort antwoord, en worden uitgevoerd op twee manieren:

1. Zit in een closed-ended vraag bij de optie “anders, leg uit…”
2. Gewoon open-ended op zichzelf, zonder vooraf ingestelde responscategorieën

Waneer kies je open-ended of closed-ended vragen? Er zijn vier algemene criteria die hierbij kunnen helpen:

1. Doel van de survey
2. Kennis van respondent over het onderwerp
3. Hoeveel aandacht de respondent schenkt aan het onderwerp
4. Motivatie van de respondent om te vraag te beantwoorden

|  |  |
| --- | --- |
| Closed-ended | Open-ended |
| Frequentie van gedrag | “Waarom”-vragen |
| Factuele informatie | Onbekend onderwerp |
| Level van overeenstemming met onderwerp | Egen opinie |
| Moeilijk onderwerp waar respondenten liever niet (uitgebreid) over willen praten | Nog niet nagedacht over onderwerp |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Attributie | Face-to-face interview | Mail of SAQ | Telefonisch interview | Internet survey |
| **Kosten** | Hoog | Laag | Gemiddeld | Laag |
| **Responspercentage** | Hoog | Laag | Hoog | Gemiddeld |
| **Controle van onderzoeker over interview** | Hoog | Laag | Gemiddeld | Gemiddeld |
| **Interviewer effects** | Hoog | Laag | Gemiddeld | Laag |

7.16 – Single items vs. Composite

Vaak bevat een onderwerp maar één vraag. Toch kunnen er ook **composite measures** zijn die meerdere items combineren. Er zijn twee typen composite measures:

1. **Index:** Somt antwoorden op survey metingen op, die de belangrijkste elementen van een bepaald concept meten
2. **Scale/schaal:** Middelt responses over een serie van gerelateerde items die een enkel concept bevatten.

7.17 – Karakteristieken van goede kwalitatieve vragen

* Duidelijk en simpel taalgebruik
* Geen dubbele negatieven gebruiken
* Geen gebruik maken van **double-barreled vragen**
* Neutraal taalgebruik bij gevoelige onderwerpen
* Ontwijk leidende vragen en morele woordkeuzes

7.18 – Nieuwe survey of al eerder gemaakte survey gebruiken?

Wanneer je gebruik maakt van andermans survey heb je een **codeboek** nodig. Een belangrijke reden voor het maken van deze keuze is vergelijking van eigen resultaten met eerder verkregen resultaten van andere samples of populaties.

Je gebruikt je originele vragen als je een nieuw onderwerp onderzoekt en er nog geen survey van bestaat/de surveys die bestaan je onderwerp niet adequaat meten. Hiervoor moet je de theoretische literatuur wel extreem goed kennen.

7.19 – De plaats van vragen

Onderzoekers besteden veel aandacht aan de **volgorde** van het stellen van vragen. Er zijn drie belangrijke factoren die hier invloed op hebben:

1. Vertrouwen opbouwen
2. Monotone vragen en response set vermijden
3. Vermijd volgorde-effecten

De eerste vraag moet altijd een makkelijke, duidelijke vraag zijn. Zorg ervoor dat de respondenten zich niet vervelen. Wissel de vragen af Er zijn twee manieren om dit te voorkomen:

Verander de antwoordcategorieën na een paar items. Gebruik positief en negatief door elkaar. Dus zeg soms: “Ik voel mij goed over mezelf” en soms “Ik denk dat ik niet veel kan”.

Vermijd volgorde-effecten. Dit is het effect dat het antwoord op een vraag beïnvloed kan zijn door een vraag die eerder is geweest. Als je bijvoorbeeld eerst iemand vraagt hoe religieus die is en die persoon zegt “heel erg” en daarna vraagt hoe vaak hij of zij naar de kerk gaat, wil de persoon graag consistent in zijn antwoord zijn en het liefst ook heel vaak invullen.

Om dit te voorkomen is het handig om de volgorde van vragen willekeurig te maken. Hier kan je weer een split-ballon design voor hebben.

Een ander volgorde-effect is het **priming effect**. Dit houdt in dat bepaalde woorden, plaatjes, vormen mensen beïnvloeden. Bijvoorbeeld wanneer mensen worden geprimed om na te denken over politieke schandalen en daarna de mate van vertrouwen moeten aangeven in de huidige verkiezingen, geven zij aan minder vertrouwen te hebben.

7.20 – Afname van de enquête

Het meest belangrijke van een enquête is het **pre-testen**. Dit is een test om te kijken of er onduidelijke vragen zijn of bijvoorbeeld de volgorde van de vragen niet goed is. De pre-test moet op precies dezelfde manier als de echte afname gaan. De methode moet hetzelfde zijn, de steekproef moet lijken op de échte steekproef en de condities van afname moeten ook hetzelfde zijn.

Er kan ook een **cognitief interview** plaatsvinden. Hierin wordt de respondent gevraagd om hardop na te denken, om zo te kijken of de vraag duidelijk was en waarom zij voor een bepaald antwoord kiezen.

Om te weten of de pre-test succesvol was en je de enquête echt kan gaan gebruiken, kan je  
“soft” en “hard” data gebruiken. “Hard” data houdt in dat er gekeken wordt naar de **frequentie distributies***.* Dat is de verdeling van de antwoorden. Als veel mensen aangeven dat ze iets niet weten, dan moet de vraag herzien worden. Soms weet men het ook echt niet, dit is dan een non-attitude. Ook als onderzoekers een genuanceerder beeld willen omdat veel mensen aan één kant van de antwoordcategorie zitten kunnen zij de antwoordcategorieën aanpassen.

Nadat de pre-test is gedaan en de enquête eventueel is aangepast, kan je je enquête afnemen. Vaak wordt er met behulp van een brief de respondenten bereikt. Dit moet één week voor de echte afname plaatsvinden. Het eerste contact met de respondent moet een duidelijke uitleg geven van het doel van de survey en moet dankbaarheid en respect tonen voor de respondent. Ook moet er **informed consent** van de respondent zijn: hij moet aangeven vrijwillig mee te doen. Ook moet de onderzoeker aangeven wat de risico’s en voordelen zijn van het meedoen. Daarnaast moet een respondent gewezen worden op dat hij vragen kan overslaan of kan stoppen met invullen.

7.21 – Ethische problemen bij enquêtes

Onderzoekers moeten hun participanten geen schade toebrengen. Ze moeten alle risico’s vermelden en hun best doen om de vertrouwelijkheid en anonimiteit van hen te waarborgen.

Er zijn verschillende strategieën om ethische gedragsregels te waarborgen:

* **Confidentialiteit**: gevoelige informatie mag niet openbaar worden gemaakt. Alleen het team van onderzoekers heeft toegang tot de informatie. Dit kan worden bereikt door elke respondent een ID-nummer te geven. Een document die de namen bevat moet goed bewaard worden.
* **Anonimiteit**: Het is moeilijk en onrealistisch om anonimiteit te beloven aan de respondenten. Vaak willen onderzoekers nog een follow-up interview doen en dan moeten zij wel weten wie welke respondent is.

Onderzoekers moeten ook rekening houden met dat vragen psychologische distress kunnen opleveren.

**Hoofdstuk 8 Morling et al.**

**Associatie claims** zijn claims die de relatie tussen twee gemeten variabelen beschrijven. Een**bivariate correlatie**wordt ook wel een **bivariate associatie**genoemd en het is een associatie die twee variabelen betreft.

1. Om associaties te onderzoeken, moet men de eerste variabele onderzoeken en vervolgens de tweede variabele en dit moet gebeuren in dezelfde groep van mensen.
2. Vervolgens worden er statistische methodes en grafieken gebruikt om de type relatie tussen de variabelen te weergeven.

Relatief veel studies zijn correlationeel.

8.1 – Hoe beschrijf je associaties tussen twee variabelen?

Nadat je alle data hebt verzameld, moet je de relatie tussen de twee gemeten variabelen beschrijven aan de hand van **een spreidingsdiagram** en de **correlatiecoëfficiënt r**. Wanneer je de twee variabelen tegen elkaar uit zet in een spreidingsdiagram en voor elk persoon de waardes als stippen noteert, dan kun je een lijn (laten) trekken door je puntenwolk.

Als je lijn van linksonder naar rechtsboven loopt, dan spreken we van een **positieve relatie**. Een positieve relatie houdt in dat hoge scores op de ene variabele samengaan met hoge scores op de andere variabele.

Wanneer de lijn van linksboven naar rechtsonder loopt, dan is er sprake van een **negatieve relatie.** Hoge scores op de ene variabele gaan dan samen met lage scores op de andere variabele.

De kracht van de correlatie kan aangeduid worden met de correlatiecoëfficiënt r. Deze loopt van -1 tot 1.

|  |  |
| --- | --- |
| Correlatiecoëfficiënt *r* | Sterkte van de correlatie |
| 0.10 of -0.10 | Zwak |
| 0.30 of -0.30 | Matig |
| 0.5 of -0.50 | Groot |

r laat dus de richting (positief of negatief) en sterkte van de relatie zien.

8.2 – Hoe beschrijf je associaties met categorische data?

Hierboven is omschreven hoe de associatie tussen twee variabelen beschreven kan worden. Echter, je moet er wel om denken dat sommige variabelen **categorisch** zijn. De waarden van een categorische variabele kunnen alleen maar in een categorie vallen.

Wanneer beide variabelen van een associatie gemeten worden met **kwantitatieve schalen**, dan is het gebruikelijk om spreidingsdiagrammen te maken. De data kan op die manier het best gerepresenteerd worden. Een spreidingsdiagram is niet handig als een van de variabelen categorisch is. De punten die personen voorstellen, komen onder elkaar te staan (verticaal dus. Het is heel lastig om bij een categorische variabele aan het spreidingsdiagram te zien of de relatie positief of negatief is.

Het is wellicht wel mogelijk om een spreidingsdiagram te maken van een categorische variabele, maar gebruikelijk is het niet. Het is handiger om een **staafdiagram** te maken. In een staafdiagram is elk persoon niet voorgesteld als een punt, maar de gemiddeldes voor elk categorie worden weergegeven. Met een staafdiagram kun je het verschil tussen de groepsgemiddelden onderzoeken.

Wanneer tenminste een van de variabelen in een associatie claim categorisch is, dan kunnen er verschillende statistische methodes gebruikt worden om de data te analyseren. Soms kan r gebruikt worden, maar het is gebruikelijker om te testen of de verschillen tussen de gemiddelden statistisch significant zijn. Dit wordt vaak gedaan met de **t-test**. Het lijkt misschien gek dat associatie claims weergeven kunnen worden met zowel spreidingsdiagrammen als staafdiagrammen of dat ze door verschillende statistische methodes beschreven kunnen worden. Het maakt niet uit welke soort grafiek of statistische maat je gebruikt; als beide variabelen gemeten zijn, dan is een studie correlationeel.

Zoals eerder besproken (hoofdstuk 3), spreken we van **een experiment** als een van de variabelen gemanipuleerd is. Experimenten zijn beter voor **causale claims**. Een associatie claim wordt niet ondersteund door een bepaalde grafiek of een bepaalde statistische maat; het wordt ondersteund door een design van een studie, waarbij beide variabelen gemeten zijn.

8.3 – Hoe onderzoek je associatie claims?

De belangrijkste **validiteiten** die onderzocht moeten worden bij associatie claims, zijn **construct validiteit en statistische validiteit**. Soms kan men ook **de externe validiteit** onderzoeken. De interne validiteit is niet belangrijk bij associatie claims.

8.3.1 – Wat houdt construct validiteit in voor associatie claims?

Omdat een **associatie claim** de relatie tussen twee gemeten variabelen beschrijft, is het belangrijk om naar de **construct validiteit** van beide variabelen te kijken. Men moet dus kijken hoe goed elk van de twee variabelen gemeten was. Je kunt je daarbij afvragen of de maat betrouwbaar is en of het meet wat het hoort te meten. Ook kun je je afvragen wat het bewijs voor **face validity, discriminant, convergent en concurrent validiteit** van de variabele is.

8.3.2 – Wat houdt statistische validiteit in associatie claims in?

Wanneer je **statistische validiteit** van een associatie claim onderzoekt, dan wil je dus eigenlijk weten of en welke factoren een invloed hebben gehad op:

* Het spreidingsdiagram
* Correlatiecoëfficiënt r
* Staafgrafieken
* Verschillen van gemiddeldes die tot je associatie claim hebben geleid

Er moet gekeken worden naar effectgrootte, uitbijters in de data, restricties en statistische significantie van de relatie.

8.4 – Wat betekent effectgrootte voor associatie claims?

**De effectgrootte** kijkt naar de sterkte van een relatie. Immers, sommige associaties zijn sterker dan andere associaties. Wanneer er twee associaties zijn, dan is de associatie met de r die dichter bij de 1 ligt sterker. Sterke effectgroottes gaan samen met nauwkeurigere voorspellingen dan zwakke effectgroottes. Je **voorspellingserro**r wordt ook lager wanneer de sterkte van effectgroottes toeneemt.

Sterkere effectgroottes zijn over het algemeen ook belangrijker dan kleinere effectgroottes. Toch zijn er uitzonderingen op deze regel. Deze uitzonderingen hangen allemaal van de context af. Soms kan een kleine effectgrootte al belangrijk zijn. Als het neerkomt op leven en dood, dan kan een klein effectgrootte al belangrijk zijn. Wanneer de uitkomst niet van levensbelang is, dan zal een kleine effectgrootte waarschijnlijk ook niet belangrijk zijn.

8.5 – Wat betekent statistische significantie voor associatie claims?

Onderzoekers kunnen natuurlijk niet alle personen uit een populatie onderzoeken en zij moeten daarom **steekproeven** gebruiken. Aan de hand van deze steekproeven worden er conclusies getrokken over de populatie. Vaak is het zo dat de resultaten van steekproeven en populatie elkaar spiegelen, maar dit is niet altijd het geval. Soms is er geen associatie tussen twee variabelen van een populatie, maar een onderzoek kan per toeval een associatie vinden in een steekproef. De correlatie van die steekproef is door toeval veroorzaakt. Dit gebeurt wel eens en we moeten ons dus altijd afvragen of er echt een associatie in de populatie is of dat er toevallig een associatie is gevonden in de steekproef.

Statistische significantie berekeningen geven een **probabilistische schatting** weer, **p**. De p zegt iets over de kans dat de associatie kwam van een populatie waarin de associatie nihil is.

→ Als de kans kleiner is dan 5%, dan kunnen we ervan uitgaan dat het heel onwaarschijnlijk is dat de resultaten kwamen uit een nihil-associatie. De correlatie wordt dan gezien als statistisch significant.

→ Wanneer de resultaten een hoge p opleveren (.05 of hoger), dan zijn de resultaten niet statistisch significant. Dan kan een onderzoeker dus niet uitsluiten dat de resultaten gekomen zijn uit een populatie waarbij de associatie tussen variabelen nihil is.

Significantie is ook gerelateerd aan effectgrootte: hoe sterker de correlatie (grote effectgrootte), hoe groter de kans dat de correlatie statistisch significant zal zijn. Statistische significantie berekeningen hangen niet alleen af van effectgrootte, maar ook van steekproefgrootte. Een kleine effectgrootte zal statistisch significant zijn als het uit een hele grote steekproef (vanaf 1000 proefpersonen) komt. Een kleine steekproef wordt makkelijker beïnvloed door kans dan grote steekproeven. Zwakke correlaties die gebaseerd zijn op kleine steekproeven zullen eerder het resultaat zijn van kans en ze zullen eerder als niet significant bestempeld worden. In wetenschappelijke artikelen kun je lezen over de significantie van een onderzoek. Je kunt een significantie zien aan de p, maar soms wordt een statistisch significant resultaat ook weergeven met **een asterix** (dat is een \*).

8.6 – Hebben uitbijters invloed op een associatie?

**Uitbijters** zijn extreme scores. Het zijn scores die totaal afwijken van de andere scores. Uitbijters kunnen soms een groot effect hebben op de correlatiecoëfficiënt r. De aanwezigheid van een uitbijter kan ervoor zorgen dat de correlatie verschuift (bijvoorbeeld van r = .26 naar r = .37). Uitbijters kunnen problemen opleveren voor associatie claims. In bivariate correlaties zijn uitbijters vooral problematisch wanneer ze extreme scores op beide variabelen hebben. Wanneer je een associatie claim onderzoekt, moet je je dus eerst afvragen of er uitbijters zijn in een steekproef. Deze uitbijters zou je kunnen vinden door te kijken naar spreidingsdiagrammen.

Uitbijters zijn vooral belangrijk om naar te kijken als je een kleine steekproef hebt. Wanneer een steekproef bestaat uit 600 proefpersonen die vrijwel allemaal in het midden scoren, dan zal één uitbijter die extreem scoort (of helemaal links of helemaal rechts) niet veel invloed hebben. Echter, wanneer je een steekproef van 16 personen hebt die in het midden scoren, dan kan een uitbijter die extreem scoort een grote invloed hebben.

8.7 – Zijn er range restricties?

Wanneer er in correlationeel onderzoek niet een heel bereik van scores is voor een van de variabelen in de associatie, dan kan de correlatie kleiner lijken dan het in werkelijkheid is. Dit wordt een **range restrictie**genoemd. Dat houdt dus in dat je niet alle waarden hebt weergeven die er beschikbaar zijn. Als onderzoekers vermoeden dat er een range restrictie is, dan kunnen zij besluiten om een statistische techniek toe te passen***,*correctie voor range restrictie** genaamd.

Range restrictie kan aanwezig zijn wanneer er, voor welke reden dan ook, weinig **variantie** is in een van de variabelen. Wanneer men kijkt naar het inkomen van ouders en de schoolprestaties van een kind, dan moet men naar alle inkomens kijken. Het is niet de bedoeling dat je alleen ouders opneemt in de steekproef die een middel inkomen hebben. Ouders met een laag en hoog inkomen zouden ook opgenomen moeten worden.

8.8 – Is de relatie curvilineair?

Als een onderzoek vermeldt dat er geen relatie is tussen variabelen, dan kan het zo zijn dat de relatie daadwerkelijk nihil is. Echter, in sommige gevallen kan het ook zo zijn dat de relatie **curvilineair** is. Dit houdt in dat de relatie tussen de twee variabelen niet als een rechte lijn voorgesteld kan worden. Het kan zijn dat de relatie in het begin positief is (hoge x-variabele gaat samen met een hoge y-variabele), maar op een gegeven moment negatief wordt (hoge x-variabele gaat samen met een lage y-variabele). Een voorbeeld hiervan is gezondheidszorg. Naarmate iemand ouders wordt, heeft hij/zij tot een bepaald punt minder gezondheidszorg (o.a. doktersbezoeken) nodig. Echter, vanaf een bepaalde leeftijd (ongeveer 60 jaar) neemt het nodig hebben van de gezondheidszorg weer toe. Er bestaat dus een curvilineaire relatie tussen leeftijd en gezondheidszorg.

8.9 – Kan er een causale inferentie worden gemaakt over een associatie?

Het is belangrijk om stil te staan bij **causaliteit**. Veel leken associëren correlaties namelijk met causaliteit. Je moet dus altijd beseffen dat correlatie geen oorzaak is! Een normale associatie kan geen oorzaak tot stand brengen. Voor causaliteit is nodig:

1. Temporele precedentie
2. Interne validiteit
3. Covariantie van oorzaak en gevolg

Bij een correlatie tussen twee variabelen weet je niet altijd welke variabele er eerst kwam en of de ene variabele de andere veroorzaakt heeft. Bovendien weet je ook niet of er nog een derde variabele was die invloed heeft uitgeoefend op een of beide variabelen. Pas als aan alle drie voorwaarden van causaliteit is voldaan, kan men spreken van een causaal verband. In een associatie claim zal er nooit voldoen kunnen worden aan alle drie voorwaarden. Causaliteit kan alleen door middel van experimenten onderzocht worden. Wanneer een derde variabele voor een correlatie zorgt tussen twee variabelen, dan spreekt men van een spurious associatie.

8.10 – Waar kan de associatie naar gegeneraliseerd worden?

Bij **externe validiteit** vraag je je af of een associatie claim gegeneraliseerd kan worden naar andere mensen, tijden en plaatsen. Het kan zo zijn dat een bivariate correlatie onderzoek geen random steekproef heeft gebruikt, maar dat wil niet zeggen dat je de associatie weg moet schrijven. De resultaten van een onderzoek kun je gewoon accepteren en het gedeelte van de generalisatie kun je aan vervolgonderzoek over laten. Veel associaties generaliseren echter wel naar de populatie.

Wanneer er sprake is van een associatie onderzoek en de relatie tussen de twee variabelen verandert doordat een andere variabele een invloed uitoefent, spreken we van een **moderator*.*** Moderatoren kunnen ons informatie geven over de externe validiteit. Wanneer een associatie door een derde variabele gemodereerd wordt, dan kunnen sommige resultaten wellicht niet gegeneraliseerd worden naar andere settings of groepen mensen.

**Field hoofdstuk 4**

4.4 Hoe kan je beginnen met SPSS?

Bij het openen van SPSS krijg je een beginscherm waar men verschillende dingen kan kiezen, zoals het openen van een al bestaand bestand of een nieuw bestand maken. De belangrijkste schermen in SPSS zijn de **data editor,** waar je je data invoert en je statistische berekeningen mee uitvoert en de **viewer**, waar je resultaten van de analyses in verschijnen.

4.5 Wat wordt bedoeld met de data editor?

Zodra je de keuze hebt gemaakt in het beginscherm, kom je op het SPSS-scherm waar alle gegevens staan (of waar je alle gegevens in kunt vullen). Dit is de **data editor**. De data editor heeft twee opties, de **data view en de variable view**. De eerste is om de cijfers in te voeren en de tweede is om te kijken wat voor variabelen gebruikt worden. Door erop te klikken kom je op het goede scherm.

Betekenissen van de termen in de balk bovenin:

* **File (bestand):** Hier kan je algemene dingen doen zoals, het opslaan van de output en het printen.
* **Edit (opmaak):** Met deze optie kan het knip- en plakwerk gedaan worden.
* **View**: Deze optie laat systeem specificaties zien, zoals grijze lijnen in de data editor.
* **Data (gegevens):** Hier kan je veranderingen in de data editor aanbrengen. Met *insert variable* kan je een nieuwe variabele invoegen, met *insert cases* kan je een nieuwe rij invoegen, met *split file* kan je de gegevens splitsen met een variabele en met *select cases* kan je een selectie maken waar de analyse op gedaan moet worden.
* **Transform**: Als je één of meerdere variabelen wilt manipuleren is deze optie handig.
* **Analyze**: In dit menu zijn de statistische procedures te vinden. Een aantal procedures:
* **Descriptive Statistics**: Voor beschrijvende statistieken, frequenties en algemene gegevens.
* **Compare Means:** Hier kan je de verschillende t-toetsen en de een weg ANOVA vinden.
* **General Lineair Model**: Dit menu is voor complexe ANOVA’s
* **Mixed Models**: Voor multilevel lineair models.
* **Correlate:** In dit menu staan de correlaties zoals die van Spearman, Pearson en Kendall.
* **Regression:** De regressietechnieken.
* **Loglineair:** Voor loglineaire analyse.
* **Data Reduction**: Hier is de factoranalyse te vinden.
* **Scale**: Hier zijn de betrouwbaarheidsanalyses te vinden.
* **Nonparametric Tests**: Verschillende nonparametrische testen zijn hier te vinden.
* **Graphs (grafieken):** Met deze optie kan je verschillende grafieken van je gegevens laten maken.
* **Window:** Hiermee kan je wisselen tussen je output scherm en het scherm met de gegevens.
* **Utilities:** Bij data-file comments kan je commentaar toevoegen om het voor jezelf makkelijker te maken.
* **Add-ons**: Hiermee kan de benodigde deelnemers ten opzichte van de power uitgerekend worden.
* **Help**: Dit spreekt voor zich.

4.6 Op welke manier wordt data in IBM SPSS statistics ingevoerd?

In SPSS zijn **de kolommen de variabelen** en **de rijen de deelnemers**. Dit betekent dat elke deelnemer zijn eigen rij heeft. Als een onderzoeker een variabele in wil voeren doe je dit eerst in de variable view. Onderstaand worden de verschillende kolommen uitgelegd:

* **Name**: Hier kan je de naam van de variabele geven. Let wel op dat het programma sommige tekens niet accepteert net als de spatie.
* **Type:** De numerieke variabele wordt het meest gebruikt. Dit houdt in dat je voor elke score één cijfer hebt. Als je de string variabele gebruikt heb je bij de score een aantal letters, bijvoorbeeld een naam. Er zijn ook valuta variabelen zoals dollars of euro’s en datum variabelen die een datum weergeven.
* **Width:** SPSS geeft standaard 8 cijfers, maar hier kan je dat veranderen als dat nodig is. De width geeft aan hoeveel tekens je maximaal kunt invullen in een antwoordveld.
* **Decimals**: Hier kan je het aantal decimalen aangeven dat je weergegeven wilt zien.
* **Label:** Hier kan je een lange beschrijving van de variabele geven die normaal niet in het korte blokje past.
* **Values:** Hier kan je nummers aan personen of groepen toekennen.
* **Missing**: Hier kan je nummers aan missende gegevens toekennen.
* **Columns:** Hier kan je de breedte van de kolommen aangeven.
* **Align**: Hier kan je kiezen of je de cijfers links, rechts of in het midden van het vakje wilt hebben.
* **Measure:** Hierin kan je het meetniveau aangeven.
* **Role**: SPSS statistics heeft diverse procedures die een poging doen om bepaalde analyses automatisch uit te voeren zonder dat je erbij na hoeft te denken wat je moet doen. SPSS denkt als het ware namens de onderzoeker -- functie: automatic linear model

Als je dit allemaal hebt gedaan voor een variabele kan je naar de data view gaan en de gegevens van die variabele in de kolom invoeren.

Er kan ook sprake zijn van **een coding variabele**. Hierbij geeft een onderzoeker een nummer aan bepaalde groepen. Bijvoorbeeld bij geslacht van personen. De mannen krijgen een 0 en de vrouwen krijgen een 1. Dit kan aangegeven worden bij **de variable view - values**. Als een onderzoeker de missing data op wil geven zijn er drie opties:

1. De eerste optie houdt in dat geen ontbrekende gegevens zijn
2. De tweede optie is dat alleen een paar waardes missen
3. De derde en laatste optie houdt in dat de waardes tussen een bepaald gebied ontbrekende waardes zijn

Naast coderende variabelen bestaan er ook **numerieke variabelen**. De numerieke variabelen zijn het makkelijkste om te maken omdat de numerieke variabele het standaard format is in SPSS. Numerieke variabelen zijn variabelen die getallen bevatten bijvoorbeeld het aantal vrienden. Er kunnen ook variabelen missen. Dit kan zijn omdat sommige deelnemers aan onderzoeken vragen missen in een vragenlijst. Daarnaast kan er bij een technische fout optreden waardoor een score niet opgeslagen wordt en meer. De cel kan leeg gelaten worden in Excel. Je kunt echter in SPSS missende scores een nummer geven zodat SPSS weet dat dit een missende score is. Klik op de kolom missing in de variable view (data/variable view). Klik op ... om de missing values te activeren.

4.8 Wat wordt bedoeld met de SPSS-viewer?

De SPSS-viewer verschijnt in een ander venster dan de data editor en geeft de uitvoer van alle procedures in SPSS weer: de tabellen met de resultaten, de grafieken, de fouten en meer. **(output)**

**Field Hoofdstuk 5**

5.3 Wat wordt bedoeld met een SPSS Chart Builder?

**Grafieken** kun je maken in de **chart builder**. Ga in SPSS naar **graphs** en klik vervolgens op chart builder. Links wordt de lijst met variabelen weergegeven. Vervolgens laat het canvas zien hoe de grafiek eruit komt te zien en de geselecteerde variabelen kan je naar de juiste drop zones slepen.

Onderaan staat de galerij. Hier kun je aangeven hoe de grafiek eruit moet komen te zien. Er zijn verschillende mogelijkheden, bijvoorbeeld:

* Een lijndiagram
* Staafdiagram
* 3D
* 2D

De **properties panel** zit aan de rechterkant. Hier kan bepaald worden wat de grafiek weer moet geven, hoe het eruit komt te zien en hoe je om moet gaan met missende variabelen.

5.4 Wat wordt bedoeld met een histogram?

**Een histogram** is een nuttige manier om te kijken naar de vorm van de gegevens en om problemen te ontdekken. Een histogram wordt ook wel **een frequentie distributie** genoemd. Wanneer bij de galerij gekozen wordt voor een histogram krijg je de keuze uit vier verschillende vormen van histogrammen.

Dit zijn de volgende vier:

* **Simple histogram**: Een simpel histogram kan gebruikt worden om de frequenties van een simpele variabele te bekijken/ visueel weer te geven.
* **Stacked histogram**: Hier worden de verschillende groepen in een histogram weergegeven. Bijvoorbeeld man en vrouw.
* **Frequency histogram**: Dit is hetzelfde als een simpel histogram. In dit geval zijn de uitkomsten echter in een lijn te zien.
* **Population pyramid**: Hier zijn twee grafieken met de onderkant tegen elkaar gezet. De frequentie is dus hier te zien in de horizontale as. Je krijgt als het ware twee grafieken die gekanteld zijn.

In een histogram en **boxplot** zie je snel wat **de uitbijters** of **uitschieters (outliers)** zijn. Dit zijn scores die ver buiten alle andere scores vallen.

5.5 Wat wordt bedoeld met een boxplot?

**Een boxplot (of een box-whisker diagram)** is een andere goede manier om de verdeling van de gegevens te bekijken. In het midden van de boxplot zit **de mediaan.** Er zijn drie verschillende boxplots waaruit je kan kiezen. Dit zijn de volgende drie:

* **1 - D boxplot:** Deze boxplot wordt gebruikt wanneer er sprake is van één variabele zonder dat er een categorische variabele geselecteerd moet worden voor de x-as.
* **Simple boxplot**: Deze boxplot wordt gebruik voor de simpele variabelen. Een onderzoeker krijgt de categorieën in aparte boxplots naast elkaar te zien.
* **Clustered boxplot**: Deze boxplot is hetzelfde als de eerste vorm alleen kan je hier de gegevens in tweeën splitsen. De boxplots zijn dan in bijvoorbeeld twee verschillende kleuren te zien.

**Een ster (\*)** in het plaatje is een uitbijter. De box zelf is de middelste 50%. Verder worden ook de laagste en hoogste scores weergegeven. Wanneer de armen vanaf de mediaan even lang zijn is de verdeling symmetrisch. Daarnaast kunnen uitbijters via SPSS met behulp van **de z-scores** gevonden worden.

5.8 Wat wordt bedoeld met een scatterplot?

Een andere mogelijk manier om om gegevens weer te geven is **een scatterplot of spreidingsdiagram.** Een scatterplot is handig om te gebruiken wanneer je wil kijken naar de relatie tussen twee variabelen. Een scatterplot laat de score op de ene variabele en de score op de andere variabele zien. Bij een scatterplot kan je ook zien wat voor soort relatie er bestaat tussen twee variabelen. Bovendien zijn de uitbijters makkelijk in de grafiek te vinden. Om een scatterplot in SPSS te maken ga je naar **chart builder -> scatter / dot** en dan zijn er acht opties waaruit gekozen kan worden. Dit zijn de volgende acht:

* **Simple scatterplot**: Deze vorm kijkt alleen naar twee variabelen. In deze grafiek kan een continue variabele tegen een andere variabele uitgezet worden. Bijvoorbeeld wanneer een psycholoog geïnteresseerd is in de effecten van tentamen stress op de resultaten van het specifieke tentamen. Bij het plotten van een scatterplot is het handig om **de regressielijn** erin zichtbaar te maken. Dit is een lijn die de relatie tussen de twee variabelen samenvat. Je krijgt de lijn door de dubbel te klikken op de grafiek en vervolgens bij **properties** voor de optie **lineair** te kiezen wanneer je een lineaire lijn wilt.
* **Grouped scatter**: Dit is hetzelfde als de simple scatter alleen kan je hier onderscheid maken tussen groepen door middel van verschillende kleuren. Bijvoorbeeld: een onderzoeker wil kijken of mannelijke en vrouwelijke studenten verschillend reageren op stress voor een tentamen. Voor een grouped scatterplot wordt er nog een derde categorische variabele aan het geheel toegevoegd, en kan je ook meerdere regressielijnen zichtbaar maken.
* **Simple 3-D scatter**: Deze opties kan je gebruiken als je een continue variabele tegen twee andere variabelen uit wilt zetten.
* **grouped 3-D scatter**: Dit is hetzelfde als de grouped scatter alleen kan je hier nog een variabele toevoegen. Het kan soms onoverzichtelijk zijn als je een 3-D scatterplot maakt. Met de optie rotation kan je de grafiek draaien om een duidelijker beeld te krijgen.
* **Summary point plot:** Deze plot is hetzelfde als het staafdiagram alleen is het hier een punt in plaats van een staaf.
* **Simple dot plot**: Dit wordt ook wel **de density plot** genoemd. Het laat de frequenties zien en elke punt is één score. Het is een manier om de verdeling van de scores te zien.
* **Scatterplot matrix:** Deze laat meerdere plots zien tussen paren van variabelen.
* **Drop-line:** Een punt laat hier het gemiddelde zien en de lijn maakt de verbinding tussen de gemiddelden van de groepen. Dit is handig voor het vergelijken van de statistieken.

**Een matrix scatterplot** laat alle mogelijke combinaties zien van alle variabelen die ingevoerd worden. Selecteer eerst alle variabelen die vergeleken moeten worden en sleep deze vervolgens naar de scattermatrix. Het is verstandiger een matrix te plotten van 2D scatterplots. Dit type zorgt ervoor dat een onderzoeker kan kijken naar de relatie tussen alle combinaties van verschillende paren

**Field Hoofdstuk 8**

8.4 Op welke manier worden relaties gemodelleerd?

Er bestaat een algemene procedure voor het opstellen van een **bivariate correlatie coëfficieënt**. Voordat een onderzoeker begint met het uitvoeren van analyses moet hij of zij checken of aan de **assumpties**voldaan is. Het gaat hier voornamelijk om de assumpties van **lineariteit en normaliteit.**

**Lineariteit** heeft betrekking op de validiteit van het model**. Normaliteit** is voor alleen kleine samples. Dit maakt alleen uit wanneer er sprake is van **significante testen**. Beide kunnen uitgevoerd worden aan de hand van grafieken. Voorbeelden van deze grafieken zijn:

* Scatterplots
* Q-Q plot
* P-P plot
* Een histogram

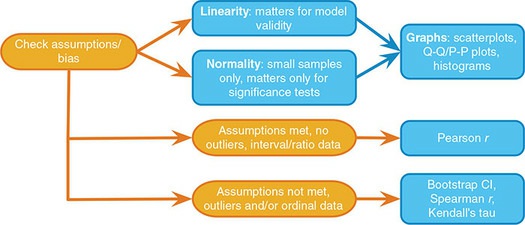
Er kan gebruik gemaakt worden van de **pearson r** wanneer:

* aan de assumpties voldaan is
* er geen outliers zijn
* de outkomst variabele gemeten wordt op interval of ratio niveau

Zo ja, dan kun je een Pearson correlatie uitvoeren. Als er uitschieters zijn, of het is niet normaal verdeeld, kun je kiezen voor:

* **En bootstrap betrouwbaarheidsinterval**
* **De Spearman correlatie**
* **Kendall’s tau (zie verderop)**

Er wordt gebruik gemaakt van de Bootstrap CI, Spearman r of de Kendall's tau wanneer niet voldaan is aan de assumpties, er sprake is van outliers en de variabelen gemeten worden op ordinaal niveau. Het rangschikken van de data maakt de kans op outliers kleiner. Er kan daarnaast gebruik gemaakt worden van de Bootstrap om de betrouwbaarheidsintervallen te bepalen.



De bivariate correlatie kan uitgevoerd worden in SPSS. Je gaat voor deze analyse in SPSS naar **analyze --> correlate --> bivariate.** De Pearson correlatie is automatisch door SPSS al aangevinkt. In **de dialog box** zijn de variabelen in **de data editor** aan de linkerkant weergegeven. Aan de rechterkant staat een leeg vlak genaamd **'variables'**. Je kunt hier zelf alle variabelen selecteren die je wilt gebruiken en met je muis naar dit lege vlak slepen.

SPSS creëert vervolgens een tabel van de correlatiecoëfficiënten voor alle combinaties van de specifieke variabelen. Deze tabel wordt **de correlatie matrix** genoemd. De correlaties staan allemaal dubbel in de matrix, dus de ene helft (diagonaal) kun je negeren. Significante correlaties geeft SPSS aan met een (**dubbele) asterisk (\*).**

In de output staat ook het **bootstrap betrouwbaarheidsinterval**. Als de variabelen niet normaal verdeeld zijn, kan hiernaar gekeken worden in plaats van naar de significantie van de variabelen. Als de 0 in het betrouwbaarheidsinterval zit, is er GEEN significante correlatie. De nul houdt in dat er geen sprake is van een effect. Valt de 0 niet in het interval is er wel sprake van een significante correlatie.

De pearson correlatiecoëfficiënt wordt bepaald door een algemene procedure te volgen in SPSS. Gebruik de test pearson en klik op OK. De matrix van resultaten wordt getoond in viewer. Er kunnen niet direct conclusies getrokken worden over de causaliteit van een correlatiecoëfficiënt. De interpretatie is belangrijk en men moet voorzichtig zijn met *causaliteit* aangezien een correlatie niet per se causaliteit betekent. Causaliteit bij correlaties kan voor twee redenen niet aangenomen worden namelijk:

* **Het derde variabele probleem**: Causaliteit tussen twee variabelen kan niet aangenomen worden omdat ook nog andere variabelen invloed kunnen hebben op de correlatie.
* **De richting van de causaliteit**: Een correlatie zegt niets over welke variabele de verandering bij de andere variabele teweegbrengt.

We gaan nog een stap verder door het te kwadrateren. Het kwadraat van de correlatie is de coëfficiënt van **determinatie R2 (R squared).** Het meet het percentage **gedeelde variantie**. Het is een maat voor hoeveel variantie de gecorreleerde variabelen delen. R2 wordt vaak gerapporteerd als dat een bepaald percentage variantie verklaard wordt door de andere variabele. Dit impliceert echter causaliteit, terwijl dat niet het geval hoeft te zijn bij een correlatie.

**Field Hoofdstuk 9 Deel 1**

9.7 Op welke manier wordt een lineair model in SPSS gebruikt?

Voordat de data geanalyseerd gaat worden is het handig als een onderzoeker eerst **een scatterplot** gaat maken van zijn of haar data om de assumptie van **lineariteit** te controleren en om te kijken of er sprake is van uitschieters. Tevens moeten de andere assumpties moet gecontroleerd worden wanneer onderzoekers willen dat het model gegeneraliseerd kan worden.

Bij **een regressie** staan de uitkomstvariabele en de voorspellers in verschillende kolommen. Daarnaast geeft elke rij de onafhankelijke waardes weer. Voor het uitvoeren van een regressieanalyse in SPSS ga je naar **analyze --> regression --> linear.** Bij **dependent** plaats je de uitkomstvariabele en bij i**ndependent(s)** plaats je de voorspeller(s). Bij **bootstrap** kun je een bootstrap betrouwbaarheidsinterval voor de regressiecoëfficiënten krijgen. Selecteer hiervoor **Bias corrected accelerated (BCa).**

9.8 Op welke manier wordt een lineair model met een voorspellende variabele geïnterpreteerd?

Bij de output vind je als eerste een tabel met de samenvatting die **R** en **R2** weergeeft. De R is bij simpele regressie de correlatie tussen de predictor en de uitkomst. Ook vind je een **ANOVA-tabel**. Daarin vind je de **MS** en de **F-ratio** en de **p-waarde** van die F-ratio. Bij significantie kunnen we zeggen dat het regressiemodel significant beter voorspelt dan alleen het gemiddelde.

De ANOVA-tabel zegt of het model in het algemeen een goede voorspelling geeft van de uitkomstvariabele, maar vertelt niets over de individuele bijdrage van de variabelen in het model. Bij simpele regressie heb je maar één variabele, dus als het een goed model is, was het ook een goede predictorvariabele.

De schattingen van de parameters in het model vind je in de tabel **coefficients.** **b0** is terug te vinden in de SPSS tabel bij **B (constant).** b1 staat ook onder B bij de predictorvariabele. Deze waarde geeft dus aan met hoeveel de uitkomstvariabele verandert als de predictor verandert met een bepaalde eenheid. Er staat een t-toets achter, die test of de b-waarde significant afwijkt van 0.

In de tabel bootstrap staat het bootstrap betrouwbaarheidsinterval. Bij BCa 95% betrouwbaarheidsinterval staan de grenzen waartussen de b-waardes waarschijnlijk liggen. Als 0 niet in dit interval zit, is er een significante relatie tussen de variabelen.

Met dit model kun je voorspellingen doen. Dit doe je door de b-waardes in te vullen in de standaard regressieformule **(y = a + bx).** Nu kun je een X-waarde invullen om een bepaalde uitkomst te voorspellen.

9.10 Op welke manier past een lineair model met twee of meer voorspellende variabele in SPSS?

Eerst kan er een scatterplot gemaakt worden om te kijken naar de relaties tussen de uitkomst variabele en de voorspellende variabelen. Het is belangrijk dat ook bij **een multipele regressie** voor de analyse de voorwaarden gecontroleerd worden. De regressieanalyse kan ook hier uitgevoerd worden via **analyze – regression – linear.**

Voor de hiërarchische methode heeft SPSS blokken voor elke stap. Bij next kom je in een volgend blok. Bij verschillende blokken is het mogelijk verschillende methoden te kiezen, in het afrolmenu bij **Method**. Je kunt twee of meer voorspellers tegelijk invoeren door de control toets ingedrukt te houden terwijl je de variabelen selecteert en naar Independents sleept.

Bij de optie statistieken kan een onderzoeker een aantal opties kiezen:

* **Estimates**: Geeft de geschatte b-waardes samen met de t-toets weer.
* **Confidence intervals**: Geeft de betrouwbaarheidsintervallen van de ongestandaardiseerde regressiecoëfficiënten weer.
* **Covariance matrix**: produceert een matrix van de covarianties, de correlatiecoëfficiënten en de varianties van de regressiecoëfficiënten van de variabelen in het model.
* **Model fit:** Deze heb je altijd nodig, en is ook al standaard aangevinkt. Het geeft de F-waarde en R weer.
* **R squared change**: Geeft de verandering van R2 weer wanneer er een nieuwe voorspeller bijkomt.
* **Descriptives:** Geeft de standaardafwijking, het gemiddelde en de n voor alle variabelen. Geeft bovendien een correlatiematrix, die nuttig is om multicollineairiteit te checken.
* **Part en partial correlations**: Geeft de zero order correlaties, de gewone bivariate Pearson correlatie, tussen de voorspellers en de uitkomstvariabele. Daarnaast geeft het de part en partial correlaties van de voorspellers.
* **Collinearity diagnostics**: Geeft VIF of tolerantie weer.
* **Durbin-Watson toets**: Toetst de assumptie van onafhankelijke meetfouten.
* **Casewise diagnostics:** Geeft de geobserveerde uitkomstwaarde, de voorspelde uitkomstwaarde, het verschil hiertussen (het residu) en het gestandaardiseerde residu.

Bij plots kan een onderzoeker verschillende grafieken maken. De meeste plots zijn ook residuals betrokken die eerder besproken zijn. Aan de linker kant staan er diverse variabelen:

* **Dependent:** dit is de uitkomst variabele.
* **\*ZPRED:** dit is de gestandaardiseerde voorspelde waarden van de uitkomst gebaseerd op het model.
* **\*ZRESOD:** dit zijn de gestandaardiseerde residuals of errors.
* **\*DRESID:** dit zijn de verwijderde residuals.
* **\*ADJPRED:** dit zijn de aangepaste voorspelde waarden.
* **\*SRESID:** dit zijn de stundentized residuals.
* **\*SDRESID:** dit zijn de studentized verwijderde residuals.

Bij **save** kan aangeven worden welke diagnostieken een onderzoeker er allemaal erbij wil hebben. Dit kan gedaan worden om bijvoorbeeld invloedrijke scores te ontdekken. Bij options kan een onderzoeker de stapmethode voor stapsgewijze regressie kiezen en aangeven hoe je met missende waarden omgaat. **Exclude cases listwise** betekent dat als een persoon op één variabele een missende waarde heeft, deze persoon uit de analyse wordt gehaald.

**Morling et al. Pagina 499 & 500**

Is de correlatie significant? Inferentiële statistieken met *r*

De correlatiecoëfficiënt is een descriptieve statistiek omdat het de richting en sterkte van een relatie tussen twee numerieke variabelen beschrijft. We kunnen ook de statistische significantie van een correlatiecoëfficiënt evalueren. Bij *t* en *F* gebruiken we de vier stappen van **NHST**:

1. Aannemen dat er geen relatie is in de populatie
2. Dataverzameling en *r* berekenen
3. Schatten de probabiliteit van het verkrijgen van de *r*, of een die meer extremer is, als de nulhypothese waar is
4. We beslissen of we de nulhypothese verwerpen of verhouden

Verwerpen we de nulhypothese? Dan nemen we aan dat onze *r* statistisch significant was.

Sampling distributie van *r*

Net als *t* en *F* heeft de correlatiecoëfficiënt sampling distributies, die gebaseerd is op de waarschijnlijke waarden van *r* die we zouden krijgen als we de studie meerdere keren met random samples van een populatie zouden uitvoeren, en waarin de nulhypothese waar is (hier zouden de waarden rondom de 0.00 liggen). Als de nulhypothese waar is met een grote sample (>30), lijkt de sampling distributie van *r* erg op de distributie van *t*. (dit geldt andersom voor een grote *r*).

De exacte *p-waarde* die uitgerekend wordt door de computer bij een *r-waarde* vertelt on de waarschijnlijkheid van het verkrijgen van die *r* of een extremere waarde als de nulhypothese waar is. wanneer de *p-waarde* kleiner is dan alfa (meestal <0.05) is het resultaat statistisch significant.

Sample size en *r*

Grote *r* is waarschijnlijker statistisch significant. Maar, dit hangt ook af van de sample size. *R* kan dus ook de meting aannemen van de effectgrootte. Bijvoorbeeld, *r* kán statistisch significant zijn (in een grote sample bijvoorbeeld), maar representeert nog steeds een hele zwakke of kleine effectgrootte.

Is die regressiecoëfficiënt significant? Inferentiële statistieken voor beta

ß is een waarde verkregen van het multipele regressie proces, die lijkt op *r*, omdat het gebruikt is als schatting van de associatie tussen twee variabelen. Maar, ß representeert vaak de relatie tussen een voorspeller (onafhankelijke) variabele, en een afhankelijke variabele, controlerend op andere variabelen.

Na het vinden van de *t-waarde* geassocieerd met de ß, kunnen we de sampling distributie van *t* gebruiken om de statistische significantie van die ß te evalueren (de sampling distributies van de *t* die gebruikt zijn voor deze evaluatie zijn hetzelfde als die voor het evalueren van het verschil tussen twee samples). Daarnaast geldt hetzelfde als voor de NHST, het vier stappenplan.

Een multipele regressie tabel in een empirisch artikel heeft vaak een kolom die een indicatie geeft of elke beta statistisch significant is, met deze uitkomsten:

* Actuele *t-waarde* die geassocieerd is met die bèta, naast de *p-waarde* geassocieerd met die *t*
* Alleen de *p-waarde* die geassocieerd is met de bèta en zijn *t*
* Alleen een asterisk (\*), die aangeeft dat de bèta statistisch significant is
* De letters n.s., die aangeven dat de bèta niet significant is

**Field Hoofdstuk 18**

18.9 Hoe wordt een betrouwbaarheidsanalyse uitgevoerd?

Het is belangrijk **de betrouwbaarheid** van de schaal te controleren wanneer een **factoranalyse** gebruikt is om vragenlijsten te valideren. Betrouwbaarheid betekent dat een meetinstrument consistent is in het meten van het construct. Dit betekent bijvoorbeeld dat wanneer een proefpersoon dezelfde vragenlijst tweemaal invult op verschillende momenten, dat deze dezelfde uitslag geven. Dit heet test-hertest betrouwbaarheid. Het betekent dat twee proefpersonen die hetzelfde zijn op het gemeten construct ook hetzelfde moeten scoren.

**De split-half betrouwbaarheid** splitst de schaal random in twee sets van items. Bij een betrouwbare schaal heeft een proefpersoon ongeveer dezelfde score op beide helften. Er wordt dan gekeken naar de correlatie tussen de twee helften. Als die correlatie voor veel personen hoog is, betekent het dat de schaal betrouwbaar is. Het nadeel van deze methode voor het meten van de betrouwbaarheid is dat de schaal op veel manieren in tweeën gesplitst kunnen worden en dat er dus verschillen uitkomsten kunnen zijn.

*Cronbach* heeft daarom een methode bedacht waarbij de betrouwbaarheid wordt gemeten door alle mogelijke split-half sets van items te berekenen en de correlatiecoëfficiënt van elke splitsing te berekenen. Het gemiddelde van deze mogelijke split-halfs wordt Cronbach’s α genoemd, de meest gebruikte statistiek om de betrouwbaarheid te meten. De volgende formule kan gebruikt worden:

**α = (N2 \* cov) / (Σs2item + Σcov item)**

Er is ook een gestandaardiseerde versie van **Cronbach’s α** en die is handig wanneer de items voor de optellen ook al gestandaardiseerd zijn. Bij een waarde van 0.7 of 0.8 of hoger wordt gezegd dat de betrouwbaarheid goed is. Maar dit kan niet zomaar aangenomen worden omdat α ook afhankelijk is van het aantal items op de schaal. Hoe meer items een schaal bevat, hoe hoger Cronbach’s α.

Vaak wordt gedacht dat α un**idimensionaliteit** meet: of de schaal één onderliggend construct heeft. Dit is echter onjuist. Dezelfde α kan verkregen worden bij verschillende factorstructuren. Als er meerdere factoren zijn moet alpha worden berekend over deze subschalen.

Bij vragenlijsten zitten er vragen tussen die andersom geformuleerd zijn om de proefpersonen wakker te houden. Voor factoranalyse maakt dit niet uit, je krijgt alleen een negatieve factorlading in plaats van een positieve. Voor Cronbach’s alpha maakt het wel uit, andersom geformuleerde vragen zorgen voor een kleinere waarde. Er kan zelfs een negatieve alpha uitkomen. In dit geval moet je nakijken of je omgekeerde vragen hebt. Deze vragen moeten dan **omgepoold** worden, andersom gecodeerd.

18.10 Hoe wordt een betrouwbaarheidsanalyse uitgevoerd in SPSS?

Voor de betrouwbaarheidsanalyse in SPSS ga je naar **Analyze –-> Scale -–> Reliability Analysis.** Sleep de vragen van de (sub)schaal naar het vak **Items**. Bij **Model** kan je de methode voor de betrouwbaarheid kiezen, maar Cronbach’s Alpha is de standaard optie. Aan de (sub)schaal kan een naam gegeven worden in Scale label.

Bij **Statistics** is de belangrijkste optie **Scale if item deleted**. Het geeft voor elke vraag aan wat Cronbach´s alpha zou zijn als dit item verwijderd wordt. Als een bepaald item een groot verschil maakt in betrouwbaarheid, moet je overwegen om deze vraag te verwijderen.

De **intraclass correlatiecoëfficiënt (ICC)** kan ook gebruikt worden. De correlatiecoëfficiënten die eerder in het boek zijn besproken berekenen de relatie tussen variabelen die verschillende dingen meten. ICC meet daarentegen de relatie tussen twee variabelen die hetzelfde meten. Het wordt veel gebruikt bij het vergelijken van gepaarde data op hetzelfde meetinstrument en het meten van de consistentie tussen juryoordelen.

18.11 Hoe ziet de output van een betrouwbaarheidsanalyse in SPSS eruit?

**De output** begint met een kleine tabel genaamd **Reliability Statistics**, waar Cronbach’s Alpha vermeld staat. Hoewel de richtlijnen dus niet blindelings moeten worden toegepast, is een score hoger dan 0.8 doorgaans goed.

In de tabel **Item-Total Statistics** is de gecorrigeerde item-totaal correlatie de correlatie tussen elk item en de totale schaal. Wanneer items niet goed correleren met de schaal (onder de .3), kan overwogen worden deze items te verwijderen. Bij Cronbach’s α if item is deleted wordt de alpha gegeven als het item niet meegenomen wordt in de berekening.

Als de alpha dan hoger is betekent het dat de betrouwbaarheid groter zal zijn als die vraag niet in de schaal zit. Er moet dan overwogen worden de vraag te verwijderen. Als een vraag een negatieve item-total correlatie heeft, kan het zijn dat deze vraag andersom geformuleerd is en nog niet is omgepoold.

18.12 Hoe wordt een betrouwbaarheidsanalyse gerapporteerd?

**De betrouwbaarheidanalyse** wordt gerapporteerd door de waarde van alpha te geven. Geef de waarde net als bij de correlatie zonder de nul ervoor (bijvoorbeeld 0.7).

**Morling et al. Pagina 244 t/m 256**

9.1 – Hoe kunnen multiple-regressie designs derde variabelen buiten sluiten?

In een onderzoek kwam naar voren dat het hebben van lange pauzes (meer dan 15 minuten) samenhing met beter gedrag in kinderen. Maar wat is de causale link? Wat kwam er eerst? Gedragen kinderen zich beter doordat ze langer pauze hebben gehad of worden brave kinderen beloond met een langere pauze? Daarnaast moet je je ook afvragen of er **derde variabelen** aanwezig zijn. Het zou zomaar kunnen dat er een of meerdere variabelen zijn die een invloed hebben gehad op de relatie tussen de duur van een pauze en goed gedrag. Met **multipele-regressie analyses** kun je enkele derde variabelen uitsluiten. Dit maakte het onderzoek een **multivariate correlationeel onderzoek**.

Met **multivariate designs** kunnen onderzoekers kijken of een relatie tussen twee variabelen blijft bestaan wanneer een derde variabel constant wordt gehouden. Je zou zo’n derde variabele op kunnen splitsen in verschillende subgroepen. Stel je neemt inkomen van ouders als derde variabele. Je zou dit op kunnen splitsen in laag inkomen, middeninkomen en hoog inkomen. Vervolgens kun je kijken of de relatie tussen gedragsproblemen en duur van pauzes blijft bestaan in elk van deze subgroepen.

9.2 – Welke statistische maten worden er gebruikt in multiple-regressie designs?

In deze designs wordt er gekeken naar drie of meer variabelen. Eerst moet men beslissen welke variabele het meest interessant is. Dit wordt de **afhankelijke variabele**of **criterion variabele** genoemd. In het onderzoek over pauzes en probleemgedrag waren onderzoekers het meest geïnteresseerd in probleemgedrag. De andere variabelen worden **onafhankelijke variabelen**of **predictor variabelen** genoemd.

Wanneer je in SPSS **een regressie** laat uitvoeren, krijg je ook een **regressie tabel.** In je regressie tabel moet je kijken naar **de bèta -waarden**. **Bèta** laat de richting en sterkte van de relatie tussen predictor en criterion variabele zien, tijdens het constant houden van de andere predictor variabelen (heel belangrijk). Het lijkt op een r,maar het voegt nog een extra dimensie toe. Een negatieve bèta duidt op een negatieve relatie en een positieve bèta duidt op een positieve relatie. Een hoge waarde betekent dat de relatie ook sterker is dan een lage waarde. Een bèta is **gestandaardiseerd** en de eenheden van de verschillende predictor variabele (bijvoorbeeld euro, minuten en centimeter) zijn allemaal gestandaardiseerd tot een maat.

Een bèta waarde kan veranderen wanneer andere predictor variabelen worden toegevoegd. Daarnaast staat er vaak in een kolom naast de bèta waarden ook wat de significantie en de p-waarde is van de bèta. Wanneer de p gelijk of hoger is dan .05, dan is bèta niet significant. Dat betekent dat de gevonden associatie tussen een predictor variabele en criterion variabele toevallig gevonden is in het onderzoek en waarschijnlijk niet bestaat in de populatie.

Wat als je naar meerdere variabelen kijkt die een invloed kunnen hebben op de relatie tussen een criterion en predictor variabele? Dan gelden dezelfde regels weer voor bèta. De bèta-waarde van een variabele zegt al iets over de relatie tussen die predictor variabele en criterion variabele, gecontroleerd voor de andere predictor variabelen die in het model opgenomen zijn. Het is handig om meerdere predictor variabelen in een onderzoek toe te voegen, zodat je met meer zekerheid kan stellen (of niet) dat een relatie niet beïnvloed wordt door een derde variabele. Wat ook handig is aan het toevoegen van meerdere predictor variabelen, is dat je aan de grootte van de bèta-waarden kunt zien welke factoren een sterkere invloed hebben op de afhankelijke variabelen. Kijk dus goed naar de bèta en haal de bèta niet door elkaar met de **ongestandaardiseerde b.** Dat is een waarde die ook vaak in een regressietabel weergegeven wordt, maar die dus kijkt naar waarden die NIET gestandaardiseerd zijn. Je zou dus niet elke variabele met elkaar kunnen vergelijken aan de hand van b, want je kunt niet euro’s met centimeter of minuten vergelijken.

In populaire tijdschriften of kranten worden vaak ook resultaten van een onderzoek uitgelicht. Vaak is het wel zo dat termen zoals ‘bèta,’ ‘p’ en ‘significantie’ niet genoemd worden. Toch kun je aan de hand van enkele termen zien dat het gaat om een multipele regressie. Termen zoals ‘**controleren voor** andere variabelen,’ ‘rekening houden met andere variabelen,’ en ‘correctie voor andere variabelen’ laten zien dat een multipele regressie in het onderzoek werd gebruikt.

9.3 – Kan regressie causaliteit vaststellen?

Ook al voeg je 20 variabelen toe die als potentiële derde variabelen gezien kunnen worden, dan betekent het niet dat je aan alle voorwaarden voor **causaliteit** hebt voldaan. Multipele regressie designs kunnen bepaalde derde variabelen uitsluiten, maar ze kunnen niet **temporele precedentie** vaststellen. Daarnaast kunnen ze ook niet controleren voor derde variabelen die niet opgenomen zijn in het onderzoek. Het kan voor komen dat onderzoekers er zich niet bewust van zijn dat er een bepaalde variabele is die invloed uit zou kunnen oefenen op de relatie tussen de criterion en predictor variabele. Deze variabele zal dan niet opgenomen worden in het onderzoek en de conclusie die men trekt aan de hand van de resultaten van het onderzoek zal vertekend zijn. Het probleem met potentiële derde variabelen kan eigenlijk alleen maar opgelost worden door experimenten uit te voeren. Door het random toeschrijven van proefpersonen aan bepaalde condities, neem je derde variabelen weg. Alleen experimenten kunnen causaliteit vaststellen.

**Morling et al. Pagina 264 t/m 267**

9.4 – Hoe zit het met de vier validiteiten in multivariate designs?

**Multipele-regressie analyses** helpen met het derde variabele probleem, **longitudinaal onderzoek** stelt temporele precedentie vast en **multivariate designs** hebben dus enig bewijs voor **interne validiteit.**

Voor multivariate designs is het ook belangrijk om de **construct validiteit** te onderzoeken door te kijken naar hoe goed elke variabele gemeten was.

Om de **externe validiteit** te onderzoeken, kan men kijken naar de proefpersonen. Zijn deze random gekozen? Zijn er personen uit verschillende lagen van de bevolking gebruikt of hebben onderzoekers bijvoorbeeld alleen maar de mensen met een laag inkomen onderzocht?

Daarnaast kan **statistische validiteit** onderzocht worden door te kijken naar de statistische data die de onderzoekers hebben aangedragen. Hoe zit het met de **effectgrootte en significantie**? Ook moet er gekeken worden naar **uitbijters en curvilineaire relaties**.

9.5 – Samenvatting hoofdstuk

*Reviewen van de drie causale criteria*

* In een multivariate design meten onderzoekers meer dan twee variabelen en kijken naar de relaties tussen deze.
* Een simpele, bivariate correlatie geeft aan dat er covariantie is, maar kan niet altijd aangeven dat er temporele precedentie of interne validiteit is, dus kan het geen causatie vaststellen.

*Vaststellen van temporele precedentie met longitudinale designs*

* Longitudinale designs starten met twee sleutel variabelen, op welke dezelfde groep mensen gemeten worden op meerdere punten in de tijd. Onderzoekers kunnen zien welke variabele eerst kwam, wat het vaststellen van temporele precedentie helpt.
* Longitudinale designs produceren **cross-sectionele correlaties** (correlaties tussen de twee sleutel variabelen op een moment in de tijd) en **autocorrelaties** (correlaties tussen een variabele en zichzelf, gedurende de tijd)
* Longitudinale designs produceren ook **cross-lag correlaties**. Door het vergelijken van de relatieve sterktes van de twee cross-lag corrrelaties, kunnen onderzoekers infereren welke van de variabelen waarschijnlijk het eerst kwam (of dat ze elkaar wederzijds versterken)

*Oorzakelijk worden met patroon en spaarzaamheid*

* Onderzoekers kunnen causale zekerheid benaderen door patroon en spaarzaamheid; ze specificeren een mechanisme voor de causale relatie en combineren de resultaten van een variatie van onderzoeksvragen. Wanneer een enkele causale theorie alle ongelijksoortige resultaten verklaart, zijn onderzoekers dichter bij het ondersteunen van een causale claim.

*Bemiddeling*

* In een bemiddelingshypothese specificeren onderzoekers een variabele die tussen twee interessante variabelen komt als een mogelijke reden dat de twee variabelen een associatie hebben. Na het verzamelen van data van alle drie de variabelen (de originele twee, en de bemiddelaar), volgen ze specifieke stappen om te evalueren hoe goed de data de bemiddelingshypothese ondersteunt.

**Field Hoofdstuk 9 deel 2**

9.2 – Hoe ziet een simpel regressie model eruit?

Om de relatie tussen twee variabelen te bekijken kan gebruik gemaakt worden van de volgende vergelijking: **outcome i = (b1 \* Xi) + error i.**

Als er wordt gewerkt met ruwe data is het belangrijk extra informatie toe te voegen over de uitkomst variabele (= de afhankelijke variabele). **Een constante variabele b0** kan worden toegevoegd. Dit wordt de intercept genoemd. Deze variabele geeft de waarde van de uitkomst wanneer de voorspellende variabele niet aanwezig is. In andere woorden: de voorspellende variabele heeft een waarde van 0. Dit resulteert in het volgende model:

**Outcome i = (b0 + b1 \* Xi) + error --> Yi = (b0 + b1 \* Xi) + e i.**

Voor Nederlanders, Brazilianen en Fransen wordt het de volgende vergelijking: **outcome i = ax + b.**

De bovenstaande vergelijkingen zijn vormen van **de straight line modellen**. **De slope** van de lijn wordt aangegeven met de letter **b1** en de **intercept** wordt aangegeven met de letter b0. De slope wordt in het Nederlands **de helling** genoemd. De slope geeft aan hoe het model eruitziet. De intercept is het snijpunt waar de lijn de x-as raakt. De intercept geeft aan waar het model zich bevindt in een grafiek.

De letters b0 en b1 zijn de **regressie coëfficiënten** die gebruikt worden in dit boek. De letter i is het **subscript**. Hiermee wordt bedoeld dat het een relatie heeft met variabele i. Als deze regressiecoëfficiënten bekend zijn, kan de uitkomst voorspeld worden. Er bestaat echter altijd een error. De error houdt in dat de voorspelling niet 100% accuraat is. Het is namelijk een voorspelde waarde.

Het is ook mogelijk een lineair model te hebben met verschillende/ meerdere voorspellende variabelen. In dit geval zijn er meerdere onafhankelijke die de uitkomst voorspellen. Wanneer er meerdere onafhankelijke variabelen zijn, is er sprake van een multipele regressie. Deze onafhankelijke variabelen hebben in deze situatie elk een eigen regressiecoëfficiënt. Het model zit er als volgt uit wanneer er twee voorspellende variabelen zijn:

**Yi = (b0 + b1 \* X1i + b2 \* X2i) + ei.**

Het model geeft een voorspelde waarde die iets afwijkt van de werkelijke data. Het verschil tussen de door het model voorspelde waardes en de werkelijke waardes wordt de error of **de residuen (residuals)** genoemd. Er bestaat een formule voor de totale error. Deze is de volgende:

**Totale error = Σ (geobserveerdi – modeli)2**

Het verschil tussen de voorspelde waarde en de werkelijke waarde kan zowel positief als negatief zijn. Als ze bij elkaar opgeteld worden kunnen ze weggelaten worden. Daarom worden ze gekwadrateerd. Omdat in een regressie gesproken wordt over residuen in plaats van een error, wordt dit **de sum of squared residuals** of **de residual sum of squares** of **residuensom** aangeduid met **SSr.** Als de SSR klein is, past de lijn goed bij de data. Als de SSR groot is, is de lijn helemaal niet representatief voor de data.

Het model dat het best past bij de data, is het model dat de kleinste SSr oplevert. Dit wordt gedaan met **de ordinary least squares method** waarbij de computer **de parameters (de b’s in je model) s**chat die de kleinste residuensom oplevert. Als je het beste model hebt, moet je kijken naar hoe goed het model is ‘the goodness of fit’.

**De goodness of fit** is de mate waarin het model past in de verzamelde data. De beste lijn kan immers nog steeds een hele slechte fit hebben met de data.

Om voorspellingen te maken hebben we het model, dus de coëfficiënten nodig dat het beste bij de gegevens past. De SSR geeft aan hoe groot de error is. Het kwantificeert de error. Het is echter belangrijk een vergelijkpunt te hebben om te kijken of het model beter is dan niets. Als je geen model hebt, is je gemiddelde de enige nuttige schatting die je kunt maken. **Het gemiddelde** wordt dus gebruikt als model waarbij er geen relatie is tussen de variabelen. Als het gemiddelde van een model genomen wordt kunnen de verschillen uitgerekend worden tussen de data en het gemiddelde. Als gevolg hiervan kan een sum of squares uitrekenen. Dit heet **de totale kwadratensom (SST),** de totale hoeveelheid verschillen wanneer het basismodel, het gemiddelde, wordt toegepast op de gegevens.

De residuensom (SSR) is het verschil tussen het regressiemodel en de data, dus de error wanneer het best mogelijke model op de data wordt toegepast. Het verschil tussen de SST en de SSR is de verbetering in voorspelling die het model biedt boven het gemiddelde. Deze verbetering is **de sum of squares van het model, de SSM.**

Een grote SSM betekent dat het regressiemodel erg verschillend is van alleen het gemiddelde gebruiken om de uitkomstvariabele te voorspellen. Met deze kwadratensommen kun je de proportie van verbetering uitrekenen:

**R2 = SSm / SSt**

**R2** is hier de proportie van verbetering en als je het vermenigvuldigt met honderd krijg je er een percentage uit. R2 is de proportie verklaarde variantie dat verklaard wordt door het model tegenover hoeveel variantie er in totaal is. Als je hieruit de wortel trekt krijg je de Pearson correlatie.

Een tweede manier om de kwadratensom te gebruiken is voor het berekenen van de **F-toets**. De F-toets is gebaseerd op de ratio van de verbetering door het model (SSM) en het verschil tussen het model en de geobserveerde gegevens (SSR). Omdat de sum of sqaures afhangen van de steekproefgrootte, wordt de gemiddelde kwadratensom (**mean square, MS**) gebruikt.

Om de mean square te bepalen wordt de SS gedeeld door het aantal v**rijheidsgraden**. Voor SSM is het aantal variabelen het aantal vrijheidsgraden. Wanneer er sprake is van een simpele regressie is **dfM** gelijk aan 1. Voor SSR is in dit geval het aantal vrijheidsgraden het aantal observaties min het aantal. De F-ratio is een meting van hoeveel het model de voorspelling van de uitkomst verbetert in vergelijking met de onnauwkeurigheid in het model.

**MSm = SSm / k**

**MSr = SSr / (N – k – 1)**

**F = MSm / MSr**

**F = (N – k – 1) \* R2 / k \* (1 – R2)**

Een goed model heeft een grote F-waarde. Dit betekent dat er een grote verbetering is in de voorspelling (een grote MSM) en dat het verschil tussen de voorspelling en de data klein is (een kleine MSR).

Elke predictor in het regressiemodel heeft een coëfficiënt (b). In simpele regressie is er sprake van slechts één predictor is. In dit geval is b de richtingscoëfficiënt van de regressielijn. De b geeft dan de verandering in uitkomst aan die komt door de verandering in de predictor. Als het gemiddelde gebruikt wordt als model, is er geen verandering in de uitkomst bij een verandering in de predictor. Wat de waarde van de predictor ook is, de geschatte uitkomstwaarde is altijd het gemiddelde. Hierbij heeft de richtingscoëfficiënt (b) een waarde van 0. De regressielijn is horizontaal.

Als een variabele significant een uitkomst wil voorspellen, dan moet het dus een b-waarde hebben die significant verschilt van 0. Dit kan getoetst worden met de t-toets. Hierbij toets je de nulhypothese dat de waarde van b gelijk is aan 0. De t-toets is net als de F-toets een ratio van de verklaarde variantie tegen de onverklaarde variantie van meetfouten. Om error van de b te schatten, wordt de standard error gebruikt. De formule wordt als volgt beschreven: **t = (b geobserveerd - b verwacht) / SEb**

Vanuit de nulhypothese wordt ervan uitgegaan dat b verwacht = 0. De formule wordt vervolgens als volgt geformuleeerd: **t = b geobserveerd / SEb**

Bij regressie is het aantal vrijheidsgraden **N – p – 1.** N is hierbij de totale steekproefgrootte en p is het aantal predictors. Bij simpele regressie wordt het als volgt: **df = N – 2**. Als er sprake is van een grote t-waarde die groter is dan de kritieke waarde (zie de tabel in de appendix) wordt de nulhypothese verworpen. In dit geval wijkt b dan significant af van 0. De predictor heeft dan een significante bijdrage in het voorspellen van de uitkomst.

9.3 – Wat wordt bedoeld met een bias in de regressiemodellen?

Nadat het model gemaakt is, zijn er twee belangrijke vragen die gesteld moeten worden:

* Wordt het model beïnvloed door een klein aantal gevallen?
* Kan het model gegeneraliseerd worden naar andere steekproeven?

Voor het beantwoorden van de eerste vraag kan gekeken worden naar uitschieters en invloedrijke gevallen. **Uitschieters (= outliers)** kunnen de schattingen van de parameters in het model sterk beïnvloeden. Uitschieters zijn te herkennen aan een groot residu. Uitschieters zorgen voor een grote afwijking van de trend. Outliers zorgen ervoor dat de lijn platter wordt en dat de intercept toeneemt. In dit geval wordt respectievelijk b1 kleiner en b0 groter.

Residuen laten de meetfout (errors) in het model zien. **De ongestandaardiseerde residuen** (de normale residuen) zijn in dezelfde schaal als de uitkomstvariabele gemeten en zijn dus moeilijk te gebruiken in andere modellen. **Gestandaardiseerde residuen** zijn residuen die tot **z-scores** zijn getransformeerd en kunnen bij meerdere modellen gebruikt worden als standaard. Het voordeel is dat er voor deze gestandaardiseerde residuen richtlijnen zijn over welke residuen acceptabel zijn en welke niet.

Er is ook nog **de studentized residu** die varieert van punt tot punt. Ze hebben dezelfde proporties als de gestandaardiseerde residuen alleen geven ze een iets preciezere schatting van de meetfout in een specifiek geval.

Soms bestaan er een of meerdere scores die de schattingen van de parameters van het model enorm beïnvloeden. Met verschillende statistieken kunnen deze scores gevonden worden. De aangepaste voorspelde waarde is zo’n statistiek. Deze wordt ook wel **de adjusted predicted value** genoemd. Deze techniek berekent het model zonder een bepaalde score. Wanneer de score weinig invloed heeft zal de aangepaste voorspelde waarde ongeveer overeenkomen met de voorspelde waarde.

**Het verwijderde residu (= the deleted residual)** is het verschil tussen de aangepaste voorspelde waarde en de geobserveerde waarde. Als dit residu gedeeld wordt door de standaarddeviatie krijg je de **studentized verwijderde residu.**

Deze techniek van het verwijderde residu geeft alleen aan hoeveel invloed de score heeft op hoe goed het model die ene score kan voorspellen. Het geeft geen informatie over hoe die score het hele model beïnvloed. **De Cook’s distance** kijkt wel naar het effect van één score op het gehele model. Wanneer de Cook’s distance groter is dan 1 is er reden tot bezorgdheid.

Een tweede manier om de invloed te meten is **leverage (hoofd waardes)** die de geobserveerde waardes boven de voorspelde waardes stelt. Het kan uitgerekend worden met **(k + 1) / n.** In dit geval staat k voor het aantal voorspellers en staat n voor het aantal gevallen. Er zijn ook nog **de Mahalanobis distances** die de afstand meet van de scores tot het gemiddelde van de predictorvariabele. Je kijkt naar de gevallen met de hoogste waardes. Hier hebben de afstanden een chi-square distributie waarbij het aantal vrijheidsgraden gelijk is aan het aantal voorspellende variabelen.

Je kunt een regressieanalyse uitvoeren met een score erbij en daarna zonder die score. Vervolgens kan dan gekeken worden hoe groot het verschil is in regressiecoëfficiënten. Het verschil tussen een schatting van de parameters met alle scores en de schattingen waarbij een score is verwijderd, heet de **DFBeta**.

Om geen invloed van de schalen te hebben wordt **de gestandaardiseerde DFBeta** gebruikt. Waardes boven de 1 hebben veel invloed op de modelparameters.

Een gerelateerde statistiek is de **DFFit**, het verschil tussen de voorspelde waarde voor een score wanneer het model is berekend met die score erbij en wanneer het model is berekend zonder die score. Wanneer een score geen invloed heeft, is de DFFit 0. Ook hier is een gestandaardiseerde DFFit mogelijk.

Als laatste is er ook nog de **covariantie ratio (CVR)**. Dit meet of een score invloed heeft op de variantie van de regressieparameters. Een score heeft weinig invloed wanneer de waarde van de CVR dicht bij 1 ligt.

**Als CVRi > 1 + [3 \* (k – 1) / n] – het verwijderen van de i-de case tast de precisie van sommige parameters in het model aan.**

**Als CVRi < 1 + [3 \* (k – 1) / n] – het verwijderen van de i-de case zorgt voor een verbetering van de precisie van sommige parameters in het model.**

In beide ongelijkheden staat k voor het aantal voorspellers, staat n voor de omvang van de steekproef en is CVRi de covariantie ratio voor de i-de deelnemer.

De hierboven genoemde statistieken zijn een manier om te kijken hoe goed het model bij de gegevens past. Het is niet een manier om te kijken welke punten handig zijn om erin te houden of erbuiten te laten om een b-waarde significant te laten zijn.

9.14 – Op welke manier worden lineaire modellen gerapporteerd?

**De summary tabel** kan gebruikt worden om te rapporteren. Deze omvat het meeste. Het minimale wat gerapporteerd moet worden zijn de beta's met de bijbehorende standaardafwijkingen en de betrouwbaarheidsintervallen. Het is ook belangrijke sommige statistieken te rapporteren die iets zeggen over de 'fit' van het model. Neem bijvoorbeeld de R2of de Bayes factor. Voor een hiërarchische regressie is het belangrijk dat de waardes per fase gerapporteerd moeten worden.