

Hoofdstuk 1

De logica van statistische vergelijkingen en analyses

‘Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write.’

-H. G. Wells

1. Inleiding: waarom data analyseren?

Ontdekking en vernieuwing zijn allicht de grote verschillen tussen de moderne bedrijvigheid en die van onze voorvaders. Sinds de Renaissance is er een nadruk komen te liggen op het leren en verbeteren van de wijze waarop we dingen doen. Wetenschappers, uitvinders en anderen die betrokken zijn in het proces van wetenschappelijk onderzoek, hebben in het verleden vaak eerbied gekregen voor hun werk. Zo worden in talloze schoolboeken onder meer Galileo, Newton, Einstein en Madam Curie nog steeds geprezen voor hun werk en uitvindingen. Zoals jullie zullen leren in deze cursus, zou dit ook zo moeten zijn voor mensen zoals Pearson en Kendall. En voor onze eigen Adolphe Quetelet, die de normaalverdeling ontdekte voor menselijke eigenschappen, zich bezig hield met statistiek als hulpmiddel om de moraliteit, de zeden en gewoonten van gebieden te bestuderen en op basis daarvan beleidsbeslissingen te nemen. Quetelet wist al heel vroeg dat mensen verschilden in de mate waarin zij betrokken zijn bij regelovertredend gedrag en dat gebieden verschilden in de mate waarin zij concentraties kenden van criminaliteit.

Wetenschap gaat onder meer over het **ontdekken van patronen en processen**. Veel patronen komen niet zomaar uit de lucht gevallen maar zijn het gevolg van het feit dat mensen nu eenmaal gewoontedieren zijn en een menselijke natuur hebben, die in samenspraak met culturele omstandigheden, attitudes, denkprocessen, morele emoties, overtuigingen over wat goed en verkeerd is, ... beïnvloeden. Het proces van **wetenschappelijk onderzoek** voorziet in een methode om zaken die ons op een systematische manier interesseren, te bestuderen. In het algemeen vereist dit proces bewijs om een argument te ondersteunen. Eén van de duidelijkste methoden om bewijs

te verkrijgen is door kwantitatieve data te bestuderen die geassocieerd zijn met objecten die bestudeerd worden. Dit gebeurt door statistische analyse.

2. Geschiedenis van de statistiek in een notendop

De vroegste vorm van wat nu gezien wordt als **statistische analyse** werd ontwikkeld door Pythagoras in de 6^{de} eeuw voor Christus (het gemiddelde). Dit was de voorloper van de **beschrijvende statistieken**. Het andere type van statistische analyse (**inferentiële statistiek**) werd -waarschijnlijk- voor het eerst ontwikkeld in het Oosten rond 200 voor Christus (Dudycha & Dudycha, 1972). Dit was in de vorm van een waarschijnlijkheidsanalyse (*probability analysis*) om na te gaan of een verwachte baby eerder een jongen of meisje zou zijn. Waarschijnlijkheidstheorie/Kansberekening (*Probability theory*), zoals het later zou noemen, evolueerde verder in de vorm van gokberekeningen (*gambling mathematics*) in het werk van Blaise Pascal (1623-1662) en Christiaan Huygens (1629-1695) (David, 1962). In de late jaren 1800 en vroege jaren 1900 werden nog vele andere beschrijvende statistieken ontwikkeld door mathematici en wetenschappers zoals Sir Francis Galton (1883), een neef van Charles Darwin, en Karl Pearson (1895).

Statistiek ging verder dan gokken en pure mathematische concepten en werd later '*political arithmetics*' genoemd. Een term dat te danken is aan de dichte associatie met diegenen die politieke onderwerpen bestudeerden, inclusief economische. Men probeerde immers op basis van gegevens met behulp van '*political arithmetics*' antwoorden op problemen van de staat te krijgen. Hier ontstond waarschijnlijk de associatie tussen leugentjes bedenken om er politiek voordeel uit te halen en liegen door statistieken te manipuleren. Inderdaad, statistieken worden wel eens gemanipuleerd door politici, maar dat is geen reden om de statistiek af te schrijven. Dat zou veeleer een reden moeten zijn om de leugenachtige politicus af te schrijven. Statistiek is vaak de pianist waarop geschoten wordt. Als we de overleveringen mogen geloven dan was de eerste politicus die van statistiek gebruik maakte een zekere John Graunt (1662). Hij gebruikte beschrijvende statistieken, zoals we ze vandaag noemen, om het sterftecijfer in Londen te bestuderen. Hoewel er een hevig debat gaande is aangaande het originele gebruik van de term '**statistiek**' of '**statistics**' (Yule, 1905), is het merendeel van de statistici het erover eens dat de term statistiek uitgevonden

is door **Eberhard August Wilhelm von Zimmerman** in de inleiding van *'A Political Survey of the Present State of Europe'* (1787).

Het moderne gebruik van de term *statistiek* (ten opzichte van *wiskunde*) wordt vaak toegeschreven aan R. A. Fisher en zijn werk *'Statistical Methods for Research Workers'* (1925), waarin hij stelt dat *"a statistic is a value calculated from an observed sample with a view to characterizing the population from which it is drawn."* Sinds die tijd hebben vele statistici bijgedragen tot de technieken die mogelijk zijn om data te analyseren. Veel procedures hebben ook de naam gekregen van deze statistici. De bijdrage aan statistische technieken gebeurt vandaag de dag nog steeds. Analysetechnieken zijn de laatste jaren enorm toegenomen en het heeft onderzoekers in de sociologie, criminologie en andere wetenschappen mogelijk gemaakt om de relatie tussen variabelen accurater te bestuderen.

3. Het gebruik van statistiek

De term 'statistiek' wordt vaak verkeerd begrepen omdat deze term eigenlijk twee praktische toepassingen kent. De eerste term, die de historiek van de term reflecteert, is de verzameling en bewaring van data, vaak uitgedrukt in een samenvattende vorm. Goede voorbeelden hiervan zijn de volkstelling- of sterftestatistieken. Volkstellingstatistieken geven een voorstelling van de karakteristieken of kenmerken van de personen die in een land wonen op een bepaald moment in de tijd en sterftestatistieken geven inzicht in het aantal personen die sterven in een land op een bepaald moment. Dergelijke statistieken worden door de administratie verzameld om zicht te krijgen op belangrijke ontwikkelingen. De misdaadstatistieken kunnen aan deze lijst worden toegevoegd. Vanuit historisch perspectief zijn de statistieken van de veroordeelden de oudste gerechtelijke statistieken. De tweede toepassing van de statistiek is het onderwerp van deze inleiding: de statistiek is een methode om data te analyseren.

Statistiek is ook in de criminologie een methode om data die verzameld werden in het proces van een wetenschappelijk onderzoek, te bestuderen. Deze methode stelt onderzoekers in staat om logisch na te denken over data, en om één of twee dingen te doen: (1) te komen tot een beknopte synthese en betekenisvolle conclusies over de data (*beschrijvende statistiek*), en/of (2) karakteristieken van grote groepen bepalen -of afleiden-, gebaseerd op data afkomstig van kleinere

delen (*steekproeven*, ‘*samples*’) van de groep (*inferentiële statistiek*). We kunnen bijvoorbeeld data verzamelen bij alle inwoners van Gent in het kader van een onderzoeksproject waarbij we de attitudes van de Gentenaars ten opzichte van de politie, het stadsbeleid, enz. willen leren kennen. Deze resultaten zouden kunnen gebruikt worden voor een beschrijvende analyse om na te gaan in welke mate de burgers het eens zijn met de wijze waarop de politie de burger behandelt. Niet iedereen heeft dezelfde opinie. Het kan dus verwacht worden dat er grote verschillen opduiken, en mogelijks dat die verschillen in attitudes groter zijn bij jongeren dan bij ouderen. De resultaten kunnen echter ook grote tendensen laten zien. Wat denkt de modale Gentenaar? Wat is de gemiddelde score die de Gentenaar geeft als hij of zij de politie een “score op 10 moet geven”? Het zou natuurlijk heel duur zijn om iedereen te bevragen. Een alternatief is om een steekproef te trekken uit de Gentse bevolking. Data uit deze steekproef worden dan gebruikt om beweringen te doen over alle inwoners van Gent. Dit betekent dat er conclusies worden getrokken (inferenties), gebaseerd op informatie uit een kleinere steekproef, getrokken uit deze groep. Dat is onder bepaalde condities mogelijk, die we in dit handboek uiteenzetten.

Statistische analyse is een werkinstrument van wetenschappelijke ontdekking en wetenschappelijke kennisopbouw. Een voorbeeld van zo een wetenschappelijk proces is het gebruik van onderzoek om het waarheidsgehalte van een criminologische theorie na te gaan. Dit wil eenvoudigweg zeggen dat wij gaan controleren of de uitspraken die in een theorie vervat zitten over criminaliteit en diens oorzaken wel overeenstemmen met de bevindingen. Een theorie kan immers van alles beweren, maar we moeten vooral niet blind zijn en nagaan of een bewering wel klopt. Dat “nagaan” of iets wel klopt, doen we door gegevens te verzamelen. Dit vereist dat empirisch bewijs (data) dat gebaseerd is op onderzoekseenheden en dat die gegevens systematisch kunnen worden bestudeerd. Het gebruik van een aantal basisvaardigheden uit de wiskunde en algebra en statistische analyse in het bijzonder, maakt onderzoekers ertoe in staat om statistische vergelijkingen op te stellen en om nieuwe informatie te ontdekken die ervoor zorgt dat men zijn/haar onderzoekseenheid beter begrijpt.

In het wetenschappelijk proces is het doel meestal om iets te ontdekken dat voordien niet gekend was, of iets te bewijzen (waar of vals) waarvan voordien werd gedacht dat het waar was, maar dat nooit gestaafd werd met hard bewijs. De wijze waarop bewijs kan verworven worden is door

informatie (data) te verzamelen en deze te onderwerpen aan statistische analyse. Men mag echter niet uitgaan van wonderen. De statistiek toont verbanden aan, maar onderzoek is meer dan verbanden aantonen. Statistieken zeggen maar iets als de data die we gebruiken kwaliteitsvol zijn. Dat wordt soms wel eens vergeten. Statistiek komt pas te pas in criminologisch onderzoek nadat we een onderzoeksvraag hebben bedacht en beslist hebben hoe we datgene dat we willen weten kunnen gaan meten.

4. Theorieconstructie in een oogopslag

Drie elementen zijn essentieel in gedegen criminologisch (statistisch) onderzoek: theoretische achtergrond, onderzoeksmethoden en kwaliteitsvolle statistische analyse. Hoewel deze elementen strikt gelinkt zijn met elkaar, is er een debat gaande -zelfs bij diegenen die het meeste achter het onderzoeksproces staan- over hun volgorde, belangrijkheid en wat precies van elk element in een boek moet opgenomen worden. Het is niet mogelijk om in dit handboek op alle elementen van wetenschappelijk kwantitatief onderzoek in te gaan. In dit basishandboek beperken we ons tot de univariate, bivariate beschrijvende statistiek, de principes van de inferentiële statistiek en een korte introductie tot de multivariate statistiek. De leerinhoud van wat in dit handboek centraal staat, moet gezien worden in het licht van de kwantitatieve criminologische methoden en technieken die in het tweede bachelorjaar van de opleiding tot bachelor in de criminologische wetenschappen gedoceerd wordt.

Wat is theorie?

Op het basisniveau bestaat een **theorie** uit beweringen (proposities) over de relaties of associaties tussen *sociale fenomenen* zoals gebeurtenissen en eigenschappen van onderzoeksobjecten, vaak individuen, soms ook groepen, buurten, steden, landen,... In de criminologie bijvoorbeeld, zijn er theorieën die stellen via welke processen individuen criminaliteit als alternatief zien en besluiten in een bepaalde omstandigheid een delict te plegen. Daarnaast zijn er theorieën over de factoren die maken dat individuen de overtreding van normen als moreel acceptabel gaan zien: dat is een kwestie van het samenspel tussen de ontwikkelingscontext (omgeving, gezin, vrienden), de sociale bindingen die men heeft, en de persoonlijke psychologische, biologische en genetische kwetsbaarheden. Theorieën bevatten stelsels van uitspraken.

We geven een heel eenvoudig voorbeeld, gebaseerd op een observatie, die we willen verklaren (het “explanandum”), een aanvangsconditie, i.e. een conditie die we waargenomen hebben bij de persoon die we observeren. Theorieën worden grondig gecontroleerd door te kijken in welke mate deze observatie door een algemene wetmatigheid kan verklaard worden. Een sociale wetmatigheid in zijn meest banale vorm is een stabiel patroon, niets meer, niets minder. Het is een beproefde samenhang. Men mag dat niet fatalistisch interpreteren.

Sociale wetmatigheid: Wanneer jongeren zwakkere morele standaarden hebben, zijn zij sneller geneigd om criminaliteit als alternatief te zien.

Aanvangsconditie: Jan Janssens vindt het moreel niet verkeerd om eigendom van anderen mee te nemen. Mensen moeten maar op hun spullen letten. Eigen schuld dikke bult.

Explanandum: Jan Janssens heeft zopas de iPad van een medeleerling gestolen

Het doel van deze beweringen is om verklaringen te ontwikkelen waarom dingen zijn zoals ze zijn (de processen van worden en zijn- ontwikkeling) en om via inzicht in processen een diepere causale verklaring te bekomen. Zonder theorie bestaan er vaak alleen vermoedens en verhaaltjes (“just so stories”). Met een criminologische theorie kunnen we beweringen of ideeën ontwikkelen die gebaseerd zijn op gedegen observatie en kunnen we misschien ook iets doen aan het probleem van de criminaliteit.

Theorie en onderzoek

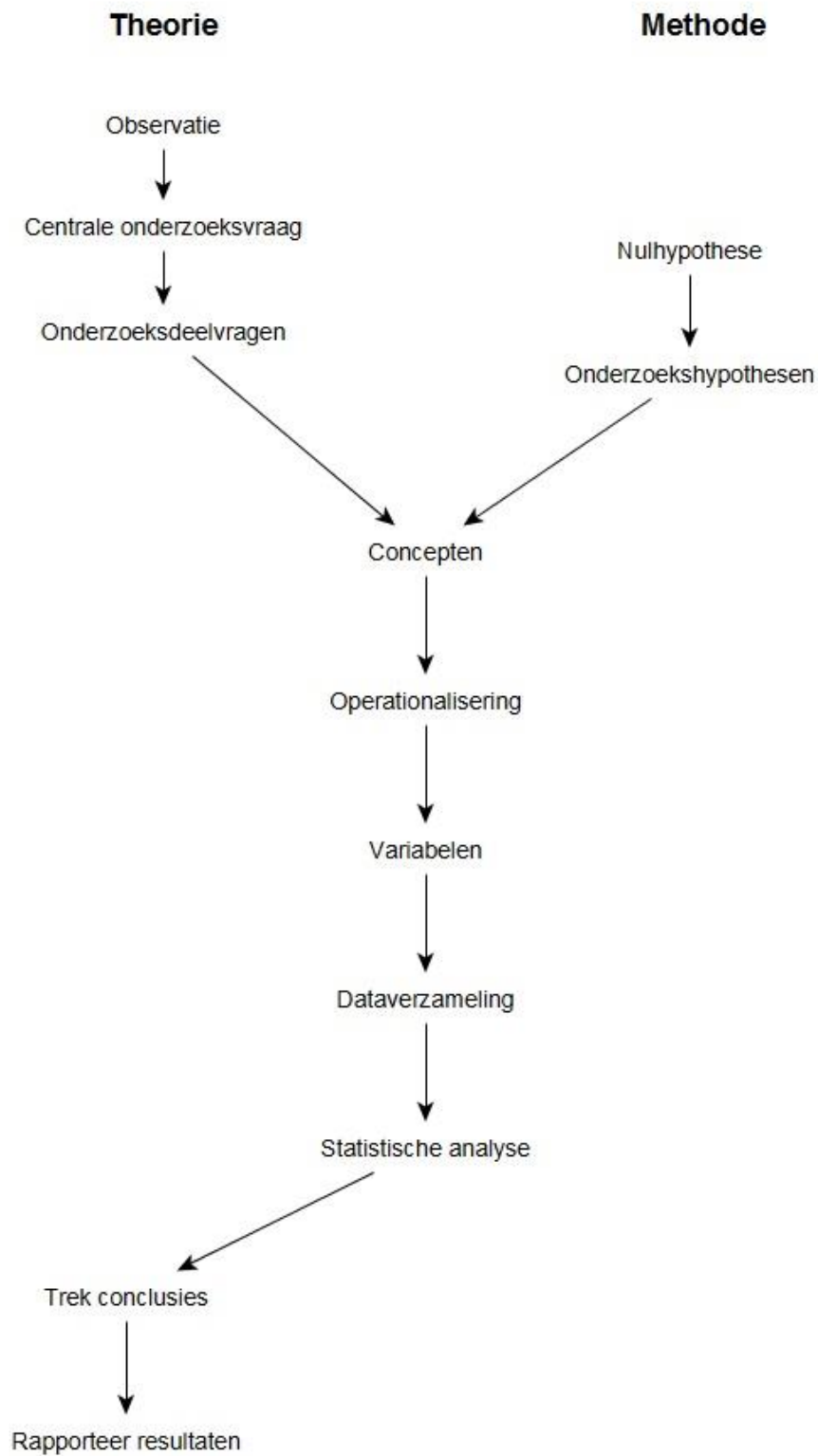
Een criminologische theorie kan op verschillende wijzen ontwikkeld worden. Ten eerste kunnen onderzoekers naar de wereld rondom hen kijken, een sociaal fenomeen nemen dat hun interesse wekt, en beweringen formuleren over waarom fenomenen op een bepaalde wijze werken. Dit wordt **inductie** genoemd. Bijvoorbeeld, een onderzoeker die criminaliteitstrends voor een aantal jaar in een stad volgt. Hij/zij kan misschien zien dat de criminaliteit een patroon volgt in de stad: de criminaliteit kan concentraties kennen in bepaalde wijken, de uitgaansbuurten, de hoerenbuurt, de verpauperde arbeiderswijken rond het centrum,... **Inductief bezig zijn is gevaarlijk.** We kijken te veel naar wat we willen zien en we veronderstellen dat we als we honderd keer hetzelfde hebben geobserveerd, dat de 101^{ste} keer ook zo observeren. Dat is absoluut geen garantie. Het feit dat de

zon al elke dag van je leven opging, kan geen garantie bieden dat ze morgen opgaat... Dat is niet bedoeld om pessimistische over te komen en je dag te vergallen, dat is een realiteit. We kunnen het niet bewijzen. We zijn soms blind voor onze eigen fouten en foutieve veronderstellingen. Daarom bestaat er een andere manier van onderzoek doen: **deductie**.

Het proces van deductie begint met een veronderstelling over hoe het nu in mekaar zit. In ons voorbeeld over de patronen van criminaliteit in de stad, kunnen we vertrekken van de Broken Windows-hypothese: een gebroken ruit of overvolle vuilnisemmers zenden signalen uit (priming): mensen zijn onbewust geneigd hun eigen afval bij de hoop te gooien die daar toch al ligt. Zo kan een buurt makkelijk afglijden als dat niet aangepakt wordt, of als de buurt niet weerbaar is. Als we het te ver laten komen, dan verhuizen de mensen die er genoeg van gehad hebben, en dan verkrot een buurt. Dat is een proces dat kan gestopt worden. De statistische samenhang tussen overlast en criminaliteit is een vrij stabiele samenhang. Maar even belangrijk als de statistiek, het hulpmiddel, is (1) de gezonde redenering, het logisch denkvermogen van de criminoloog die kritisch denkt, zich geen oor laat aannaaien bij het lezen van een tabel waar statistieken in staan, en (2) het beschikken over kwaliteitsvolle gegevens. In de praktijk is een criminologisch onderzoek een combinatie van inductie en deductie,

5. Het proces van wetenschappelijk onderzoek

Het proces van wetenschappelijk onderzoek (met gebruik van deductie) wordt getoond in **Figuur 1-1**. Zoals te zien is in dit diagram, is de theorie het startpunt van het proces. Theorie wordt gedreven door observaties en leidt onderzoekers tot het initiëren van het onderzoeksproces door het formuleren van een centrale **onderzoeksvraag** en onderzoeksdeelvragen. Het is uit dit proces van theorievorming dat onderzoekers komen tot het proces gaande van de ontwikkeling van een nulhypothese tot het communiceren van resultaten.



Figuur 1-1 Onderzoeksproces: Theorie, onderzoeksmethoden en statistische analyse

Observatie en nieuwsgierigheid

De eerste stappen in het proces van wetenschappelijk onderzoek worden vaak over het hoofd gezien, terwijl ze wel belangrijk zijn, nl. **observatie en nieuwsgierigheid**. Zo zijn vele onderzoeksprojecten nooit uitgevoerd omdat onderzoekers het initiële onderwerp plots niet langer interessant vonden, waardoor een onderzoek een andere richting is uitgegaan. Een criminoloog zou het idee kunnen gehad hebben om het fenomeen van de jeugdcriminaliteit te bestuderen vanuit het perspectief van de economische deprivatie, maar na het lezen van vele onderzoeksrapporten toch maar besluiten om een andere richting in te slaan. Zoiets gebeurt, zeker wanneer eerdere onderzoeken aantonen dat er geen eenduidig verband bestaat. Wat drijft onderzoekers? Kennis? Honger naar kennis? Of een honger naar het onbekende? Kennis is steeds feilbaar (net als het menselijke redeneervermogen), maar dat wil niet zeggen dat alles zo maar even relatief of waar is. Wie dat zegt, is een even grote charlatan als diegene die zegt de waarheid in petto te hebben. Het is vaak de theorie die observatie en wetenschappelijk onderzoek stimuleert. Als je onderzoeksmateriaal doorneemt waarin je geïnteresseerd bent, kan het zijn dat je misschien denkt een betere manier te kennen om het zelfde onderzoeksprobleem te benaderen. Bijvoorbeeld: denk terug aan het probleem van de economische deprivatie en criminaliteit. Het verband is vaak zwak, maar het is mogelijk dat het verband wel bestaat voor bepaalde vormen van criminaliteit of dat het verband enkel in bepaalde omstandigheden of voor een beperkt aantal personen geldig is: het verband zou bijvoorbeeld afwezig kunnen zijn op jonge leeftijd, maar het verband zou sterker kunnen zijn in de volwassenheid. Uiteraard moet je al heel wat kennis hebben over theorieën vooraleer je dergelijke denkoefeningen kan maken. Statistische analyses doe je dus altijd best gewapend met een goede kennis van de criminologie, dit wil zeggen de bestaande theorie en de kritieken er op, de stand van zaken met betrekking tot het bestaande onderzoek. Het gebruik van een gestructureerd wetenschappelijk proces om jouw criminologisch relevante observaties te evalueren en criminologische verklaringen te formuleren is de basis van goede theorieontwikkeling.

Een voorbeeld van inductieve theorievorming kan getoond worden in Robert Burgess zijn zonale theorie. Studenten aan de universiteit van Chicago maakten kaarten van Chicago waarop de verschillende karakteristieken van buurten werden vermeld zoals welzijn, kindersterfte en huisvesting. Burgess observeerde deze kaarten en zag dat deze een zeer gelijkaardig patroon

volgden doorheen de stad. Zijn observaties leidden tot de ontwikkeling van een theorie over hoe steden groeien en veranderen. Deze theorie stelt dat steden groeien in concentrische ringen zoals de ringen die ontstaan wanneer je een steen in het water gooit. In deze configuratie zal de ring het dichtste tot het centrum van de stad gekenmerkt worden door de meeste verloedering, het hoogste level van kindersterfte en andere sociale ziektes in vergelijking met de ringen erbuiten. Dit alles is bedacht door simpelweg het bestuderen van kaarten die gemaakt werden door studenten en door het gebruik van inductieve theorievorming.

Centrale onderzoeksvragen

De **centrale onderzoeksvraag** is de drijvende gedachte achter een onderzoeksproject. Ze moet de reden voor de statistische criminologische studie weergeven. Centrale onderzoeksvragen zijn belangrijk omdat criminologisch onderzoek vaak geëvalueerd wordt op basis van hoe goed de onderzoeker erin is geslaagd de centrale onderzoeksvraag te beantwoorden. De centrale onderzoeksvraag moet voorzichtig worden geformuleerd zodat deze exact de focus van de studie weergeeft. Bijvoorbeeld in onderzoek naar het gebruik van dodelijk geweld door politie zou een mogelijke centrale onderzoeksvraag de volgende kunnen zijn:

“Welke persoonlijkheidskenmerken, attitudes en omgevingskenmerken tijdens een arrestatie zijn het meest van invloed op het gebruik van dodelijk politiegeweld tijdens een arrestatie?” Deze vraag is breed en ietwat vaag, maar deze kan perfect het doel van het onderzoeksproject weergeven.

Onderzoeksdeelvragen

Vaak is het zo dat de centrale onderzoeksvraag theoretisch, vaag en bijna onmogelijk direct te beantwoorden is. **Onderzoeksdeelvragen** delen de centrale onderzoeksvraag in subproblemen die gemakkelijker te behandelen zijn en zorgen ervoor dat de centrale onderzoeksvraag toetsbaar wordt met behulp van onderzoek. Als de centrale onderzoeksvraag het doel van het onderzoek omvat, dan suggereren de onderzoeksdeelvragen hoe dit doel kan verwezenlijkt worden.

Een onderzoeker die geïnteresseerd is in sociale en psychologische determinanten van jeugdcriminaliteit kan de volgende vragen stellen

- Hoe veel delicten rapporteren jongeren op jaarbasis?

- Hoe groot is het empathisch vermogen van jongeren?
- Hoeveel jongeren behoren tot een straatgroepje
- Is er een verband tussen het behoren tot een straatgroepje en de hoeveelheid delicten die men rapporteert?
- Wat is de relatie tussen de empathische vermogens van de jongere en de frequentie waarmee hij of zij geweld pleegt?
- Wat is het verband tussen het behoren tot een straatgroepje en het empathisch vermogen?

Deze onderzoeksdeelvragen delen de centrale onderzoeksvraag in kleinere delen die makkelijker kunnen onderzocht worden. De antwoorden op deze onderzoeksdeelvragen zijn afgeleid van het onderzoeksproces en statistische analyse en stellen de onderzoeker in staat om de centrale onderzoeksvraag te beantwoorden.

‘Doing research is like defusing a bomb. When you begin, you are all excited and focused on the end result. If you have a good plan, know the layout, and work the plan, you will typically get the results you sought. If you run in and start cutting without a plan, it is likely to blow up in your face.’

6. Onderzoek: bewegen van theorie naar data en terug

Theorie kan niet op zich bestaan, net zoals onderzoek of statistiek. Theorie zonder onderzoek en statistische analyse is niet meer dan een fabel. Onderzoek zonder theorie is zoals het bouwen van een huis zonder plannen, en onderzoek zonder statistische analyse is zoals het bouwen van een huis zonder nagels: het is mogelijk en het is al gebeurd, maar het zou effectiever zijn met. Statistische analyse zonder theorie en onderzoeksmethodologie als gids voor onderzoek is dan weer zoals het hebben van intercontinentale ballistische raketten: leuk om te hebben en het zou je burens kunnen imponeren, maar niet echt bruikbaar. Wat de techniek ook is voor het ontwikkelen van theorie, **onderzoek** is de methode om een theorie te toetsen en valideren. In zijn puurste vorm is onderzoek een wetenschappelijke, systematische studie om nieuwe informatie te ontdekken of om de validiteit van theorieën die eerder werden ontwikkeld, te toetsen. Het voornaamste doel van onderzoek is ontdekken. Afhankelijk van het gebruik van een inductief of deductief proces, is onderzoek een systematische manier om observatie en statistische analyse om te zetten in theorie (inductie) of het toetsen van theorie met behulp van statistische analyse (deductie). Hoewel er geen exacte stappen

bestaan die moeten gevolgd worden bij het doen van onderzoek, zijn er wel een aantal algemene richtlijnen die zouden gevolgd moeten worden om zeker te zijn dat er niks ontbreekt in de studie. Deze stappen worden in dit hoofdstuk besproken en geven weer waar onderzoek en statistiek in het volledige proces van wetenschappelijk onderzoek passen.

Hypothesen formuleren

Eens de onderzoeksdeelvragen werden ontwikkeld, moet er beslist worden wat het onderzoek probeert te bepalen. Hypothesen zijn vragen of verklaringen waarvan de antwoorden de theoretische proposities van het onderzoek ondersteunen of weerleggen. Hypothesen zijn in het algemeen onder te verdelen in onderzoekshypothesen en nulhypothesen. Een **onderzoekshypothese** is een propositie, een veronderstelling die je wilt toetsen. Bij het gebruik van onderzoekshypothesen wordt de relatief abstracte bewoording van theorieontwikkeling omgezet in een meer concrete en toetsbare vorm, geschikt voor statistische analyse.

Eén van de vaak moeilijk begrijpbare, maar essentiële elementen van statistische analyse en hypothesetoetsing is dat onderzoek alleen niets kan bewijzen. Ook als onderzoekers een groot stuk van bevestiging vinden voor een associatie tussen twee variabelen, kan het zijn dat deze resultaten worden gevonden omdat er informatie ontbreekt of omdat het model wat gebrekkig is. Mogelijks kunnen andere onderzoekers de bevindingen weerleggen door bijkomstig onderzoek te doen. Als statistisch onderzoek niets definitiefs kan bewijzen, waarvoor kan het dan wel gebruikt worden? Dit werd gevonden met behulp van een **nulhypothese**, dat in het algemeen de vorm van één van de volgende voorbeelden aanneemt:

- Er is geen **statistisch significant** tussen twee kenmerken in een steekproef.
- De verschillen tussen de groepen bestaan door toevallige fouten.

Het doel van een nulhypothese is om een verband te weerleggen. Weerleggen dat er geen relatie is (het weerleggen van de nulhypothese), helpt het ondersteunen van een conclusie dat er een relatie is tussen de bestudeerde fenomenen.

Constructie van het onderzoeksdesign

Eens een beslissing is gemaakt van wat precies bestudeerd zal worden, kan het plannen van het actuele onderzoek beginnen. Als onderzoeker dien je voorzichtig te zijn en niet te snel te springen van deze stap naar andere stappen in het onderzoeksproces. Zo start je niet aan het bouwen van een huis zonder eerst andere huizen te bekijken en te bedenken hoe jouw huis er uit moet zien. Dus, waarom zou je een onderzoeksproject starten zonder een degelijke overweging van wat je wilt doen en vinden? Activiteiten in deze stap zijn het bepalen van onder andere de onderzoeksmethode (experiment, survey of een andere methode) en hoe het onderzoek in het algemeen moet benaderd worden. Als de onderzoeker data moet verzamelen, moeten beslissingen gemaakt worden aangaande hoe deze data dienen verzameld te worden, van welke groep deze data moeten komen en andere parameters. De beslissingen die hier worden gemaakt sturen de rest van het project, dus ze moeten zeer zorgvuldig worden gemaakt. Deze stap in het onderzoeksproces is ook opgelegd door het type van verzamelde data, dat op zich ook de statistische analyses bepaalt die gebruikt zullen worden.

Conceptualisering

Eens de onderzoeksdeelvragen en hypothesen zijn geformuleerd, moeten ze onderverdeeld worden in meer handelbare delen. Dit wordt gedaan door concepten uit de vragen en hypothesen te halen. **Concepten** zijn termen waarover in het algemeen een consensus bestaat en kan betrekking hebben op een kenmerk, fenomeen of een groep van onderling gerelateerde fenomenen. Concepten kunnen heel abstract (delinquentietolerantie) zijn of net concreet (man-vrouw).

Operationalisering

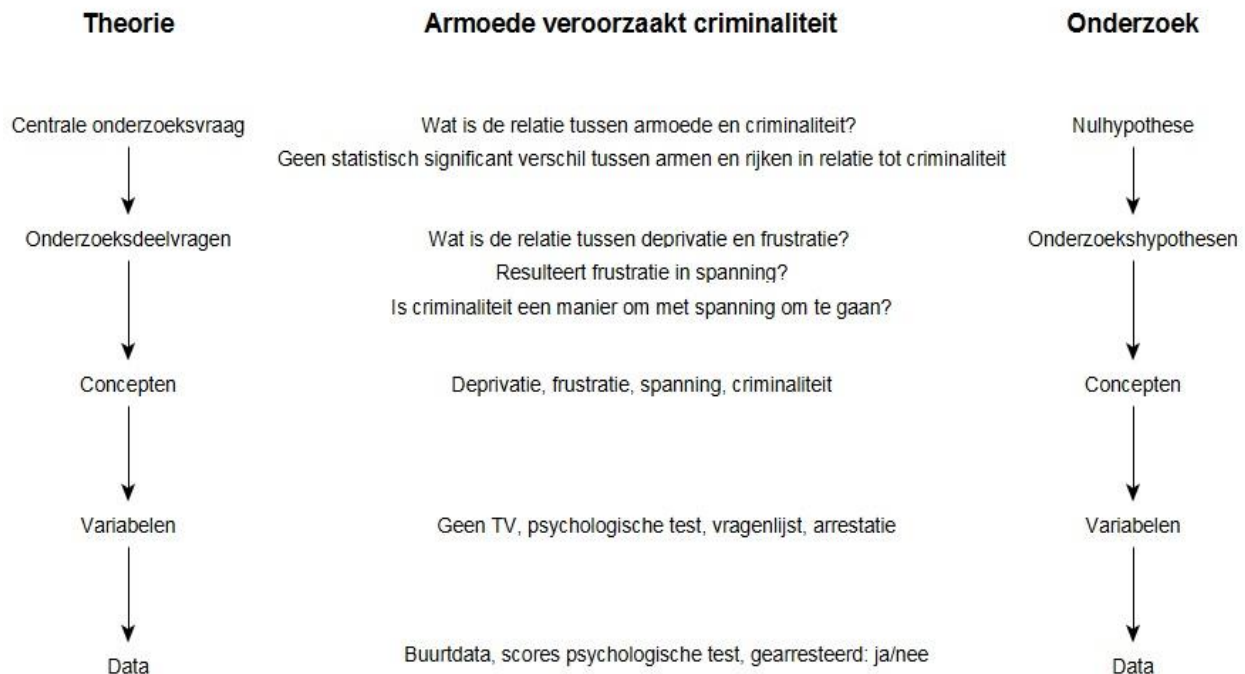
Om concepten te gebruiken in statistische analyse, moeten ze zo geformuleerd worden dat ze mathematisch kunnen worden geanalyseerd. Dit gebeurt door middel van **operationalisering**. Dit is het proces van het vertalen van een concept, dat abstract en woordelijk is, in een variabele die kan gezien en getoetst worden, door het omschrijven hoe het concept kan gemeten worden. Het proces van transformeren van concepten in variabelen demonstreert een kritiek punt in operationalisering: de geoperationaliseerde definities die gebruikt worden in onderzoek, zijn de onderzoeker zijn/haar definities en deze hoeven niet overeen te stemmen met de definities die anderen gebruiken of definities die dezelfde onderzoeker gebruikt in ander onderzoek.

Bijvoorbeeld, in dit onderzoeksproject is het mogelijk dat een politiemann/-vrouw geoperationaliseerd werd als een inspecteur van politie binnen de lokale politie die belast is met interventie. Anderen zouden politiemann/-vrouw anders kunnen definiëren en er bijvoorbeeld ook rechercheurs aan toevoegen en/of anderen met een politiefunctie.

Data verzamelen

Dataverzameling is de stap in het onderzoeksproces waar de meeste mensen eigenlijk willen starten en eigenlijk de laatste is waarmee zou moeten begonnen worden. Wanneer we terugkeren naar het voorbeeld van het bouwen van een huis, zou het starten met het verzamelen van data te vergelijken zijn met het beslissen om een huis te bouwen, 5 zakken cement en een ton nagels te bestellen en aan de slag gaan, zonder te werk te gaan met een degelijk ontwikkeld plan. Het kan zijn dat je effectief tot een huis komt, zeker wanneer je een expert bent, maar het zou beter zijn wanneer je begint met het zorgvuldig opmaken van een plan. Op dit punt zouden alle beslissingen betreffende het onderzoek moeten gemaakt zijn. De onderzoeker zou een weloverwogen theoretisch model moeten hebben evenals een duidelijk en compleet onderzoeksdesign waarin staat hoe de data zullen verzameld en geanalyseerd worden. De concepten zouden al geoperationaliseerd moeten zijn in variabelen die accuraat gemeten kunnen worden. Het enige wat nog moet gedaan worden is data verzamelen volgens het onderzoeksdesign. Deze analyseren volgens het design en de resultaten rapporteren.

Het onderzoeksproces betreffende het gebruik van dodelijk geweld door politie wordt weergegeven in **figuur 1-2**. Het toont elke stap in een deductief proces. De onderzoeker heeft een onderzoeksplan opgesteld volgens de stappen in het onderzoeksproces gaande van theorie naar data en van analyse naar de publicatie van conclusies. Dit wetenschappelijk proces is zowel bruikbaar in academisch als praktisch onderzoek. Let wel dat tussen de stappen ‘concepten’ en ‘het trekken van conclusies’ deze stappen liggen die direct van invloed zijn op de statistische analyse. Het is dus van het grootste belang dat deze stappen zorgvuldig doorlopen zijn aangezien ze bepalend zijn voor de statistische analyses.



Figuur 1-2 Wetenschappelijk onderzoeksproces: voorbeeld van theorie naar data

Figuur 1-2 geeft een illustratie van het proces van wetenschappelijk onderzoek. Hoewel niet alle stappen zijn weergegeven, toont de figuur wel de verschillende stappen dat men moet doornemen en het verschil tussen theorie en onderzoek. In de figuur kunnen ook de types van werkproducten afgelezen worden die na elke stap worden bekomen. In dit voorbeeld wordt vertrokken vanuit de vraag of het armoede criminaliteit veroorzaakt. De onderzoeksdeelvragen en hypothesen zetten de abstracte centrale onderzoeksvraag om in toetsbare stellingen. De concepten delen de onderzoeksdeelvragen en hypothesen verder op in sleutelelementen die gemeten moeten worden. De variabelen die de operationalisering van de concepten betreffen, zijn die elementen waarover data verzameld moeten worden. Tenslotte kunnen de verzamelde data (de aantallen en andere informatie) worden bestudeerd door middel van statistische analyse.

Conclusies trekken

Velen denken dat het proces van statistische analyse, en zelfs wetenschappelijk onderzoek, stopt aan het einde van de analyse. Niets is minder waar. Statistici en onderzoekers onderscheiden zich in de interpretatie van analyses en de conclusies die getrokken kunnen worden. Dit is over het algemeen ook het moeilijkste deel van statistische analyse. Deze stap houdt in of de resultaten van

de statistische analyse de hypothesen die bij de start van het onderzoeksproces zijn ontwikkeld, worden ondersteund. In deze stap van het onderzoeksproces stoppen we met de statistische analyse en methodologische kwesties en wordt er teruggekeerd naar theorie. Als de onderzoeker gebruik maakt van een deductief proces, is dit het punt waarop de theorie die in de eerste stappen werd uitgelijnd, vergeleken wordt met de resultaten. Nu valt het verdict of de theorie ondersteund of weerlegd wordt. In een inductief proces is dit het punt waarop de onderzoeker de eerste conclusies trekt over wat hij of zij heeft gezien. Kennelijk verlaat de theorie nooit echt het onderzoeksproces, net zoals methodologische en statistische kwesties belangrijk zijn in elke stap, van het opzetten van een theorie tot het trekken van conclusies.

Een student veranderen in een statistisch geschoold sociaal wetenschapper vereist het verscherpen van vaardigheden met betrekking tot het interpreteren van analyses. In staat zijn om problemen uit te werken of de computer ertoe te brengen om een antwoord te bieden op een specifieke analyse is één ding; het uitvoeren van die analyse, het terugkoppelen in het proces van wetenschappelijk onderzoek en het interpreteren van wat je hebt gevonden op zo een manier dat het nieuwe inzichten brengt in het topic dat wordt bestudeerd, is een andere zaak.

Communiceren van resultaten

De laatste stap in het proces van wetenschappelijk onderzoek is het communiceren van de resultaten van het onderzoek. Deze stap wordt ook vaak over het hoofd gezien. Velen geloven dat de bevindingen van het onderzoek niet gecommuniceerd moeten worden omdat ze het niet waard zijn, tenzij de bevindingen de hypothesen ondersteunen of een gigantische ontdekking inhouden. Hoewel het inderdaad zo is dat veel van de meest prestigieuze tijdschriften terugschrikken van het publiceren van negatieve resultaten, betekent het niet dat deze niet gecommuniceerd moeten worden. Het is belangrijk voor de praktiserende criminoloog in het werkveld dat zelfs negatieve resultaten worden gecommuniceerd. Dit moet mensen behoeden dezelfde fouten opnieuw te maken en zou moeten vermijden dat geld verspild wordt door iets te onderzoeken dat al eens eerder is onderzocht. Deze finale en essentiële stap van het onderzoeksproces kan op verschillende manieren worden verwezenlijkt. Het meest wenselijke is om de resultaten van een statistische studie te publiceren in een academisch tijdschrift of boek. De resultaten kunnen ook gecommuniceerd worden in meer praktisch gerichte publicaties (vb. laagdrempelige publicaties voor politie), in

paper presentaties op professionele conferenties. Echter, ook een masterproef kan gezien worden als een eerste stap in deze richting. Daartoe zijn de vaardigheden die hier getraind zullen worden essentieel. Het komt er immers op neer wanneer de resultaten van een statistische analyse uitgerekend zijn, om deze ook nog helder neer te schrijven zodat de informatie een didactische waarde verkrijgt. Alleen oefening baart kunst.