

voor onze illustratie nemen we de cijferreeksen van Cruijff e.a. (2002) over, die uitgedrukt worden als kansen die optellen tot 1.

³ Poisson regressie, een eenvoudige uitbreiding van de Poisson verdeling in formule (3), ziet er als volgt uit:

$$p(y(i); m(i)) = \frac{e^{-m(i)} m^{y(i)}}{y(i)!}$$

waarbij $\log m(i) = b_0 + b_1 x_{i1} + b_2 x_{i2} + \dots + b_k x_{ki}$ (4)

Elk individu heeft dus een eigen Poisson parameter $m(i)$, en deze parameter is voor elk individu dus een functie van een aantal verklarende variabelen $x(i)$. Doordat deze verklarende variabelen voor elk individu van invloed zijn op de Poisson parameter, bepalen ze ook de kans dat een bepaald individu 'gepakt' of 'herpakt' wordt ($y(i)$). Ook hier worden de Poisson parameters uiteindelijk ook $f(0)$ opnieuw geschat op basis van een computerprogramma voor afgeknotte Poisson regressie (Cruijff e.a., 2002).

⁴ Echter, ook voor ongeobserveerde heterogeniteit, wat kan leiden tot 'overdispersion', worden oplossingen geformuleerd: "For the truncated Poisson regression model a Lagrange multiplier test on overdispersion was developed by Gurmu (1991). It compares the model fit of the Poisson model against alternative models with an extra dispersion parameter included, such as the negative binomial model" (van der Heijden e.a., 2003:297).

⁵ Ook hiervoor biedt Poisson regressie een 'oplossing' (zie formule (4)). De uitval, tijdelijk of definitief, van individuen maakt de $y(i)$ voor die individuen onbetrouwbaar; de $x(i)$ waarden zijn dit echter wel. Nu is het mogelijk om voor die individuen waarvan we de $y(i)$ waarden niet betrouwen, Poisson parameters te berekenen enkel op basis van zij waarvoor we $x(i)$ en $y(i)$ wel betrouwen. Die specifieke Poisson parameters worden tenslotte gebruikt voor het berekenen van de kansen. Cruijff e.a. waarschuwen echter dat "niettemin een dergelijke schatting altijd met enige voorzichtigheid moet worden beschouwd" (2002:53).

⁶ Stefan Bogaerts is seniorhoofdonderzoeker bij IVA Tilburg, Universiteit van Tilburg en bijz. hoofddocent aan de Faculteit Psychologie, Universiteit Gent. Hij is onder meer promotor van

het onderzoek 'Aard en omvang vuurwapencriminaliteit in Nederland 2001-2003', uitgevoerd door het IVA in opdracht van het WODC (Ministerie van Justitie, Nederland), waarvan sprake in deze bijdrage.

⁷ Stefaan Pleysier is wetenschappelijk onderzoeker (FWO) aan de Afdeling Strafrecht, strafverordering en criminologie, Faculteit Rechtsgeleerdheid, van de Katholieke Universiteit Leuven. Hij zetelt als methodoloog in de WODC begeleidingscommissie van het bovenvermelde onderzoek 'Vuurwapencriminaliteit'.

⁸ Stijn Vanheule is doctorassistent aan de Faculteit Psychologie, Universiteit Gent.

VALKUILEN BIJ DE ANALYSE VAN PARKETDATA

In 2003 voerde de onderzoeksgroep *Sociale Veiligheidsanalyse* (UGent - VUB) een studie uit naar de exploitatiemogelijkheden van parketdatabanken¹. Tijdens het analyseren van de parketdata werden wij, de onderzoekers², geconfronteerd met enkele belangrijke methodologische valkuilen. Deze valkuilen hebben betrekking op keuzes die de onderzoeker dient te maken. Het gaat om belangrijke keuzes met consequenties voor de inhoudelijke interpretatie van de resultaten. In deze bijdrage leggen we de klemtoon op de aard van deze problemen en trachten we deze eveneens te illustreren.

Een *eerste* en ongetwijfeld belangrijkste valkuil betreft de *teleenheid* die moet gekozen worden om de data te analyseren. De teleenheid is de statistische eenheid (unit of analysis) op basis waarvan uitspraken zullen worden gedaan. Parketdata bevatten echter uiteenlopende info. De onderzoeker dient hier een expliciete keuze te maken. Moeten feiten, processen verbaal (PV), betrokkenen (verdachten en slachtoffers) of tenlastenleggingen als teleenheid gekozen worden of zijn er nog andere mogelijkheden? Een *tweede* valkuil betreft de doorstroom naar de parketten van de gegevens die op het niveau van de politiediensten worden gevat. Zolang deze doorstroom niet op een geautomatiseerde wijze gebeurt, blijft er een onnodige foutenbron bestaan die invloed heeft op de accuraatheid van de parketdata. Een *derde* caveat betreft het vergelijken van parketten op basis van cijfergegevens. Zo

registreren niet alle parketten op een gelijkaardige manier. *verkorte processen verbaal* (VPV). Deze VPV's worden door diverse parketten verschillend behandeld. Sommige parketten nemen deze mee op in hun digitaal bestand, anderen niet. Dit gegeven dient men steeds in het achterhoofd te houden wanneer men parketten op basis van cijfers wil gaan vergelijken. Een vierde probleem is dat parketgegevens *an sich* slechts beperkte analyses mogelijk maken omwille van de beperkte informatie die in de parketdatabanken is opgenomen. Wanneer politionele en parketdatabanken aan elkaar gekoppeld worden, levert dit een informatievoordeel op. Door deze twee databronnen samen te voegen, wordt het mogelijk analyses uit te voeren die bijkomende inzichten leveren in de werking van het strafrechtstelsel. Hieronder gaan we dieper in op deze diverse knelpunten.

I. Teleenheid

Wanneer men een analyse van parketdata wil uitvoeren, dient men zich bij aanvang de vraag te stellen 'Wat wil ik *precies* weten?'. Wanneer men geïnteresseerd is in het vervolgingsbeleid van een parket, meer bepaald in welke mate een parket uiteenlopende *zaken* vervolgd, daarbij eventueel onderscheid makend naar delicttypes, kan dit gebeuren met PV's als teleenheid. Wanneer men echter wil nagaan hoeveel verdachten er effectief vervolgd werden dan zal men logischerwijs *verdachten* als teleenheid gebruiken. Met andere woorden, afhankelijk van de vraagstelling zal men een andere teleenheid dienen te gebruiken en operationaliseren.

Wanneer PV's als teleenheid gebruikt worden, moet de gebruiker opletten voor bijkomende valkuilen. In eerste instantie mag men het aantal PV's niet verwarren met het aantal *tenlasteleggingen*. Het is immers zo dat er per PV tot negen tenlasteleggingen door het parket geregistreerd kunnen worden. Dit impliceert dat de onderzoeker of analist die het aantal tenlasteleggingen verkeerdelijk gelijk stelt aan het aantal PV's onvermijdelijk tot dubbeltellingen van PV's zal komen. Het hoeft geen verder betoog dat dit leidt tot een vertekening van de resultaten. Een zelfde opmerking kan gemaakt worden met betrekking tot feiten. Sommige PV's bevatten

immers meerdere feiten waardoor ook hier dubbeltellingen kunnen gebeuren. Hoewel feitcodes (ISLP³ codes) normaliter niet worden opgenomen in parketdatabanken, achten we het noodzakelijk dat dit probleem expliciet vermeld wordt. Wij pleiten immers voor het koppelen van politie en parketdatabanken, omdat deze operatie toelaat om op parketniveau analyses uit te voeren op basis van feiten of criminele figuren⁴ (zie verder). Voorlopig is dit evenwel nog niet mogelijk.

Wanneer *feiten* als teleenheid worden gekozen, moet men er zich van bewust zijn dat niet steeds alle feiten zullen opgenomen worden. Dit is afhankelijk van de beoordeling van de registrerende politieagent. Laat ons dit illustreren aan de hand van een sprekend voorbeeld. Een persoon huurt een taxi. Wanneer hij moet betalen, betwist hij het bedrag. Tijdens de daaropvolgende twist bedreigt de persoon de taxichauffeur met een mes. Hij verkoopt vervolgens de taxichauffeur enkele rake klappen. Wanneer de taxichauffeur aangifte doet zal er door de politie één PV opgesteld worden. Het PV kan dan volgende feiten bevatten: flessentrekkerij - huurrijtuig, opzettelijke slagen en verwondingen en wederrechtelijk dragen van verboden wapens. Afhankelijk van de dienstdoende agent zullen of alle of meerdere feiten of enkel het door de agent als voornaamste gepercipieerde feit in het politiebestand (ISLP) terecht komen.

Wanneer informatie over de betrokkenen (slachtoffers en verdachten) wordt bestudeerd, zijn de gegevens die in de parketbestanden terug te vinden zijn, vaak heel summier. Immers, niet voor alle PV's wordt door alle parketten alle beschikbare informatie betreffende de betrokkenen in het parketdatabestand opgenomen. Vaak is dit afhankelijk van de gevolgen die aan een PV worden gegeven: met ander woorden of er vervolgd wordt of niet. Ook hier kan een koppeling tussen politionele en parketdatabanken soelaas brengen. In de politionele databanken is informatie over de betrokkenen (slachtoffers en verdachten) immers wel aanwezig. Let wel, noch de parketdatabanken noch de politionele databanken garanderen de gebruiker een accuraat beeld van noch de slachtoffer- noch de verdachtenpopulatie. Sommige slachtoffers doen immers om

verschillende redenen⁵ nooit aangifte van een vergrijp waarvan zij slachtoffer waren. Een niet onbelangrijk deel van gepleegde criminaliteit wordt dus nooit aangegeven. Voor deze feiten hebben we dan ook noch informatie over het slachtoffer(s) noch over de dader(s). Tevens is voor een groot deel van de bij de politie aangegeven feiten de verdachte(n) onbekend. Dit varieert evenwel sterk naar delicttype. Voor winkeldiefstal bijvoorbeeld was er in ons onderzoek in 80% van de gevallen een gekende verdachte terwijl voor woninginbraak dit slechts 7% bedroeg. Wanneer deze onbekende slachtoffers of verdachten systematisch verschillen van de gekende slachtoffers of verdachten kunnen er geen veralgemeningen gebeuren op basis van de bekende slachtoffers of verdachten. Immers de kans dat dan een vertekend beeld van de realiteit geschetst wordt is niet onbestaande. Analyses betreffende verdachten of slachtoffers op basis van politionele gegevens of parketgegevens leren ons met andere woorden hoofdzakelijk iets over de bij de politie en parket gekende slachtoffers en verdachten.

II. Doorstroom politionele gegevens

Processen verbaal die op een parket aankomen worden tot op vandaag de dag manueel door vatters⁶ ingeput, dit op een moment dat de digitale overdracht van gegevens tussen politie en parket, wanneer rekening houdend met de huidige technologieën die voor handen zijn, geen enkel probleem meer zou mogen vormen. Deze digitale overdracht zou op de eerste plaats heel wat werk besparen, immers PV's moeten dan niet meer gevat worden. Daarnaast zou het registratiefouten die nu gemaakt worden bij het vatten elimineren. Anderzijds zou het ook toelaten informatie zoals feitcodes en informatie over betrokkenen (slachtoffers en verdachten) op een eenvoudige wijze in het parketdatabestand op te nemen (zie verder). De digitale overdracht van gegevens tussen politie en parket zou met de implementatie van het Phenix⁷-project moeten gerealiseerd worden. In dit project is immers de link tussen Justitie en externe partners zoals politie voorzien.

Het manueel vatten van gegevens is trouwens een van de belangrijkste redenen waarom in

sommige parketten VPV's niet in het digitale bestand van het parket opgenomen worden. Het invoeren van deze VPV's verhoogt de werklust immers danig daar deze VPV's tot een derde van alle op een parket binnenkomende PV's kunnen bedragen. Dit betekent dat het werk voor de vatters in sommige gevallen met een factor anderhalf zou toenemen.

Naast dit probleem van de doorstroom van gegevens tussen politie en parket is er nog een ander informatica gerelateerd probleem welke een extra drempel vormt wanneer men parketdata wil analyseren. Door een gebrekkige capaciteit van de huidige informaticasystemen bij de parketten is het voor de dagelijkse gebruikers van het systeem, de parketmagistraten en bedienden, vaak moeilijk om hun werk op een efficiënte manier uit te voeren. Het extraheren van grotere datasets, waarop analyses uitgevoerd kunnen worden, is een extra belasting van deze systemen. Een extra belasting die kan leiden tot een verdere vermindering van de preformantie van het systeem, erger nog, tot een falen.

III. Vergelijken van statistische info over diverse parketten

Wanneer men databanken van parketten met elkaar wil gaan vergelijken, moet men steeds rekening houden met het feit dat de registratiewijze niet altijd dezelfde is. Het is dan ook aan te raden bij aanvang, wanneer men een vergelijking wil maken, na te gaan op welke wijze deze bij de verschillende parketten gebeurt. Indien men dit niet doet, bestaat de kans dat appels met citroenen vergeleken worden. Zoals reeds aangegeven worden *verkorte processen verbaal* (VPV's) niet door alle parketten in hun digitaal databestand geregistreerd. Wanneer men op basis van de digitale databank van parket A, welke de VPV's wel digitaal registreert, en de digitale databank van parket B, die deze VPV's niet registreert, de mate waarin deze twee parketten PV's seponeren met elkaar wil vergelijken op basis van de gegevens in hun digitale databestanden, dan is het niet ondenkbeeldig foute conclusies te trekken betreffende de mate waarin beide parketten seponeren, wanneer geen rekening gehouden wordt met het feit dat het ene parket wel en het andere niet de VPV's digi-

taal registreert. Laten we dit illustreren aan de hand van een voorbeeld. Wanneer parket A zowel PV's als VPV's registreert welke als volgt verdeeld zijn: 10% VPV's en 90% PV's⁸ en men weet dat 1 op 2 PV's geseponeerd wordt, kan men het totaal aantal sepot's voor parket A berekenen dit door de VPV's mee in rekening te brengen. Ervan uitgaande dat deze VPV's allen geseponeerd worden betekent dat in totaal 55% (45% (=1/2 van de PV's) + 10% (VPV's)) van alle PV's (PV's + VPV's) in parket A zonder gevolg geklasseerd worden. Parket B registreert enkel de PV's. Deze PV's maken 70% uit van alle PV's dat is VPV's en PV's samen. Veronderstellen we dat ook voor parket B 1 op de 2 (50%) PV's geseponeerd worden dan zou men kunnen besluiten, wanneer men enkel gebruik maakt van het digitale bestand, dat de seponeringsgraad op parket B lager is dan op parket A, doch dit is niet zo. Immers de VPV's die niet in het digitale parketdatabestand vervat zaten, werden niet in de analyse opgenomen. Wanneer dit wel gebeurt, veronderstellend dat alle VPV's geseponeerd worden dan kan genoteerd worden dat 65% (35% (=1/2 van de PV's) + 30% (VPV's)) van alle PV's (PV's + VPV's) in parket B zonder gevolg geklasseerd worden. De seponeringsgraad van parket B is in werkelijkheid hoger dan deze van parket A. Met andere woorden: wanneer men wil vergelijken moet men steeds zorgen met een gelijke noemer te werken, dat is het totaal aantal PV's (PV's + VPV's). Zoniet worden appels met citroenen vergeleken. Het is trouwens aan te raden wanneer sepotpercentages berekend worden steeds de VPV's mee in rekening te brengen.

IV. Koppelen van politionele en parketdatabanken: meerwaarde

Een van de belangrijkste problemen waarmee men als onderzoeker of analist geconfronteerd wordt, is dat de benodigde informatie om de gewenste analyses uit te voeren vaak over verschillende databanken verspreid is. Wanneer men bijvoorbeeld wil nagaan welke gevolgen door het parket (sepot, vervolging, minnelijke schikking, ...) aan welke feiten worden gegeven, moet men op het niveau van het parket over informatie betreffende feiten beschikken. De informatie is echter enkel voorhanden in het

politionele databestand. Door de koppeling te maken tussen politionele en parketdatabanden kan men op parketniveau over gegevens betreffende feiten beschikken. Het koppelen van politionele en parketdatabanken levert ons een informatievoordeel op.

Door het koppelen van de politionele en parketdatabanken beschikt de onderzoeker zowel op het niveau van de politie als het parket over informatie betreffende feitcodes. Dit laat toe om ook op het niveau van het parket met feiten te werken. Meer nog, het stelt de onderzoeker in staat om zowel op het niveau van het parket als op het niveau van de politie *criminele figuren* te hanteren welke in tegenstelling tot parketcodes en voornaamste tenlasteleggingen, die een juridische connotatie hebben, veeleer een criminologische realiteit schetsen die dichter aansluit bij zowel de leefwereld van beleidsmakers als politieagenten.

Het werken met *criminele figuren* is trouwens een mooi voorbeeld waarbij gegevens uit een derde databank worden gebruikt om parketdata te analyseren. Immers de conversietabel die toelaat feiten, meer bepaald ISLP feitcodes, om te zetten in *criminele figuren* is een aparte databank⁹. Wanneer gebruik gemaakt wordt van de feitcodes die in ISLP vervat zitten, wordt de analist geconfronteerd met een veelheid aan categorieën die weinig zinvolle interpretaties toelaten. *Criminele figuren* die feiten op een hoger niveau aggregeren laten meer zinvolle interpretaties toe.

Een tweede voordeel van het koppelen van politie en parketdatabanken is, zoals reeds werd aangehaald, dat er meer informatie betreffende verdachten en slachtoffers op het niveau van het parket aanwezig is. Het beschikken over gegevens betreffende betrokkenen laat toe, rekening houdend met de beperking van de gegevens, om inzichten over verdachten en slachtoffers te verwerven. Zo zou men kunnen nagaan voor hoeveel daders of verdachten hun dossier geseponeerd werd, voor hoeveel van hen hun dossier door een rechter ten gronde werd behandeld, enz.

Het koppelen van databanken biedt ook de mogelijkheid om de doorstroom van PV's in de gerechtelijke molen in kaart te brengen. Op die

manier kan men vragen zoals 'Hoelang duurde het na de registratie door de politie vooraleer het PV op het parket terecht kwam?'; 'Hoeveel tijd nam het in beslag vooraleer er door het parket een (voorlopig¹⁰) definitieve beslissing werd genomen?'.

V. Besluitend

Wanneer men met parketstatistieken werkt, zal men in functie van de vraag die men als analist of onderzoeker stelt de teleenheid moeten kiezen. Afhankelijk van de vraag die men zich stelt kan het noodzakelijk zijn om de parketdata aan te vullen met informatie uit andere bronnen zoals politionele databanken.

Criminele figuren kunnen als alternatief gebruikt worden voor feiten of voornaamste tenlasteleggingen. Zij bieden de mogelijkheid tot zinnvollere meer criminologisch relevante interpretaties.

BART GOEMINNE¹¹.

Noten

¹ Het onderzoek 'Analyse van exploitatiemogelijkheden in databanken van politie en parket ten behoeve van het opsporings- en vervolgingsbeleid' werd uitgevoerd door de interuniversitaire onderzoeksgroep 'Sociale Veiligheidsanalyse' (Vrije Universiteit Brussel - Universiteit Gent) in de periode november 2002 - november 2003. Het onderzoek stond onder toezicht van en werd gefinancierd door de Dienst Strafrechtelijk Beleid.

² Bart Goeminne, Paul Ponsaers (UGent), Els Enhus (VUB).

³ ISLP staat voor 'Integrated System for the Local Police'. ISLP is een gespecialiseerd computersysteem dat door lokale politiediensten gebruikt wordt en waarin politiemensen die tot een van deze lokale politiediensten behoren hun PV's opmaken. ISLP bevat alle registraties met betrekking tot de taken van de bestuurlijke en gerechtelijke politie, inclusief verkeer.

⁴ Dit is een juridisch feit dat gekenmerkt wordt door een specificatie die niet in het Strafwetboek omschreven staat. Bijvoorbeeld: juridisch feit + welbepaalde plaats (winkeldiefstal), juridisch feit + object (handtasdiefstal), juridisch feit +

specifiek slachtoffer (weerspanningheid en sla-gen aan agent).

⁵ Redenen waarom slachtoffers geen aangifte doen zie : Goeminne, B., Enhus, E., Ponsaers, P., 2003, *Criminaliteit in de publieke ruimte, een inleiding in de hotspot analyse.*, Politeia, Brussel, p. 29.

⁶ Vatters zijn personen die informatie over binnenkomende pv's in het digitale parketdata-bestand inputten.

⁷ Doel van het Phenix-project is de 13 verschillende informatietoepassingen binnen Justitie op één lijn te brengen. Centraal in het Phenix-project staat het digitaal dossier. Doorheen de chronologie van een procedure zou bijvoorbeeld een rechter, maar ook een advocaat de voortgang in een dossier kunnen opvolgen.

⁸ Dit betreft zowel PV's als APA's (Autonome Politionele Afhandeling)

⁹ Deze conversietabel werd door ons, de onderzoekers, ontwikkeld. Als basis werd een nomenclatuur gebruikt die wordt voorgesteld door de federale politie. Deze nomenclatuur is een catalogus van feiten waarvoor de politie een proces-verbaal moet opstellen. Naast de feiten bevat de nomenclatuur een reeks criminele figuren die vaak voorkomen in de politiepraktijk. Deze criminele figuren zijn een aggregaat van feiten. Dit aggregaat bevat echter niet enkel criminele figuren in de strikte zin van het woord maar ook feiten en fenomenen. Dit is echter voer voor een andere discussie.

¹⁰ Bijvoorbeeld sepot.

¹¹ De auteur is als onderzoeker verbonden aan het Instituut voor Sociaal en Politiek Opiniononderzoek (ISPO), K.U.Leuven. en was voorheen werkzaam als onderzoeker bij de onderzoeksgroep *Sociale Veiligheidsanalyse*, UGent.

CIJFERS INZAKE DE UITHANDENGEVING

Naar aanleiding van het kernkabinet van 16 mei 2002, wilde de toenmalige minister van Justitie exact weten hoe vaak jeugdrechters een beroep doen op de procedure van uithandengeving¹. Daartoe wendde hij zich tot de Dienst voor het Strafrechtelijk beleid, die in samenwerking met het departement criminologie van het NICC, relevante gegevens voor de jaren 1999 - 2001 ver-