



Universiteit
Leiden
The Netherlands

Current challenges in statistical DNA evidence evaluation

Cereda, G.

Citation

Cereda, G. (2017, January 12). *Current challenges in statistical DNA evidence evaluation*. Retrieved from <https://hdl.handle.net/1887/45172>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/45172>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/45172> holds various files of this Leiden University dissertation.

Author: Cereda, G.

Title: Current challenges in statistical DNA evidence evaluation

Issue Date: 2017-01-12

Samenvatting

In een justitiële zaak kan de *likelihood ratio* als een statistiek gebruikt worden om een bewijsstuk op waarde te schatten. Namelijk, de mate waarin het bewijsstuk de hypothesen van de aanklager ondersteunt ten opzichte van de hypothesen van de verdediging. Een type bewijsstuk dat vaak gevonden wordt op de plaats van het delict is DNA. Dit proefschrift richt zich op de eigenschappen van de *likelihood ratio* wanneer als bewijsstuk DNA gebruikt wordt.

Het statistisch op waarde schatten van DNA-mengels (sporen met DNA van verschillende personen) is ingewikkeld omdat verschillende combinaties van DNA-profielen van bekende en/of onbekende personen compatibel kunnen zijn met het gevonden DNA-mengsel. Bovendien, als het kwantificeerbare aandeel van de DNA van één van de personen minder is dan 10% van de totale hoeveelheid, dan wordt het DNA-profiel van de persoon “gemaskeerd” door het DNA-profiel van de andere personen / of persoon. Het is erg moeilijk om met de klassieke werkwijzen van genotypering en de standaard aan statistische technieken om het genotype van deze gemaskeerde kleine hoeveelheden van DNA te detecteren. Dergelijke uiterst onevenwichtige mengsels van DNA komen echter vaak voor, bijvoorbeeld in het geval van seksueel geweld waar het DNA-mengsel wordt overheerst door dat van het slachtoffer. Oftewel, er is een dringende vraag naar een oplossing om deze kleine hoeveelheden aan DNA betrouwbaar en gemakkelijker te kunnen detecteren. Met de komst van nieuwe technologie zoals het DIP-STR (*Deletion Insertion Polymorphisms - Short Tandem Repeats*) markersysteem wordt er een oplossing geboden voor dit uiterst onevenwichtige DNA mengsel probleem.

Het oorspronkelijke doel van dit proefschrift, vervuld in hoofdstuk 3, was om een Bayesiaans statistisch model te ontwikkelen dat DIP-STR resultaten op waarde zou kunnen schatten in het licht van de belangrijke en concurrerende hypothesen; een essentieel element voor het weergeven van de potentie van deze nieuwe techniek in de toepassing voor beoefenaars. In hoofdstuk 4, hebben we vanuit een statistisch en forensisch oogpunt, en met betrekking tot toepasbaarheid en gebruikersgemak, de DIP-STR markers vergeleken met die van traditionele markersystemen, zoals klassieke STR en Y-STR markers.

Tijdens de voortgang van ons onderzoek, zijn we enkele delicate methodologische vraagstukken tegenkomen voor de forensische statistiek. Een eerste bevinding was dat wat in de literatuur een Bayesiaanse methode genoemd wordt beter gezien kan worden als een ad hoc benadering van de volledige Bayesiaanse oplossing. Vervolgens kwamen we in aanraking met het zeldzaam type match probleem: de situatie waarin er een match wordt gevonden tussen de kenmerken van bepaalde teruggewonnen DNA-materiaal en die van het DNA-controle materiaal, dit terwijl bij eerder verzamelde monsters deze match nog niet was waargenomen (dat wil zeggen,

de match was niet aanwezig in de op dat moment beschikbare database). Het zeldzaam type match probleem is in het bijzonder aanwezig in situaties waar gebruik wordt gemaakt van Y-STR (of mitochondriaal) DNA profielen, of wanneer er met de genotyperings-technieken wordt gewerkt, zoals DIP-STR markers, waarvoor de beschikbare database nog maar beperkt is in grootte.

In ons onderzoek richten we ons op de Y-STR data voor het bestuderen van zowel de nieuwe en huidige oplossingen voor het zeldzaam type probleem. In Hoofdstuk 6 hebben we de klassieke Bayesiaanse methoden (met beta-binomiale en Dirichlet-multinomiale verdelingen) herzien, en vervolgens vergeleken met een non-parametrische Bayesiaanse benadering die speciaal was ontwikkeld voor de zeldzaam Y-STR match probleem, zie Hoofdstuk 7).

Twee Frequentistische oplossingen voor het zeldzaam type probleem worden geanalyseerd in hoofdstuk 3: de discrete Laplace-methode en een nieuwe oplossing op basis van de Good-Turing schatter. Tijdens het bestuderen van oplossingen vanuit het Frequentistisch perspectief, zijn we erachter gekomen dat verschillende methodes gebaseerd zijn op data reductie, en dat dat zelden aan bod komt in de Forensische literatuur. Bovendien zijn er verschillende niveaus van onzekerheid welke in acht genomen dienen te worden. Door te werken aan beide de Frequentistische en de Bayesiaanse methodes hebben we het verschil tussen de twee benaderingen beter leren te begrijpen, en het verschil tussen de volledige - en de plug-in Bayesiaanse benadering. Om de volledige Bayesiaanse *likelihood ratio* te verkrijgen onder verschillende regulariteits-condities, hebben we een lemma bewezen.

Ter afsluiting van het project is een van de ontwikkelde Bayesiaanse methodes voor het zeldzaam type match probleem ook toegepast op de DIP-STR data in Hoofdstuk 8. Dit model dat geconstrueerd is voor de DIP-STR data, en nog in haar kinderschoenen staat, is verder verbeterd door het uit te breiden op een manier waarbij het wordt toegestaan om de onzekerheid van de parameters op te nemen in het model op een consistente Bayesiaanse manier.