

VU Research Portal

Beyond the average

Aarts, E.

2016

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Aarts, E. (2016). *Beyond the average: Choosing and improving statistical methods to optimize inference from complex neuroscience data*.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

Nederlandse samenvatting

Binnen de neurowetenschappen zijn er de afgelopen decennia als gevolg van technologische ontwikkelingen meer geavanceerde experimentele technieken beschikbaar gekomen om het zenuwstelsel te bestuderen. Deze technieken maken het bijvoorbeeld mogelijk om nauwkeuriger te meten en continue metingen langduriger uit te voeren, en om dieper in te zoomen op biologische processen van het zenuwstelsel: van het niveau van een dier naar het celniveau, naar mechanismes binnen de cel, tot mechanismes die plaats vinden op subcellulair niveau. Ook kunnen metingen geautomatiseerd en gestandaardiseerd worden, en kan de totale doorlooptijd van studies makkelijker verkort worden, bijvoorbeeld door geautomatiseerd meerdere experimenten parallel uit te voeren.

Deze technologische innovaties dragen bij aan de enorme kennisontwikkeling binnen de neurowetenschappen, maar zorgen er ook voor dat datastructuren steeds complexer worden. Momenteel wordt de data die verzameld wordt in experimenten nog vaak samengevat en beschreven middels klassieke statistische technieken. Deze technieken, zoals t-test, ANOVA en regressie, zijn geschikt voor experimentele designs die resulteren in een beperkt aantal (tot maximaal enkele honderdtallen) onafhankelijke waarnemingen, en vormen niet altijd een optimale match met de complexe datastructuren die voortkomen uit de nieuwste technologieën. Het gebruik van klassieke analyse technieken leidt dan tot suboptimaal gebruik van de informatie, en doet geen recht aan de complexiteit van de data. Dit is zonde: enerzijds kunnen de betrouwbaarheid en validiteit van de onderzoeksconclusies in het gedrang komen, anderzijds kan waardevolle informatie onbenut blijven.

In het kader van de technologische innovaties binnen de neurowetenschappen, is het doel van dit proefschrift om te onderzoeken welke statistische methodes recht doen aan de complexiteit van de datastructuren die daaruit voortvloeien, en waar nodig nieuwe statistische methodes te ontwikkelen die de informatie binnen deze complexe datastructuren optimaal benutten.

In **Hoofdstuk 2** en **Hoofdstuk 3** richtten we ons op de vraag wat de optimale analysemethode is als data voortkomt uit onderzoeksdesigns waarbij meerdere metingen aan hetzelfde onderzoeksobject gedaan worden. Dit gebeurt bijvoorbeeld wanneer we meerdere neuronen van hetzelfde dier meten, of aan elke neuron meerdere

metingen doen, bijvoorbeeld aan meerdere synapsen of dendriten. Metingen die van hetzelfde onderzoeksobject afkomstig zijn lijken vaak meer op elkaar dan op metingen uit andere onderzoeksobjecten: ze zijn "geclusterd" of "genest". Derhalve kunnen deze metingen niet als onafhankelijke metingen beschouwd worden. Klassieke analyse technieken gaan er echter wel vanuit dat metingen onafhankelijk van elkaar zijn. Middels een literatuurstudie in 5 gerenommeerde wetenschappelijke tijdschriften demonstreerden we in **Hoofdstuk 2** dat een dergelijk "genest" onderzoeksdesign vaak voorkomt binnen de neurowetenschappen (53% van de 314 bekeken wetenschappelijke artikelen maakte gebruik van dit onderzoeksdesign). In **Hoofdstuk 2** beperkten we ons tot het onderzoeksdesign waarin alle metingen aan hetzelfde onderzoeksobject tot dezelfde experimentele conditie behoren (Design A). Bijvoorbeeld, alle neuronen uit hetzelfde dier hebben hetzelfde genotype (wild type, knock-out) of krijgen dezelfde experimentele behandeling. In dat hoofdstuk lieten we zien dat de validiteit van de getrokken onderzoeksconclusies in het gedrang komt wanneer er geen rekening gehouden wordt met dergelijke afhankelijkheid in de data. Als we twee groepen met elkaar vergeleken liep de kans dat we onterecht concludeerden dat experimentele condities van elkaar verschillen (Type I fout) soms op tot wel 80% in plaats van de 5% die normaal gesproken als een maatstaf wordt aangehouden. Ook lieten we in dit hoofdstuk zien hoe je statistisch gezien rekening kan houden met deze afhankelijkheid door multilevel analyse toe te passen, en hoe onderzoekers de statistische power van hun studie (de kans om een bestaand experimenteel effect ook statistisch te detecteren) kunnen maximaliseren door de verdeling van het aantal observaties per onderzoeksobject, en het aantal onderzoeksobjecten per experimentele conditie, te optimaliseren.

In **Hoofdstuk 3** richtten we ons op een onderzoeksdesign waarbij eveneens meerdere metingen aan hetzelfde onderzoeksobject gedaan worden, maar deze metingen binnen een onderzoeksobject tot verschillende experimentele condities kunnen behoren (Design B). Waar in Design A het gemiddelde van de afhankelijke variabele kan variëren over onderzoeksobjecten, kan in Design B ook het experimentele effect variëren over onderzoeksobjecten. Ook in dit hoofdstuk lieten we zien dat het belangrijk is om een statistische methode te kiezen die rekening houdt met de geneste structuur van de data, zoals multilevel analyse. Net als in Design A, kan het negeren van de geneste structuur bij Design B data (en dus het onterecht toepassen van klassieke analyse technieken) resulteren in een vergrootte Type I fout, namelijk als het effect van de experimentele manipulatie inderdaad varieert over onderzoeksobjecten. Echter, als alleen de gemiddelde waarde van de controle conditie (baseline) varieert over onderzoeksobjecten, dan resulteert het negeren van de geneste structuur van de data in een veel lagere statistische power. Beide gevolgen – verhoogde Type I fout en verlaagde statistische power – zijn zeer onwenselijk voor wetenschappelijk onderzoek. Daarnaast lieten we zien dat multilevel analyse bij Design B data waardevolle inzichten kan geven over de generaliseerbaarheid van het experimentele effect over (biologisch intrinsiek) verschillende condities, en over de oorzaak van de variatie in het experimentele effect over onderzoeksobjecten.

In **Hoofdstuk 4** en **Hoofdstuk 5** richtten we ons op een systeem om het gedrag van muizen te fenotyperen door middel van geautomatiseerde en langdurige observatie

in de thuishooi: de PhenoTyper. Binnen de neurowetenschappen zijn diermodellen een waardevolle manier om de relatie tussen genotype (de genetische opmaak van dieren) en fenotype (het gedrag van dieren) te bestuderen, aangezien diermodellen een gecontroleerde experimentele setting bieden waarbinnen de effecten van genen en omgeving op (variatie in) gedrag bestudeerd kunnen worden. Om de neurale mechanismes van complex gedrag beter te begrijpen, net als afwijkingen in gedrag die kenmerkend zijn voor neurologische en psychiatrische aandoeningen, zijn muizen verreweg het meest gebruikte diermodel. Door gebruik te maken van observatie in de thuishooi kunnen we multidimensionale aspecten van het gedrag bestuderen op een geautomatiseerde manier over een langdurige periode, terwijl mens-dier interacties geminimaliseerd worden. Onder multidimensionale aspecten van gedrag verstaan we bijvoorbeeld spontaan (basaal) gedrag, gedrag wanneer het dier op de proef gesteld wordt, en de progressie of verandering van gedrag over tijd.

In **Hoofdstuk 4** richtten we ons op een optimale methode om angst te testen binnen de PhenoTyper. De PhenoTyper is een geïnstrumenteerde thuishooi waarin het gedrag van de muis door middel van een op video gebaseerd observatie systeem automatisch gemonitord wordt. Binnen deze angsttest, de "light spot test", maakten we gebruik van de natuurlijke neiging van muizen om actief te worden tijdens het begin van de nacht en dan te gaan eten, en de aversie van muizen tegen (fel) licht. Tijdens de light spot test werd, na een gewenningsperiode van 3 dagen in de thuishooi, 15 minuten na de start van de nacht een lichtstraal gericht op de voederbak voor de periode van een uur. Analyses lieten zien dat muizen tijdens dit uur significant minder tijd doorbrengen buiten hun schuilplaats in vergelijking met dezelfde tijdsperiode op andere testdagen. Als de lichtstraal weer werd uitgezet vertoonden de muizen weer het normale geobserveerde gedrag. Wanneer we de muizen een angstremmend middel toedienden, werd de geobserveerde gedragsreactie op de lichtstraal significant verminderd. Dit sterkt het idee dat de test inderdaad angst weerspiegelt.

In **Hoofdstuk 5** richtten we ons op het analyseren van het patroon van het gedrag van muizen over tijd dat men kan observeren tijdens langdurige observaties in de thuishooi. We hebben op basis van Markov-modellen een statistische methode ontwikkeld en geïmplementeerd om de organisatie van gedrag over tijd te beschrijven. Om het mogelijk te maken om gedragspatronen van groepen muizen te beschrijven en statistisch te vergelijken was een uitbreiding van bestaande modellen nodig. Door middel van een simulatie studie hebben we onderzocht of het uitgebreide model goed werkt. Daarnaast hebben we het nieuwe uitgebreide geavanceerde model gebruikt om gedragspatroon van jonge en oude muizen te vergelijken. Deze resultaten hebben we vergeleken met de uitkomsten die verkregen worden als klassieke analyse methodes gebruikt worden, die niet naar het gedragspatroon kijken maar gedragsmaten samenvatten over de hele observatie periode. Hoewel onze simulatie studie liet zien dat het door ons ontwikkelde model nog verdere optimalisatie vereist, liet het echte data voorbeeld al wel duidelijk het voordeel zien van het gebruik van ons geavanceerde model boven de klassieke analyse technieken. Door met de ontwikkelde statistische methode naar het patroon van het gedrag te kijken, ontdekten we verschillen in gedragspatronen tussen jonge en oude muizen die met klassieke analyse methodes onopgemerkt bleven. Bovendien konden we met onze nieuwe methode ook verklaren

waar geobserveerde verschillen in gedrag vandaan kwamen.

Tenslotte vatten we in **Hoofdstuk 6** de resultaten van het gehele PhD-project over optimale statistische en methodologische technieken binnen de neurowetenschappen samen. Daarnaast bereekten we de uitdagingen en kansen wanneer er meerder metingen aan hetzelfde onderzoeksobject gedaan worden, de validiteit van geïnstrumenteerde thuiscoöien en mogelijke ontwikkelpunten, en de toegevoegde waarde van en het toekomstig onderzoek aan de door ons ontwikkelde statistische methode om de organisatie van gedrag over tijd te beschrijven. De resultaten in dit proefschrift onderstrepen het cruciale belang van het gebruik van optimale statistische en methodologische methodes: methodes die de kansen welke de nieuwe experimentele technieken binnen de neurowetenschappen bieden, volledig benutten.