

Statistiek: het onzichtbare zichtbaar maken

Frank van der Meulen
Vrije Universiteit Amsterdam
24 Januari 2025

Inhoudsopgave

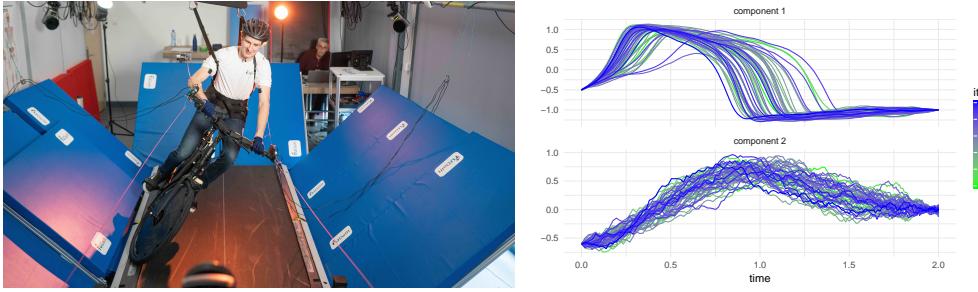
1	Introductie	1
2	Onderzoek	5
2.1	Een toy-example	7
2.2	Een voorbeeld uit de fylogenie	14
2.3	Mathematische Statistiek aan de VU	15
3	Onderwijs	16
3.1	De rol van onderwijs en professionalisering aan de universiteit	16
3.2	Mijn onderwijs innovatie voorstel	17
3.3	Studierendement en de rol van de student	17
4	Taal	19
5	Dank	20

Geachte rector, geachte aanwezigen

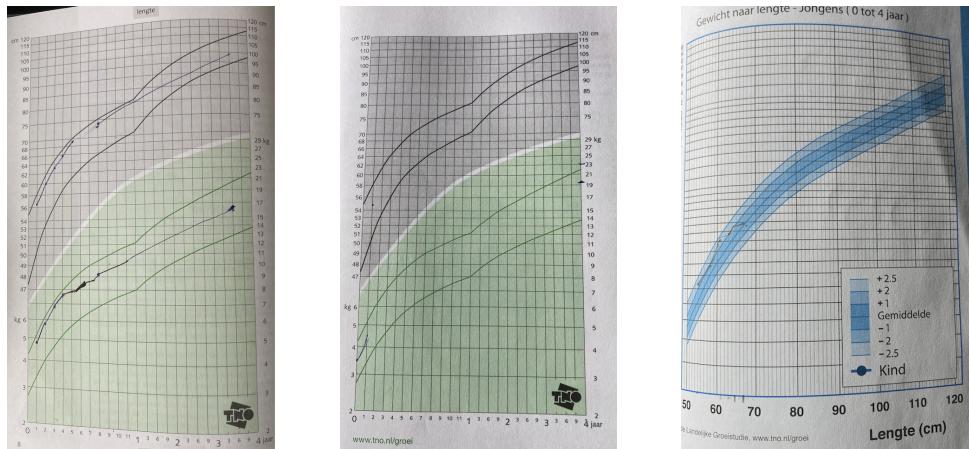
1 Introductie

Het is een voorrecht om hier vandaag te staan. De leerstoel die ik vandaag officieel aanvaard heet “mathematische statistiek”.

Als familie of vrienden me vragen waar mijn onderzoek over gaat, het liefst even in één zin – ik ben echt niet dom dus je kunt het me wel vertellen – bemerk ik altijd een aarzeling bij mezelf. Na wat aandringen kies ik een voorbeeld wat hopelijk enigszins tot de verbeelding spreekt, en sluit het onderwerp af. Dat is natuurlijk geen goede reclame voor mijn vak. Een collega hoorde ik eens iets zeggen in de strekking van “Je moet de uitleg over je onderzoek op de achterkant van een bierviltje kunnen schrijven.”. Ik knikte wat mee en dacht daarbij, misschien als ik heel klein schrijf dan lukt me dat ook wel. En ook, “dan zal het wel niet zo heel veel voorstellen”, wat natuurlijk niet waar is.



Figuur 1: Twee figuren voortkomende uit eigen onderzoek. Links: [2]. Rechts: [3].



Figuur 2: Groeiboekjes.

De linker afbeelding die u ziet heeft betrekking op een voorbeeld dat ik soms gebruik om iets te vertellen over mijn vakgebied. Maar eigenlijk vind ik de rechter afbeelding en de wiskundige vraagstukken die daarachter schuilen minstens zo interessant. Het verhaal daarachter is alleen wat lastiger uit te leggen. Vandaag heb ik de mogelijkheid om langer uit te wijden over mijn vakgebied en zo mathematische statistiek bij een breed publiek onder de aandacht te brengen. Ik zal zometeen terugkomen bij deze figuren.

Eerst het vakgebied statistiek, waar gaat dit over? In het Engels kan het kort samengevat worden: "Making sense of data". Dit past overigens gerust op de achterkant van een bierviltje maar is niet specifiek. Het is bovendien niet in het Nederlands, dus dat past goed in de ontwikkeling van de Nederlandse universiteiten van het afgelopen decennium. "Making sense of data": dat klinkt in eerste instantie als een weinig spannend vakgebied. Toch zijn er vele interessante statistische vraagstukken. Laat ik eerst twee praktische voorbeelden geven.

Voorbeeld 1. Neem bijvoorbeeld het groeiboekje (Figuur 2). Hierin zien we eigenlijk 2 grafieken in een gedrukt, waarbij de lengte en gewicht van een van mijn zoons zijn ingevuld. Naast het feit dat 2 figuren in 1 plaatje zijn gezet, is ook de horizontale as opmerkelijk. Dit is een voorbeeld van beschrijvende statistiek. Er zijn echter al twee vetgedrukte curves toegevoegd: één in het groen en één in het zwart. Wat betekenen deze curves eigenlijk en hoe zijn ze vastgesteld? De medewerksters van het JGZ konden me geen antwoord geven. Vooruitlopend op mijn onderzoeksinteressen: in feite zien we hier een voorbeeld van een stochastisch proces. Iets dat over de tijd varieert en waar onzekerheid op zit. We meten een karakteristiek op enkele tijdstippen om de groei van het kind te kunnen beoordelen. We gebruiken dit echter ook om te interpoleren en zouden het ook kunnen

gebruiken om te extrapoleren: een voorspelling van de lengte op latere leeftijd. Daar zit natuurlijk onzekerheid op. Hoe kunnen we deze kwantificeren? Is er een optimale methode? Wat bedoelen we met optimaal? Het vanuit wiskundig oogpunt bestuderen van statistisch methoden is exact waar het om gaat in mathematische statistiek. Zouden we verder de onzekerheid kunnen reduceren door karakteristieken van de ouders of grootouders mee te nemen?

Het gekozen voorbeeld is bewust eenvoudig, maar in plaats van lengte en gewicht zou de meting ook een indicatie van het stadium van een ziekte kunnen zijn. In dat geval, stel dat we van andere patiënten, met vergelijkbare eigenschappen, het effect weten van verschillende behandelplannen, zou dat niet kunnen helpen om een effectief behandelschema op te stellen? Dit zijn typische vraagstukken waarbij statistiek van groot belang is.

De getoonde grafiek is van mijn oudste zoon, die van mijn andere zonen zal ik nu tonen. Het moge duidelijk zijn dat een liefde voor mathematische statistiek zeker niet samen hoeft te gaan met het zelf verzamelen van data. Verder is ook te zien dat TNO ergens tussen 2011 en 2015 besloten heeft dat het niet handig is om de twee grafieken voor gewicht en leeftijd gecombineerd in 1 figuur weer te geven.

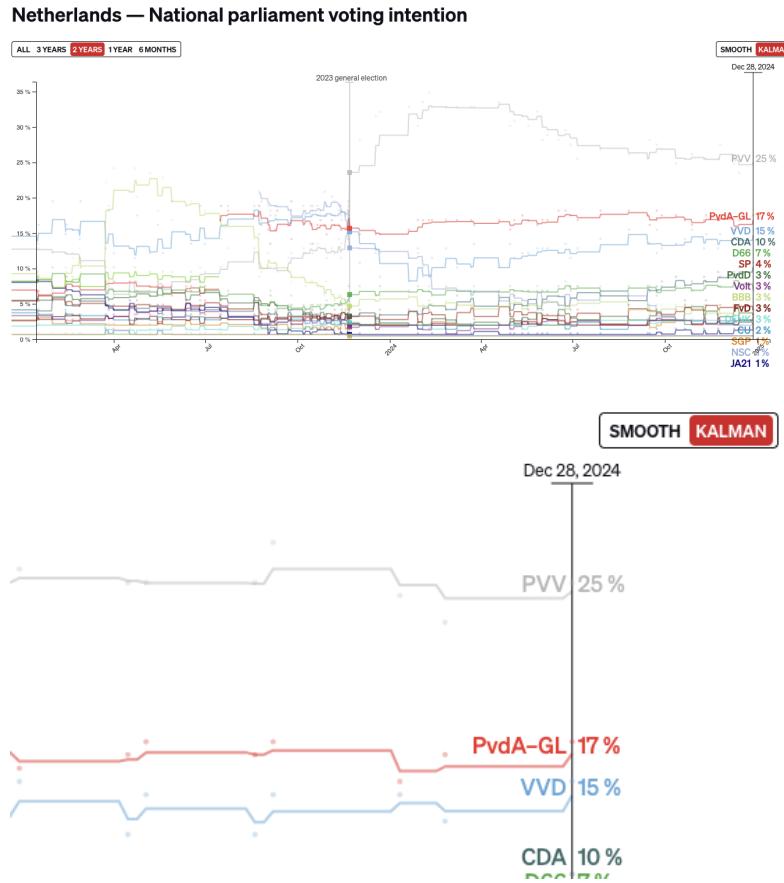
Voorbeeld 2. Als tweede voorbeeld zou ik de zogenaamde "poll of the polls" van de website <https://www.politico.eu/> willen noemen. Voor alle EU-landen worden hier verschillende opiniepeilingen gecombineerd. Iedere peiling, in het Engels "poll", geeft aan welk percentage van de kiezers in een land op een specifieke partij stemt. Deze percentages zijn enigszins lastig waar te nemen punten in de figuur, die we beter zien als we inzoomen. Zie Figuur 3.

Nu worden er over de tijd genomen meerdere peilingen gemaakt door verschillende bureaus. Die peilingen rapporteren doorgaans verschillende getallen, en de vraag rijst op welke getallen we af moeten gaan. Daarnaast is niet iedere peiling gebaseerd op een even grote steekproef. Zoals omschreven op de website "One poll can be misleading, so POLITICO aggregates data from multiple sources to give a more accurate picture." In feite wordt hier op grond van Kalman filtering uit de individuele peilingen een geaggregeerde peiling geconstrueerd. Het onderliggende algoritme voert terug tot 1960 en vormt de basis voor veel problemen uit signaalverwerking, maar heeft ook talloze toepassingen in andere gebieden, zoals dit voorbeeld illustreert.

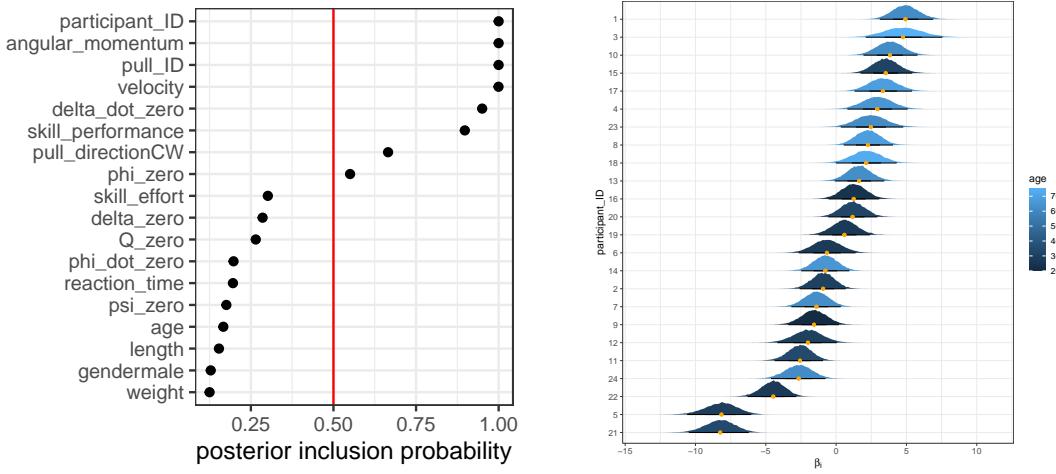
Alvorens verder te gaan, mogelijk is het u opgevallen dat de gegeven voorbeelden twee karakteristieken bevatten:

1. het betreft sequentiële data, met andere woorden data verkregen over opeenvolgende tijdstippen;
2. voordat we de data verkregen, was er onzekerheid over de waarden die we zouden krijgen. Zo wist ik voordat mijn zoon geboren werd, niet na hoeveel maanden hij welk gewicht en welke lengte zou hebben. Hetzelfde geldt voor opiniepeilingen. Deze onzekerheid wordt ook wel stochasticiteit genoemd.

Dus een herformulering is dat we kijken naar stochastische processen, waar we iets over willen leren, bijvoorbeeld voorspellen, op grond van data waargenomen op zekere tijdstippen. Deze data geven ons beperkte informatie, over iets wat we niet direct waar kunnen nemen. Iets wat in feite niet zichtbaar is. Bij de poll of the polls bijvoorbeeld zouden we iedere dag alle stemgerechtigden in een land kunnen vragen wat hun politieke voorkeur is. Dit is echter niet praktisch uitvoerbaar en de taak van de statisticus is nu om deze daadwerkelijke politieke voorkeur te achterhalen met behulp van beperkte informatie. Deze beperkte informatie is in dit geval de informatie verstrekt door meerdere opiniepeilingen over de tijd.



Figuur 3: Peilingen politieke partijen Nederland rond Juli 2024. De onderste figuur zoomt in op een stukje van de bovenste figuur teneinde de individuele peilingen beter zichtbaar te maken. Bron: [website politico.eu](http://politico.eu).



Figuur 4: Links: de aposterior kans dat een voorspeller in the verklarende model zit. Rechts: visualisatie van de vaardigheid van testpersoon om de fiets in evenwicht te houden.

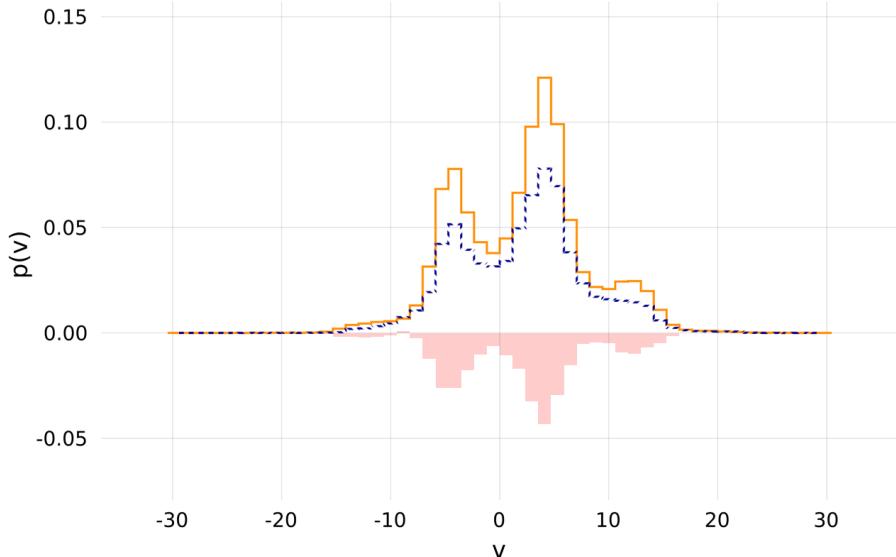
2 Onderzoek

Nu ik deze voorbeelden gegeven heb, wil ik terugkomen op de twee figuren waarmee ik gestart ben. (Figuur 1). Deze zijn beide voortkomend uit onderzoek dat ik verricht heb.

1. Links ziet u een persoon in een testopstelling op een fiets. Deze foto is onderdeel van onderzoek uitgevoerd tezamen met Arend Schwab, Marco Reijne en Frans van der Helm van de TU Delft ([2]). Aan het experiment doen meerdere personen mee. Onder verschillende omstandigheden neemt iedere deelnemer aan het experiment plaats op de fiets en wordt vervolgens enigszins uit balans gebracht. Er wordt dus een verstoring aangebracht waarbij geregistreerd wordt of de persoon kan corrigeren of valt. Er zijn meerdere sensoren aangebracht, deze meten onder andere de hellingshoek, stuurhoek en zijwaardse positie op het moment van de verstoring. Niet iedere deelnemer doet exact hetzelfde experiment: als iemand bij een klein zetje al valt, dan gaan we niet moedwillig nog harder duwen. De hoofdonderzoeksraag is “Bij welke waarde van het impulsmoment (angular momentum) is de kans op vallen 50%?”. En welke factoren hebben daar wel of geen invloed op? Hier valt te denken aan lengte, gewicht en leeftijd bijvoorbeeld. Is het verder mogelijk de deelnemers aan het experiment te ordenen naarmate ze stabiel op de fiets kunnen blijven zitten? In Figuur 5 ziet u een screenshot van wat ik aangeleverd heb gekregen, dit zijn de ruwe data. In Figuur 4 ziet u twee figuren die voortgekomen zijn uit de statistische analyse. In de linker figuur zien we welke factoren van invloed zijn, of beter gezegd, de kans dat we deze factoren mee moeten nemen. In de rechter figuur zien we gekwantificeerd hoe goed een ieder van de deelnemers inherent in staat is om op de fiets te blijven zitten. Hierbij is gecorrigeerd voor het feit dat niet alle deelnemers exact hetzelfde experiment hebben ondergaan als ook verschillen in bijvoorbeeld leeftijd, lengte en gewicht van deelnemers. Dit is een voorbeeld van onderzoek dat direct door een toepassing gemotiveerd is: onderzoekers uit een ander vakgebied doen een experiment, verzamelen data, en hebben daarbij concrete vragen. De data “an sich” bevatten informatie, maar deze is verborgen, oftewel niet zichtbaar.
2. Rechts in Figuur 1 ziet u een afbeelding van een gesimuleerde diffusie brug voor het zogenaamde FitzHugh-Nagumo model. Het is een wiskundig model dat gebruikt wordt voor het beschrijven en verklaren van neurologische activiteit. Het plaatje laat zien hoe we tussen

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	participant	gender	age	weight	length	skill_perform	skill_effort	velocity	pull_ID	initial_search	pull_direc	outcome	pull_force	angular_mom	reaction_time	phi_phi_zero	phi_dot_phi_zero
2	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	NA	TRUE	CW	0	20	4.23E+06	3.49E+06	-1.35E+06	1.65E+05
3	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	NA	TRUE	CW	1	50	9.99E+06	4.50E+04	-1.27E+06	-2.28E+03
4	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	1	FALSE	CW	0	30	6.76E+06	4.50E+03	-1.18E+06	1.94E+05
5	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	2	FALSE	CW	1	40	9.26E+06	1.90E+04	-9.79E+05	1.63E+05
6	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	3	FALSE	CW	1	60	1.19E+07	7.60E+04	-1.42E+06	-1.35E+04
7	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	4	FALSE	CW	0	20	2.94E+06	9.90E+04	-9.77E+05	8.10E+04
8	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	5	FALSE	CW	1	40	8.27E+06	4.90E+04	-5.80E+04	3.56E+04
9	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	6	FALSE	CW	0	20	3.71E+06	6.55E+04	-2.05E+06	3.69E+03
10	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	7	FALSE	CW	1	50	9.84E+06	3.86E+04	8.63E+04	9.41E+04
11	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	8	FALSE	CW	1	40	8.62E+06	7.90E+04	-9.05E+05	4.40E+04
12	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	9	FALSE	NA	0	0	0	NA	NA	NA
13	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	10	FALSE	CW	1	40	8.14E+06	4.85E+04	-9.74E+05	-3.93E+04
14	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	11	FALSE	NA	0	0	0	NA	NA	NA
15	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	12	FALSE	CW	0	10	1.31E+06	1.53E+05	-1.91E+06	2.22E+04
16	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	13	FALSE	CW	0	30	6.12E+06	5.35E+04	-6.23E+04	-4.83E+04
17	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	14	FALSE	CW	0	30	6.73E+06	5.60E+04	-8.21E+05	6.26E+04
18	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	15	FALSE	CW	1	60	1.27E+07	4.05E+04	-8.75E+05	4.27E+04
19	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	16	FALSE	CW	1	60	1.18E+07	4.55E+04	-1.78E+06	-2.15E+04
20	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	17	FALSE	CW	1	60	1.21E+07	6.40E+04	-1.52E+06	1.97E+04
21	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	18	FALSE	CW	0	10	8.27E+05	5.00E+02	-7.39E+05	5.92E+04
22	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	19	FALSE	CW	0	30	6.54E+06	2.45E+04	-1.96E+03	5.45E+04
23	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	20	FALSE	CW	1	30	1.00E+07	1.17E+04	-1.00E+06	3.95E+04
24	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	21	FALSE	NA	0	20	4.25E+06	8.20E+04	-8.89E+05	7.75E+04
25	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	22	FALSE	NA	0	20	4.08E+07	4.70E+04	-1.45E+06	4.89E+04
26	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	23	FALSE	NA	1	80	1.56E+07	1.75E+04	-1.80E+05	4.57E+04
27	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	24	FALSE	CW	1	60	1.16E+07	4.00E+04	-3.14E+05	1.28E+05
28	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	25	FALSE	CW	1	80	1.55E+07	2.50E+03	-3.49E+05	2.78E+04
29	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	26	FALSE	CW	0	30	6.72E+06	7.95E+04	-3.38E+05	4.57E+04
30	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	27	FALSE	CW	0	50	1.07E+07	2.20E+04	-1.19E+06	1.84E+04
31	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	28	FALSE	CW	1	90	1.70E+07	2.80E+04	-2.10E+04	-3.68E+04
32	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	29	FALSE	CW	0	50	9.39E+06	4.50E+03	-1.20E+06	7.88E+04
33	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	30	FALSE	CW	0	50	1.02E+07	4.85E+04	-6.30E+04	2.49E+04
34	1	female	61	8.36E+07	1.61E+06	8.13E+07	3.80E+06	12	31	FALSE	CW	0	60	1.18E+07	5.19E+07	-2.67E+06	3.95E+04

Figuur 5: Screenshot van de data geanalyseerd in [2].



Figuur 6: Figuur overgenomen uit [3]. Het bevestigt een vermoeden over absolute continuïteit van kansmaten.

2 observaties kunnen interpoleren. Uitleggen waarom dit plaatje een stukje van een lastig statistisch vraagstuk oplost, is natuurlijk veel lastiger dan in voorgaande gevallen. Het getoonde plaatje kan ik vervangen door vele andere, bijvoorbeeld voor diffusiebruggen op een manifold, of geconditioneerde processen gedefinieerd door een stochastische partiële differentiaalvergelijking.

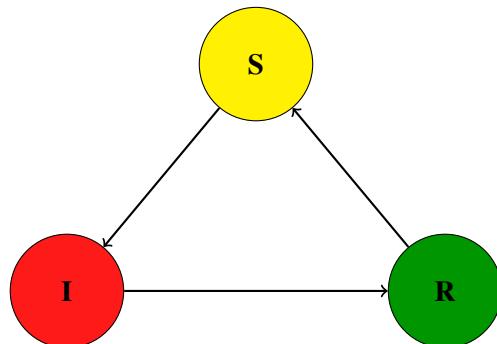
Uit dezelfde publicatie ([3]), komt Figuur 6. Dit is naar mijn mening misschien nog wel interessantste plaatje: het illustreert een vermoeden dat ik vele jaren had dat bepaalde voorwaarden noodzakelijk zijn voor absolute continuïteit van zekere kansmaten. Het bevestigt in feite een vermoeden wanneer iets niet werkt. En net zoals we in wiskunde proberen te bewijzen wanneer iets wel werkt, willen we ook precies weten wanneer iets niet werkt. Helaas is het getoonde plaatje natuurlijk geen bewijs en voor dit specifieke probleem zou ik graag exact willen weten wanneer iets werkt en wanneer niet.

Als u bij het laatste voorbeeld denkt, waar heeft hij het over, dan bevindt u zich in goed gezelschap. Hedendaagse wetenschap is zeer gespecialiseerd en alleen al het uitleggen wat het probleem is, of waarom het interessant is, kan gerust een uur duren, ook voor vakgenoten. Om die reden zal ik een zogenaamd “toy-example” bespreken om iets meer te vertellen over mijn onderzoek. Hiermee bedoel ik een voorbeeld dat niet al te moeilijk is en waar we van kunnen leren, maar evenwel vaak net te simpel om direct toepasbaar te zijn. Bovendien, het is een toy, dus we kunnen ermee spelen. Zoals John Cleese al zei “We can play, and that is what allows our natural creativity to surface.” (Creativity In Management <https://www.youtube.com/watch?v=Pb5oIIP062g>). Zulke toymodels worden veelvuldig gebruikt in de wiskunde om intuïtie te ontwikkelen voor meer realistische modellen.

2.1 Een toy-example

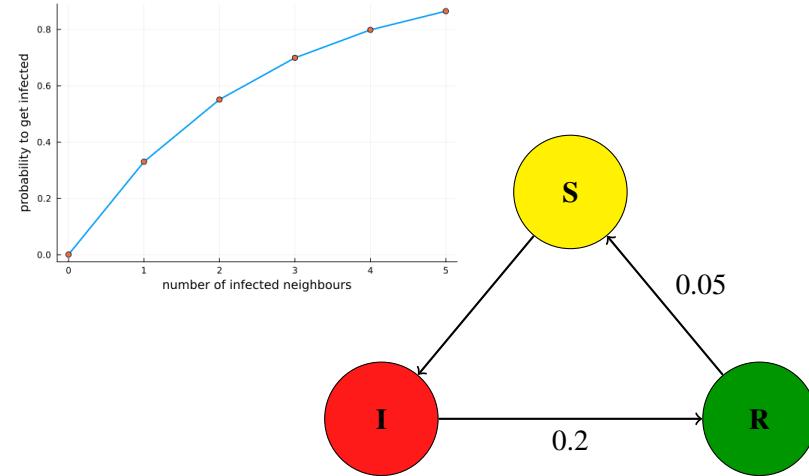
Het voorbeeld dat ik beschouw is karakteristiek voor heel veel statistische problemen waar ik me mee bezighoud. Iedere wiskunde student heeft waarschijnlijk tijdens zijn of haar studie hiervan een variant gezien, zij het meestal zonder statistische component. Het voorbeeld voert terug tot werk van Kermack & McKendrick in 1927 [4], die een wiskundig model voor het rondgaan van een virus onder mensen bestudeerden, bijvoorbeeld het influenza virus. Het doel was om te onderzoeken hoe zekere factoren de grootte van de epidemie beïnvloeden. Wanneer is het bijvoorbeeld zo dat aan het einde van de epidemie iedereen geïnfecteerd is geweest? Dit hoeft natuurlijk niet. Ik zal een variant van hun model geven.

Een eenvoudige beschrijving van het rondgaan van een virus is als volgt: eerst ben je vatbaar voor de ziekte (susceptible), vervolgens kun je geïnfecteerd raken (infected) and uiteindelijk herstel je (recovered), ervan uitgaande dat dat altijd gebeurt. Vervolgens kun je misschien weer opnieuw vatbaar worden, etc. We kunnen dit als volgt visualiseren, waarbij we ook meteen een kleurenschema vastleggen:



Laten we eerst naar een klein voorbeeldje kijken hoe het verloop van het virus kan gaan met slechts 6 personen die niet van locatie wisselen. Zie Figuur 7. Ik heb dit voorbeeld handmatig geconstrueerd.

Hoe modelleren we het rondgaan van een virus op een gestructureerde wijze? Kunnen we een model maken? Het ligt voor de hand dat het model meeneemt dat iemand die veel in contact komt met geïnfecteerde individuen, een hogere *kans* heeft zelf ook geïnfecteerd te raken. Stel dat we de status van alle personen op tijdstip t weten. De figuur hieronder geeft weer hoe we nu voor ieder persoon zijn of haar status op het volgende tijdstip bepalen en is een voorbeeld van een *stochastisch model*.



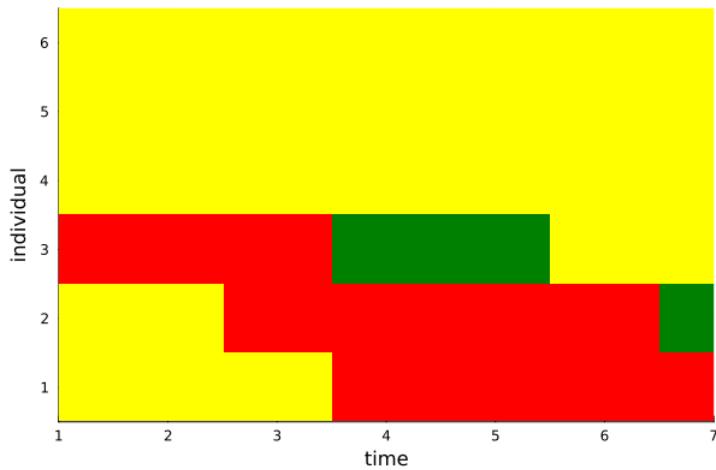
De pijl van I naar R geeft aan dat iemand die geïnfecteerd is met kans 0.2 het volgende tijdstip hersteld is. Net zo geeft de pijl van R naar S aan dat iemand die hersteld is opnieuw vatbaar wordt met kans 0.05. De transitie van S naar I is iets gecompliceerder: het ligt voor de hand dat de kans op deze transitie afhangt van het aantal geïnfecteerde buren. We zullen in het vervolg aannemen dat personen tot maximale afstand 2 tot buren gerekend worden. De grafiek linksboven geeft exact weer hoe de kans op transitie vanuit status vatbaar naar geïnfecteerd afhangt van het aantal geïnfecteerde buren. Alhoewel het wat lastig te zien is, als er geen geïnfecteerde buren zijn, dan is er een kleine kans van 0.1% dat een individu geïnfecteerd raakt.

We zouden dit zo in deze zaal na kunnen spelen. Laten we aannemen dat iedereen vatbaar is aanvankelijk en misschien 2 personen in het midden van de zaal geïnfecteerd zijn. Iedereen die vatbaar is, telt zijn of haar aantal geïnfecteerde buren. Vervolgens gooien alle vatbare personen, de meesten hier, een muntje op dat bepaald of ze daarna geïnfecteerd zijn. Dat muntje is niet zuiver, de kans op kop – laten we zeggen dat deze uitkomst betekent dat je geïnfecteerd raakt – is niet voor iedereen hetzelfde. De 2 geïnfecteerden gooien een ander type muntje, zij blijven ofwel geïnfecteerd, ofwel herstellen. Als iedereen zijn muntje heeft opgegooid, herhalen we ditzelfde proces vele keren.

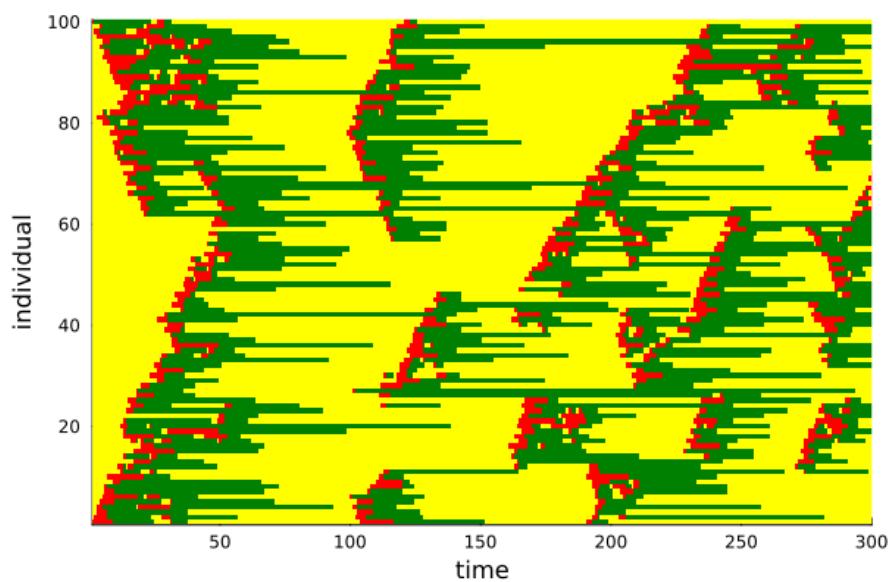
Het is veel sneller om dit op een computer na te bootsen. Dit simuleren heet *stochastische simulatie* en berust op het genereren van toevalsggetallen: willekeurige getallen tussen 0 en 1. In Figuur 8 zien we het resultaat van één zo'n simulatie met 100 personen over 300 tijdstappen. Als we de kans om van S naar R te gaan van 0.05 naar 0.15 veranderen, dan verwachten we dat gemiddeld de groene balkjes korter zijn. Dit is inderdaad het geval, zoals de stochastische simulatie in Figuur 9 laat zien.

Het betreft een toymodel, we maken geen onderscheid tussen personen: iedereen heeft dezelfde afweer. Bovendien blijft iedereen op zijn plek, precies zoals we hier zitten.

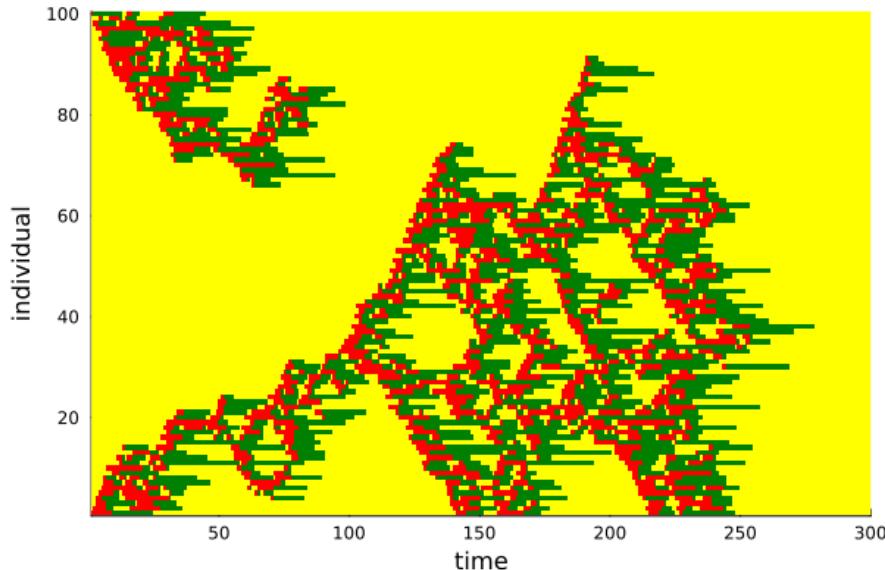
Uiteraard kunnen we niet iedereen op ieder tijdstip meten. Van vele personen weten we hun status



Figuur 7: Verloop van besmetting van virus. Op tijdstip pen 1 en 2 is alleen individu 3 geïnfecteerd, terwijl alle andere individuen vatbaar zijn. Op tijdstip 3 wordt individu 2 geïnfecteerd. Op tijdstip 4 is individu 3 hersteld, is individu 2 nog steeds geïnfecteerd en raakt ook individu 1 geïnfecteerd.



Figuur 8: Een voorbeeld met 100 individuen over 300 tijdstappen. De “buren” van individu 12 zijn individuen 10, 11, 13 en 14. Aanvankelijk zijn er 2 geïnfecteerden.



Figuur 9: Een voorbeeld met 100 individuen over 300 tijdstappen. De “buren” van individu 12 zijn individuen 10, 11, 13 en 14. Aanvankelijk zijn er 2 geïnfecteerden.

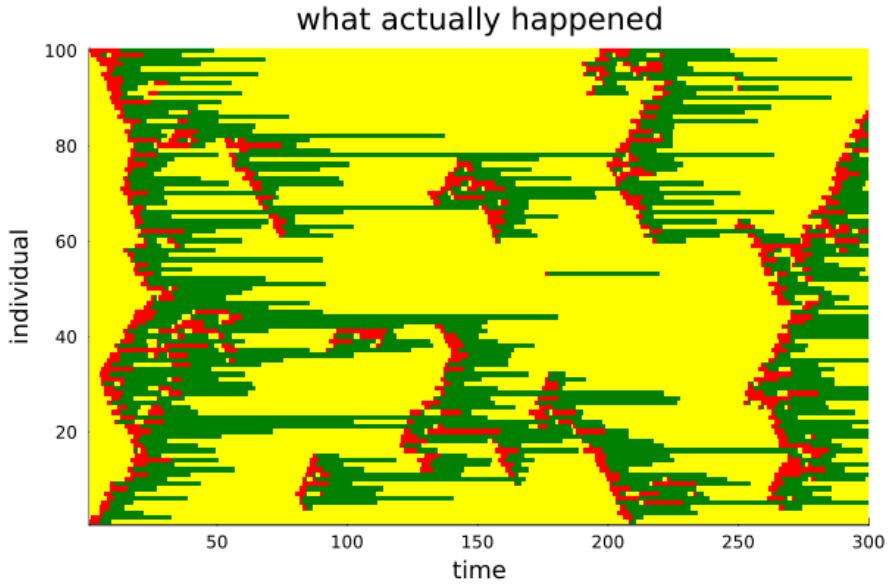
dus niet. Bovendien kunnen testen ook een verkeerde uitkomst en geven. Het is dus realistischer dat we een observatie hebben zoals in Figuur 11. Alle tijd-persoons combinaties in het zwart nemen we niet waar, deze zijn niet zichtbaar. In de bovenste figuur zijn die tijdstippen willekeurig gekozen; in de onderste figuur meten we op een aantal vaste tijdstippen. In beide figuren meten we slechts 4% van het echte verloop van het virus. De volgende vragen rijzen:

1. Kunnen we achterhalen hoe de infectie zich verspreid heeft?
2. Als we een methode hebben om dit te doen, kunnen we garanties geven hoe goed deze methode werkt? Bijvoorbeeld, met kans tenminste 0.95 reconstrueren we tenminste 90% van de figuur correct.
3. Stel dat we 100 metingen mogen doen, op welke tijdstippen doen we dan metingen en bij welke personen?
4. Mensen blijven niet op 1 plek, stel dat we GPS-gegevens zouden verkrijgen, kunnen we daarmee een realistischer model verkrijgen?
5. Niet ieder persoon is hetzelfde, is het mogelijk het model uit te breiden en individuele karakteristieken meenemen daarvan? En welke karakteristieken zijn dan van belang?

Ik zal hier later nog enkele vragen aan toevoegen.

Het invullen van de missende waarden is een complex *computationeel* probleem. Ieder missend vakje kan 3 waarden aannemen, dus als ik 100 personen heb, en het process over 300 tijdsintervallen evolueert, en ik daarvan slechts 4% observeer, dan heb ik $100 \times 300 \times 0.96 = 28800$ vakjes in te vullen. Het aantal mogelijkheden daartoe is 3^{28800} , een onvoorstelbaar groot aantal. Het aanschaffen van een zeer geavanceerde computer zal echt niet helpen. Is er een oplossing? Hieraan heb ik gewerkt met masterstudent Daniël Brus en Moritz Schauer. *Kunnen we het onzichtbare zichtbaar maken?* *Kunnen we data omzetten in kennis teneinde betere beslissingen te maken onder onzekerheid?* Ik denk dat dit de belangrijkste taak is van mijn vakgebied.

In Figuur 12 toon ik een mogelijke reconstructie van Figuur 10 op grond de observatie in de onderste figuur van Figuur 11. De rekentijd is circa een minuut op mijn laptop. Ben ik echt alle mogelijkheden nagegaan? Nee natuurlijk niet. Het onderliggende idee is “message passing”, het

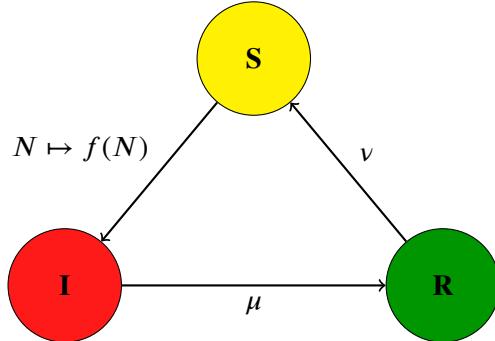


Figuur 10: Een voorbeeld met 100 individuen over 300 tijdstappen.

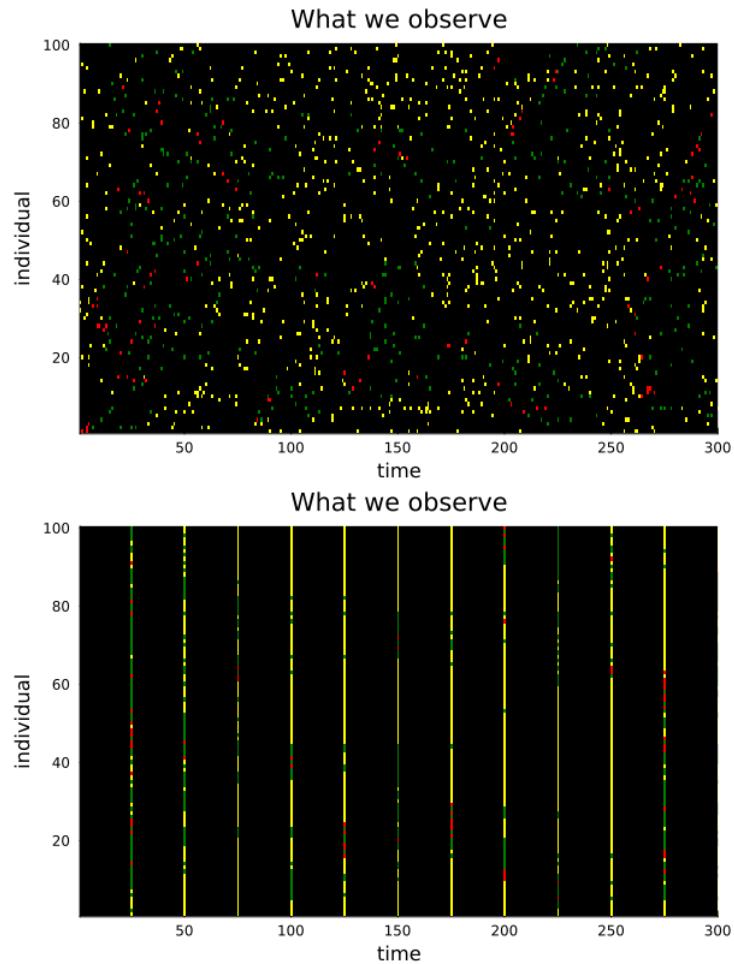
versturen van berichten. We beginnen bij het laatste tijdstip, en iedere persoon stuurt informatie over zijn staat van gezondheid (S, I of R) één tijdstap terug. Dit herhalen we tot we helemaal terug bij tijdstip 1 zijn. Vervolgens kunnen we de informatie gebruiken om de kleurplaat in te vullen, waarbij we ook uitrekenen hoe waarschijnlijk de ingevulde kleurplaat is. Het voert hier te ver om op details in te gaan.

Ik vind het fascinerend dat een eenvoudig model zoals dit al een lastig computationeel probleem oplevert, wat vervolgens op grond van enkele inzichten “op te lossen” is. Is de reconstructie niet te goed, heb ik me schuldig gemaakt aan “cherry-picking” hier? Cruciaal is dat ik heb aangenomen dat er aanvankelijk rond de 2 infecties zijn. Iets preciezer, voor iedere persoon neem ik aan dat de kans dat de persoon geïnfecteerd is gelijk is aan 0.02 en de kans dat de persoon vatbaar is gelijk is aan 0.98. Zonder het meenemen van deze kennis wordt het al een stuk lastiger.

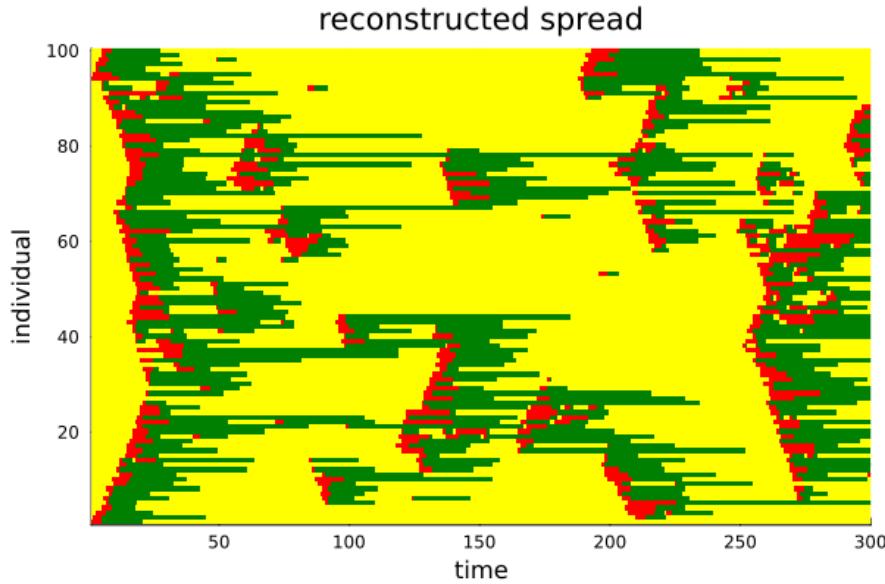
Is het probleem hiermee opgelost? In werkelijkheid weten we de kansen in het model niet. Een algemener model wordt verkregen door aan te nemen dat over een tijdsinterval ter lengte h de kansen op transitie als volgt zijn:



Hierin is N het aantal geïnfecteerde buren en zijn f , μ , ν onbekende parameters in het model. Deze zouden we ook graag willen achterhalen, ofwel *schatten*, op grond van de data. Tot dusver heb ik ze als bekend verondersteld. De volgende onderzoeks vragen liggen dus voor de hand:



Figuur 11: We krijgen niet alle informatie. Beide figuren tonen dezelfde virusverspreiding als in Figuur 10. Alle zwartgemaakte velden corresponderen met tijdstippen waarop individuen niet geobserveerd worden. In beide figuren is slechts 4% van de “volledige” data observeerbaar. In de bovenste figuur zijn willekeurig cellen gekozen die geobserveerd worden; in de onderste figuur nemen we alle individuen op vaste tijdstippen waar.

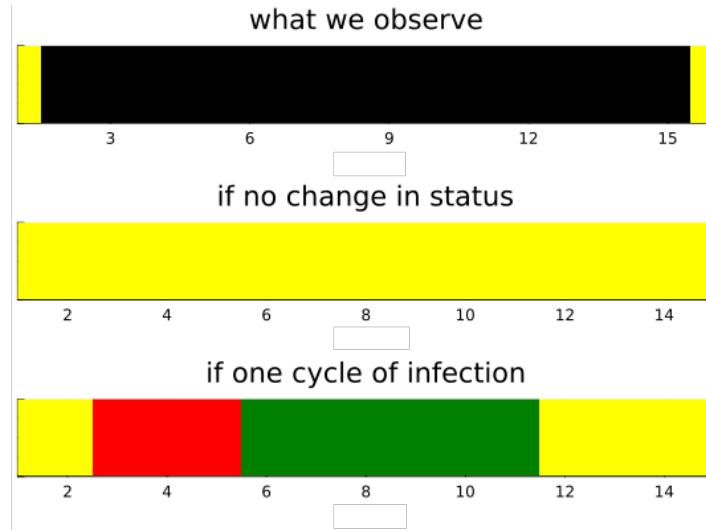


Figuur 12: Een reconstructie van Figuur 10. We observeren op vaste tijdstippen als weergegeven in de onderste figuur van Figuur 11.

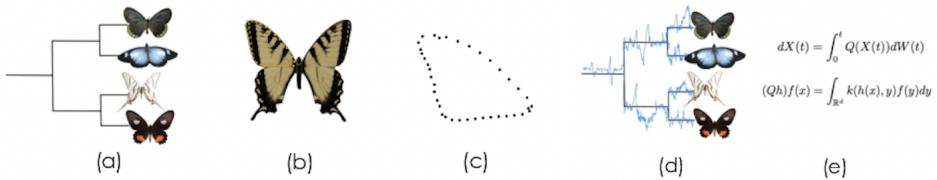
1. Kunnen we achterhalen wat de waarden van de parameters (f, μ, v) zijn?
2. Hoe goed kunnen we deze parameters achterhalen?
3. Is er een optimale methode?

Het probleem is complex. Om daar een idee van te geven, stel dat ik van een persoon weet dat hij op tijdstip 1 en 16 vatbaar is. Afhankelijk van de parameters, kan het waarschijnlijk zijn dat deze persoon de gehele tijd vatbaar was. Maar het kan ook dat er deze persoon een cyclus van infectie ($I \mapsto R \mapsto S \mapsto I$) heeft ondergaan. Zie Figuur 13. Wat het meest waarschijnlijk is, hangt sterk af van de parameters. Er kunnen dus meerdere plausibele verklaringen zijn en idealiter geeft onze methode alle mogelijke verklaringen, met daarbij hoe waarschijnlijk ze zijn. Zelfs voor dit toyexample is dit nog een hele uitdaging.

Tot slot, het model is discreet in tijd; in feite komt het voort uit een discretisatie van een model dat continu in tijd is. Het stochastische model is dan complexer en lastiger uit te leggen.



Figuur 13: Bovenste figuur: we observeren de status S (susceptible/vatbaar) op tijdstippen 1 en 16. De middelste en onderste figuren geven 2 mogelijke reconstructies weer.

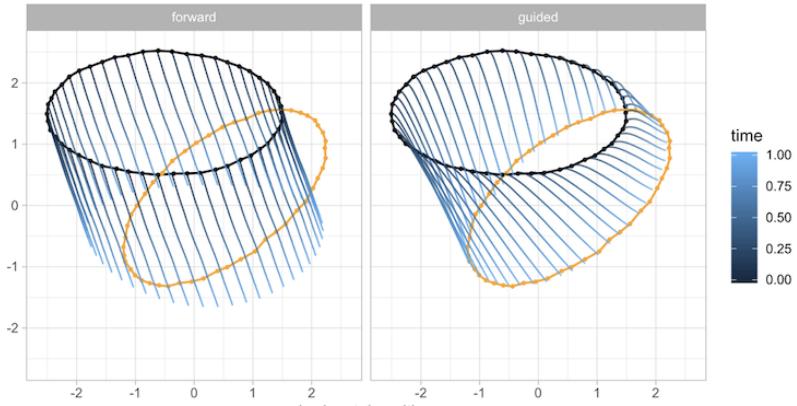


Figuur 14: (a) Fylogentische boom voor evolutie van vlinders. (c) toont landmark van de linkervoortvleugel van de vlinder in (b). (d): visualisatie van de statistische methode: een stochastisch model voor de verandering van de landmarks op de linkervoortvleugel. (e): stochastische differentiaalvergelijking voor het proces van de landmarks. Figuur gemaakt door Sofia Stroustrup.

2.2 Een voorbeeld uit de fylogenie

Nu we het vorige voorbeeld begrepen hebben, kan ik overgaan op een wat realistischer model uit de fylogenie. Dit is de studie naar evolutionaire relaties tussen soorten. Ik kijk hier naar vlinders. Biologen zijn bijvoorbeeld geïnteresseerd of de voorvleugels sneller evolueren dan de achtervleugels. Kunnen we dit achterhalen? Linksboven in Figuur 14 zien we links een visualisatie van een fylogenetische boom: een schema hoe bepaalde vlindersoorten ontstaan zijn door evolutie. We veronderstellen dat we de fylogenetische boom weten, en van bestaande vlindersoorten kunnen we de vorm van de linkervoortvleugel vastleggen door zogenaamde *landmarks*. Een stochastisch model specificeert hoe deze landmarks veranderen over de tijd. Dat is gevisualiseerd in panel (d). We kunnen vervolgens proberen de landmarks van vlindersoorten uit het verleden te reconstrueren, aangezien deze niet zichtbaar zijn.

Onderliggend de analyse is een stochastisch model hoe vormen veranderen. Hoe modelleren we hoe één vorm overgaat in een andere vorm? Dit is iets waar ik gedurende mijn verblijf in Kopenhagen aan gewerkt heb. Zie Figuur 15 Het stochastisch model is hier een stuk ingewikkelder, maar de methoden die ik hiertoe voorstel zijn analoog aan die in het toyexample.



Figuur 15: Voorbeeld van een stochastisch model voor deformatie van vormen. Links: een realisatie van het voorwaardse model. Recht: een realisatie van een process dat de vormen interpoleert.

2.3 Mathematische Statistiek aan de VU

Het gekozen voorbeeld valt binnen mijn eigen onderzoek. Het onderzoek van statistici aan de VU is natuurlijk veel breder. Zo doen we bijvoorbeeld ook onderzoek in hoog-dimensionale statistiek en is er sinds vele jaren samenwerking met levenswetenschappen, bijvoorbeeld via het Amsterdam Medical Center. Ik hoop de komende jaren de zichtbaarheid van mijn vakgebied binnen de VU verder te vergroten door het uitbreiden van samenwerkingen met andere vakgebieden. Dit te meer om in deze tijden het belang van de wiskunde afdeling en de opleidingen die zij verzorgt te onderstrepen.

3 Onderwijs

3.1 De rol van onderwijs en professionalisering aan de universiteit

Graag wil ik ook ingaan op de rol van onderwijs. Naast onderzoek is kennisoverdracht namelijk een zeer belangrijke taak van de universiteit. Excellent onderwijs, of beter “excellence in teaching” is wat iedere universiteit natuurlijk wil. Het woord “innovatief” komt vele keren voor in het kwaliteitsplan onderwijs 2019-2024 van de VU. Er zijn natuurlijk vele innovatieve ideeën in onderwijsland. Voor wie daar meer over leren, ook in historisch perspectief, raad ik het zeer lezenswaardige boek “Nog wat geleerd vandaag?” van Maarten Huygen aan ([5]). Het laat zien dat onderwijsvernieuwers graag uitgaan van de ideale student, ik zal hier later op terugkomen met betrekking tot het begrip “studierendement”. Het boek wordt besproken in de column “Onderwijsvernieuwing dient vaak alleen de ideale student” van Henk Verhoeven (Scienceguide van 21 December 2023). Het grote probleem is dat het merendeel van de studenten niet ideaal zijn. Ik sluit me graag aan bij het volgende uit zijn slotalinea:

Wat werkt wel? Simpel: directe instructie, kennisoverdracht, vaak herhalen, toetsen wat overgekomen is, oefenen in vaardigheden en basisroutines drillen.

[Henk Verhoeven in scienceguide, 21 December 2023.](#)

Terecht wordt ook aan de VU veel belang gehecht aan goed onderwijs. Niet voor niets is er een “centre for teaching & learning”. Hier worden vele cursussen aangeboden om docenten te professionaliseren, waaronder

- start-to-teach day
- basiskwalificatie onderwijs
- mixed classroom in practice
- Pioneers of Education’ programma voor junior docenten
- theatraaldigheden voor docenten
- begeleiding en beoordeling van afstudeerwerken
- become an expert in educational design
- oriëntatie op leiderschap in onderwijs
- inclusive leadership
- leiderschap in de tussenruimte: morele ambitie in het hoger onderwijs
- art & nature based learning

Ik noem slechts enkele van de mogelijkheden, er zijn er veel meer.

Ik wil hier niet oordelen over specifieke cursussen en zeker niet claimen dat deze geen positief effect kunnen hebben, maar wel een kanttekening maken. Ik vermoed dat de docenten die mij tijdens mijn studie geïnspireerd hebben geen van deze opleidingen gevolgd heeft, en betwijfel of hun onderwijs veel veranderd zou zijn na het volgen van deze cursussen. Is de kwaliteit van colleges wezenlijk veranderd ten opzichte van mijn studietijd (1995-2001)? Universiteiten antwoorden graag bevestigend, we hebben immers prestatie-afspraken en we professionaliseren docenten. Ik betwijfel het. De ondergrens is waarschijnlijk wat opgeschoven. Het antwoord op mijn vraag zullen we niet weten omdat “goed onderwijs” heel moeilijk te kwantificeren is. Daarnaast veranderen schoolprogramma’s. Er waren goede docenten en minder goede, en de goede docenten gaven zeker niet op dezelfde manier les. Zelf heb ik ook college en instructie gekregen van zeer bekwame docenten die evenwel nooit gepromoveerd zijn. Vele goede docenten gebruikten slechts een bord en krijtje en gaven directie instructie. Zou het kunnen dat lesgeven een vaardigheid is die in de één inherent meer aanwezig is dan de ander, net zoals bij onderzoek? Bij leerlingen op school zien we



Figuur 16: De collegezaal.

ook dat de ene leerling heel makkelijk tegen een groep praat als hij of zij een spreekbeurt geeft, terwijl dat voor de ander heel lastig is.

Goed en waardevol onderwijs begint en eindigt bij de docenten.

aldus het kwaliteitsplan onderwijs 2019-2024 van de VU (pagina 5). De docent staat dus centraal. Maar als dit zo is, hoe kan het dan dat we bij het werven van nieuwe posities zelden of nooit specifiek de voorkeur geven aan een kandidaat met talent voor onderwijs? Idealiter is de kandidaat heel goed in zowel onderzoek, onderwijs als valorisatie en vertelt hij of zij de pers op aansprekende wijze over de bijdragen van het onderzoek aan de maatschappij. Wat is er mis met het aanstellen van universitair docenten die zich primair op onderwijs richten? Graag zou ik zien dat de er meer gedaan wordt aan het vasthouden van onderwijs talent, in plaats van tijdelijke JUDO contracten. We laten nu simpelweg talent lopen.

3.2 Mijn onderwijs innovatie voorstel

Graag wil ik zelf ook een “innovatie” voor beter onderwijs inbrengen. In het voornemen van het kwaliteitsplan staat een investering van 24.1 miljoen euro. Je zou denken dat voor die investering de collegezalen geheel naar de wensen van docenten en studenten ingericht zijn. De meeste collegezalen bij de VU zien er mijns inziens in eerste instantie ook prima uit, er is meestal zelfs een krijtbord. Ik heb begrepen dat de wiskunde afdeling hiervoor gestreden heeft. Mijn innovatievoorstel is dat we ook zorgen dat deze borden schoongemaakt kunnen worden. In iedere zaal dus graag een spons, trekker en emmer met water. Bij het ontwerp van nieuwe zalen is er namelijk niet over nagedacht dat er een kraan in de zaal moet zijn. Ik hoop dat mijn “innovatie” nog gerealiseerd kan worden ondanks de huidige bezuinigingen.

3.3 Studierendement en de rol van de student

Ik wil het ook nog kort hebben over studierendement. Daartoe neem ik u graag mee naar de collegezaal via een foto, zie Figuur 16.

De dingen die mij opvallen in dit plaatje zijn de volgende

1. Er ontbreekt diversiteit, ik zie geen studenten of docent met een donkere huidskleur. De verhouding man-vrouw is niet gelijk.

2. De studenten lijken allen buitengewoon geïnteresseerd.
3. Eén van de studenten staat. Of is dat de docent?
4. De gezichten kloppen niet; het is AI-gegenererd.

Als opleidingsdirecteur heb ik geleerd dat er drie heel belangrijke zaken zijn:

1. studierendement,
2. instroom van studenten,
3. efficiëntie in onderwijs.

Vooropgesteld, ik denk niet dat de VU hier anders in is dan andere universiteiten; het is een rechtstreeks gevolg van financiering vanuit de overheid. Het bespreken van dit onderwerp is dus zeker niet bedoeld om af te geven op het faculteitsbestuur of ondersteunende diensten. Ik wil eerst ingaan op de term “studierendement”: In de Van Dale online is de definitie van rendement:

1. opbrengst van een economische activiteit; = winst
2. nuttig effect van iets, bijvoorbeeld van een apparaat.

Zoeken op “studierendement” gaf evenwel "Geen resultaat voor 'studie-rendement' ". Het woord bestaat dus niet. Ik vermoed dat daar een goede reden voor is.

Nu het gewenste scenario, dat is als volgt: zoveel mogelijk studenten doen de studie binnen de gestelde tijd, idealiter met niet te veel personele inzet, maar we geven wel iedere student persoonlijk aandacht. Of, zoals ik intern bij de VU heb vernomen, we moeten gaan voor “nominaal = normaal”. Maar is dit eigenlijk wel normaal? Hoort falen niet bij leren? Zou het verder kunnen dat sommige studenten

- een verkeerde studiekeus gemaakt hebben?
- het niveau op de universiteit hoger is dan dat op het VWO en daardoor sommige studenten zelfs met grote inzet en excellent onderwijs het niet halen?
- sommige studenten naast hun studie ook nog veel moeten werken en simpelweg niet de tijd in de studie kunnen stoppen die nodig is?

Er lijkt een zwijgen te liggen op deze vragen. Als opleidingsdirecteur krijg ik soms het gevoel dat dit geheel mijn verantwoordelijkheid is. Terugkomend op Figuur 16, de wasmachine in het plaatje represeneert de ideale student. Het is een apparaat, en we kunnen praten over het “nuttig effect” van dit apparaat. We kunnen het apparaat zelfs instellen zodat alles in 1 keer goed gaat. Laten we zeggen dat de wasmachine nominaal presteert. Met welke inspanning van opleiding of docent ook, het is een utopie om te denken dat studenten als wasmachines zijn. Ze zijn niet ideaal, en dat is ook niet reëel om te verwachten.

Daarmee wil ik niet suggereren dat niet iedere opleiding kritisch naar zijn programma moet kijken en genoemde bezwaren klakkeloos als excuus kan opvoeren. Maar er gebeurt al veel en ik vraag me af of we hier de grens niet bereiken. Ik wil graag nogmaals het kwaliteitsplan onderwijs 2019-2024 (pagina 8) aanhalen:

Het opleidingsmanagement en de docenten zijn verantwoordelijk voor kwalitatief goede en studeerbare academische opleidingen, de student is vervolgens verantwoordelijk voor de eigen studieloopbaan en het eigen studiesucces.

Laten we deze verantwoordelijkheden ook zo naleven in ons beleid.

4 Taal

In een prachtig nummer zingt de Surinaams-Nederlandse zanger Kenny Bron, beter bekend als Kenny B. in het nummer “Parijs”

Praat Nederlands met me, even Nederlands met me.

Het leren van Nederlands is een keuze. De discussie over het belang, of niet, van de Nederlandse taal is de afgelopen jaren toegenomen. Helaas blijkt deze discussie snel te verzanden in een discussie over een nationalistisch dan wel kosmopolitisch wereldbeeld, en dat betreft ik. Er is wat mij betreft geen discussie dat het niveau van de Nederlandse universiteiten versterkt wordt door mensen voor wie de Nederlandse taal nieuw is. Onderzoek is internationaal en, afhankelijk van het vakgebied, publiceren we in het Engels en dat lijkt me prima.

Er zijn goede redenen om Nederlands te leren, los van het feit dat het de taal is die in Nederland gesproken wordt. Het is mijn ervaring dat nieuwe medewerkers vaak wel beginnen met een cursus, maar dat het daar bij blijft. Ze kunnen daarna geen vergadering in het Nederlands volgen, als ook geen documenten in het Nederlands lezen op het benodigde niveau. Die één of twee gevulde cursussen zijn dus grotendeels verspilde inspanning en weggegooid geld. Beginnende UDs zijn zich onbewust dat als je uiteindelijk betrokken wil zijn bij het nemen van beslissingen aan de universiteit, op een gegeven moment Nederlands de voertaal is. Het belang wordt dus onvoldoende duidelijk gemaakt wat soms tot de reactie “I don’t need it, why can’t all be in English?” leidt.

Laten we ook kijken naar het standpunt van de VU op dit onderwerp. Het handboek onderwijskwaliteit taalbeleid (versie 2019) stelt het volgende:

KWALITEITSEISEN TAALBELEID GECONCRETISEERD VOOR WETENSCHAPPELIJK PERSONEEL.

Aannamebeleid

Het aannamebeleid van wp is gericht op tweetaligheid. In alle communicatie naar nieuwe medewerkers wordt aangegeven dat we een tweetalige universiteit zijn. In personeelsadvertenties wordt opgenomen dat nieuwe medewerkers bij voorkeur minimaal twee talen (Nederlands en Engels) beheersen. Zie ook hierboven bij obp. Internationaal wp zou hierdoor niet afgeschrikt moeten worden, maar juist **aangemoedigd om ook Nederlands te leren**. De VU moet daarvoor ondersteuning aanbieden.

Taalvaardigheidseisen docenten

Voor docenten gelden de volgende eisen aan taalvaardigheid:

- wp dat doceert in Nederlandstalige opleiding of vakken: minimaal C1-niveau Nederlands
- wp dat doceert in Engelstalige opleiding of vakken: minimaal C1-niveau Engels

Voor een docent binnen een Engelstalige opleiding betekent dit dus er een aanmoediging is om Nederlands te leren. Een aanmoediging bij mijn kinderen om elke dag eerst hun kamer op te ruimen en dan hun huiswerk te maken werkt doorgaans niet... Gezien de inspanning die het vergt om een nieuwe taal te leren, laat het resultaat van dit beleid zich makkelijk voorspellen. Wat ook niet helpt is om direct naar het Engels te “switchen”. Buitenlandse collega’s beginnen vaak enthousiast, maar worden denk ik gedemotiveerd doordat mensen niet de moeite nemen ze op weg te helpen. De quote waarmee ik dit onderwerp ben begonnen heb ik namelijk meermaals gehoord van collega’s uit het buitenland en is dus niet alleen van toepassing op mijzelf, wat u wellicht aanvankelijk dacht.

Ik zou graag zien dat de VU een duidelijker standpunt inneemt ten aanzien van het belang van de Nederlandse taal aan de universiteit: ofwel laat het volledig zitten, als dat te motiveren valt, ofwel onderneem daadwerkelijk actie.

5 Dank

Allereerst wil ik mijn collega's van de wiskunde afdeling, in het bijzonder de statistici, danken voor het warme welkom op de VU. Verder dank ik het College van Bestuur en het bestuur van de faculteit der Bèta wetenschappen.

Er was geen plan van mijn kant om op deze positie te komen. Er was geen plan voor een academische carrière. Die is er wel gekomen, en daar wil ik een aantal mensen voor danken:

1. Mijn co-auteurs, in het bijzonder Moritz Schauer, Stefan Sommer, Shota Gugushvili, Peter Spreij, Joris Bierkens en Aad van der Vaart.
2. Harry van Zanten: ik wil je danken voor je steun om na een uitstap in de industriële statistiek mijn weg binnen de mathematische statistiek weer te vinden, onze gezamenlijke onderzoeks-projecten en jouw adviezen bij mijn start aan de VU.
3. Mathisca de Gunst en Jan Bouwe van den Berg: ik dank jullie voor jullie hulp om mijn weg binnen de afdeling te vinden.
4. Mijn grootste dank in mijn wetenschappelijke ontwikkeling is aan Geurt Jongbloed. Geurt, vanaf het begin van mijn promotie-onderzoek heb ik van jou geleerd, heb je me vertrouwen gegeven en mij gesteund. Zonder dat had ik hier nu niet gestaan. Jouw betrokkenheid bij mensen is ongeëvenaard en is een voorbeeld voor mij.

Tot slot, er is meer dan werk, wil ik me richten tot mijn gezin. Micha, Hugo, Boele, Lidi: ik dank jullie voor het gezin dat wij vormen. Lidi: dit pad had ik niet kunnen bewandelen zonder jou, daar ben ik je enorm dankbaar voor. Zonder jou was dit niet gelukt.

Ik heb gezegd.

Referenties

1. A. Arnaudon, F.H. van der Meulen, M.R. Schauer and S. Sommer (2022) *Diffusion bridges for stochastic Hamiltonian systems and Shape Evolutions*, SIAM Journal on Imaging Sciences, **15**(1), 293–323
2. M.M. Reijne, F.H. van der Meulen, F.C.T. van der Helm and A.L. Schwab (2024) *A Model Based On Cyclist Fall Experiments Which Predicts The Maximum Allowable Handlebar Disturbance From Which A Cyclist Can Recover Balance* Submitted
3. J. Bierkens, F.H. van der Meulen and M. Schauer (2020) *Simulation of elliptic and hypo-elliptic conditional diffusions* Advances in Applied Probability **52**(1), 173–212
4. W.O. Kermack and A.G. McKendrick (1927) *A Contribution to the Mathematical Theory of Epidemics*
5. M. Huygen (2023) *Nog wat geleerd vandaag?* Boom.