

Statistiek begrijpen zonder formules

Een intuïtieve benadering

Dustin de Leeuw

Auteur: Dustin de Leeuw

Foto auteur: door Dustin de Leeuw gegenereerd met behulp van ChatGPT

Uitgever: Bookmundo

ISBN: 9789403817217

Eerste editie: februari 2026

© Copyright: Dustin de Leeuw

Alle rechten voorbehouden. Dit boek en de inhoud ervan zijn beschermd door auteursrecht. Het is uitdrukkelijk verboden om de inhoud van dit boek, geheel of gedeeltelijk, te reproduceren, kopiëren, distribueren, of op enige andere wijze te gebruiken voor commerciële doeleinden zonder voorafgaande schriftelijke toestemming van de auteur.

Voor feedback, vragen en suggesties:

stuur een e-mail naar info@dustindeleeuw.nl

www.dustindeleeuw.nl

Voorwoord

Veel docenten zijn zo gepassioneerd over de schoonheid van statistiek (en terecht!) dat ze vergeten om een stapje terug te doen en met een frisse blik studenten te verwelkomen in deze wondere wereld. Na ruim 20 jaar bijles aan studenten van allerhande studierichtingen, merk ik dat ze vaak de connectie missen tussen de wiskundige modellen en de inhoudelijke vragen waar ze zich tijdens de rest van hun studie mee bezig houden. Wat herkenbare voorbeelden en praktische toepassingen van statistiek kunnen voor veel studenten in een enkele bijles al een hoop helderheid verschaffen. In dit boek heb ik geprobeerd mijn favoriete bijles samen te vatten, waarin ik de achtergrond van statistiek intuïtief uitleg en probeer om het doel van statistiek tastbaar en inzichtelijk te maken.

Hopelijk helpt deze op papier gezette bijles ook jou om beter te begrijpen wat statistiek is en waarom je het nodig hebt voor je opleiding. Zo kun je je door de pittige en soms droge formules heen werken zonder het grote plaatje uit het oog te verliezen. Ik hoop van harte dat dit boek je helpt om plezier in statistiek te ontdekken! De theorie wordt afgewisseld met veel alledaagse, grappige of prikkelende voorbeelden. Een beetje absurditeit helpt mij om moeilijke concepten makkelijker te onthouden.

Dit boek is geen volledige statistiekcursus, maar is vooral bedoeld als introductie op en ter aanvulling van een statistiekmethode. Dit boek kan gebruikt worden als een Hoofdstuk Nul ter completering van bestaande lesmethodes. Het kan je helpen om een beter beeld van de stof te krijgen en de concepten in je hoofd te laten klikken.

Dit boek hoopt ook om studenten die niet zelf met statistiek aan de slag gaan te helpen om kritische vragen te stellen over hoe we aan de kennis komen die zij in colleges, leerboeken en wetenschappelijke artikelen aangeboden krijgen. Wist je bijvoorbeeld dat de voedingsadviezen in de Schijf van Vijf allemaal op keiharde statistiek gebaseerd zijn? Hoewel statistiek vooral over kwantitatief onderzoek gaat, zijn de onderliggende filosofische principes ook fundamenteel voor kwalitatief onderzoek. Omdat dit boek geen formules gebruikt, hoop ik dat het de principes van statistiek ook voor studenten van meer praktijkgerichte studies toegankelijk maakt.

Je kunt het boek ook gebruiken als naslagwerk om af en toe in terug te bladeren als je context nodig hebt, bijvoorbeeld als je bezig bent met het schrijven van je scriptie. Daarvoor is soms een dieper begrip van statistiek nodig dan voor een eerste kennismaking. Tijdens scriptiebegeleiding ga ik soms dieper in op de stof. Op dezelfde conceptuele manier als voor de basiskennis, maar deze stof is wat pittiger en in een eerste lezing niet noodzakelijk. Deze paragrafen zijn gemarkeerd met een asterisk (*) en kun je overslaan, de kennis van deze paragrafen is niet nodig voor een goed begrip van de rest van het boek. En als je een lichthartige introductie wilt met nóg minder moeilijke termen, lees dan alleen de voorbeelden in de grijze kaders.

Inhoudsopgave

Voorwoord	3
Inhoudsopgave	5
1 Waar zijn we nou eigenlijk helemaal mee bezig?	7
2 De juiste vraag stellen	9
2.1 Operationalisatie van variabelen	9
2.2 Het soort antwoord dat je kunt krijgen	12
2.3 Beperkingen van statistiek	15
3 Kies de geschikte techniek	22
3.1 Het verschil tussen toetsen	23
3.2 De overeenkomsten tussen toetsen*	24
4 Associaties met meerdere variabelen	26
4.1 Meerdere determinanten*	27
4.2 Mediatie	29
4.3 Confounding	31
4.4 Collider bias	38
4.5 Effectmodificatie (moderatie)	39
4.6 De 'werkelijke' associatie tussen X en Y	42
5 Vertekening van de associatie (<i>bias</i>)	43
5.1 Experiment of observatie?	43
5.2 Bij elke stap loert gevaar	44
6 Selecteren en gebruiken van meetinstrumenten	46
6.1 Validiteit	46
6.2 Betrouwbaarheid	47
6.3 Meetniveau*	48
6.4 Informatiebias	51

7	Steekproeftrekking	54
7.1	Aselect	55
7.2	Representatief	56
7.3	Voldoende groot*	57
8	Statistische analyse	60
8.1	Selectie van de juiste toets	60
8.2	Assumpties*	61
8.3	Keuzes	64
8.4	Table 1 Fallacy & Table 2 Fallacy*	66
9	Conclusie	68
9.1	Type I en Type II fouten	68
9.2	Overgeneralisatie	73
9.3	Statistisch niet significant is óók interessant	73
9.4	Statistisch significant is niet hetzelfde als praktisch relevant	76
10	Ten slotte	77
	Referenties	79

1 Waar zijn we nou eigenlijk helemaal mee bezig?

Je neemt waar dat er variatie optreedt in de waarde van een of andere grootte en je vraagt je af hoe dat kan. Is dat puur toeval, of valt er iets zinnigs over te zeggen? Dit is de hele reden om statistiek te doen, deze vraag is waar alles om draait. Je begint dus met een observatie in de echte wereld, die leidt tot een vraag of hypothese (verwachting, veronderstelling) en je wilt daar een antwoord op dat betrekking heeft op de wereld om je heen. Waarschijnlijk heb je ook al een verwachting: misschien hangt die variatie in de ene grootte samen met een andere grootte.

Statistiek is een verzameling technieken die we gebruiken om de associatie tussen twee of meer grootheden te onderzoeken. We komen in de analyse allemaal uitdagingen tegen, we gaan het hebben over assumpties en limitaties, we hebben vragen over validiteit en betrouwbaarheid, we lijken het hele Griekse alfabet te gebruiken en elke keer als je denkt dat je wel genoeg toetsen kent, blijkt er weer een nieuwe te bestaan. Je kunt gaandeweg best afgeleid of geïntimideerd worden door alle formules die daarbij komen kijken, maar houd voortdurend in gedachten dat al dat gegoochel met getallen maar één doel heeft: je wilt iets kunnen zeggen over de associatie tussen X en Y . Meestal veronderstellen we daarbij dat het niet zomaar een associatie is, maar dat X invloed heeft op Y . In andere woorden: X is een determinant of bepalende factor voor een uitkomst Y . We zeggen ook wel dat Y afhangt van X , dus dat Y de afhankelijke en X een onafhankelijke variabele is.

Een tip om je te helpen de statistische technieken concreter en praktischer te maken: kies voor jezelf een aansprekend voorbeeld waar jij graag onderzoek naar zou willen doen. Misschien vraag je je af wat de associatie is tussen roken en longkanker. Tussen het eten van appels en gezondheid. Tussen reclame en aankoopintentie. Tussen marketingbudget en jaaromzet. Tussen het aantal uren studeren en je tentamencijfer. Tussen je trainingsschema en je *one rep max* voor de bench press. Tussen je kleur oogschaduw en je succes in de kroeg. Elke keer als je iets nieuws leert in je statistiekcursus, probeer dit dan toe te passen in de context van jouw onderzoek. Hoe kan deze techniek jou helpen?