

Hans van Maanen

Fokke Simonszstraat 92
1017 TK Amsterdam
hans@vanmaanen.org

Maatschappij

Worden baby's door ooievaars gebracht?

De statistici onder ons worden er op feestjes vast dikwijls mee geconfronteerd: het wijdverbreide misverstand dat de toename van het ene fenomeen wordt verklaard door de gelijktijdige toename van iets anders. Een klassiek voorbeeld wordt in dit artikel ontzenuwd door wetenschapsjournalist Hans van Maanen.

De klassieke manier om studenten duidelijk te maken dat er een verschil is tussen een statistisch verband en een oorzakelijk verband maakt gebruik van ooievaars en baby's. Als er in een dorp meer ooievaars komen en het aantal geboorten gaat eveneens omhoog, dat betekent dat nog niet dat ooievaars de baby's brengen. Er is slechts sprake van correlatie, niet van causaliteit.

Het voorbeeld duikt, voor zover valt na te gaan, het eerst op in het invloedrijke Amerikaanse statistiekboek van Box, Hunter en Hunter, *Statistics for experimenters*, gepubliceerd in 1978 [2]. Meteen op pagina 8 waarschuwen de auteurs tegen het verwarren van correlatie en causaliteit en geven zij een 'plot van de bevolking van Oldenburg aan het eind van het jaar afgezet tegen het aantal waargenomen ooievaars in dat jaar, 1930–1936', hier weergegeven als figuur 2.

Oldenburg is een stadje in Noord-Duitsland. Als bronnen geven de schrijvers de *Ornithologische Monatsberichte* 44 (1936) en 48 (1940) en het *Statistisches [sic] Jahrbuch Deutscher Gemeinde* 27–33 (1932–1938). Het voorbeeld

en de grafiek zijn inmiddels gereproduceerd door andere leerboeken, en het thema wordt bijna achteloos en vanzelfsprekend gebruikt in talloze artikelen en statistiekboeken [2, 7], die op hun beurt ook weer achteloos en vanzelfsprekend worden geciteerd. Dat geeft ons genoeg reden om dit voorbeeld eens precies na te lopen.

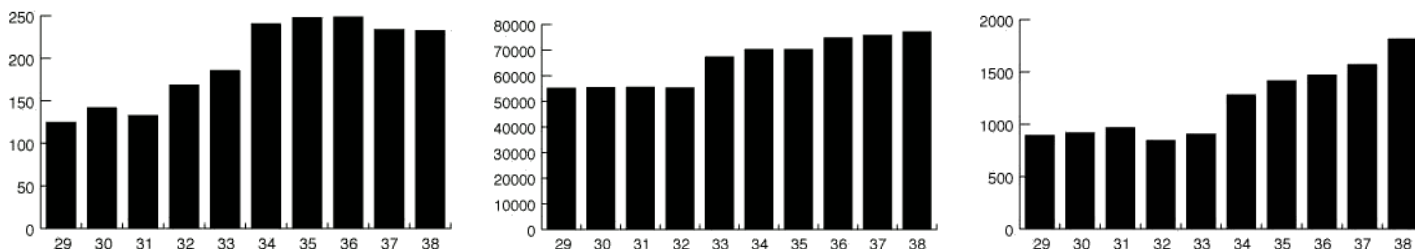
De artikelen in de *Ornithologische Monatsberichte* zijn van de hand van Ernst Schüz van de Vogelwarte Rossitten. Het is meteen al onduidelijk waarom het eerste artikel, uit 1936, wordt genoemd: het tweede artikel, 'Bewegungen im Bestand des Weissen Storches seit 1934', bevat alle relevante gegevens en wat verbeteringen ten opzichte van het eerste. Het *Statistisches Jahrbuch* is een jaarlijkse productie van het Duitse bureau voor statistiek; tot 1934 heette het *Statistisches Jahrbuch Deutscher Städte*.

De grafiek van figuur 2 is lastig te interpreteren. In de eerste plaats is de hypothese niet dat ooievaars inwoners brengen, maar dat zij baby's brengen. Hiermee wordt de oorspronkelijke theorie in het geheel niet ge-

toetst. In de tweede plaats spreekt het artikel van Schüz over 'nestparen' en niet over 'ooievaars'. Dat maakt natuurlijk geen verschil in de correlatie, maar het is onduidelijk waarom de Amerikaanse auteurs het aantal ooievaars in Oldenburg halveren. In de derde plaats loopt de tabel die Schüz geeft van 1928 tot 1939, niet van 1930 tot 1936, en het is, al-



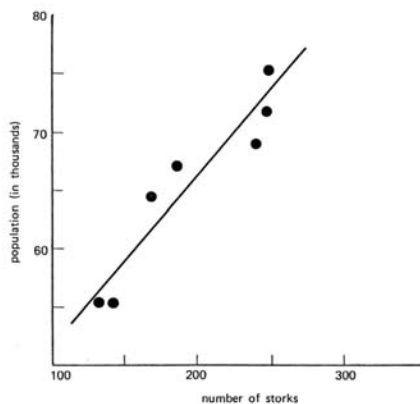
Ooievaar van glas door Ed. Hazebroek (www.glasblazerijhazebroek.nl)



Figuur 1 Aantallen ooievaarsparen (links), inwoners (midden) en geboorten (rechts) in Oldenburg (1929–1938)

weer, onduidelijk waarom hier data zijn weggegooid.

In de vierde plaats is het goed om, voordat er een correlatie wordt vastgesteld, de afzonderlijke gegevens te bekijken. In figuur 1, geven wij links de tabel van Schüz, voor het gemak als staafdiagram, in het midden de aantallen inwoners van Oldenburg uit de oude



Figuur 2 Het verband volgens Box et al.

statistische jaarboeken. Van 1928 waren de inwonertallen niet meer te achterhalen, vandaar dat we beginnen bij 1929. In beide bestanden lijkt een duidelijke cesuur op te treden rond 1934, en dan is een lineaire regressie misschien niet zo geschikt. Die breuk blijft enigszins verborgen in de data van alleen 1930 tot 1936, en al helemaal in een regressie-analyse. De sprong in aantallen ooievaarsparen is raadselachtig, zoals Schüz zelf ook zegt in zijn artikel. Maar hij voegt eraan toe dat ook elders in Duitsland de ooievaarspopulaties nog wel eens met horten en stoten groeien en slinken.

De plotse groei in het aantal inwoners in 1934 is eenvoudiger, maar onprettiger te verklaren. De nazi's waren toen net een paar maanden aan de macht, en om hun vernietigingsmachine zo efficiënt mogelijk te laten draaien, organiseerden zij alvast in juni 1933 een Volkszählung. Dit leidde in Oldenburg tot een bovenwaartse correctie van ongeveer 12000 inwoners.

Vóór 1933 waren de bevolkingsaan-

tallen van Oldenburg stabiel terwijl de ooievaarspopulatie groeide, na 1933 groeide de bevolking maar bleef de ooievaarspopulatie stabiel. Doordat Box en zijn co-auteurs alleen de correlatie en de regressielijn tonen, en alleen de zes datapunten van 1930 tot 1936, wordt de argeloze lezer op het verkeerde been gezet. Het voorbeeld, beschouwd als grappig en aansprekend, hoort op deze manier verteld eigenlijk in geen enkel statistiekboek thuis.

Voor de volledigheid geeft het rechter diagram in figuur 1 het aantal geboorten in Oldenburg van 1929 tot 1939. Het aantal baby's steeg sterk na 1933 door het aanmoedigingsbeleid van het regime. Met ooievaars had het inderdaad bitter weinig te maken.

Dankwoord

Met dank aan Klaus Wodarzik van de Statistikstelle Oldenburg.

Referenties

- G.E.P. Box, W.G. Hunter en J.S. Hunter (1978), *Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building*, New York: John Wiley, p. 8
- G.V. Glass (1984), *Statistical Methods in Education and Psychology*, New York: Prentice-Hall, pp. 104–105
- R.D. De Veaux, P.F. Velleman en D.E. Bock (2006), *Intro Stats*, Boston: Pearson, p. 153
- T. Höfer, H. Przyrembel en S. Verleger (2004), 'New evidence for the theory of the stork', *Paediatric and perinatal epidemiology* **18**(1), pp. 88–92
- R.M. Matthews (2000), 'Storks deliver babies ($p = 0.008$)', *Teaching statistics* **22**(2), pp. 36–38
- S. Wirth (2000), 'King Kong, storks and birth rates', *Teaching statistics* **25**(1), pp. 29–32
- H. Sies (1988), 'A new parameter for sex education', *Nature* **332**(6164), p. 495