

Stellingen

behorende bij het proefschrift

Essays on Neural Network Sampling Methods and Instrumental Variables

Lennart Frank Hoogerheide
Erasmus Universiteit Rotterdam
29 juni 2006

I

Het meten van het effect van het genoten onderwijs van individuen op hun verdiende inkomen is van groot maatschappelijk belang. Het probleem is echter dat de niet waargenomen intellectuele capaciteiten invloed hebben op het onderwijs en ook direct het inkomen beïnvloeden. Hierdoor zou een eenvoudige regressie van het verdiende inkomen op het aantal jaren onderwijs tot vertekende conclusies leiden. Verder is het lastig om instrumentele variabelen te vinden die gecorreleerd dienen te zijn met onderwijs maar niet met capaciteiten. Angrist en Krueger (1991)¹ gebruiken Amerikaanse data met instrumenten op basis van het kwartaal waarin iemand geboren is en concluderen dat het probleem van de (systematische) vertekening van een eenvoudige regressie meevalt. Echter dit resultaat wordt vrijwel volledig bepaald door data van de zuidelijke staten. Als het effect van onderwijs op inkomen verschillend is voor de andere regio's van de V.S., wat niet onmiddellijk uitgesloten kan worden gegeven de grote economische verschillen tussen de regio's van de V.S., dan is het geschatte effect van onderwijs op inkomen niet representatief voor het gemiddelde effect van onderwijs op inkomen in de V.S.

(Hoofdstuk 5 van dit proefschrift)

II

De basis van de modellen van Angrist en Krueger (1991) wordt gevormd door de interactie tussen leerplicht wetten en kwartalen waarin personen geboren zijn. De conclusie van Bound, Jaeger en Baker (1995)² dat deze interactie niet veel bruikbare informatie over het effect van onderwijs op inkomen geeft, is te sterk. Het gebruik van de methode van maximale aannemelijkheid of een Bayesiaanse benadering met een zwak informatieve prior kansverdeling in plaats van de twee-staps kleinste kwadraten schatter, die door Angrist en Krueger (1991) en Bound, Jaeger en Baker (1995) gebruikt wordt, in een eenvoudig aangepaste versie van een model van Angrist en Krueger (1991) kan bruikbare informatie geven over het causale effect van onderwijs op inkomen in regio's van de V.S.

(Hoofdstuk 5 van dit proefschrift)

¹Angrist, J.D., Krueger, A.B., 1991. Does compulsory school attendance affect schooling and earnings? Quarterly Journal of Economics 106, 979-1014.

²Bound, J., Jaeger, D.A., Baker, R.M., 1995. Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. Journal of the American Statistical Association 90, 443-450.

III

Doordat op het gebied van flexibele structuren (zoals neurale netwerken) veel verschillende namen gebruikt worden om dezelfde begrippen aan te duiden, is de literatuur op dit gebied minder toegankelijk dan gewenst. Verschillen in modellen/methoden tussen verschillende publicaties lijken vaak groter dan deze in werkelijkheid zijn. Naast meer uniformiteit in naamgeving zou een woordenboek waarin synoniemen voor begrippen gegeven worden voor opheldering kunnen zorgen.

(Hoofdstuk 2 van dit proefschrift)

IV

De schaal van het gebruik van neurale netwerken wordt door menigeen onderschat. De menselijke hersenen vormen een biologisch neuraal netwerk, zodat iedereen voortdurend zijn/haar eigen neuraal netwerk gebruikt. Voorts geldt dat elke wiskundige functie waarbij argumenten en functiewaarden scalair, vectoren of matrices zijn en waarbij de functiewaarden een expliciete uitdrukking in termen van de argumenten zijn, beschouwd kan worden als een kunstmatig neuraal netwerk van mogelijk hogere orde.

(Hoofdstuk 2 van dit proefschrift)

V

Het specificeren van een neurale netwerk functie als een belangfunctie (of ‘importance’ functie) in de Importance Sampling methode of als een kandidaat dichtheidsfunctie in het Metropolis-Hastings algoritme kan een verbetering opleveren ten opzichte van het geval waarbij een normale of Student-t kandidaat verdeling gebruikt wordt. Ook ten opzichte van Gibbs sampling kunnen sneller resultaten met een bepaalde precisie berekend worden. Vooral als men met hoge precisie eigenschappen wenst te berekenen van een sterk niet-elliptische kansverdeling (bijv. een multimodale verdeling), dan kan het verschil in snelheid tussen methoden op basis van neurale netwerken en andere methoden van substantiële betekenis zijn.

(Hoofdstukken 2 en 3 van dit proefschrift)

VI

Door de ontwikkeling van steeds meer modellen waarin te verklaren variabelen en/of storingen een niet-normale kansverdeling bezitten, en door de immer toenemende capaciteit van computers die ertoe leidt dat een steeds grotere variëteit aan modellen in de praktijk gebruikt kan worden, is het gebruik van de normale verdeling in de econometrie steeds minder normaal te noemen.

VII

Aangezien de persoonlijkheid van een mens door de werking van zijn/haar hersenen bepaald wordt, kan het begrip ‘onthoofden’ beter vervangen worden door ‘ontrompen’.

VIII

Metafysische overtuigingen kunnen tot gewetensbezwaren leiden bij het uitvoeren van econometrische analyses. De eis dat instrumentele variabelen exogeen moeten zijn, betekent bijvoorbeeld dat een holist, die aanneemt dat alles met alles samenhangt, niet een schattingsmethode met instrumentele variabelen behoort te gebruiken.

IX

Het gebruik van de uitdrukking ‘twee dezelfde’ in de zin van twee gelijke voorwerpen is onjuist. Als twee voorwerpen exact gelijk zouden zijn (op dezelfde tijd, op dezelfde plaats, in exact gelijke hoedanigheid), dan betreft dit er slechts één.

X

Omdat bij verkiezingen op een bepaald niveau (bijv. gemeente of land) voor een deel van de kiezers de mening over politieke partijen die ook op andere niveaus actief zijn, beïnvloed wordt door het functioneren van deze partijen op de andere niveaus, kan de uitslag van verkiezingen enkel als een beoordeling van het functioneren van partijen op het betreffende niveau beschouwd worden als alle deelnemende partijen uitsluitend op dit niveau actief zijn, dus wanneer o.a. landelijke partijen niet bij gemeenteraadsverkiezingen toegelaten worden.

XI

Het tegelijkertijd strenger optreden tegen roken op de werkplek en het plaatsvinden van een reorganisatie op de Faculteit der Economische Wetenschappen van de Erasmus Universiteit kan ervoor gezorgd hebben dat indexcijfers voor de kwaliteit van de atmosfeer en de sfeer gelijktijdig respectievelijk toe- en afnamen. Als men simpelweg een tijdreeks betreffende de kwaliteit van de sfeer op een tijdreeks met betrekking tot de atmosfeer zou regresseren, dan zou dit een zogenaamde schijnregressie kunnen betreffen, waarbij ook als er geen direct effect van een sigaretje op de gezelligheid is, er schijnbaar sprake is van een significant positief effect van roken op de sfeer.