

STANDAARDISATIE EN HET PRESENTEREN VAN KRUISTABELLEN IN PANELONDERZOEK

Frank J.R. van de Pol en Paul van Beek *)

Samenvatting

Bij de analyse van doorstromingstabellen is het vaak interessant te toetsen of er een trend is en of de doorstroming en de trend verschillen tussen bepaalde subgroepen. Standaardisatie kan een nuttig hulpmiddel zijn voor het presenteren van een verschil in verandering tussen subgroepen. Dit verschil wordt namelijk pas goed zichtbaar wanneer gecontroleerd is voor verschillen tussen de subgroepen in de beginverdeling van het onderzochte kenmerk. De methode wordt beschreven aan de hand van een voorbeeld. Daarnaast wordt enige aandacht geschonken aan loglineaire modeltoetsen waarmee men kan onderzoeken of deze standaardisatie al dan niet zinvol is. Tenslotte wordt, om de verschillen tussen subgroepen te analyseren, een berekeningswijze gepresenteerd voor de varianties van de gestandaardiseerde celfrequenties.

STANDARDIZING TURNOVER TABLES

Summary

In the analysis of turn-over tables the emphasis is often on model selection and the presentation of model parameters. However in some cases the impact of an independent variable may be demonstrated more clearly by simply standardizing the influence of intervening variables. When turnover tables from several populations are to be compared there should be a correction for differences in the composition of these populations. Furthermore the standardized table may be summarized in categories like 'change' and 'no change'. An example is given and standard deviations are computed for the standardized cell frequencies.

*) De in dit artikel weergegeven opvattingen zijn die van de auteurs en komen niet noodzakelijk overeen met het beleid van het Centraal Bureau voor de Statistiek. De auteurs danken P.M. de Jong, P. Kroonenberg, J. Pannekoek, D. Sikkels, L.J.M. Thijssen, A. Verbeek en een anonieme beoordelaar voor commentaar en advies.

Adres van de auteurs:
Centraal Bureau voor de Statistiek
Hoofdafdeling Statistische Methoden
Postbus 959
2270 AZ Voorburg

Telefoon:
070 - 694341

1. Inleiding, probleemstelling

In sociaal-wetenschappelijk onderzoek wordt veelal gebruik gemaakt van dwarsdoorsnede of cross-sectioneel onderzoek. Hiertoe wordt een steekproef getrokken uit de desbetreffende doelpopulatie. Presentatie van de resultaten geschiedt door het geven van populatieschatters, bijvoorbeeld in de vorm van percentages of associatiematen. Als deze schatters op een gegeven moment verouderd dreigen te raken wordt, soms, besloten tot het houden van een tweede dwarsdoorsnede onderzoek. Dit onderzoek levert nieuwe schatters en maakt het bovendien mogelijk conclusies te trekken over de verandering ten opzichte van het eerste onderzoek.

Bij de interpretatie van de verandering van populatieschatters moet men altijd bedenken dat het om verandering van geaggregeerde grootheden gaat. Wanneer de interesse uitgaat naar veranderingen van individuen op een bepaald kenmerk kan niet worden volstaan met de vergelijking van twee populatieschatters van het genoemde type. Het is immers mogelijk dat de schatters in de tijd gelijk blijven terwijl op individueel niveau veel verandering optreedt. Wil men individuele veranderingen onderzoeken dan is het noodzakelijk dezelfde individuen op twee of meer tijdstippen te ondervragen.

De uitkomsten van een dergelijk panelonderzoek kunnen worden gepresenteerd door middel van zogenaamde turnover- of doorstromingstabellen. Een doorstromingstabel geeft het verband weer tussen een bepaald kenmerk op het eerste meetmoment (A) en

Tabel 1. Aantal uren vrijetijdsbesteding naar onderzoeksjaar. Nederlanders die in 1976 55 - 64 jaar oud waren en toen werkten

	1976 (A)	1982 (B)		
		1. weinig	2. gemiddeld	3. veel
1. weinig	46	64	27	137
2. gemiddeld	26	126	70	222
3. veel	2	19	30	51
totaal	74	209	127	410

hetzelfde kenmerk op het tweede meetmoment (B). Tabel 1 is een voorbeeld van een doorstromingstabel, ontleend aan Claessens en Thijssen (CBS, 1985).

Bij de interpretatie van doorstromingstabellen hebben vele auteurs gewezen op het verschijnsel regressie naar het gemiddelde (Blomqvist, 1977, Dessens en Jansen, 1982, Goldstein, 1979, Groenland en Van de Stadt, 1985, Labouvie, 1982, Markus, 1980, Nesselroade e.a., 1980, Van de Pol, 1985). Met regressie naar het gemiddelde wordt bedoeld dat van mensen die op een kenmerk op het eerste tijdstip extreem scoren, verwacht mag worden dat ze op het tweede tijdstip minder extreem scoren, dat wil zeggen dichter bij de modale categorie c.q. het gemiddelde komen te liggen.

Wanneer het om een kenmerk gaat waarvan de marginale verdeling min of meer constant is in de tijd dan kan het niet anders of er doet zich regressie naar het gemiddelde voor, althans als het verband tussen opeenvolgende metingen niet perfect is of, anders gezegd, wanneer de individuele scores veranderen. De extreem scorende mensen op het eerste tijdstip kunnen immers op het tweede tijdstip slechts minder of even extreem scoren, niet nog extremer, tenzij het aantal categorieën en de variantie toenemen. Uiteraard is er behalve deze regressie van extremen naar het gemiddelde ook een compenserende 'digressie' vanuit het midden naar die extremen. Anders zouden alle scores uiteindelijk in het midden uitkomen.

Bij een kenmerk dat op intervalniveau wordt gemeten, zoals het inkomen, manifesteert regressie naar het gemiddelde zich als een negatief verband tussen enerzijds de score op het eerste tijdstip, A, en anderzijds de verandering in de scores tussen tijdstip 1 en 2, B-A (Groenland en Van de Stadt, 1985). De correlatie tussen aanvangsscore A en verschilscore B-A is namelijk negatief, $\rho(A, B-A) = \{[\text{cov}(A, B) - \text{var}(A)] / c^2\} < 0$, als de covariantie van A en B kleiner is dan de variantie van A. Het is eenvoudig in te zien dat dit het geval is wanneer $\text{var}(A) = \text{var}(B)$, tenzij er geen verandering is, $\text{cov}(A, B) = \text{var}(A)$. Wanneer de variantie van B groter is dan die van A dan hoeft $\rho(A, B-A)$ niet negatief te zijn. In tabel 2 wordt geïllustreerd dat ook voor de gegevens van tabel 1 geldt dat $\rho(A, B-A) < 0$, althans wanneer men bereid is om de variabelen A en B even op te vatten als interval-variabelen, waarvoor verschilcores berekend mogen worden.

Tabel 2. Verandering in vrijetijdsbesteding tussen 1976 en 1982 naar de uitgangssituatie (1976), de correlatie $\rho(A, B-A) = -0,55$

1976 (A)	verschil 1982 - 1976 (B-A)				
	-2	-1	0	1	2
1. weinig	.	.	46	64	27
2. gemiddeld	.	26	126	70	.
3. veel	2	19	30	.	.

De inhoud van een doorstromingstabel hangt dus samen met de beginverdeling. Stel nu dat een gedeelte van een populatie geconfronteerd wordt met een verandering zoals pensionering en een ander deel, de "nonequivalent control group" niet (Cook en Campbell, 1979). Dan willen we het eventuele effect van pensionering op de doelvariabele zo zuiver mogelijk schatten, los van verschillen die vantevoren al bestonden tussen de experimentele en de contrôlegroep. Doorgaans zal tenminste de verdeling van de doelvariabele in de uitgangssituatie niet hetzelfde zijn in beide groepen, zoals wordt geïllustreerd in de tabel 3. Er wordt onderscheid gemaakt tussen ouderen die tussen 1976 en 1982 zijn blijven werken en ouderen die zijn uitgetreden uit het arbeidsproces. Uiteraard zou men ook meer dan twee condities met elkaar kunnen vergelijken.

Tabel 3. Verandering in het aantal uren vrijetijdsbesteding voor ouderen die nog werken en ouderen die inmiddels zijn uitgetreden (met standaardafwijkingen)

1976 (A)	1982 (B)							
	werkenden (C=1)				uitgetreden (C=2)			
	weinig	gemiddeld	veel	totaal	weinig	gemiddeld	veel	totaal
	$\bar{x}$							
1. weinig	16(4)	16(4)	8(3)	40	10(2)	16(2)	6(1)	31
2. gemiddeld	7(3)	35(5)	9(3)	51	6(1)	29(3)	20(2)	55
3. veel	0(0)	7(2)	2(1)	8	1(0)	4(1)	9(2)	14
totaal	23	57	20	100	16	49	35	100
aantal resp.	25	61	21	107	49	148	106	303

Uitgangspunt bij de interpretatie van tabel 3 is de vraag of er een verschil bestaat in de mate van verandering tussen de twee groepen op het kenmerk vrijetijdsbesteding. Om het verschil in verandering beter te kunnen zien is uit tabel 3 een nieuwe tabel geconstrueerd die per groep aangeeft hoe vaak de vrijetijdsbesteding is toegenomen, afgenomen dan wel gelijk gebleven (tabel 4). In termen van de verschilvariabele B-A uit tabel 2 staat toename voor 1 en 2, afname voor -1 en -2 en 'gelijk' voor 0. Bij variabelen met meer categorieën valt te overwegen de toename en afname verder uit te splitsen.

Tabel 4. Werkenden en uitgetreden en naar verandering in aantal uren vrijetijdsbesteding (met standaardafwijkingen)

	toename	gelijk	afname	totaal	totaal
	$\bar{x}$				abs.
werkenden	34 (5)	52 (6)	14 (3)	100	107
uitgetreden en	41 (3)	48 (4)	11 (2)	100	303
totaal	39	49	11	100	410

In tabel 4 is te zien dat de vrijetijdsbesteding van uitgetreden en meer toenam dan die van werkenden. We hebben zojuist echter gezien dat de mate van verandering in een doorstromings-tabel afhankelijk is van de verdeling van de scores op het eerste tijdstip. Wanneer de ene groep aanvankelijk overwegend lage scores heeft en de andere groep overwegend hoge scores, dan doet zich meestal bij de volgende peiling een terugkeer voor naar de gemeenschappelijke verdeling van beide groepen samen. Men kan zich in zeldzame gevallen echter ook een data-genererend proces indenken waarbij dit niet het geval is (Van de Stadt en Wansbeek, 1985, Nesslerode e.a., 1980).

Deze verandering die samenhangt met een afwijkende beginverdeling wordt veelal triviaal gevonden. Wat men wil weten is of de verandering van beide groepen ook zou verschillen, als de beginverdeling voor beide groepen hetzelfde was geweest. Dat kan eenvoudig door de tabel zodanig te wijzigen dat de verdeling op het eerste tijdstip hetzelfde is voor de beide deelgroepen, werkenden en uitgetreden en. Als er een tendens was tot regressie naar een gemeenschappelijk gemiddelde dan is die nu voor beide groepen gelijk. De verschillen die overblijven hangen samen met het onderscheid tussen beide groepen.

In paragraaf 2 wordt de standaardisatie-methode behandeld, waarmee de invloed van de beginverdeling wordt geëlimineerd. In paragraaf 3 worden varianties gegeven voor gestandaardiseerde waarden. Het artikel eindigt met een discussie.

2. Standaardisatie

2.1 Algemeen

In tabel 3 zijn afzonderlijke doorstromingstabellen berekend voor twee deelpopulaties, de werkenden en de uitgetreden. Wanneer die tabellen aanzienlijk lijken te verschillen dan wil men dat verschil zo duidelijk mogelijk presenteren. In de inleiding is er op gewezen dat de doorstroming samenhangt met de verdeling op het begintijdstip. De belangstelling van het onderzoek waaraan het materiaal is ontleend gaat echter uit naar verschillen in verandering, niet naar verschillen in niveau, op wat voor tijdstip dan ook. Er moet dus gecontroleerd worden voor de beginsituatie om de verandering beter te doen uitkomen (CBS, 1985).

Dit kan gerealiseerd worden door standaardisatie toe te passen voor de interactie tussen de beginsituatie (A) en de verklarende variabele (C), $A \times C$. Dat wil zeggen dat de marginalen van A voor werkenden, $C=1$, gelijk gemaakt worden aan die voor uitgetreden, $C=2$, eventueel op een constant verhoudingsgetal c na. In de hele steekproef geldt de verhouding 33%, 54%, 12% voor A. Het ligt voor de hand deze verhouding op te leggen binnen de beide categorieën van de verklarende variabele C. De verhouding c wordt hier dus op 1 gesteld. Het vergelijken wordt vergemakkelijkt als het totaal in beide groepen gelijk is, bijvoorbeeld 100 (Mosteller, 1968). Deze procedure is een bijzonder geval van directe standaardisatie (Fleiss, 1973, Bishop e.a., 1975, Israëls en De Ree, 1983).

Veel panelonderzoekers presenteren doorstromingstabellen in de vorm van overgangskanstabellen. Ieder rijtotaal wordt op 100 gesteld; de inhoud van een cel $A_i B_j$ in zo'n tabel is het percentage mensen dat gegeven een bepaalde uitgangssituatie i terecht komt in de nieuwe situatie j . Deze manier van presenteren werd onder meer toegepast door Clausen (1972). Plewis (1985) geeft een model van Koch e.a. (1977) voor het beschrijven van de marginalen van de doorstromingstabellen. Wanneer de proporties niet getrans-

formeerd worden tot logits dan zijn de parameters eenvoudig te interpreteren. Andere artikelen over "typische panelvragen" werden geschreven door Duncan (1979) en Hagenaars (1983).

In dit artikel is het echter de bedoeling om de doorstromingstabel samen te trekken tot een veranderingstabel met minder cellen, zoals in tabel 4. In zo'n geval kan niet iedere rij hetzelfde gewicht 100 hebben, maar moet een gewicht overeenkomstig de omvang van die categorie gehanteerd worden.

Sommige auteurs zien loglineaire analyse als een alternatief voor standaardisatie (Fienberg, 1980). Het is inderdaad juist dat men met loglineaire modellen prima kan toetsen of bepaalde interacties significant zijn in een uitgesplitste doorstromingstabel. Wanneer men echter aan de lezer wil tonen wat er in de tabel aan de hand is, dan vinden we een gestandaardiseerde tabel aantrekkelijker dan de loglineaire modelparameters omdat standaardisatie bij de lezer geen voorkennis vereist. Helaas kunnen gestandaardiseerde tabellen, zoals wordt uiteengezet in de appendix, niet berekend worden met de loglineaire modelparameters (tenzij zich een uitzonderingssituatie voordoet). In de context van standaardisatie van drie- en meerweg tabellen moet het gebruik van deze modellen daarom exploratief blijven.

In deelparagraaf 2.2 wordt standaardisatie voor verschillen in de beginverdeling ($A \times C$) besproken. In deelparagraaf 2.3 wordt tweezijdige standaardisatie besproken, zowel voor verschillen in de beginverdeling ($A \times C$) als voor verschillen in de eindverdeling ($B \times C$).

2.2 Eenzijdige standaardisatie

Zoals gezegd wordt de beginmeting A genoemd, de tweede meting B en de verklarende variabele C. In tabel 5 is tabel 3 zodanig gestandaardiseerd, dat de verdeling van A hetzelfde is voor werken (C=1) en voor uitgetreden (C=2). Deze verdeling is nu in beide groepen gelijk aan die in de hele steekproef. In de onderste rij wordt het ongestandaardiseerde aantal respondenten gegeven.

Deze bewerking kan zonder standaardprogramma worden uitgevoerd, maar voor grote tabellen kan men desgewenst ECTA gebruiken. De elementen van iedere A_i -de rij van de oorspronkelijke ge-

observeerde frequentietabel f_{ijk}^{ABC} worden gedeeld door het rij-totaal f_{i+k}^{ABC} van de k-de groep van C en vermenigvuldigd met het gemiddelde rijtotaal van alle groepen f_{i++}^{ABC}/K . K is het totaal aantal groepen in C; in het voorbeeld is K gelijk aan 2. Vervolgens wordt het totaal voor iedere groep weer op 100 gebracht, net als in tabel 3.

Tabel 5. Verandering in het aantal uren vrijetijdsbesteding (AxB) voor ouderen die nog werken en ouderen die inmiddels zijn uitgetreden (C) gestandaardiseerd voor de AxC interactie (met standaardafwijkingen)

1976 (A)	1982 (B)							
	werkenden (C=1)				uitgetreden (C=2)			
	weinig	gemiddeld	veel	totaal	weinig	gemiddeld	veel	totaal
$\bar{x}$								
1. weinig	13(3)	13(3)	7(2)	33	10(2)	17(2)	6(1)	33
2. gemiddeld	8(3)	36(4)	10(3)	54	6(1)	29(2)	19(2)	54
3. veel	0(0)	10(2)	3(2)	12	1(0)	4(1)	8(1)	12
totaal	21	59	20	100	17	49	34	100
aantal resp.	25	61	21	107	49	148	106	303

Het totaal van de gestandaardiseerde celfrequenties x_{ijk}^{ABC} moet zijn $n' = 100 K$, dus er wordt gedeeld door $n = f_{i++}^{ABC}$ en vermenigvuldigd met 100 K. (Rekenend vanuit tabel 3 moet de factor $100/n$ buiten beschouwing worden gelaten.)

$$x_{ijk}^{ABC} = \frac{f_{ijk}^{ABC}}{f_{i+k}^{ABC}} \cdot \frac{f_{i++}^{ABC}}{K} \cdot \frac{100 K}{n} = \frac{100}{n} \cdot \frac{f_{i++}^{ABC}}{f_{i+k}^{ABC}} \cdot f_{ijk}^{ABC} \quad (1)$$

De gestandaardiseerde tabel 5 maakt het in principe mogelijk voor iedere rij, dat wil zeggen voor iedere uitgangssituatie, de werkenden te vergelijken met de uitgetreden. Vanuit de theorie over regressie naar het gemiddelde weten we dat bij de rij van ouderen met veel vrijetijdsbestedingen, x_{3jk}^{ABC} , een afname in vrijetijdsbestedingen is te verwachten, een terugkeer naar het gemiddelde. Gezien tabel 5 is die afname geringer voor de uitgetreden dan voor de werkenden. Dit wordt niet veroorzaakt door de samenhang tussen de verandering en de uitgangssituatie, want er is gecorrigeerd voor de verschillen in uitgangssituatie. In paragraaf 3.2 wordt berekend dat het verschil tussen x_{331}^{ABC} en x_{332}^{ABC} juist significant is op 5% niveau.

In tabel 4 werd de verandering in hoeveelheid vrijetijdsbesteding bondig weergegeven in drie categorieën: toename, geen verandering en afname. In tabel 3 is echter te zien dat de uitgangssituatie voor werkenden en uitgetreden en verschilt. De mensen die zijn blijven werken besteedden op het eerste tijdstip al minder tijd aan vrijetijdsbestedingen. Hoe zou tabel 4 eruit zien als er gecontroleerd wordt voor dit verschil in beginverdeling tussen werkenden en uitgetreden en? Tabel 6, een samentreking van tabel 5, geeft hierop het antwoord.

Tabel 6. Werkenden en uitgetreden en naar verandering in hoeveelheid vrijetijdsbesteding, gestandaardiseerd voor de $A \times C$ interactie (met standaardafwijkingen tussen haakjes)

	toename	gelijk	afname	totaal	totaal
	$\bar{x}$				abs.
werkenden	30 (4)	52 (5)	18 (3)	100	107
uitgetreden en	43 (3)	47 (3)	10 (2)	100	303
totaal	39	49	11	100	410

De verschillen in verandering blijken in dit voorbeeld groter te zijn wanneer men rekening houdt met de verschillen in uitgangssituatie. Hoewel de uitgetreden en oorspronkelijk al meer tijd besteedden aan hun hobbies is juist bij hen de toename in vrijetijdsbestedingen het grootst. Het is dus niet het verschil in samenstelling van beide groepen, maar de uittreding uit het arbeidsproces (en wat daarvan de oorzaak is) die een toename van vrijetijdsbestedingen bewerkstelligt.

2.3 Tweezijdige standaardisatie

Alle respondenten samen gingen in de beschouwde 6 jaar meer tijd besteden aan vrijetijdsbestedingen (tabel 1). Blijkens de tabellen 3 en 5 deed deze toename zich met name voor onder de mensen die in deze periode ophielden met werken. Dit zou men een $(A \times B) \times C$ interactie kunnen noemen. Wat betreft de drieweg interactie $A \times B \times C$ kan men zich afvragen of de geringere terugkeer naar het gemiddelde onder uitgetreden actieve hobbyisten, x_{3j2}^{ABC} vergeleken met x_{3j1}^{ABC} , verklaard wordt door deze toename van vrijetijdsbestedingen onder uitgetreden en ($x_{432}^{ABC} > x_{431}^{ABC}$). Om dat te beoordelen moet de rij van de actieve hobbyisten gezien worden in verhouding tot de totaalrij. Voor de lezer is het echter makkelijker als de tabel voor dit doel niet alleen gestandaardiseerd

wordt voor de AxC interactie, maar ook voor het verschil in de trend, de BxC interactie.

Zo'n standaardisatie naar twee marginalen moet iteratief tot stand komen omdat de ene aanpassing (aan de gewenste BxC marginalen) de andere aanpassing (aan de gewenste AxC marginalen) gedeeltelijk weer teniet doet. Alvorens het iteratief aanpassen van de marginalen te starten wordt eerst het totaal van de tabel op 100 K gebracht (voor K=2 wordt het tabeltotaal n'= 200). Het iteratieve proces kan als volgt worden beschreven.

$$x_{ijk}^{ABC}(0) = f_{ijk}^{ABC} \frac{100 K}{n}$$

$$x_{ijk}^{ABC}(1) = \frac{x_{ijk}^{ABC}(0)}{x_{i+k}^{ABC}(0)} \frac{x_{i++}^{ABC}(0)}{K} \quad (\text{AxC})$$

$$x_{ijk}^{ABC}(2) = \frac{x_{ijk}^{ABC}(1)}{x_{+jk}^{ABC}(1)} \frac{x_{+j+}^{ABC}(0)}{K} \quad (\text{BxC}) \quad (2)$$

$$x_{ijk}^{ABC}(3) = \frac{x_{ijk}^{ABC}(2)}{x_{i+k}^{ABC}(2)} \frac{x_{i++}^{ABC}(0)}{K} \quad (\text{AxC})$$

etc.

De vierde stap is weer een (BxC) aanpassing en er wordt verder geïtereerd totdat de tabel de gewenste marginalen heeft verkregen. Dit is een vorm van iteratief proportioneel fitten, waarvan bekend is dat de procedure convergeert. De verdeling van het aantal uren vrijetijdsbesteding op het tweede tijdstip (B) in de hele steekproef is: 18% weinig, 51% gemiddeld en 31% veel. In tabel 7 is de B- verdeling voor C=1 en C=2 vervangen door deze verdeling.

Men ziet in tabel 7 dat de geringere terugkeer naar het gemiddelde onder uitgetreden actieve hobbyisten, x_{3j2}^{ABC} , slechts voor een klein deel verklaard wordt door de grotere toename van vrijetijdsbestedingen onder uitgetreden (de trend van x_{i+2}^{ABC} naar x_{+j2}^{ABC}), die in tabel 3 nog te zien was maar in tabel 7 is weggestandaardiseerd. Er wordt uiteraard steeds vergeleken met de werkenden, de actieve hobbyisten x_{3j1}^{ABC} , c.q. de trend van x_{i+1}^{ABC} naar x_{+j1}^{ABC} . Het in tabel 5 geconstateerde verschil blijft grotendeels, maar niet geheel, bestaan. Een inhoudelijke

Tabel 7. Verandering in het aantal uren vrijetijdsbesteding (AxB) voor ouderen die nog werken en ouderen die inmiddels zijn uitgetreden (C) gestandaardiseerd voor de AxC en BxC interacties

1976 (A)	1982 (B)							
	werkenden (C=1)				uitgetreden (C=2)			
	weinig	gemiddeld	veel	totaal	weinig	gemiddeld	veel	totaal
χ^2								
1. weinig	11	11	11	33	11	17	6	33
2. gemiddeld	7	32	16	54	6	30	18	54
3. veel	0	8	4	12	1	4	8	12
totaal	18	51	31	100	18	51	31	100
aantal resp.	25	61	21	107	49	148	106	303

interpretatie hiervan valt moeilijk te geven. Misschien betreft het hier een afwijking die het gevolg is van de steekproeftrekking van de respondenten. Het effect in tabel 5 was niet ver over de grens van significantie; het zwakkere effect in tabel 7 kan aan het toeval worden toegeschreven.

Tweezijdige standaardisatie kan nuttig zijn om de verandering vanuit een bepaalde uitgangssituatie los te zien van de trend. Aangezien de uitgangssituatie (A) niet meer voorkomt in een veranderingstabel, is het minder zinvol om tabel 7 samen te trekken tot een veranderingstabel (tabel 8).

Tabel 8. Werkenden en uitgetreden naar verandering in hoeveelheid vrijetijdsbesteding, gestandaardiseerd voor de AxC en BxC interacties

	toename	gelijk	afname	totaal	totaal
	χ^2				abs.
werkenden	38	47	15	100	107
uitgetreden	40	49	11	100	303
totaal	39	48	12	100	410

Na standaardisatie voor verschillen tussen beide randtotalen blijkt in tabel 8 dat er nog steeds verschillen bestaan tussen de beide groepen, ook al zijn die in dit geval te klein om significant te zijn. Het bestaan van verschillen is niet verwonderlijk aangezien tabel 8 niet alleen afhankelijk is van de trend, dat wil zeggen van de randtotalen van de doorstromingstabellen, maar

ook van de doorstroming in die tabellen. Tabel 8 kan men opvatten als het verschil in verandering tussen de beide groepen, conditioneel op het verschil in trend.

Het verschil tussen werkenden en uitgetreden en in tabel 8 is veel kleiner dan in tabel 6. Het grootste deel van het verschil in tabel 6 moet dus worden toegeschreven aan het verschil in trend, dat wegvalt als zowel voor $A \times C$ als voor $B \times C$ interacties gecontroleerd wordt. Een kleiner deel van het verschil in tabel 6 kan worden toegeschreven aan de drieweginteractie $(A \times B) \times C$, althans de samentrekking daarvan die we $V \times C$ zullen noemen. Deze nuancering is zoals gezegd alleen relevant als methodologische exercitie. Voor de presentatie van verandering geven we de voorkeur aan tabel 6, waar op overzichtelijke wijze zowel de trend als iets van de doorstroming getoond wordt.

3. Varianties van gestandaardiseerde frequenties

Wanneer een tabel is gestandaardiseerd voor bepaalde effecten die niet relevant zijn voor de probleemstelling dan blijft de vraag of de effecten die in de gestandaardiseerde tabel getoond worden wel significant zijn. Deze vraag kan globaal beantwoord worden door loglineaire of logit-modellen op de gegevens te passen (Duncan, 1979, Hagenaars, 1983, Plewis, 1985). Op het niveau van de celfrequenties zijn de effecten meer in detail zichtbaar. De significantie van deze effecten kan beoordeeld worden door standaardafwijkingen van de gestandaardiseerde frequenties te berekenen.

Wanneer we de (relatief kleine) covarianties verwaarlozen dan geven de standaardafwijkingen aan welke verschillen tussen de cellen zo groot zijn dat men ze mag interpreteren en welke verschillen zich niet onderscheiden van toevalsfluctuaties als gevolg van de steekproeftrekking. De varianties van de in tabel 3 gegeven waarden zijn niet de gebruikelijke n.p.q van de binomiaalverdeling, want in tabel 3 worden niet de geobserveerde frequenties weergegeven maar percentages ten opzichte van het totaal aantal respondenten in iedere groep $(C_1, \dots, C_K)$, f_{++k}^{ABC} . De varianties van de quotiënten in tabel 3 kunnen worden benaderd met een Taylor expansie. (Het superscript ABC wordt in het vervolg weggelaten.)

$$\text{var}\left(\frac{100}{f_{++k}} f_{ijk}\right) \approx \frac{100^2}{n p_{++k}^2} p_{ijk} \left[1 - p_{ijk} - \frac{p_{ijk}}{p_{++k}} (1 - p_{++k}) \right] \quad (3)$$

waarin $p_{ijk} = E(f_{ijk}/n)$. De afleiding is opvraagbaar bij de auteurs. Men zou een iets grotere variantie vinden wanneer de totalen van de deelgroepen, f_{++k} , (ten onrechte) als een niet-stochastische constante werden opgevat, $f_{++k} = n p_{++k}$. Dan zou tussen de grote haken in formule (3) alleen $[1 - p_{ijk}]$ staan.

De voor AxC gestandaardiseerde celinhouden van tabel 5, x_{ijk} , zijn berekend volgens formule 1. Wanneer we schrijven $w = f_{i++}/f_{i+k}$ en $y_{ijk} = w f_{ijk}$ dan is de variantie van x_{ijk}

$$\text{var}(x_{ijk}) = \text{var}\left[\frac{100}{n} \frac{f_{i++}}{f_{i+k}} f_{ijk}\right] = \frac{100^2}{n^2} \text{var}(w f_{ijk}) = \frac{100^2}{n^2} \text{var}(y_{ijk}) \quad (4)$$

De variantie van y_{ijk} wordt zodanig uitgeschreven dat de wegingsfactor (gedeeltelijk) buiten haakjes gehaald kan worden (met dank aan A.Verbeek). De variantie wordt ontbonden in twee delen, conditioneel op f_{i+k} . Zie onder meer Mood e.a., 1974, blz.159.

$$\text{var}(y_{ijk}) = E[\text{var}(y_{ijk} | f_{i+k})] + \text{var}[E(y_{ijk} | f_{i+k})] \quad (5)$$

De uitwerking van formule (5) zou teveel ruimte innemen. Wanneer het resultaat wordt ingezet in (4) wordt voor de variantie van x_{ijk} gevonden

$$\text{var}(x_{ijk}) \approx \frac{100^2}{n^2} \text{var}(y_{ijk}) = \frac{100^2 p_{i++}^2}{n p_{i+k}^2} p_{ijk} \left[1 - p_{ijk} - \frac{p_{ijk}(p_{i++} - p_{i+k})}{p_{++k} p_{i+k}} \right] \quad (6)$$

Deze varianties vallen weer wat kleiner uit dan die van formule (3) omdat nu niet alleen voor f_{++k} , maar ook voor f_{i+k} gecontrôleerd wordt. Voor het vergelijken van twee gestandaardiseerde cellen x_{ijk} en $x_{(ijk)}$, moet men eigenlijk ook de covariantie van deze twee waarden kennen. Immers

$$\text{var}(x_{ijk} - x_{(ijk)}) = \text{var}(x_{ijk}) + \text{var}(x_{(ijk)}) - 2\text{cov}(x_{ijk}, x_{(ijk)}) \quad (7)$$

De covariantie tussen twee geobserveerde frequenties f_1 en f_2 is $-np_1p_2$, terwijl de varianties $np_1(1-p_1)$ en $np_2(1-p_2)$ zijn. Bij een tabel met een groot aantal cellen zoals hier is de covariantie relatief klein en negatief. Wanneer we de covariantie in (7) verwaarlozen wordt een lichte onderschatting verkregen van de gezochte variantie van $x_{ijk} - x_{(ijk)'}.$

De standaardafwijking van het verschil tussen x_{331} en x_{332} , de diagonaalelementen in de derde rij van tabel 5, met waarden 3 en 8, wordt benaderd met $\sqrt{(1,8^2 + 1,4^2)} \approx 2,3$. Het verschil wijkt significant af van 0 op 5% niveau ($8 - 3 = 5 > 4,4 = 1,96 \times 2,3$). Het getal 1,96 is ontleend aan de normale verdeling.

De standaardafwijkingen in de tabellen 4 en 6 zijn verkregen door evenals bij formule (7) de covarianties van de opgetelde x_{ijk} 's te verwaarlozen. Omdat in dit geval niet een verschil maar een som van x_{ijk} 's wordt berekend zouden de covarianties moeten worden opgeteld in plaats van afgetrokken zoals in (7). Aangezien de covarianties die worden verwaarloosd negatief zijn, verwaarlozen we een negatieve bijdrage aan de variantie; de standaardafwijkingen in tabel 4 en 6 zijn lichte overschattingen van de gezochte varianties.

De vraag of de toename in vrijetijdsbestedingen (na standaardisatie voor de uitgangssituatie) significant groter is geweest bij de uitgetreden dan bij de werkenden beantwoorden we als volgt. Uit tabel 6 blijkt dat het verschil $43 - 30 = 13$ is. De standaardafwijking van het verschil is bij benadering $\sqrt{(4,4^2 + 3,3^2)} \approx 5,5$. Het verschil in toename is significant op 5% niveau ($43 - 30 = 13 > 10,8 = 1,96 \times 5,5$).

Voor sociologisch onderzoek naar doorstroming zoals dat van Claessens en Thijssen (CBS, 1985), waar dit voorbeeld aan is ontleend, kan men concluderen dat het verklaren van een doorstromingstabel door naar deelgroepen te kijken flinke steekproef-aantallen vraagt. Zeker als men de verandering in een trichotoom kenmerk bestudeert is een steekproefomvang wenselijk van enige duizenden respondenten, meer dan de enkele honderden van het voorbeeld. Verder lijkt het erop, dat de meer globale loglineaire toetsen krachtiger zijn dan de toetsen op het verschil tussen gestandaardiseerde frequenties.

4. Discussie

Regressie naar het gemiddelde is een heel normaal verschijnsel bij herhaalde metingen aan dezelfde objecten. Niettemin zijn uit menige analyse van panelgegevens al conclusies getrokken die aanvechtbaar waren, omdat geen rekening werd gehouden met de terugkeer naar het gemiddelde. Fraaie voorbeelden worden besproken in het artikel van Dessens en Jansen (1982). Dit verschijnsel speelt ook een rol bij het onderwerp van dit artikel, het vergelijken van doorstromingstabellen van verschillende groepen.

In de inleiding is uiteengezet dat de aanvangsverdeling in een doorstromingstabel negatief gecorreleerd is met de verandering in het onderzochte kenmerk. Het vergelijken van de verandering in groepen met een verschillende aanvangsverdeling is derhalve problematisch, want een zinvolle interpretatie kan alleen gegeven worden conditioneel op het verschil in aanvangsverdeling. In de literatuur over de analyse van veranderingsscores in kenmerken die gemeten worden op intervalniveau wordt dan ook aanbevolen om nooit met verschilscores $B-A$ te werken maar alleen met 'gain' scores, de residuen van de regressie van $B-A$ op aanvangsscore A , of kortweg de residuen van B gegeven A (Cronbach en Furby, 1970, Kessler, 1977).

Dit is ook aan te bevelen als men de verandering van verschillende groepen vergelijkt. Wanneer men voor verschillende groepen gemiddelde 'gain' scores vergelijkt in plaats van gemiddelde verschilscores dan wordt gecorrigeerd voor eventuele verschillen in aanvangsgemiddelde tussen de groepen. In dit artikel ging de aandacht uit naar doorstromingstabellen van een kenmerk dat wordt gemeten op een lager niveau dan intervalniveau. Dan moet niet gecorrigeerd worden voor verschillen in aanvangsgemiddelde, maar voor verschillen in de aanvangsfrequentieverdeling. De in dit artikel voorgestelde standaardisatie voor verschillen in aanvangsverdeling bij data van ordinaal of lager niveau sluit dus goed aan bij de methoden die bekend zijn voor intervalgegevens.

Standaardisatie voor zowel de begin- als de eindverdeling kan enig gemak bieden in aanvulling op de reeds genoemde standaardisatie voor de beginverdeling. Het stelt de lezer in staat om de verandering vanuit een bepaalde aanvangssituatie los te zien van het verschil in trend tussen de beschouwde groepen. (De trend is de verandering in de frequentieverdeling.)

Het samentrekken van een doorstromingstabel in categorieën als 'veranderd' en 'niet veranderd' is niet alleen efficient bij kleine steekproef aantallen maar ook overzichtelijk. Ook het vergelijken van de ontwikkelingen in verschillende groepen op grond van zulke veranderingstabellen kan alleen zinvol gebeuren als wordt gestandaardiseerd voor de verschillen in de aanvangsverdeling. Zulke tabellen tonen overigens ook de trend, de verandering in de frequentieverdeling en het verschil tussen beide groepen in dit opzicht. In principe kan de verandering gegeven het verschil in trend getoond worden door de doorstromingstabellen samen te trekken die gestandaardiseerd zijn voor zowel de aanvangs- als de eindverdeling. Doorgaans zal men echter volstaan met een tabel die de doorstroming inclusief de trend weergeeft.

In de derde paragraaf werd aangegeven hoe getoetst kan worden of een discrete verklarende variabele samenhangt met de doorstroming. Bovendien werd aangegeven hoe de standaardafwijking berekend kan worden voor een cel die gestandaardiseerd is voor de beginverdeling. De steekproefomvang van $n = 410$ in de voorbeeldtabel werd aan de kleine kant bevonden. Voor het vergelijken van doorstromingstabellen van een trichotoom kenmerk zijn eigenlijk enige duizenden respondenten nodig. Een grotere steekproef is vooral noodzakelijk wanneer men niet alleen globale uitspraken wil doen over het bestaan van bepaalde interacties maar ook de waarden in de gestandaardiseerde tabel wil kunnen vergelijken. En om dit laatste was het begonnen. Want standaardisatie van doorstromingstabellen is een nuttige techniek om te laten zien wat de invloed is van een verklarende variabele op het veranderingsproces.

BIJLAGE De relatie tussen loglineaire modellen en standaardisatie

Men zal zich wellicht afvragen of de loglineaire modelparameters mogelijkheden bieden om de gewenste standaardisatie uit te voeren. Leidt het weglaten (op één stellen) van de AxC parameters bij het berekenen van de onder het model verwachte frequenties niet tot de gewenste standaardisatie? Om dit te onderzoeken schrijven we het loglineaire model als multiplicatief model.

$$f_{ijk}^{ABC} = \tau^O \tau_i^A \tau_j^B \tau_k^C \tau_{ij}^{AB} \tau_{ik}^{AC} \tau_{jk}^{BC} \tau_{ijk}^{ABC} \quad (A1)$$

De standaardisatie voor AxC volgens formule (A1) betekent in loglineaire termen dat alleen in de bivariate AxC tabel de AxC interactieparameters één zijn. Dat wil echter niet zeggen dat dan in de complete AxBxC tabel de loglineaire AxC interactieparameters eveneens één zijn.

Dit kan worden aangetoond met het 'collapsibility' theorema (Bishop e.a., 1975, Fienberg, 1980). Het collapsibility theorema kan voor dit geval geformuleerd worden als: 'De AxC interactie in een AxBxC tabel is gelijk aan de AxC interactie in de bivariate AxC tabel als de AxBxC interactie afwezig is en ofwel de AxB interactie afwezig is ofwel de BxC interactie afwezig is.' Ervan uitgaande dat de AxB interactie, de stabiliteit, wel aanwezig is, moet de BxC interactie dus afwezig zijn, wil men met de loglineaire modelparameters kunnen standaardiseren. Maar in dat geval is het uitsplitsen van de AxB tabel naar C niet interessant, want er is geen samenhang tussen C en de andere variabelen, behalve wellicht de ongewenste verschillen in uitgangssituatie, AxC. Fienberg (1980) wijst er verder op dat het over B 'collapsen' van een tabel met behoud van dezelfde loglineaire parameters voor AxC ook mogelijk is, als voor iedere waarde van B de verwachte proporties van de AxC tabel, $f_{+jk}^{ABC} f_{ij+}^{ABC} / f_{++j+}^{ABC}$, hetzelfde zijn. Zo'n perfecte regelmaat is echter een uitzondering.

Een nare consequentie van het bovenstaande is dat het loglineaire model niet mooi aansluit op de standaardisatie. De VxC interactie in het loglineaire model wordt berekend gegeven de AxC interactie in de trivariate AxBxC tabel, maar de VxC interactie in tabel 6 geldt gegeven de AxC interactie in de bivariate AxC tabel. Loglineaire modeltoetsen hebben in dit verband alleen een

globale, explorerende functie. Als men zekerheid wil hebben over de significantie van het verschil tussen twee gestandaardiseerde waarden bieden standaardafwijkingen uitkomst.

Het bovenstaande mag niet de indruk wekken dat het principiële onmogelijk zou zijn de gewenste standaardisatie uit te voeren met de parameters van een multiplicatief model. Er moeten dan echter bijzondere restricties worden opgelegd aan de parameters. Stel dat we gestandaardiseerde waarden willen vinden door het weglaten van r_k^C en r_{ik}^{AC} en door het tabeltotaal op 100 k te brengen in plaats van n,

$$x_{ijk}^{ABC} = r^O \frac{100 K}{n} r_i^A r_j^B r_{ij}^{AB} r_{jk}^{BC} r_{ijk}^{ABC} \quad (A2)$$

Een eerste restrictie die moet worden opgelegd is dat het totaal van iedere doorstromingstabel voor $C=1, \dots, K$ steeds 100 is.

$$x_{++k}^{ABC} = 100 \quad (A3)$$

Dit impliceert volgens (A2) dat

$$\sum_i \sum_j r^O \frac{100 K}{n} r_i^A r_j^B r_{ij}^{AB} r_{jk}^{BC} r_{ijk}^{ABC} = 100 \quad \text{voor iedere } k. \quad (A4)$$

Verder moeten zodanige restricties aan de parameters van (A1) worden opgelegd, dat de marginale verdeling van $A \times C$ in de gestandaardiseerde driewegtabel x_{i+k}^{ABC} , dezelfde verdeling van A heeft voor alle waarden van C. Anders gezegd: x_{i+k}^{ABC} mag alleen van i afhangen voor iedere k.

$$x_{i+1}^{ABC} = x_{i+2}^{ABC} = \dots = x_{i+k}^{ABC} \quad (A5)$$

Dit impliceert volgens (A2) dat moet gelden:

$$\sum_j r_j^B r_{ij}^{AB} r_{j1}^{BC} r_{ij1}^{ABC} = \sum_j r_j^B r_{ij}^{AB} r_{j2}^{BC} r_{ij2}^{ABC} = \dots \quad (A6)$$

Men ziet dat de restricties die deze standaardisatie oplegt aan de parameters van het model niet zo simpel zijn als in het log-lineaire geval waar eenvoudige restricties kunnen worden opgelegd als $\Pi_j r_j^B = 1$. Deze restricties zijn niet lineair in de logaritmie en er is voor zover bekend geen literatuur waarin beschreven staat hoe model (A1) geschat zou moeten worden onder de restricties (A4) en (A6).

Literatuur

- Bishop, Y.M.M., S.E. Fienberg en P.W. Holland, 1975, *Discrete multivariate analysis: Theory and practice* (the MIT Press, Cambridge, Massachusetts).
- Blomqvist, N., 1977, On the relation between change and initial value, *Journal of the American Statistical Association* 72, blz. 746-749.
- CBS (Centraal Bureau voor de Statistiek), 1985, *De leefsituatie van de Nederlandse bevolking van 55 jaar en ouder, 1982. Deel 4, gevolgen van uittreding uit het arbeidsproces, een panelanalyse 1976-1982* (samengesteld door L. Claessens en L.J.M.Thijssen) (Staatsuitgeverij, 's-Gravenhage).
- Clausen, J.A., 1972, *Studies of the postwar plans of soldiers*. In: P.F. Lazarsfeld, A.K. Pasanella en M. Rosenberg (eds.), *Continuities in the language of social research* (the Free Press, New York/ Collier-Mcmillan, London), blz. 363-367.
- Cook, T.D. en D.T.Campbell, 1979, *Quasi-experimentation, design and analysis issues for field settings* (Houghton Mifflin, Boston).
- Cronbach, L.J. en L. Furby, 1970, How should we measure change - or should we? *Psychological bulletin* 74, blz. 68-80.
- Dessens, J. en W. Jansen, 1982, Terugval naar het gemiddelde; een methodologische notitie, *Sociologische Gids* 29, blz. 410-419.
- Duncan, O.D., 1979, Testing key hypotheses in panel analysis. In: K.F. Schuessler (ed.), *Sociological Methodology 1980* (Jossey Bass, San Francisco), blz.279-289.
- Fienberg, S.E., 1980, *The analysis of cross-classified categorical data, second edition* (the MIT Press, Cambridge, Massachusetts).
- Fliss, J.L., 1973, *Statistical methods for rates and proportions* (Wiley, New York).
- Groenland, E.A.G. en H. van de Stadt, 1985, *Sociaal economisch panelonderzoek onder huishoudens. CBS Select 3, (Centraal Bureau voor de Statistiek, Voorburg)*, blz. 89-103.
- Hagenaars, J.A.P., 1983, *De beantwoording van 'typische panelvragen' met behulp van loglineaire modellen*, *Sociale wetenschappen* 26, blz. 147-171.
- Healy, M.J.R. en H.Goldstein, 1978, *Regression to the mean*, *Annals of Human Biology* 5, blz. 277-280.
- Israels, A.Z. en S.J.M. de Ree, 1982, *Standaardisatietechnieken. CBS-Select 2* (Centraal Bureau voor de Statistiek, Voorburg), blz.237-246.
- Keller, W.J. en A. Verbeek, 1984, *ANOTA: Analysis of tables*, *Kwantitatieve methoden* 15, blz. 28-44.
- Kessler, R.C., 1977, The use of change scores as criteria in longitudinal survey research, *Quality and Quantity* 11, blz. 43-66.
- Koch, G.G., J.R.Landis, J.L.Freeman jr. en R.G.Lehnen, 1977, A general methodology for the analysis of experiments with repeated measures of categorical data, *Biometrics* 33, blz. 133-158.
- Labouvie, E.W., 1982, The concept of change and regression toward the mean, *Psychological Bulletin* 92, blz. 251-257.
- Markus, G.B., 1980, *Analyzing panel data*, *Sage university papers* 18 (Sage, London).
- Mood, A.M., F.A. Graybill en D.C. Boes, 1974, *Introduction to the theory of statistics*, third ed. (McGraw-Hill, London / Kogakusha, Tokyo).
- Mosteller, F., 1968, Association and estimation in contingency tables, *Journal of the American Statistical Association* 63, blz. 1-28.
- Nesselroade, J.R., S.M.Stigler en P.B.Baltes, 1980, Regression toward the mean and the study of change, *Psychological Bulletin* 88, blz. 622-737.
- Plewis, I., 1985, *Analysing change. Measurement and explanation using longitudinal data* (Wiley, New York).
- Pol, F.J.R.van de, 1985, *Better measurement with a panel design? Te verschijnen in W.E.Saris en I.N.Gallhofer, eds., "Sociometric research 1985" (in voorbereiding).*
- Stadt, H.van de en T.J.Wansbeek, 1985, *Regression effects in tabulating from panel data, voorlopig rapport* (Rijksuniversiteit Groningen).