

Een toepassing van gemengde modellen voor binaire gegevens

Vaclav Fidler^{*}

Samenvatting

Molenaar (1982) beschrijft een experiment uit de sociale wetenschappen en stelt de statistische analyse van gegevens uit dit experiment ter discussie. In dit artikel wordt een analyse van het probleem van Molenaar voorgesteld. Deze analyse is gebaseerd op een gemengd variantie-analyse model voor binaire gegevens zoals behandeld in Fidler en Nagelkerke (1982).

* Centrum voor Medische Informatica en Statistiek,
Faculteit der Geneeskunde,
Bloemsingel 10, 9712 KZ Groningen.

1. Probleemstelling

In het experiment worden m opstellen in een ongecorrigeerde en in een gecorrigeerde vorm gebruikt. Een groep van $2n$ beoordelaars wordt in twee even grote groepen gesplitst; een groep kijkt de ongecorrigeerde opstellen na en de andere groep de gecorrigeerde opstellen. Het resultaat van het nakijken van een opstel wordt op zijn beurt beoordeeld en kan juist of onjuist zijn. De vraag is nu of de gecorrigeerde opstellen minder vaak door de beoordelaars goed worden beoordeeld.

In het originele experiment zoals beschreven door Molenaar waren er nog twee groepen van beoordelaars en twee groepen van opstellen. Het volledige experiment wordt in hoofdstuk 4 behandeld.

Laat $\{X_{ijk}\} = \{X_{ijk}, i = 1, \dots, n, j = 1, \dots, m, k = 1, 2\}$ de uitkomsten van het experiment zijn, X_{ijk} is gelijk aan één indien de i -de beoordelaar ($i = 1, \dots, n$) uit de k -de groep ($k = 1$: ongecorrigeerd, $k = 2$: gecorrigeerd opstel) het j -de opstel ($j = 1, \dots, m$) juist heeft beoordeeld en X_{ijk} is gelijk aan nul indien deze beoordeling fout was. De vraagstelling kan nu vertaald worden in het toetsen van

$$H_0: P(X_{ij1} = 1) = P(X_{ij2} = 1) \text{ tegen}$$

$$H_1: P(X_{ij1} = 1) > P(X_{ij2} = 1),$$

waarbij i en j een aselekt gekozen beoordelaar respectievelijk een aselekt gekozen opstel aanduiden.

2. Kansmodel

Het in hoofdstuk 1 vermelde experiment kan geclassificeerd worden als een gemengde variantie-analyse model met drie indelingen en met binaire uitkomsten. De indeling naar het al dan niet gecorrigeerd zijn van een opstel introduceert een vast effect, opstellen en beoordelaars introduceren toevallige effecten waarbij beoordelaars genest zijn in correctiegroepen. Typerend voor een gemengd model is de afhankelijkheid tussen diverse groepen van waarnemingen. In Fidler en Nagelkerke (1982) wordt aangegeven hoe deze afhankelijkheid gemodelleerd kan worden. Voor ons experiment kan het volgende model in aanmerking komen:

$$P(\{X_{ijk}\} = \{x_{ijk}\}) = C \cdot \exp \left\{ \alpha \cdot (x_{++1} - x_{++2}) + \sum_{j=1}^m \sum_{k=1}^2 \sum_{s=0}^n a_s \cdot \delta(s, x_{+jk}) + \right. \\ \left. + \sum_{j=1}^m \sum_{s=0}^{2n} b_s \cdot \delta(s, x_{+j+}) + \sum_{i=1}^n \sum_{k=1}^2 \sum_{s=0}^m c_s \cdot \delta(s, x_{i+k}) \right\},$$

waarbij C een normaliserende constante is, $\{a_s\}$ zijn parameters die de afhankelijkheid tussen n beoordelaars -die een identiek opstel zien- modelleren, $\{b_s\}$ zijn parameters die de afhankelijkheid tussen $2n$ beoordelaars -die een op de correctie na identiek opstel zien- modelleren, $\{c_s\}$ zijn parameters die de afhankelijkheid tussen m opstellen -die allen door een beoordelaar gezien worden- modelleren, $\delta(i,j)$ is de functie van Kronecker (gelijk aan 1 als $i = j$ en anders 0) en α is de correctie-effectparameter. Het plusteken op de plaats van een index geeft sommatie over deze index aan.

De asymmetrie in het modelleren van afhankelijkheid is te verklaren door het feit dat elke beoordelaar maar een groep van m -gecorrigeerde of niet gecorrigeerde-opstellen bekijkt terwijl elk opstel -op de correctie na- door $2.n$ verschillende beoordelaars gezien wordt.

Het toetsingsprobleem is nu $H_0: \alpha = 0$ tegen $H_1: \alpha > 0$. Laat

$$V = X_{++1} - X_{++2};$$

$$A_s = \sum_{j=1}^m \sum_{k=1}^2 \delta(s, X_{+jk}), \quad s = 0, 1, \dots, n;$$

$$B_s = \sum_{j=1}^m \delta(s, X_{+j+}), \quad s = 0, 1, \dots, 2n;$$

$$C_s = \sum_{i=1}^n \sum_{k=1}^2 \delta(s, X_{i+k}), \quad s = 0, 1, \dots, m.$$

De stochast A_s , $0 \leq A_s \leq 2m$ is gelijk aan het aantal opstellen (waarbij gecorrigeerde en niet gecorrigeerde opstellen apart worden geteld) dat door precies s beoordelaars juist wordt beoordeeld. De stochast B_s , $0 \leq B_s \leq m$ is gelijk aan het aantal paren van gecorrigeerde en niet gecorrigeerde versies van opstellen dat door s beoordelaars juist wordt beoordeeld. De stochast C_s , $0 \leq C_s \leq 2n$, geeft het aantal beoordelaars aan dat precies s opstellen juist heeft beoordeeld. Er geldt $\sum A_s = 2m$, $\sum B_s = m$, $\sum C_s = 2n$, $\sum s A_s = \sum s B_s = \sum s C_s = X_{+++}$. Voor de identificeerbaarheid van parameters volgt hieruit dat vijf van de parameters door middel van een bijvoorwaarde vastgelegd moeten worden, bv. a_0, b_0, b_1, c_0, c_1 .

De standaardtoetsingstheorie leidt tot de uniform meest onderscheidende toets van H_0 onder alle zuivere toetsen. De niet-gerandomiseerde versie van deze toets gebruikt als overschrijdingskans

$$P(V \geq v | \{A_s\}, \{B_s\}, \{C_s\}, \alpha = 0).$$

Teneinde deze kans te berekenen, moeten wij het aantal n_{ABC} $2m.n$ -dimensionale vectoren $\{X_{ijk}\}$ bepalen waarvoor $\{A_s\}$, $\{B_s\}$ en $\{C_s\}$ de waargenomen waarden hebben en het aantal $n_{V,ABC}$ van zulke vectoren waarvoor

$X_{++1} - X_{++2}$ minstens even groot is als de waargenomen waarde van V ; de overschrijdingskans is dan gelijk aan $P = n_{V,ABC}/n_{ABC}$.

Rest ons deze toets op de gegevens van Molenaar toe te passen. Hiervoor zouden wij over somscores $\{x_{i+k}\}$ en $\{x_{+jk}\}$ moeten beschikken. Afgezien hiervan zou een directe berekening nogal veel werk met zich meebrengen. Een handig rekenalgoritme bij kleine aantallen en een simulatie-algoritme of een benaderde berekeningswijze bij grote aantallen zou welkom zijn. Teneinde de toets toch te illustreren zullen wij een denkbeeldig voorbeeld bespreken.

3. Rekenvoorbeeld

In de onderstaande tabel staan de resultaten van het experiment met $m = 2$ opstellen en $n = 3$ beoordelaars per correctiegroep.

beoordelaar (correctie) $i(k)$	opstel 1	j 2	x_{i+k}
1(1)	0	1	1
2(1)	1	1	2
3(1)	1	1	2
x_{+j1}	2	3	
1(2)	0	0	0
2(2)	1	0	1
3(2)	0	1	1
x_{+j2}	1	1	
x_{+j+}	3	4	

Uit de tabel volgt $(A_0, A_1, A_2, A_3) = (0, 2, 1, 1)$, $(B_0, B_1, \dots, B_6) = (0, 0, 0, 1, 1, 0, 0)$, $(C_0, C_1, C_2) = (1, 3, 2)$ en $V = 3$. Beschouw allereerst het effect van conditioneren op $\{C_s\}$. Dit houdt in dat het aantal vectoren $\{x_{ijk}\}$ beperkt wordt tot die waarvoor de vector $\{x_{i+k}\}$ op een permutatie na gelijk is aan $(2, 2, 1, 1, 1, 0)$. De toegestane vectoren $\{x_{i+k}\}$ zijn gegeven in de onderstaande tabel op de permutaties binnen de correctiegroepen na. Voor elk van deze vectoren wordt in deze tabel ook het aantal $n_{V,ABC}$ van verschillende vectoren $\{X_{ijk}\}$ vermeld. De berekening hiervan wordt verderop toegelicht.

$i(k)$	1(1)	2(1)	3(1)	1(2)	2(2)	3(2)	$v=x_{++1}-x_{++2}$	$n_{v.ABC}$
x_{i+k}	2	2	1	1	1	0	3	36
	2	1	1	2	1	0	1	36
	2	2	0	1	1	1	1	0
	1	1	1	2	2	0	-1	0
	2	1	0	2	1	1	-1	36
	1	1	0	2	2	1	-3	36

Wij zullen nu een vector $\{x_{ijk}\}$ schrijven als een matrix met rijen corresponderend met opstellen ($j=1$ en $j=2$) en met kolommen $i(k)$ in de volgorde van de bovenstaande tabel. De eerste vector $\{x_{i+k}\}$ uit deze tabel, waarvoor $v = 3$, laat op de permutaties binnen de correctiegroepen na de volgende zes matrices $\{x_{ijk}\}$ toe:

1. $\begin{pmatrix} 110 & 000 \\ 111 & 110 \end{pmatrix}$
2. $\begin{pmatrix} 110 & 010 \\ 111 & 100 \end{pmatrix}$
3. $\begin{pmatrix} 110 & 110 \\ 111 & 000 \end{pmatrix}$
4. $\begin{pmatrix} 111 & 000 \\ 110 & 110 \end{pmatrix}$
5. $\begin{pmatrix} 111 & 010 \\ 110 & 100 \end{pmatrix}$
6. $\begin{pmatrix} 111 & 110 \\ 110 & 000 \end{pmatrix}$

Het conditioneren op $\{B_s\}$ houdt in dat (x_{+1+}, x_{+2+}) gelijk moet zijn aan (3,4) of (4,3). Hieraan voldoen matrices nr. 2,3,4 en 5. Het conditioneren op $\{A_s\}$ houdt in dat de vector $\{x_{+jk}\}$ een permutatie van (1,1,2,3) dient te zijn. Dit beperkt de toegestane matrices verder tot nrs. 2 en 5. Voor elk van deze matrices is het aantal verschillende permutaties van kolommen (beoordelaars) binnen correctiegroepen gelijk aan $3.3! = 18$. Er zijn aldus $n_{3.ABC} = 18+18 = 36$ verschillende vectoren $\{x_{ijk}\}$ met de gegeven waarde van $\{A_s\}$, $\{B_s\}$, $\{C_s\}$ en met $V = 3$.

Op een soortgelijke wijze vindt men waarden van $n_{v.ABC}$ zoals gegeven in de tabel. Uit deze tabel volgt aldus $P = P(V \geq 3 | H_0) = 36/144 = 0,25$.

4. Volledig experiment

Het in hoofdstuk 1 aangegeven experiment werd in feite uitgevoerd met twee soorten beoordelaars en twee soorten opstellen. Het complete experiment kan als een variantie-analyse model met de volgende vijf indelingen worden geclassificeerd:

- A: beoordelaars ($i=1, \dots, n$; $n=24$; nested in $C \times E$);
- B: opstellen ($j=1, \dots, m$; $m=7$; nested in D);
- C: opstelcorrectie ($k=1, 2$; ja-neen);
- D: opstelsoort ($u=1, 2$; laag-hoog);
- E: beoordelaarsoort ($v=1, 2$; LK-PA).

De eerste twee indelingen introduceren toevallige effecten, de overige indelingen vaste effecten. Een binaire waarneming is nu x_{ijkuv} en een mogelijk model is te schrijven als

$$P(\{x_{ijkuv}\} = \{x_{ijkuv}\}) = C \cdot \exp \{\text{som van een aantal termen}\},$$

waarbij C een normaliserende constante is en de in het model opgenomen termen zijn:

1. $\sum_u \sum_j \sum_k \sum_{s=0}^{2n} a_s \cdot \delta(s, x_{+jku+})$, $\{a_s\}$ zijn parameters die de afhankelijkheid tussen $2n$ beoordelaars -die een identiek opstel zien- modelleren;
2. $\sum_u \sum_j \sum_v \sum_{s=0}^{4n} b_s \cdot \delta(s, x_{+j+uv})$, $\{b_s\}$ zijn parameters die de afhankelijkheid tussen $4n$ beoordelaars -die een op de correctie na identiek opstel zien- modelleren;
3. $\sum_k \sum_v \sum_i \sum_{s=0}^{2m} c_s \cdot \delta(s, x_{i+k+v})$, $\{c_s\}$ zijn parameters die de afhankelijkheid tussen $2m$ door dezelfde beoordelaar opstellen modelleren;
4. $\alpha(x_{++1++} - x_{++2++})$, α is de correctie-effectparameter;
5. $\beta_1(x_{+++1+} - x_{+++2+})$, β_1 is de opstelsoorteffectparameter;
6. $\beta_2(x_{++++1} - x_{++++2})$, β_2 is de parameter die het verschil tussen twee soorten beoordelaars aangeeft;
7. $\beta_{12}(x_{+++11} - x_{+++12} - x_{+++21} + x_{+++22})$,
 β_{12} modelleert de opstelsoort-beoordelaarsoort interactie.

De toets van $H_0: \alpha=0$ wordt nu gebaseerd op de voorwaardelijke kansverdeling van $x_{++1++} - x_{++2++}$ gegeven

$$\{\sum_u \sum_j \sum_k \delta(s, x_{+jku+})\}, \{\sum_u \sum_j \sum_v \delta(s, x_{+j+uv})\},$$

$$\{\sum_k \sum_v \sum_i \delta(s, x_{i+k+v})\}, \{x_{+++uv}\} \text{ en } \alpha = 0.$$

Ook kan men bv. de hypothese van geen $D \times E$ interactie (het vaste opstel-effect is gelijk voor beide beoordelaarsoorten), $\beta_{12} = 0$, toetsen.

Merk op dat in het geval dat de vaste effecten D en E en hun interactie afwezig zijn, $\beta_1 = \beta_2 = \beta_{12} = 0$, de toets voor $\alpha = 0$ dezelfde is als de in hoofdstuk 2 afgeleide toets.

In het experiment probeerde elke beoordelaar m opstellen van elk

(hoog, laag) milieu te herkennen. Indien van tevoren goed geïnstrueerd, had de beoordelaar er voor kunnen zorgen dat hij m opstellen als laag en evenveel als hoog classificeerde. In zo'n situatie zou $x_{+++1+} - x_{+++2+}$ gelijk aan nul zijn en dus zou men de storende parameter β_1 door het conditioneren in de opzetfase, in plaats van na het experiment, elimineren. Hetgeen gezien de vraagstelling te verkiezen is (zie in dit verband de klassieke "tea tasting lady" van Fisher (1960)).

5. Conclusies en discussie

De toetsen uit hoofdstukken 2 en 4 zijn mogelijke oplossingen voor het gestelde toetsingsprobleem. De conclusies van deze toetsen hebben betrekking op de populatie van beoordelaars en op de populatie van opstellen, dit in tegenstelling tot toetsen uitgevoerd per opstel of over beoordelaars zoals in Molenaar (1982).

Een aantal varianten van het gebruikte kansmodel is mogelijk, zoals een model met een eenvoudiger afhankelijkheidsstructuur (merk op dat in de gebruikte modelversie het aantal afhankelijkheidsparameters van n en m afhangt) of een andere schaal voor het uitdrukken van het effect van de opstelcorrectie.

In een wat complexer experiment, zoals in het onderhavige geval, is het niet op het eerste gezicht duidelijk welke permutationale toets op zijn plaats is. Het construeren van een kansmodel (in het onderhavige geval een model uit de klasse van gemengde modellen voor binaire gegevens zoals geïntroduceerd door Fidler en Nagelkerke (1982)) kan helpen zo'n toets te vinden.

Nawoord

Mijn dank aan Professor Dr.Ir. L. de Pater en aan Drs. N.J.D. Nagelkerke voor het kritisch lezen van het manuscript en voor hun commentaar.

Referenties

- Fidler, V. en Nagelkerke, N.J.D. (1982). Mixed model for binary observations. KM 8 (1982) p. 50-57
- Fisher, R.A. (1960). The design of experiments. Oliver and Boyd. Zevende editie.
- Molenaar, I.W. (1982). Opstelbeoordeling naar sociaal milieu: een dubbele generalisering, KM 6 (1982) p. 2-11