

## Tadeusz SOZAŃSKI

### PRZYKŁAD BUDOWY SKALI GUTTMANA

Materiał ilustracyjny do ćwiczeń z "Technik badań socjologicznych"  
r. ak. 1999/2000, aktualizacja: r.ak. 2003/2004

W notatce wykorzystane zostaną dane z badań Hany Šlechtovej (studentki z Brna) wykonanych w 1997 roku w Krakowie (patrz rozdział na zajęciach kwestionariusz "Czy aborcja powinna być legalna?") na populacji 287 studentów medycyny, prawa i socjologii. W bazie przygotowanej do ćwiczeń niektóre braki odpowiedzi uzupełniono odpowiedziami zgodnymi z modelem Guttmana (postąpiono tak w 14 przypadkach, w których respondent nie udzielił odpowiedzi na tylko jedną z 10 pozycji), a niektóre przypadki odrzucono (19 respondentów, którzy nie odpowiedzieli na co najmniej 2 pozycje). Ostatecznie do analizy zakwalifikowano 268 przypadków.

Pozycjom przypisano zmienne P1\_1,...,P1\_10, sprowadzone do postaci zerojedynkowej przez zastąpienie wartości 2 (odpowiedź "Nie" na pytanie o dopuszczalność aborcji w danej sytuacji) liczbą 0. Częstości zaobserwowania wartości (odpowiedzi "Tak") wyrażone w procentach (100%=268) dla 10 zmiennych wyglądają następująco:

P1_1	'Zagrozenie zycia'	89.9
P1_2	'Zagrozenie zdrowia'	81.3
P1_5	'Ciaza wynikiem gwaltu'	75.7
P1_4	'Uposledzenie psychiczne'	60.4
P1_3	'Uposledzenie fizyczne'	57.8
P1_6	'Sytuacja materialna'	42.2
P1_10	'Odrzucenie macierzynstwa'	26.9
P1_8	'Brak akceptacji bliskich'	25.4
P1_7	'Sytuacja psychologiczna'	19.8
P1_9	'Sytuacja zawodowa'	14.6

Dla każdej pary zmiennych (par takich jest 45) sprawdzono istotność różnicy między dwoma częstościami za pomocą testu dla dwu wartości średnich w jednej próbie. Różnice między zmiennymi P1\_3 i P1\_4 oraz między zmiennymi P1\_8 i P1\_10 okazały się nieistotne (p odpowiednio równe .13 i .53). Dla tych dwu ostatnich pozycji także analiza łącznego rozkładu reakcji ujawnia brak "pola zerowego", a więc nie można stwierdzić, że jedna z tych pozycji poprzedza drugą na kontinuum cechy. Dla ilustracji sytuacji diametralnie przeciwnej zamieszczam obok drugi przykład analizy zależności dwu pozycji (w obu przypadkach zastosowano procentowanie do liczebności całej próby).

		P1_10		P1_9	
		0	1	0	1
P1_8	0	178	22	27	27
		66.4	8.2	10.1	10.1
	1	18	50	202	39
		6.7	18.7	75.4	14.6
		196	72	229	39
		73.1	26.9	85.4	14.6
		268	100.0	268	100.0

Pozycje P1\_1 i P1\_9 leżą na przeciwnych końcach "continuum" postawy wobec aborcji, nic więc dziwnego, że mamy tu do czynienia z idealnie zerowym polem. Ktoś kto potwierdza pozycję P1\_9 (przyzwala na aborcję ze względu na sytuację zawodową), musi oczywiście potwierdzić także pozycję P1\_1 (dopuszczać aborcję w sytuacji zagrożenia życia). P1\_9=1 implikuje zatem P1\_1=1 i przypadek P1\_9=1 wraz z P1\_1=0 nie może wystąpić.

Ostatecznie do konstrukcji skali użyto 8 pozycji, oznaczonych X1,...,X8 i ponumerowanych według malejącej częstości ich potwierdzania.

X1	'Zagrozenie zycia'	89.9	P1_1
X2	'Zagrozenie zdrowia'	81.3	P1_2
X3	'Ciaza wynikiem gwaltu'	75.7	P1_5
X4	'Uposledzenie plodu'	63.1	P1_3 ∨ P1_4
X5	'Sytuacja materialna'	42.2	P1_6
X6	'Ciaza nieakceptowana'	33.6	P1_8 ∨ P1_10
X7	'Sytuacja psychologiczna'	19.8	P1_7
X8	'Sytuacja zawodowa'	14.6	P1_9

Zmienne zerojedynkowe X1, X2, X3, X5, X7, X8 to pozostawione bez zmian pozycje z pierwotnej listy. Zmienna X4 jest *sumą boolowską* zmiennych P1\_3 i P1\_4; jest to zmienna zerojedynkowa, która przyjmuje wartość 1 wtedy i tylko wtedy, gdy przynajmniej jedna z dwu zmiennych przyjęła wartość 1 (respondent odpowiedział "tak" przynajmniej na jedno z dwu pytań). Podobnie zmienna X6 jest sumą boolowską zmiennych P1\_8 i P1\_10.

Utwórzmy sumę arytmetyczną 8 zmiennych, czyli zmienną  $X = X1 + X2 + X3 + X4 + X5 + X6 + X7 + X8$ , przyjmującą wartości od 0 do 8. A oto jej rozkład w naszej próbie.

0	20	7.5	=====
1	16	6.0	=====
2	30	11.2	=====
3	37	13.8	=====
4	55	20.5	=====
5	28	10.4	=====
6	32	11.9	=====
7	13	4.9	=====
8	37	13.8	=====
---	-----		
268	100.0		

Jest on dość symetryczny, daleki jednak od normalności ze względu na widoczną tendencję do polaryzacji. Obok najliczniejszego środka (X=4 - 20.5%), który tworzą zwolennicy obecnej regulacji prawnej, silne są umiarkowane skrzydła: proaborcyjne (X=5,6 -razem 22.3%) i antyaborcyjne (X=2,3 - razem 25%). Dość liczni są też "ekstremiści" po obu stronach: radykalni proaborcjonisci (X=7,8 - 18.7%) i radykalni antyaborcjonisci (X=0,1 - 13.5%).

Im wyższa wartość X, tym szerszy zakres sytuacji, w których wedle respondenta powinna być dopuszczalna aborcja. Wartość 8 oznacza uznanie nieograniczonego żadnymi warunkami prawa do aborcji, wartość 0 oznacza akceptację bezwarunkowego zakazu przerywania ciąży. Jeśli 8 pozycji tworzy układ Guttmanowski, wówczas z wartości zmiennej X można jednoznacznie odtworzyć konfigurację odpowiedzi na 8 pozycji. W 8-wymiarowym systemie dychotomicznym liczba wszystkich możliwych konfiguracji jest równa  $2^8 = 256$ . 9 spośród nich jest zgodnych z modelem Guttmana. Są to konfiguracje:

X=8	11111111	X=3	11100000
X=7	11111110	X=2	11000000
X=6	11111100	X=1	10000000
X=5	11111000	X=0	00000000
X=4	11110000		

W próbie konfiguracje te wystąpiły w 225 przypadkach, co stanowi 85% populacji. A oto pełne zestawienie wszystkich zaobserwowanych 31 konfiguracji.

$i$	$n_i$	X	X1	X2	X3	X4	X5	X6	X7	X8	$e_i$	$e_i n_i$
1	37	8	1	1	1	1	1	1	1	<u>1</u>		
2	11	7	1	1	1	1	1	1	<u>1</u>	0		
3	2	7	1	1	1	1	1	1	0	<b>1</b>	1	2
4	30	6	1	1	1	1	1	1	0	0		
5	1	6	1	1	1	1	<b>0</b>	<u>1</u>	<b>1</b>	0	1	1
6	1	6	1	1	1	1	1	0	<b>1</b>	0	1	1
7	21	5	1	1	1	1	1	0	0	0		
8	2	5	1	1	1	<b>0</b>	1	<b>1</b>	0	0	1	2
9	1	5	1	1	1	<b>0</b>	1	0	<b>1</b>	0	2	2
12	1	5	<b>0</b>	1	1	1	<u>1</u>	<b>1</b>	0	0	1	1
10	2	5	1	1	1	1	0	<b>1</b>	0	0	1	2
11	1	5	1	<b>0</b>	1	1	0	<b>1</b>	<b>1</b>	0	2	2
13	46	4	1	1	1	1	0	0	0	0		
14	2	4	1	<b>0</b>	1	1	<b>1</b>	0	0	0	1	2
15	1	4	<b>0</b>	1	1	<u>1</u>	0	<b>1</b>	0	0	2	2
16	3	4	1	1	1	0	<b>1</b>	0	0	0	1	3
17	2	4	1	1	1	0	0	<b>1</b>	0	0	1	2
18	1	4	1	1	1	0	0	0	<b>1</b>	0	1	1
19	26	3	1	1	1	0	0	0	0	0		
20	3	3	1	<b>0</b>	<u>1</u>	<b>1</b>	0	0	0	0	1	3
21	7	3	1	1	0	<b>1</b>	0	0	0	0	1	7
22	1	3	1	1	0	0	<b>1</b>	0	0	0	1	1
23	21	2	1	<u>1</u>	0	0	0	0	0	0		
24	6	2	1	0	<b>1</b>	0	0	0	0	0	1	6
25	1	2	1	0	0	<b>1</b>	0	0	0	0	1	1
26	1	2	<b>0</b>	0	<b>1</b>	<b>1</b>	0	0	0	0	2	2
27	1	2	<b>0</b>	0	0	<b>1</b>	<b>1</b>	0	0	0	2	2
28	13	1	<u>1</u>	0	0	0	0	0	0	0		
29	1	1	0	<b>1</b>	0	0	0	0	0	0	1	1
30	2	1	0	0	<b>1</b>	0	0	0	0	0	1	2
31	20	0	0	0	0	0	0	0	0	0		

—  
48

$$4 \quad 7 \quad 9 \quad 16 \quad 8 \quad 9 \quad 5 \quad 2 = 60$$

Dopasowanie modelu okaże się jeszcze lepsze, gdy weźmiemy pod uwagę liczbę błędnych reakcji w konfiguracjach niezgodnych z modelem. Liczba błędów w takiej konfiguracji w naszych danych jest równa 1 lub 2. Rozważmy np. konfigurację nr 9: 11101010. Aby przekształcić ją w jedną z konfiguracji modelowych (11111000, 11100000, 11111110) potrzebne są zmiany w dwu miejscach: X4 i X7, X5 i X7 lub X4 i X6 (pierwszy sposób poprawienia błędu wydaje się najlepszy, bo nie zmienia wartości X). Mnożąc liczbę błędów w danej konfiguracji ( $e_i$ ) przez liczbę powtórzeń tej konfiguracji w próbie

( $n_j$ ), a następnie dodając iloczyny dostajemy całkowitą liczbę błędów, równą 48. Całkowita liczba reakcji jest równa  $8 \cdot 268 = 2144$ , reakcje błędne stanowią około 2% tej liczby, a stąd współczynnik odtwarzalności jest równy .98.

Do przedstawionych wyżej danych zastosowano także metodę identyfikacji błędów zwaną "techniką Cornell" (patrz Meinz, Holm, Huebner 1985, ss. 75-81). Jako błąd (zaznaczony wytłuszczeniem) traktujemy wtedy każdą reakcję 0 nad linią podziału (podkreślenie) i każdą reakcję 1 pod tą linią wstawioną w takim miejscu, by liczba zidentyfikowanych w ten sposób reakcji niezgodnych z modelem była minimalna. Metoda ta daje nieco większą liczbę błędów (60), ale umożliwia jednoznaczne "poprawienie" błędnej konfiguracji i ustalenie rozkładu błędów wg pozycji. W naszym przykładzie błędy wystąpiły najliczniej dla pozycji 4 (16 przypadków). Widać wtedy także, które konfiguracje są najbardziej nietypowe (dużo błędów w wierszu, w przykładzie po 3 błędy mają konfiguracje nr 11, 26 i 27).

System dychotomiczny można jeszcze bardziej uprościć, wprowadzając 5 zmiennych określonych następująco:  $Y1=X1 \vee X2$  ('zagrożenie życia lub zdrowia'),  $Y2=X3$  ('gwałt'),  $Y3=X4$  ('upośledzenie płodu'),  $Y4=X5 \vee X6$  ('ważne powody społeczne'),  $Y5=X7 \vee X8$  ('inne powody'). Częstości potwierdzania tych pozycji będą wówczas odpowiednio równe: 91.0, 75.7, 63.1, 44.8, 20.5. Jeśli pominąć zmienne  $Y3$  i  $Y5$ , są to wyniki dość podobne do podanych przez E. Babbie, który w swym znanym podręczniku (*Badania społeczne w praktyce*, tł. polskie, 2003, s. 194) rozważa podobny przykład. Trzy uwzględnione przez niego uzasadnienia prawa do aborcji, "poważne zagrożenie zdrowia kobiety", "gwałt" i "ciąża pozamałżeńska", w liczącej 1231 jednostek próbie ogóln amerykańskiej zostały potwierdzone odpowiednio przez 92, 86 i 48% badanych, zaś liczba przypadków niezgodnych z modelem Guttmana dla trzech pozycji wyniosła tylko 27.

W badaniach pani Hani dla określonych wyżej pięciu pozycji otrzymano następujący rozkład konfiguracji:

i	$n_i$	Y	Y1	Y2	Y3	Y4	Y5
1	53	5	1	1	1	1	1
2	57	4	1	1	1	1	0
3	1	4	1	1	0	1	1
4	49	3	1	1	1	0	0
5	7	3	1	1	0	1	0
6	1	3	1	1	0	0	1
7	32	2	1	1	0	0	0
8	8	2	1	0	1	0	0
9	1	2	1	0	0	1	0
10	1	2	0	1	1	0	0
11	1	2	0	0	1	1	0
12	35	1	1	0	0	0	0
13	2	1	0	1	0	0	0
14	20	0	0	0	0	0	0

Ćwiczenie: przeprowadzić samodzielnie analizę dopasowania modelu Guttmana do powyższych danych, zastosować obie metody i obliczyć współczynnik odtwarzalności.

