A CONTINGENCY TABLE MODELING APPROACH TO USING PARITY PROGRESSION RATIOS IN THE ANALYSIS OF THE NUMBER OF CHILDREN EVER BORN

Lester R. Curtin, I. Joel Frockt, and Gary G. Koch
University of North Carolina at Chapel Hill

1. Introduction

Recently, Namboodiri (1972,1974), Simon (1975a,1975b), and others have empirically studied fertility as a sequential decision-making process. It has been observed that a number of factors may have different effects at different birth orders. In particular, income may have both positive and negative effects, and the net effect of income on mean fertility or total fertility would then depend on the parity composition of the sample. For this reason, analysis of the effects of several factors may be obscured if the dependent variable is chosen to be some aggregate variable such as the mean number of children ever born (CEB).

Data on the number of CEB can be used to study the sequential thesis of fertility. Analytical methods which can be employed include regression analysis and discriminant function analysis. A contingency table modeling approach can also be used on the same type of data. This paper indicates how weighted least squares fitting procedures can be used to analyze parity progression ratios derived from the number of CEB. An example looks at the effects of religion, education, and income on the parity progression ratios.

2. Data

The data for the example are from the 1965 National Fertility Study (Princeton). The twelve sub-populations of white women are based on religion (Catholic, non-Catholic), education (less than high school, high school or better), and husband's income (low, medium, high). All women were in their first marriage and have had fifteen years or more exposure to the risk of pregnancy since marriage. As such, these women have effectively completed their childbearing experience.

3. Analysis

Data on the number of CEB for cohorts of women who have reached the end of their childbearing experience can be formulated in terms of a contingency table. Let n_{ij} be the number of women in the i-th sub-population $(i=1,2,\ldots,s)$ with parity j $(j=0,1,\ldots,r)$. The proportion of women in the i-th sub-population with parity j is

$p_{ij} = n_{ij}/n_i$, where $n_i = \sum_{j=0}^{r} n_{ij}$, and an unbiased estimate of the mean number of CEB is

$\hat{\mu}_i = \sum_{j=1}^{r} j\, p_{ij}$. An algebraically equivalent estimate can be derived from the parity progression ratios.

The parity progression ratios are defined for women in the i-th sub-population as

$a_{i0} = p_{i1}+p_{i2}+\ldots+p_{ir}$,

$a_{i1} = (p_{i2}+p_{i3}+\ldots+p_{ir})/(p_{i1}+p_{i2}+\ldots+p_{ir})$, ...,

$a_{ik} = (p_{ik+1}+\ldots+p_{ir})/(p_{ik}+p_{ik+1}+\ldots+p_{ir})$, ...,

$a_{ir-1} = (p_{ir})/(p_{ir-1}+p_{ir})$. Under the assumption that fertility has been constant, the k-th parity progression ratio for the i-th sub-population, $a_{ik}(k=0,1,\ldots,r-1)$, can be interpreted as the conditional probability of having a (k+1)-th birth. It can be seen that $a_{i0}a_{i1} = p_{i2}+\ldots+p_{ir}$, until $a_{i0}a_{i1},\ldots a_{ir-1} = p_{ir}$. Thus

$a_{i0}+a_{i0}a_{i1}+\ldots+a_{i0}a_{i1}\ldots a_{ir-1} = \sum_j jp_{ij} = \hat{\mu}_i$,

the estimate of the mean number of CEB. A closed form expression for the variance of a_{ik} is not readily available, but it has been asserted that $\mathrm{cov}(a_{ik},a_{ik'}) = 0,\ k \neq k'$.

The compound function formulation of Forthofer and Koch (1974) can be used to calculate the parity progression ratios and to obtain estimates of their covariance structure. The GSK weighted least squares method outlined in Grizzle, Starmer, Koch (1969) can now be applied to several models for the parity progression ratios. One strategy is to fit an incremental model to the parity classes within each of the sub-populations. A second strategy is to fit a factorial model across sub-populations for each parity class. The two models can be combined to account for both the variation across parities and the variation across sub-populations. For this method, an incremental model is fit across parities to the differences within the effects in the factorial model. For the data in Table 1, with the twelve sub-populations, this model has the parameterization displayed in Table 2.

The model is then reduced. The estimated parameters for the reduced model are reported in Table 4. The goodness of fit for the reduced model is summarized in Table 5. A more complete documentation of the analysis is given in Curtin et al (1976).

4. Discussion

The data used in the example are somewhat limited in scope but some interesting results can be noted by examining the predicted values in Table 6 which were obtained via the weighted least squares procedures. In general, the probability of an additional birth is greater for Catholics than for non-Catholics. However, for high parities, the low income, low education, non-Catholics have a greater probability of an additional birth than the corresponding sub-population of Catholics.

Education has no effect for the 0→1 parity transition and a positive effect for the 1→2 transition. Here a positive effect means that increased education is related to an increased probability of an additional birth. For the transition 2→3, and for higher parity transitions, education has a negative effect for all sub-populations except for the Catholic-high income sub-population which has a positive education effect for the transitions 1→2, 2→3, and 3→4 and a negative effect for the transitions 4→5 and 5→6. The education effect becomes more negative (less positive) as parity increases.

At all parities and sub-populations, income has a negative effect; that is, additional income decreases the probability of an additional birth, with the exception of low education Catholics at high parities. For this group, the probability of an additional birth decreases as income goes from low to medium, but then increases greatly for the high income sub-population at parities four or five. The income effects for non-Catholics increase as parity increases. No such generalization can be made for Catholics as there is substantial fluctuation across parities for the income effects.

ACKNOWLEDGMENTS

This research was in part supported by the U. S. Bureau of the Census through Joint Statistical Agreement JSA 76-93. The authors would like to thank Bea Parker for her conscientious typing of this manuscript.

REFERENCES

Bernhardt, Eva M. (1972). Fertility and economic status--Some recent findings on differentials in Sweden. *Population Studies 26*, 175-184.

Curtin, L.R., Frockt, I.J., and Koch, G.G. (1976). On the mixed effects of income and education on fertility. Submitted to *Demography*.

Grizzle, J.E., Starmer, C.F., Koch, G.G. (1969). Analysis of categorical data by linear models. *Biometrics 25*, 489-504.

Forthofer, R.N. and Koch, G.G. (1973). An analysis for compound functions of categorical data. *Biometrics 29*, 143-157.

Namboodiri, K. (1972). Some observations on the economic framework for fertility analysis. *Population Studies 26*, 185-206.

Namboodiri, K. (1974). Which couples at given parities expect to have additional births. *Demography 11*, 45-56.

Seiver, D.A. (1976). Which couples at given parities expect to have additional births. Paper presented at the annual meeting of the Population Association of America, Montreal, Canada.

Simon, Julian (1975a). The mixed effects of income upon successive births may explain the convergence phenomenon. *Population Studies 29*, 109-122.

Simon, Julian (1975b). Puzzles and further explorations in the interrelationships of successive births with husband's income, spouses' education, and race. *Demography 12*, 259-274.

TABLE 1

PARITY DISTRIBUTION BY RELIGION, EDUCATION, AND HUSBAND'S INCOME

Sub-population			Parity 0	1	2	3	4	5	6 or more	Total
Cath	less	low	3	6	19	11	17	10	18	84
Cath	less	med	5	15	12	18	21	8	12	91
Cath	less	high	2	2	7	4	2	1	7	25
Cath	more	low	2	2	6	6	9	3	6	34
Cath	more	med	4	2	34	22	15	15	13	105
Cath	more	high	7	18	30	17	20	9	13	114
Non	less	low	23	36	63	49	49	29	161	410
Non	less	med	15	34	63	52	28	24	35	251
Non	less	high	4	11	19	20	7	10	7	78
Non	more	low	14	12	34	31	17	11	23	142
Non	more	med	20	34	100	69	48	19	12	302
Non	more	high	33	47	108	99	50	20	12	369

TABLE 2

PARAMETERIZATION OF THE INCREMENTAL-FACTORIAL MODEL
WITHIN EACH PARITY CLASS

Source of Variation	Estimated Incremental Parameter	Indicator Variable
Mean	b_1	$x_1 = 1$ always
Main effect: Religion (R)	b_2	$x_2 = \begin{cases} 1 & \text{Catholic} \\ -1 & \text{Non-Catholic} \end{cases}$
Main effect: Education (E)	b_3	$x_3 = \begin{cases} 1 & \text{less than high school} \\ -1 & \text{high school of better} \end{cases}$
Main effect: Income (I)	b_4, b_5	$x_4 = \begin{cases} 1 & \text{low} \\ 0 & \text{medium} \\ -1 & \text{high} \end{cases}$, $x_5 = \begin{cases} 0 & \text{low} \\ 1 & \text{medium} \\ -1 & \text{high} \end{cases}$
Interaction: R × E	b_6	$x_6 = x_2 x_3$
Interaction: R × I	b_7, b_8	$x_7 = x_2 x_4$, $x_8 = x_2 x_5$
Interaction: E × I	b_9, b_{10}	$x_9 = x_3 x_4$, $x_{10} = x_3 x_5$
Interaction: R × E × I	b_{11}, b_{12}	$x_{11} = x_2 x_3 x_4$, $x_{12} = x_2 x_3 x_5$

TABLE 3

OBSERVED PARITY PROGRESSION RATIOS AND MEAN NUMBER CEB
WITH THEIR STANDARD ERRORS

Sub-population			Parity 0	1	2	3	4	5	Not Truncated Mean	Truncated Mean
Cath	less	low	0.964 (0.020)	0.926 (0.029)	0.747 (0.050)	0.804 (0.053)	0.622 (0.072)	0.643 (0.091)	3.929 (0.252)	3.607 (0.192)
Cath	less	med	0.945 (0.024)	0.826 (0.041)	0.831 (0.044)	0.695 (0.060)	0.488 (0.078)	0.600 (0.110)	3.264 (0.204)	3.176 (0.183)
Cath	less	high	0.920 (0.054)	0.913 (0.059)	0.667 (0.103)	0.714 (0.121)	0.800 (0.126)	0.875 (0.117)	3.840 (0.599)	3.320 (0.407)
Cath	more	low	0.942 (0.040)	0.938 (0.043)	0.800 (0.073)	0.750 (0.088)	0.500 (0.118)	0.667 (0.157)	3.618 (0.330)	3.500 (0.296)
Cath	more	med	0.962 (0.019)	0.980 (0.014)	0.656 (0.048)	0.661 (0.059)	0.652 (0.073)	0.464 (0.094)	3.375 (0.236)	3.324 (0.122)
Cath	more	high	0.939 (0.022)	0.832 (0.036)	0.663 (0.050)	0.712 (0.059)	0.524 (0.077)	0.591 (0.105)	3.051 (0.182)	2.912 (0.162)
Non	less	low	0.944 (0.011)	0.907 (0.015)	0.820 (0.020)	0.830 (0.022)	0.795 (0.026)	0.847 (0.026)	4.798 (0.161)	3.941 (0.100)
Non	less	med	0.940 (0.015)	0.856 (0.023)	0.688 (0.032)	0.626 (0.041)	0.678 (0.050)	0.593 (0.064)	3.457 (0.118)	3.020 (0.111)
Non	less	high	0.949 (0.025)	0.851 (0.041)	0.698 (0.052)	0.545 (0.075)	0.708 (0.093)	0.412 (0.119)	3.079 (0.244)	2.936 (0.186)
Non	more	low	0.901 (0.025)	0.906 (0.026)	0.707 (0.042)	0.622 (0.054)	0.667 (0.066)	0.676 (0.080)	3.190 (0.178)	3.056 (0.154)
Non	more	med	0.934 (0.014)	0.879 (0.019)	0.597 (0.031)	0.534 (0.041)	0.390 (0.055)	0.387 (0.087)	2.679 (0.087)	2.649 (0.082)
Non	more	high	0.911 (0.015)	0.860 (0.019)	0.626 (0.028)	0.453 (0.037)	0.390 (0.054)	0.375 (0.086)	2.542 (0.077)	2.526 (0.074)

TABLE 4

ESTIMATED INCREMENTAL-FACTORIAL PARAMETERS
WITH THEIR STANDARD ERRORS

Parameter	Parity 0	1	2	3	4	5
b_1	0.942 (0.005)	-0.050 (0.007)	-0.179 (0.013)	-0.050 (0.007)	-0.050 (0.007)	-0.016 (0.028)
b_2	0.006 (0.002)	0.006 (0.002)	0.006 (0.002)	0.033 (0.016)	-0.034 (0.021)	NS
b_3	NS	-0.009 (0.007)	0.045 (0.008)	NS	0.045 (0.008)	NS
b_4	NS	0.030 (0.005)	0.030 (0.005)	0.030 (0.005)	-0.030 (0.027)	0.056 (0.034)
b_5	NS	-0.010 (0.006)	-0.010 (0.006)	NS	-0.050 (0.025)	NS
b_6	NS	-0.009 (0.003)	-0.009 (0.003)	NS	-0.009 (0.003)	NS
b_7	NS	NS	NS	-0.066 (0.015)	NS	NS
b_8	NS	0.009 (0.006)	0.009 (0.006)	NS	-0.025 (0.024)	NS
b_9	NS	NS	NS	NS	-0.031 (0.020)	NS
b_{10}	NS	-0.036 (0.009)	0.051 (0.015)	NS	-0.036 (0.009)	NS
b_{11}	NS	NS	NS	NS	NS	NS
b_{12}	NS	-0.025 (0.009)	0.041 (0.015)	NS	-0.062 (0.023)	NS

TABLE 5

ANALYSIS OF VARIANCE FOR THE FINAL MODEL
FOR THE PARITY PROGRESSION RATIOS

Source of Variation	d.f.	x^2
Model	23	1198.80
Error	48	28.60

TABLE 6

PREDICTED PARITY PROGRESSION RATIOS AND MEAN NUMBER CEB
WITH THEIR STANDARD ERRORS

Sub-population			Parity 0	1	2	3	4	5	Mean
Cath	less	low	0.948 (0.006)	0.916 (0.013)	0.810 (0.020)	0.757 (0.030)	0.648 (0.043)	0.687 (0.046)	3.635 (0.146)
Cath	less	med	0.948 (0.006)	0.824 (0.023)	0.778 (0.030)	0.761 (0.032)	0.540 (0.051)	0.524 (0.057)	3.179 (0.152)
Cath	less	high	0.948 (0.006)	0.917 (0.026)	0.660 (0.042)	0.679 (0.051)	0.865 (0.070)	0.794 (0.075)	3.383 (0.263)
Cath	more	low	0.948 (0.006)	0.952 (0.013)	0.773 (0.021)	0.720 (0.033)	0.600 (0.054)	0.640 (0.059)	3.545 (0.158)
Cath	more	med	0.948 (0.006)	0.981 (0.013)	0.679 (0.030)	0.662 (0.033)	0.563 (0.048)	0.547 (0.053)	3.291 (0.127)
Cath	more	high	0.948 (0.006)	0.831 (0.022)	0.685 (0.031)	0.704 (0.037)	0.560 (0.055)	0.489 (0.066)	2.972 (0.138)
Non	less	low	0.935 (0.005)	0.909 (0.011)	0.808 (0.014)	0.822 (0.180)	0.798 (0.024)	0.838 (0.024)	3.867 (0.082)
Non	less	med	0.935 (0.005)	0.849 (0.016)	0.707 (0.022)	0.624 (0.024)	0.664 (0.037)	0.648 (0.041)	3.024 (0.097)
Non	less	high	0.935 (0.005)	0.878 (0.020)	0.727 (0.029)	0.546 (0.038)	0.644 (0.061)	0.572 (0.072)	3.009 (0.142)
Non	more	low	0.935 (0.005)	0.909 (0.012)	0.701 (0.017)	0.714 (0.023)	0.645 (0.043)	0.685 (0.046)	3.269 (0.097)
Non	more	med	0.935 (0.005)	0.869 (0.016)	0.601 (0.022)	0.517 (0.024)	0.397 (0.041)	0.381 (0.046)	2.628 (0.069)
Non	more	high	0.935 (0.005)	0.858 (0.015)	0.618 (0.021)	0.437 (0.026)	0.418 (0.044)	0.346 (0.057)	2.572 (0.062)