Special Issue on Family Demography, Canadian Studies in Population
Vol. 27(1), 2000, pp. 195-207

Modeling and Projection of Age and Sex Specific Marriage Rates

N. Lalu

Population Research Laboratory
Department of Sociology
University of Alberta
Edmonton, Alberta

Abstract

Age and sex specific marriage rates and crude marriage rates of Canada and Provinces from 1975 to 1995 are analyzed in this paper. Single value decomposition method is applied to the matrix of age specific rates for a 21 year period. It is found that two components (out of 13 orthogonal components) is sufficient to reproduce the observed values. This model could be used to project the age specific marriage rates. A regression analysis of the crude marriage rates for the provinces of Canada show that there is significant regional variation and Quebec stands out from the rest of the provinces. There is a linear trend of declining marriage rates over time. It is seen that the odds of getting married (as indicated by the crude rate) declined by about two percent per year in the last 20 years. It is suggested that single value decomposition should be under taken for provinces of Canada for the purpose of projecting the marriage rates for a longer period than five years.

Résumé

Le présent article analyse les taux bruts de nuptialité, ainsi que par sexe et par groupe d'âge, du Canada et des provinces (1975-1995). La méthode de décomposition à une seule valeur est appliquée à une matrice de taux par groupe d'âge sur une période de 21 ans. Elle révèle que deux composantes (sur 13 composantes orthogonales) suffisent à reproduire les valeurs observées. Ce modèle pourrait servir à projeter les taux de nuptialité par groupe d'âge. Une analyse de régression des taux bruts pour les provinces canadiennes montre une variation régionale significative et le Québec se démarque du reste du Canada. Il existe une tendance linéaire vers des taux fléchissants au fil du temps – les probabilités de mariage ayant baissé d'environ 2 p. 100 par an au cours des 20 dernières années. La méthode ci-dessus pourrait permettre de prévoir les taux de nuptialité des provinces canadiennes sur une période de plus de cinq ans.

Key words: marriage rates, single value decomposition, projection.

Introduction

Information on marriage rate by age and sex shows the extent to which people take part in family formation which in turn has great impact on population change. In the developed countries rates of marriage, divorce and widowhood have changed dramatically in the last 20 to 30 years. When marriage was universal, and divorces were rare, as observed in most traditional societies, it was not necessary to dwell on the future changes in marriage rates. But now marriage rates vary between countries as well as within the same country from one region to another. In order to chart the future course of marriage rates it is necessary to develop models that can adequately explain the variations in the past data. Keyfitz noted that (Keyfitz, 1985) understanding the past is the key to the success of any forecasting attempt. In this paper we have developed two mathematical models to represent the marriage rates and their change over time. Crude marriage rates of regions are modeled applying regression analysis to log odds of marrying and single value decomposition is used to model the age specific rates and their changes over time.

Age specific marriage rates and mortality rates are necessary to apply mutistate life table techniques in nuptiality analysis. Espenshade(1983), Nour and Suchindran (1984, 1985) Willekens *et al* (1982), Adams and Nagnur (1988), and others have used multistate life table techniques in nuptiality analysis. This type of analysis can give the average time spent in married state and the number of times a person visits married state. The projected marriage rates may be used to construct nuptiality tables for future years.

Data and Methodology

The data are taken from Statistics Canada (1991 and 1996). The age specific marriage rates from 1975 to 1995 were available in these publications for Canada and Provinces. There are 13 age groups starting at age group 20 and under to 75 and over with a class interval of five years. The marriage rates for males and females were considered separately for this study. For the first part of the analysis, the age specific marriage rates for Canada are utilized from the above mentioned publications. In the second part the crude marriage rate for the provinces for the same 21 year period from the above source is used.

The time series of age specific marriage rates for 13 age groups for a period of 21 years can be arranged as a matrix (rectangular array) of 13 rows and 21 columns. The matrix (say **R**) is decomposed into three matrices, the product of which will exactly reproduce the matrix R. The is done by the "Single Value Decomposition" in Matrix procedure in the Statistical Package for Social Sciences (SPSS). The procedure is similar to principal component analysis in factor analysis.

$$\mathbf{R} = \mathbf{U} . \mathbf{Q} . \mathbf{V} \qquad (1)$$

where **U** and **V** are square matrices of order 13 and 21 respectively and **Q** is a 13 by 21 matrix with non-zero values in the principal diagonal and zeros elsewhere. These diagonal values are similar to the eigen values in a factor analysis. The diagonal values are arranged in the descending order of magnitude. It should also be pointed out that the matrices **U** and **V** are "unitary" meaning that the transpose of each is its own inverse. The representation of **R** as **U Q V** can be used to decompose the matrix **R** into 13 independent components. Let us represent **U** by 13 column vectors α_1, α_2, .. α_{13} and **V** by 21 row vectors β_1, β_2, .. β_{21}. Let the non zero diagonal elements of **Q** be λ_1, λ_2 . . . λ_{13}. Then we can write the (see Keyfitz, 1968, p. 60 for similar decomposition of a projection matrix) single value decomposition as:

$$\mathbf{R} = \alpha_1 \lambda_1 \beta_1 + \alpha_2 \lambda_2 \beta_2 + \ldots + \alpha_{13} \lambda_{13} \beta_{13} \qquad (2)$$

It should noticed that $\lambda_1, \lambda_2, \ldots \lambda_{13}$ are scalars and therefore $\alpha_i \lambda_i \beta_i$ is equivalent to $\lambda_i \alpha_i \beta_i$. The product $\alpha_i \beta_i$ is a matrix product resulting in a 13 by 21 matrix. Equation (2) shows how the matrix **R** can be expressed as the sum of 13 components, which will exactly reproduce **R**. These 13 components are independent of one another (or orthogonal) and the relative importance depends on the relative values of λ's. If the first few components of (2) can reproduce **R**, we will have fewer parameters and will achieve a reduction in complexity of the data matrix **R**. The α's represent effects of age, while the β's represent the effect of period (year). We will choose the minimum number of components (here also, one may note the parallelism with factor analysis) that will explain most of the variation of **R** and use them to estimate the rates for a few selected years, and compare them with the observed rates to test the adequacy of the model.

Results.

In Table 1 the diagonal elements of the matrix **Q**, which show the importance and relative values of the 13 components are shown. As mentioned earlier, the analysis was carried out for males and females separately. The percentages and cumulative percentages show the relative adequacy of the components. The patterns for males and females are similar. The first component alone more than 80 percent of the variation.

Table 1
Diagonal Values Q Matrix Indicating the Relative Importance of each of the 13 Components of the Decomposition

Index of the Diagonal Element of Q Matrix (I)	Male			Female		
	λ_i	%	Cum %	λ_i	%	Cum %
1	836.2	81.2	81.2	856.4	88.7	88.7
2	95.4	9.3	90.5	77.7	8.0	96.8
3	58.6	5.7	96.2	10.5	1.1	97.9
4	13.6	1.3	97.5	7.5	0.8	98.6
5	9.0	0.9	98.4	3.8	0.4	99.0
6	4.5	0.4	98.8	2.5	0.3	99.3
7	2.7	0.3	99.1	1.8	0.2	99.5
8	2.5	0.2	99.3	1.6	0.2	99.7
9	2.3	0.2	99.6	1.2	0.1	99.8
10	1.6	0.2	99.7	0.8	0.1	99.9
11	1.3	0.1	99.8	0.6	0.1	99.9
12	1.1	0.1	99.9	0.4	0.0	100.0
13	0.6	0.1	100.0	0.2	0.0	100.0
Total	1029.2	100.0		965.1	100.0	

It is clear that two or three components are enough to reproduce the observed rates. Here we have taken only the first two α's, β's and λ's. In Table 2 the columns of **U** (α_1 and α_2) represent the age effect . The first column of **U** indicates the average age pattern of marriage rates which is prevalent through out the 21 year period. There is a sharp peak for males in the age group 25 to 29 and for females, the values for 20 to 24 and 25 to 29 are almost equal (.578 and .593 respectively). The second part (α_2) with its negative values mainly represent the peculiarities of older ages.

Table 2
The Age Effect Implied in the First two Components of the Decomposition (the first two columns (α_1 and α_2) of the U Matrix) in Separate Analysis of Male and Female Marriage Rates

	Age Effect			
Age	Male		Female	
	Component 1 α_1	Component 2 α_2	Component 1 α_1	Component 2 α_2
<20	0.028	0.078	0.120	0.430
20-24	0.375	0.473	0.578	0.694
25-29	0.581	0.255	0.593	-0.373
30-34	0.430	-0.827	0.402	-0.262
35-39	0.358	0.125	0.265	-0.187
40-44	0.270	0.028	0.187	-0.176
45-49	0.222	-0.037	0.139	-0.189
50-54	0.179	-0.044	0.093	-0.140
55-59	0.144	-0.017	0.058	-0.061
60-64	0.119	-0.015	0.037	-0.020
65-69	0.097	-0.015	0.024	-0.014
70-74	0.070	-0.013	0.013	-0.018
75 +	0.068	-0.026	0.008	-0.013

Table 3 shows the parameters corresponding to the time variable represented in the first two components (β_1 and β_2). The values decline over time, and are similar for males and females. This factor indicates the trend in the general decline marriage rates over time for both sexes. If we examine the actual rates (which is not shown here) there are some age groups where rates have gone up or at least have not declined as much as in most age groups. The second component of the time variable shows some negative and some positive values. The ups and downs of the marriage rates by age groups are implied in the second to 13 th rows of the **V** matrix. Of course, the relative importance of the

Table 3
The Year Effect Implied in the First two Components of the Decomposition (the first two rows of the V Matrix) in Separate Analysis of Male and Female Marriage Rates

	Male		Female	
Year	Component 1 β_1	Component 2 β_2	Component 1 β_1	Component 2 β_2
1975	0.317	0.125	0.311	0.410
1976	0.250	0.914	0.294	0.264
1977	0.297	0.020	0.289	0.231
1978	0.289	-0.004	0.281	0.192
1979	0.287	-0.033	0.278	0.154
1980	0.285	-0.075	0.277	0.096
1981	0.233	-0.024	0.223	0.095
1982	0.222	-0.046	0.214	0.067
1983	0.213	-0.079	0.206	-0.014
1984	0.209	-0.089	0.202	-0.065
1985	0.200	-0.087	0.196	-0.094
1986	0.185	-0.072	0.182	-0.106
1987	0.190	-0.110	0.188	-0.198
1988	0.189	-0.113	0.191	-0.221
1989	0.184	-0.102	0.188	-0.235
1990	0.175	-0.111	0.181	-0.259
1991	0.157	-0.105	0.163	-0.261
1992	0.146	-0.120	0.153	-0.281
1993	0.138	-0.119	0.145	-0.283
1994	0.135	-0.111	0.143	-0.281
1995	0.132	-0.113	0.139	-0.292

factors are very small indeed as indicated by the corresponding magnitude of the λ's.

Now the first two parts of the decomposition are used to estimate the rates by age and sex for 1994 and 1995. The estimated values are obtained by the equations:

$$\hat{R}_1 = \alpha_1 \lambda_1 \beta_1 \quad (3)$$

$$\hat{R}_2 = \alpha_1 \lambda_1 \beta_1 + \alpha_2 \lambda_1 \beta_2 \quad (4)$$

The predicted values and observed rates are given in Table 4. Extreme old ages and lower ages show more deviations than the rest. In Figure 1 the observed and predicted values for 1995 are plotted and shows the remarkable equality of the two. It is clear that two components are enough to reproduce the observed rates. To use these components as tools for projections it is necessary to do a time series analysis of the components representing the period. (Lee and Carter 1992).

Table 4
Predicted and Observed Marriage Rates (per 1000) for Males for 1994 and 1995

Age	Predicted Rate for Males using the First Component		Predicted Rate for Males using the First and Second Component		Observed Rate for Males	
	1994	1995	1994	1995	1994	1995
< 20	3.2	3.1	2.4	2.3	1.5	1.4
20-24	42.4	41.3	37.4	36.2	30.4	29.1
25-29	65.7	64.0	63.0	61.3	66.1	64.8
30-34	48.6	47.4	57.3	56.3	53.5	52.4
35-39	40.5	39.5	39.2	38.2	38.3	36.9
40-44	30.5	29.7	30.2	29.4	30.4	29.8
45-49	25.1	24.5	25.5	24.9	27.5	26.7
50-54	20.2	19.7	20.7	20.2	24.9	24.5
55-59	16.2	15.8	16.4	16.0	21.0	20.3
60-64	13.4	13.1	13.6	13.3	15.8	16.2
65-69	10.9	10.6	11.1	10.8	12.8	12.7
70-74	7.9	7.7	8.1	7.9	10.7	9.9
75 +	7.6	7.4	7.9	7.7	11.3	11.2

Table 5
Predicted and Observed Marriage Rates (per 1000) for Females for 1994 and 1995

Age	Predicted Rate for Females using the First Component		Predicted Rate for Females using the First and Second Component		Observed Rate for Females	
	1994	1995	1994	1995	1994	1995
<20	14.7	14.3	5.3	4.6	6.5	6.0
20-24	70.6	68.9	55.5	53.2	54.7	52.4
25-29	72.4	70.7	80.6	79.2	81.4	79.4
30-34	49.0	47.9	54.8	53.8	55.3	54.7
35-39	32.3	31.6	36.4	35.8	35.9	35.6
40-44	22.8	22.3	26.7	26.3	25.8	25.2
45-49	17.0	16.6	21.2	20.9	20.0	20.4
50-54	11.4	11.1	14.4	14.3	14.3	14.6
55-59	7.1	6.9	8.4	8.3	8.7	9.0
60-64	4.5	4.4	4.9	4.9	5.3	5.0
65-69	2.9	2.8	3.2	3.1	3.4	3.1
70-74	1.6	1.5	2.0	1.9	2.0	2.0
75 +	1.0	0.9	1.3	1.2	1.4	1.5

Analysis of Crude Rates for Canada and Provinces

The rates were transformed log-odds of getting married by the transformation:

$$y = Ln\frac{\text{Rate}}{(1-\text{Rate})}$$

At first a regression equation with time (year) as independent variable was fitted to assess the linear trend. The Multiple R is 0.654 and R-Square is 0.427, which is highly significant. As shown in the following Table 6 the regression coefficient is -.0189 showing a downward trend. If we take the e^B we get the value of 0.98 showing a 2 percent decline in the odds getting married per year in the period 1975 to 1995.

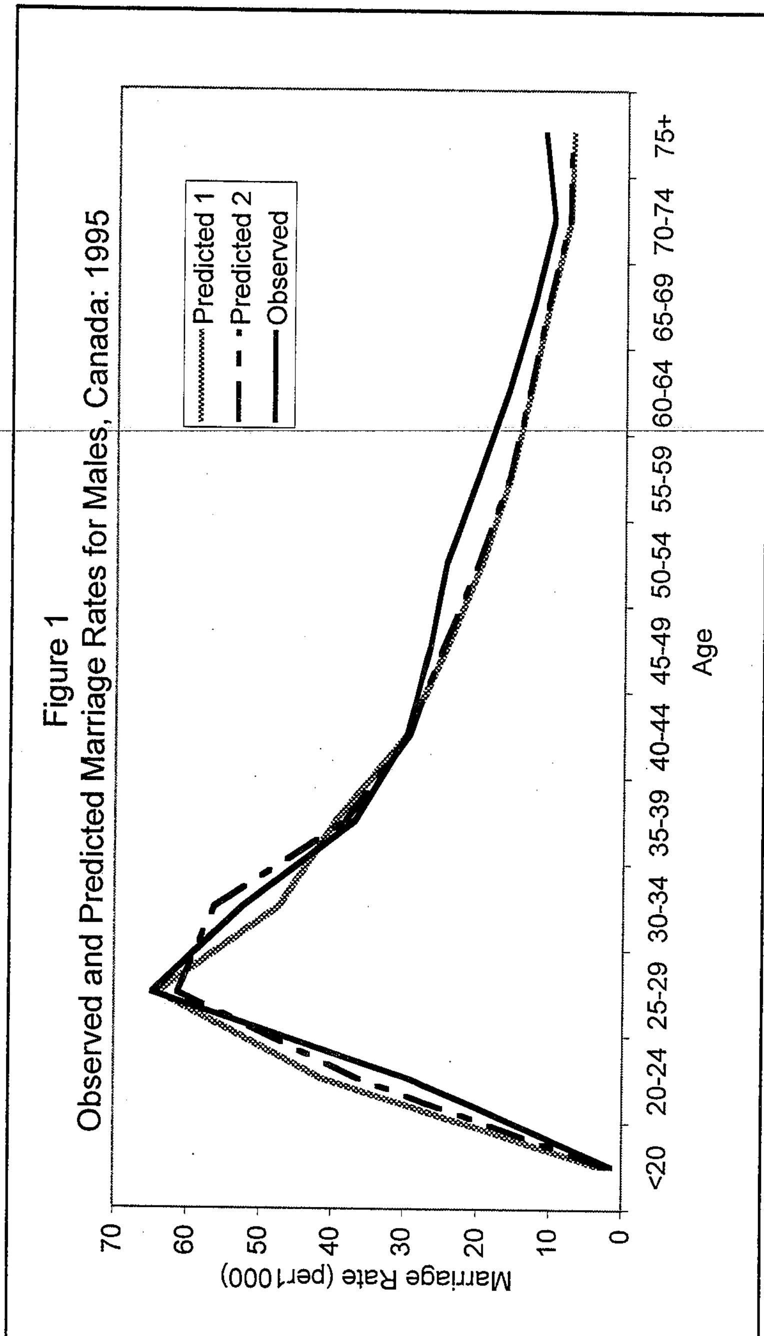

Figure 1
Observed and Predicted Marriage Rates for Males, Canada: 1995

Table 6
Regression Analysis Using Time as an Independent Variable

Variable	B	SEB	Beta	T	Sig T
YEAR	-.019	.001	-.654	-13.070	.0000
(Constant)	32.581	2.871		11.345	.0000

Now we fit a regression of log-odds on ten dummy variables created to represent each of the provinces with Canada as a whole as the reference category, Here the regression coefficients represent the deviation from the Canadian average. The multiple R in this case is .62883 and R Square is .395 with an F value of 14.389 which is highly significant. This means that there is considerable variation between provinces. Table 7 shows the results of the regression analysis.

Table 7
Regression Analysis Using Provinces as Independent Variables

Variable	B	SEB	Beta	T	Sig T
Nfld	-.096	.043	-.158	-2.238	.02
PEI	.044	.043	0.73	1.030	.30
NS	.038	.043	.063	.885	.37
NB	-.005	.043	-.008	-.117	.90
Quebec	-.277	.043	-.455	-6.446	.00
Ontario	.078	.043	.128	1.816	.07
Manitoba	.036	.043	.060	.849	.39
Saskatchewan	-.016	.043	-.026	-.372	.71
Alberta	.1615	.043	.265	3.753	.00
BC	.101	.043	.166	2.343	.02
(Constant)	-4.962	.030		-162.909	.00

Quebec, Alberta and B.C. are the provinces with large regression coefficients. In the case of Quebec the negative coefficient indicates lower that average marriage rate. Alberta and B. C. have positive coefficients showing higher than average rates. It is not surprising to find Quebec has significantly different pattern given the changes that have been taking place there since 1960.

Table 8 shows the results of a regression analysis with both "Provinces" and time (Year) entered as independent variables. The Multiple R is .907 and R Square is 0.82269 with F value of 92.374 which is significant. It is clear that the effect of the Provinces or that of time has not changed from the previous two regression analyses. It means that their effects are not confounded by one another. In view of the high multiple correlation already achieved it is clear that interaction between the variables (i.e. the set of dummy variables representing Provinces and the time variable) is not meaningful to be included or tested even though one may find statistically significant interaction between the two because of the large value of the degrees of freedom associated with the test.

Table 8
Regression Analysis Using Provinces and Time as Independent Variables

Variable	B	SEB	Beta	T	Sig T
Nfld	-.096	.023	-.158	-4.124	.0001
PEI	.044	.023	.072	1.898	.0590
NS	.038	.023	.062	1.631	.1043
NB	-.005	.023	-.008	-.215	.8297
Quebec	-.277	.023	-.455	-11.875	.0000
Ontario	.078	.023	.128	3.346	.0010
Manitoba	.036	.023	.060	1.565	.1191
Saskatchewan	-.016	.023	-.026	-.686	.4936
Alberta	.162	.023	.265	6.915	.0000
BC	.101	.023	.165	4.317	.0000
YEAR	-.018	.001	-.653	-22.972	.0000
(Constant)	-4.772	.0185		-258.429	.0000

The regression coefficient for time is virtually unchanged and is -.0189 and the exponentiated value is 0.98 showing a decline of 2 percent per years in the odds of marriage even after controlling for the variation between provinces. Quebec has shown a negative effect and so does Newfoundland, N.B., and Saskatchewan, even though to a lower degree. The Western provinces, in general, show higher rates. It should be noted that the negative coefficient of Saskatchewan is not significant.

Conclusion

This analysis shows that there is considerable heterogeneity among provinces and hence any meaningful modeling of marriage rate should include the variation among provinces. It may be enough to consider four regions of Canada- Atlantic region, Quebec, Central Canada and Western Provinces.

It was suggested by one of the reviewers that the estimation of the parameters in equations (3) and (4) should be made from the first 19 years of data and then the model should be tested with the observed data for the last two years. The parameters were recalculated for females with the marriage rates for 19 years. It was found that the λ's differed from the previous ones by less than five percent. The differences in the predicted values were only in the third decimal place. This shows that a reanalysis along the suggested lines will not alter the findings of this paper.

Projection of age specific marriage rates by age and sex for Canada as whole may not be appropriate. Even though there is no conclusive proof, it is probable that short term projections (up to five years) may not be very much off the mark if we simply took the nation as a whole. If it is desired to project marriage rate for longer period it imperative that we include the regional variations also into account.

References:

Adams, O. B. and D. N. Nagnur 1988. *Marriage, Divorce and Mortality: A Life Table Analysis for Canada and Regions, 1980-1982.* Statistics Canada (Catalogue 84-536E)

Cohen, J. E. 1985, "Mathematical Demography: Recent Developments in Population Projections," in International Population Conference, Florence, 1985, 5-12 June. Congress International de la Population.. Liege, Belgium, International Union for the Scientific Study of Population, 1985. Vol. 4: 179-189.

Espenshade, T. J. 1983. "Marriage, Divorce, and Remarriage from Retrospective Data: a Multiregional Approach," *Environmental Planning*, 15: 1633-1652.

Nour, E. S. and C. M. Suchindran, 1984. "The Contributions of Multistate Life tables: Comments on the article by Willekens *et. al.*," *Population Studies*, 38: 325-328.

Nour, E. S. and C. M. Suchindran, 1985. "Multistate Life Tables: Theory and Applications," in *Biostatistics: Statistics in Biomedical, Public health and Environmental Sciences*, edited by P. K.Sen. North Holland Publishers Amsterdam.

Keyfitz, Nathan 1985. *Applied Mathematical Demography*. Springer-Verlag, New York.

Keyfitz, Nathan 1968. *Introduction to the Mathematics of Population*. Addison-Wesley Publishing Company, Reading, Massachusetts.

Lee, R. D. and L.R. Carter. 1992 , "Modeling and Forecasting US Mortality," *Journal of the American Statistical Association*. 87: 659-671

Schopflocher, Donald and Yan Jin. 1998 "Life Expectancy as a Health Status Indicator: Alberta and its Health Regions 1986- 1996" Paper presented at the Kalbach Conference at the University of Alberta, March , 1998.

Smith, D.P. 1992. *Formal Demography*. Plenum Press New York

Statistics Canada 1996. *Marriages 1995*. Catalogue 84-212. Ottawa: Statistics Canada.

Statistics Canada 1991. *Marriages 1990*. Health Reports Supplement # 16, 3:4, Catalogue 82-003 S16. Ottawa: Statistics Canada.

Willekens, F. J., I. Sha, and P. Ramachandran. 1982 "Multistate Analysis of Marital Status Life Tables: Theory and Application," *Population Studies*. 36: 129-144.

Received February 1999; Revised November 1999