Ann. Inst. Statist. Math.
38 (1986), Part B, 353-370

BAYESIAN COHORT MODELS FOR GENERAL COHORT TABLE ANALYSES

TAKASHI NAKAMURA

(Received Oct. 12, 1982; revised Jan. 16, 1986)

Summary

New Bayesian cohort models designed to resolve the identification problem in cohort analysis are proposed in this paper. At first, the basic cohort model which represents the statistical structure of time-series social survey data in terms of age, period and cohort effects is explained. The logit cohort model for qualitative data from a binomial distribution and the normal-type cohort model for quantitative data from a normal distribution are considered as two special cases of the basic model. In order to overcome the identification problem in cohort analysis, a Bayesian approach is adopted, based on the assumption that the effect parameters change gradually. A Bayesian information criterion ABIC is introduced for the selection of the optimal model. This approach is so flexible that both the logit and the normal-type cohort models can be made applicable, not only to standard cohort tables but also to general cohort tables in which the range of age group is not equal to the interval between periods. The practical utility of the proposed models is demonstrated by analysing two data sets from the literature on cohort analysis.

1. Introduction

Birth cohort, which may be defined as a set of people who were born within the same time interval, is an important concept in the study of social change (Ryder [13], Glenn [4], Mason and Fienberg [9]). Cohort analysis is concerned with the separation of the age, period and cohort effects from observations of some quantity classified by age group and survey period. It has been pointed out, however, that there exists an identification problem in cohort analysis, that is, the age, period and cohort effects cannot be separated without some prior in-

---

Key words: Bayesian cohort model, cohort analysis, ABIC, identification problem, logit model, general cohort table.

formation (Mason and others [10], Pullum [12], Fienberg and Mason [3]).

Nakamura [11] proposed a Bayesian logit cohort model for a standard cohort table in which the range in years covered by each age group equals the interval in years between successive survey periods, extending the basic idea of the general Bayesian modeling for linear problems developed by Akaike [1]. This modeling successfully solved the identification problem in cohort analysis.

This paper presents Bayesian cohort models for general cohort table analyses. These models will be applicable to a wide variety of cohort data sets: quantitative as well as qualitative, and not only sets arranged in the standard cohort table but also those arranged in the general cohort table in which the range of each age group does not equal the interval of survey periods.

First, the basic cohort model is described in Section 2. Next the likelihood of the cohort model is obtained in Section 3, assuming the binomial sampling model for qualitative data. In Section 4, a Bayesian approach to the cohort analysis is adopted and, following Akaike [1], a Bayesian information criterion ABIC for selecting the optimal model is defined. The practical utility of the proposed model is demonstrated in Section 5, by analysing two data sets: the data of homicide rates and the data of percentages of political independents in the United States. The sampling model for quantitative data and the corresponding cohort model are also considered in Section 6. The details of the derivation of ABIC and the parameter estimates, and the APL programs are shown in the Appendices.

The usefulness of the Bayesian approach is now being verified by other successful applications by Ishiguro and Akaike [5], Kashiwagi [7], Ishiguro and Sakamoto [6], and others.

2. The basic cohort model

In the cohort analysis model, we assume that y_{ij}, a parameter which characterizes the i-th age group in the j-th survey period, can be decomposed into the form

$$(2.1) \qquad y_{ij}=\mu+\mu_i^A+\mu_j^P+\mu_k^C\,, \qquad i=1,\cdots,I;\; j=1,\cdots,J;\; k=1,\cdots,K;$$

where μ is the grand mean effect, μ_i^A the effect of the age group, μ_j^P the effect of the survey period, and μ_k^C the effect of the birth cohort. The μ's in (2.1) are subject to the following constraints:

$$(2.2) \qquad \sum \mu_i^A=\sum \mu_j^P=\sum \mu_k^C=0\,.$$

Note that there is a functional relationship among i, j and k due to the inherent structure of cohorts; for example, $k=j-i+I$ and $K=I$

$+J-1$ in standard cohort table analysis.

We can rewrite the model (2.1) with the constraints (2.2) in vector and matrix notation as

$$\boldsymbol{y}=D\boldsymbol{\mu}\,, \tag{2.3}$$

where $\boldsymbol{y}=(y_{11}, y_{21}, \cdots, y_{IJ})'$, $\boldsymbol{\mu}=[\mu\ \boldsymbol{\mu}_*']'$, $\boldsymbol{\mu}_*=(\mu_1^A, \cdots, \mu_{I-1}^A, \mu_1^P, \cdots, \mu_{J-1}^P, \mu_1^C, \cdots, \mu_{K-1}^C)'$, $D=[\boldsymbol{1}\ D_*]$, $\boldsymbol{1}=(1, \cdots, 1)'$, D_* is an appropriate $IJ\times L$ design matrix expressing the relationship among the three demographic factors, and $L=I+J+K-3$. The prime ($'$) denotes the transpose of a vector or a matrix. For an example of the design matrix, see Section 5.

3. Logit model and the likelihood

Here we consider the binomial sampling model for dichotomous response data. The other sampling model for quantitative data is mentioned in Section 6.

Under the binomial sampling model, it is natural to take $\boldsymbol{q}$, the vector of logits, for $\boldsymbol{y}$ in (2.3). Then we can write the model (2.3) as

$$\begin{aligned}\boldsymbol{q}&=(\log p_{11}/(1-p_{11}), \log p_{21}/(1-p_{21}), \cdots, \log p_{IJ}/(1-p_{IJ}))'\\ &=D\boldsymbol{\mu}\,,\end{aligned} \tag{3.1}$$

where p_{ij} is the cell probability of a positive response for the (i, j) cell of the cohort table, and log denotes the natural logarithm. We may call it as *the logit cohort model.*

Let N_{ij} be the sample size and n_{ij} be the observed frequency of positive response for the (i, j) cell, then the likelihood of this model (3.1) is given by

$$f(\boldsymbol{n}\,|\,\boldsymbol{p}(\boldsymbol{\mu}))=\prod\prod\binom{N_{ij}}{n_{ij}}p_{ij}^{n_{ij}}(1-p_{ij})^{N_{ij}-n_{ij}} \tag{3.2}$$

where $\boldsymbol{n}=(n_{11}, n_{21}, \cdots, n_{IJ})'$ and $\boldsymbol{p}=(p_{11}, p_{21}, \cdots, p_{IJ})'$.

Usually we use the maximum likelihood procedure to obtain the estimate of the parameter $\boldsymbol{\mu}$, but in the cohort model we cannot uniquely identify the estimate since the column vectors of the design matrix D_* are linearly dependent. More information on the parameters is necessary in order to completely specify the cohort model.

4. A Bayesian approach to the cohort model

To overcome the difficulty of the identification problem in cohort analysis, we adopt a Bayesian approach discussed by Akaike [1]. As age, period and cohort effects have the ordinal property, it is natural to expect that the successive parameters of the age groups, survey periods, or birth cohorts are not so different, or change gradually. In other words, the first order differences of the successive effect parameters,

$$\mu_i^A-\mu_{i+1}^A\,,\quad \mu_j^P-\mu_{j+1}^P\,,\quad \text{and}\quad \mu_k^C-\mu_{k+1}^C\,,$$

are close to zero for $i=1, \cdots, I-1$; $j=1, \cdots, J-1$; and $k=1, \cdots, K-1$.

In order to explicitly describe the expectation stated above, we assume that the parameter vector $\boldsymbol{\mu}_*$ has the prior distribution defined by

$$\text{(4.1)}\quad \pi(\boldsymbol{\mu}_*\,|\,\sigma_A^2, \sigma_P^2, \sigma_C^2)\propto(2\pi)^{-L/2}|D_s'S^{-1}D_s|^{1/2}\exp\left[-(1/2)\boldsymbol{\mu}_*'D_s'S^{-1}D_s\boldsymbol{\mu}_*\right]$$

for the logit cohort model, where D_s is an $L\times L$ matrix expressing the first order differences of the parameters (see Section 5 for an example of this matrix), and S is an $L\times L$ diagonal matrix, the first to the $(I-1)$-th diagonal elements of which are equal to σ_A^2, the next $J-1$ diagonal elements are equal to σ_P^2, and the last $K-1$ diagonal elements are equal to σ_C^2, i.e., $S=\operatorname{diag}(\sigma_A^2, \cdots, \sigma_A^2, \sigma_P^2, \cdots, \sigma_P^2, \sigma_C^2, \cdots, \sigma_C^2)$.

If we specify the values of σ_A^2, σ_P^2 and σ_C^2, it is reasonable to define the estimate of $\boldsymbol{\mu}$ by the mode of the posterior distribution of $\boldsymbol{\mu}$, the distribution which is proportional to $f(\boldsymbol{n}\,|\,\boldsymbol{\mu})\cdot\pi(\boldsymbol{\mu}_*\,|\,\sigma_A^2, \sigma_P^2, \sigma_C^2)$. We denote this estimate by $\hat{\boldsymbol{\mu}}$, or, more explicitly, by $\hat{\boldsymbol{\mu}}(\sigma_A^2, \sigma_P^2, \sigma_C^2)$.

The remaining problem is how to determine the values of σ_A^2, σ_P^2 and σ_C^2. As a criterion for the determination of σ_A^2, σ_P^2 and σ_C^2, we adopt ABIC (A Bayesian Information Criterion) proposed by Akaike [1]. ABIC is defined by

$$\text{(4.2)}\quad \begin{aligned}\text{ABIC}&=-2\log(\text{marginal likelihood})\\&=-2\log\int f\cdot\pi\, d\boldsymbol{\mu}_*\,.\end{aligned}$$

We select the values of σ_A^2, σ_P^2 and σ_C^2 so as to minimize ABIC.

In our Bayesian logit cohort model, the value of ABIC is evaluated approximately by

$$\text{(4.3)}\quad \begin{aligned}\text{ABIC}=&\sum\sum\left[n_{ij}\log n_{ij}/\hat{n}_{ij}+(N_{ij}-n_{ij})\log(N_{ij}-n_{ij})/(N_{ij}-\hat{n}_{ij})\right]\\&+\hat{\boldsymbol{\mu}}_*'D_s'S^{-1}D_s\hat{\boldsymbol{\mu}}_*-\log|D_s'S^{-1}D_s|+\log|D_v'V^{-1}D_v+D_s'S^{-1}D_s|\,,\end{aligned}$$

where $\hat{n}_{ij}=N_{ij}\hat{p}_{ij}=N_{ij}/[1+\exp(-\hat{q}_{ij})]$, $\hat{\boldsymbol{q}}=(\hat{q}_{11}, \hat{q}_{21}, \cdots, \hat{q}_{IJ})'=D\hat{\boldsymbol{\mu}}$, V is an $IJ\times IJ$ diagonal matrix, the $(i+(j-1)I)$-th diagonal element of which is $[N_{ij}\hat{p}_{ij}(1-\hat{p}_{ij})]^{-1}$, and $D_v=D_*-(\mathbf{1}'V^{-1}\mathbf{1})^{-1}\mathbf{1}\mathbf{1}'V^{-1}D_*$. For further details of the derivation of the estimate $\hat{\boldsymbol{\mu}}$ and ABIC in this Bayesian logit cohort model, see Appendix A and B.

As ABIC is a non-linear function of σ_A^2, σ_P^2 and σ_C^2, we use the grid search technique to obtain the smallest ABIC. In practice, we may compare the values of ABIC for every combination of $(\sigma_A^2, \sigma_P^2, \sigma_C^2)$ drawn from, for example, $\{16\tilde{\sigma}^2, 4\tilde{\sigma}^2, \tilde{\sigma}^2, \tilde{\sigma}^2/4, \tilde{\sigma}^2/16\}$, where $\tilde{\sigma}^2$ is a geometric mean of $[n_{ij}(N_{ij}-n_{ij})/N_{ij}]^{-1}$'s for the logit cohort model.

5. Applications

Analysis of homicide rates

The Bayesian estimation procedure for the cohort model proposed in Section 4 is applied first to the data presented in Table 1. The analysis is an example of standard cohort table analyses while it is not clear why the periods of the original table are labeled '1952–1956', etc. The data representing homicide frequencies per 100,000 crosstabulated by age and period were given by Smith [14] cited in Knoke and Burke ([8], 57–63). Though the assumption of sample size 100,000 for each cell seems to be unrealistic to use the logit cohort model, we will use it temporarily for the purpose of comparing our results with Knoke and Burke's.

Table 1. Age-period-cohort crosstabulation of homicide frequencies per 100,000

	Period					
Age group	1952-1956	1957-1961	1962-1967	1967-1971	1972-1976	Cohort
1. 15-19	6.2	7.5	8.6	15.1	17.1	11
2. 20-24	11.8	13.6	14.2	22.9	25.5	10
3. 25-29	12.4	11.9	13.6	19.3	22.2	9
4. 30-34	10.8	10.6	10.9	15.5	16.9	8
5. 35-39	9.4	8.8	9.1	12.5	13.4	7
6. 40-44	7.7	6.8	7.1	9.6	10·2	6
7. 45-49	6.1	5.7	5.5	7.3	7.4	
Cohort	1	2	3	4	5	

Source: Smith [14] cited by Knoke and Burke [8].

Figure 1 displays the result of age-period-cohort decomposition of homicide rates per 100,000 by using the ordinary logit model with the identification specification such that the effect of the 40–44 age level is equal to the effect of the 45–49 age level. This model is almost identical to the log-linear model used by Knoke and Burke. Examining the result of their analysis, Knoke and Burke stated: it could be seen that period effects were decreasing over time, contrary to first impressions of the table.

Figure 2 shows the result obtained by the proposed Bayesian logit cohort model. We can see that period effect is increasing over time, as opposed to the result by the ordinal logit model with the identifica-

```
< GRAND MEAN >

           -9.0626
           (   .01 )

< PERIOD >         -2.0      -1.0       .0       1.0       2.0
                     +---------+---------+---------+---------+
  1952-56    .4290   |         |         |   *     |         |
  1957-61    .1478   |         |         |*        |         |
  1962-67   -.1124   |         |        *|         |         |
  1968-71   -.0813   |         |        *|         |         |
  1972-76   -.3831   |         |     *   |         |         |
                     +---------+---------+---------+---------+
<  AGE    >
                     +---------+---------+---------+---------+
   15-19   -1.1336   |        *|         |         |         |
   20-24    -.2231   |         |       * |         |         |
   25-29     .0665   |         |         |*        |         |
   30-34     .2156   |         |         | *       |         |
   35-39     .3356   |         |         |  *      |         |
   40-44     .3695   |         |         |   *     |         |
   45-49     .3695   |         |         |   *     |         |
                     +---------+---------+---------+---------+
< COHORT >
                     +---------+---------+---------+---------+
    1      -1.4406   |     *   |         |         |         |
    2      -1.2160   |       * |         |         |         |
    3      -1.0059   |         *         |         |         |
    4       -.7410   |         |  *      |         |         |
    5       -.4510   |         |    *    |         |         |
    6       -.1688   |         |       * |         |         |
    7        .1834   |         |         | *       |         |
    8        .5373   |         |         |    *    |         |
    9        .9715   |         |         |         *         |
    10      1.4255   |         |         |         |   *     |
    11      1.9056   |         |         |         |        *|
                     +---------+---------+---------+---------+
```

Figure 1. Estimated effect parameters for homicide rates per 100,000 in the United States presented in Table 1, based on the ordinary logit model with identification specification such that $\mu^A_{(40\text{-}44)} = \mu^A_{(45\text{-}49)}$.

```
< HYPER-PARAMETERS AND ABIC >        HYPER-
                                   PARAMETER

                        PERIOD =     .5000
                           AGE =    1.0000
                        COHORT =     .0039

                          ABIC =   31.1891    (SIGMA=0.092039)
< GRAND MEAN >

           -9.1094
           (   .01 )

< PERIOD >        -2.0      -1.0        .0       1.0       2.0
                     +---------+---------+---------+---------+
  1952-56   -.2165   |         |       * |         |         |
  1957-61   -.1928   |         |       * |         |         |
  1962-67   -.0964   |         |        *|         |         |
  1968-71    .1962   |         |         | *       |         |
  1972-76    .3094   |         |         |  *      |         |
                     +---------+---------+---------+---------+
<  AGE   >
                     +---------+---------+---------+---------+
  15-19      .0282   |         |         *         |         |
  20-24      .3865   |         |         |   *     |         |
  25-29      .3220   |         |         |  *      |         |
  30-34      .1337   |         |         |*        |         |
  35-39     -.0665   |         |        *|         |         |
  40-44     -.3015   |         |      *  |         |         |
  45-49     -.5025   |         |    *    |         |         |
                     +---------+---------+---------+---------+
< COHORT >
                     +---------+---------+---------+---------+
     1      -.0005   |         |         *         |         |
     2      -.0007   |         |         *         |         |
     3      -.0014   |         |         *         |         |
     4      -.0024   |         |         *         |         |
     5      -.0032   |         |         *         |         |
     6      -.0034   |         |         *         |         |
     7      -.0023   |         |         *         |         |
     8      -.0000   |         |         *         |         |
     9       .0031   |         |         *         |         |
    10       .0052   |         |         *         |         |
    11       .0057   |         |         *         |         |
                     +---------+---------+---------+---------+
```

Figure 2. Estimated effect parameters for homicide rates per 100,000 in the United States presented in Table 1, based on the Bayesian logit cohort model.

```
< GRAND MEAN >

          -9.0626
           (   .01 )

< PERIOD >         -2.0      -1.0        .0       1.0       2.0
                     +---------+---------+---------+---------+
  1952-56   -.2298   |         |       * |         |         |
  1957-61   -.1816   |         |       * |         |         |
  1962-67   -.1124   |         |        *|         |         |
  1968-71    .2481   |         |         | *       |         |
  1972-76    .2758   |         |         |  *      |         |
                     +---------+---------+---------+---------+
<  AGE   >
                     +---------+---------+---------+---------+
   15-19    -.1453   |         |        *|         |         |
   20-24     .4357   |         |         |   *     |         |
   25-29     .3959   |         |         |   *     |         |
   30-34     .2156   |         |         | *       |         |
   35-39     .0062   |         |         *         |         |
   40-44    -.2893   |         |      *  |         |         |
   45-49    -.6188   |         |   *     |         |         |
                     +---------+---------+---------+---------+
< COHORT >
                     +---------+---------+---------+---------+
     1       .2066   |         |         | *       |         |
     2       .1017   |         |         |*        |         |
     3      -.0176   |         |         *         |         |
     4      -.0822   |         |        *|         |         |
     5      -.1216   |         |        *|         |         |
     6      -.1688   |         |       * |         |         |
     7      -.1460   |         |        *|         |         |
     8      -.1216   |         |        *|         |         |
     9      -.0168   |         |         *         |         |
    10       .1078   |         |         |*        |         |
    11       .2585   |         |         |  *      |         |
                     +---------+---------+---------+---------+
```

Figure 3. Estimated effect parameters for homicide rates per 100,000 in the United States presented in Table 1, based on the ordinary logit model with identification specification such that $\mu_5^C=\mu_8^C$.

tion specification in Figure 1. Age effect starts in the middle, peaks in the 20-24 age level, and then decreases. Cohort effect is almost constant. All these results are consistent to the first impression of the data.

Our procedure proposed here can give us a satisfactory explanation of the data almost automatically without any identification specification, the indefiniteness of which is likely to mislead us. If Knoke and Burke had added the identification specification such that the parameter of cohort 5 is equal to that of cohort 8, they would have obtained the different result with period effect not decreasing (see Figure 3). It is difficult to determine which identification specification should be selected, since the impression of the result depends largely on where the specification is imposed.

Analysis of percentages of political independents

The second application of our procedure is an example of general cohort table analyses. We reanalyse the data of percentages of political independents in the United States presented in Table 2, which is adapted by Glenn ([4], p. 56). As the table is constructed in the form of cohort×period table and the range of age group is not equal to the interval between periods, the situation is a little complicated. We can use, however, the Bayesian logit cohort model by introduction of overlapping age effect parameters involving some hypothetical ones.

Table 2. Percentage of independents in five 10-year birth cohorts, 1945 to 1957 (Ns are in parenthese).

	Year			
	1945	1949	1953	1957
Cohort 1—Ages:	20-29	24-33	28-37	32-41
	25.6 (649)	21.4 (1909)	22.6 (1520)	23.0 (1536)
Cohort 2—Ages:	30-39	34-43	38-47	42-51
	17.6 (666)	19.4 (2046)	20.2 (1321)	19.0 (1287)
Cohort 3—Ages:	40-49	44-53	48-57	52-61
	19.8 (662)	19.7 (1727)	21.2 (1114)	17.0 (943)
Cohort 4—Ages:	50-59	54-63	58-67	62-71
	16.1 (490)	17.9 (1209)	18.1 (750)	16.9 (672)
Cohort 5—Ages:	60-69	64-73	68-77	72-81
	20.0 (307)	13.8 (727)	13.4 (382)	13.1 (291)

Source: American Institute of Public Opinion Surveys, Cited by Glenn ([4], p. 56).

Since the method is not necessarily restricted to age×period tables, we consider the vector $\boldsymbol{q}$ as $(q_{11}, \cdots, q_{KJ})'$, the $(k+(j-1)K)$-th element of which is the logit of the expected cell probability for the (k, j) cell of the cohort×period table like Table 2. Let

$$\boldsymbol{\mu}=(\mu, \mu^A_{(20\text{-}29)}, \mu^A_{(22\text{-}31)}, \mu^A_{(24\text{-}33)}, \cdots, \mu^A_{(70\text{-}79)}, \mu^P_{(1945)}, \mu^P_{(1949)}, \mu^P_{(1953)}, \mu^C_1, \cdots, \mu^C_4)'\,,$$

$$D_*=\begin{array}{cccccccccccccccc|c}
\multicolumn{9}{c}{\leftarrow\qquad I-1\qquad\rightarrow} & \multicolumn{3}{c}{\leftarrow J-1\rightarrow} & \multicolumn{4}{c}{\leftarrow\ K-1\ \rightarrow} & \\
1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & \uparrow \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & K \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 & -1 & -1 & -1 & \downarrow \\
0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & \uparrow \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & K \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & \\
0 & 0 & 0 & 0 & \cdots & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & -1 & -1 & -1 & \downarrow \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & \uparrow \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & K \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & -1 & -1 & -1 & \downarrow \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & -1 & -1 & -1 & 1 & 0 & 0 & 0 & \uparrow \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & -1 & -1 & -1 & 0 & 1 & 0 & 0 & \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & -1 & -1 & -1 & 0 & 0 & 1 & 0 & K \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & -1 & -1 & -1 & 0 & 0 & 0 & 1 & \\
-1 & -1 & -1 & -1 & \cdots & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & -1 & -1 & -1 & \downarrow
\end{array}$$

and

$$D_s=\begin{array}{cccccccccccccccc|c}
\multicolumn{9}{c}{\leftarrow\qquad I-1\qquad\rightarrow} & \multicolumn{3}{c}{\leftarrow J-1\rightarrow} & \multicolumn{4}{c}{\leftarrow\ K-1\ \rightarrow} & \\
1 & -1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \uparrow \\
0 & 1 & -1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \\
0 & 0 & 1 & -1 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \\
0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \\
 & & \vdots & & \ddots & & & & & & \vdots & & & \vdots & & & I-1 \\
0 & 0 & 0 & 0 & \cdots & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \\
0 & 0 & 0 & 0 & \cdots & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \\
1 & 1 & 1 & 1 & \cdots & 1 & 1 & 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \downarrow \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & \uparrow \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & J-1 \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 1 & 1 & 2 & 0 & 0 & 0 & 0 & \downarrow \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & \uparrow \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & K-1 \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 2 & \downarrow
\end{array}$$

```
< HYPER-PARAMETERS AND ABIC >      HYPER-
                                   PARAMETER

                        PERIOD =     .0039
                           AGE =     .0039
                        COHORT =    2.0000

                          ABIC =  39.1325      (SIGMA=0.005799)
< GRAND MEAN >

           -1.3518
           ( 20.56 )

< PERIOD >         -2.0      -1.0       .0        1.0       2.0
                     +---------+---------+---------+---------+
   1945      .0001   |         |         *         |         |
   1949     -.0001   |         |         *         |         |
   1953      .0004   |         |         *         |         |
   1957     -.0004   |         |         *         |         |
                     +---------+---------+---------+---------+
<  AGE    >
                     +---------+---------+---------+---------+
   20-29     .0024   |         |         *         |         |
   22-31     .0022   |         |         *         |         |
   24-33     .0020   |         |         *         |         |
   26-35     .0022   |         |         *         |         |
   28-37     .0023   |         |         *         |         |
   30-39     .0020   |         |         *         |         |
   32-41     .0012   |         |         *         |         |
   34-43     .0006   |         |         *         |         |
   36-45     .0004   |         |         *         |         |
   38-47     .0003   |         |         *         |         |
   40-49     .0001   |         |         *         |         |
   42-51     .0002   |         |         *         |         |
   44-53     .0002   |         |         *         |         |
   46-55     .0002   |         |         *         |         |
   48-57     .0002   |         |         *         |         |
   50-59     .0000   |         |         *         |         |
   52-61    -.0002   |         |         *         |         |
   54-63    -.0003   |         |         *         |         |
   56-65    -.0006   |         |         *         |         |
   58-67    -.0008   |         |         *         |         |
   60-69    -.0015   |         |         *         |         |
   62-71    -.0020   |         |         *         |         |
   64-73    -.0020   |         |         *         |         |
   66-75    -.0021   |         |         *         |         |
   68-77    -.0022   |         |         *         |         |
   70-79    -.0023   |         |         *         |         |
   72-81    -.0024   |         |         *         |         |
                     +---------+---------+---------+---------+
< COHORT >
                     +---------+---------+---------+---------+
  1916-25    .2040   |         |         | *       |         |
  1906-15    .1151   |         |         |*        |         |
  1896-05   -.0632   |         |        *|         |         |
  1886-95   -.0756   |         |        *|         |         |
  1876-85   -.1803   |         |       * |         |         |
                     +---------+---------+---------+---------+
```

Figure 4. Estimated effect parameters for percentages of political independents in the United States presented in Table 2, based on the Bayesian logit cohort model.

Then the rest of the procedure can be carried out in the same manner as described in Section 4. Note that seven age effect parameters such as $\mu^A_{(22-31)}$, $\mu^A_{(26-35)}$, $\mu^A_{(36-45)}$, $\mu^A_{(46-55)}$, $\mu^A_{(56-65)}$, $\mu^A_{(66-75)}$ and $\mu^A_{(70-79)}$ are hypothetical.

Figure 4 shows the result of cohort analysis for the data of percentages of independents by using the Bayesian logit cohort model. Cohort effect is dominant and monotonically decreases. These results are consistent with what Glenn stated on the data earlier. Period and age effect cannot be observed.

6. Normal sampling model

In this section, we consider the other sampling model, the normal one for quantitative data.

Under the normal sampling model, let x_{ij} be the observed quantity for the (i, j) cell. The likelihood of the model (2.3) is then given by

$$f(\boldsymbol{x}\,|\,\boldsymbol{\mu}, \sigma^2)=(2\pi\sigma^2)^{-IJ/2}\exp[(-1/2\sigma^2)(\boldsymbol{x}-D\boldsymbol{\mu})'W^{-1}(\boldsymbol{x}-D\boldsymbol{\mu})] \tag{6.1}$$

where σ^2 is an unknown constant. W is a known $IJ\times IJ$ positive definite diagonal matrix such that $\det(W)=1$, and $\boldsymbol{x}=(x_{11}, x_{21}, \cdots, x_{IJ})'$. We may call it as *the normal-type cohort model*. The procedure for estimating parameters is the same as is described for the logit cohort model in Section 3. We assume the prior distribution of $\boldsymbol{\mu}_*$, defined by

$$\begin{aligned}&\pi(\boldsymbol{\mu}_*\,|\,\sigma^2_A, \sigma^2_P, \sigma^2_C, \sigma^2)\\ &\quad\propto(2\pi\sigma^2)^{-L/2}|D_s'S^{-1}D_s|^{-1/2}\exp[-(1/2\sigma^2)\boldsymbol{\mu}_*'D_s'S^{-1}D_s\boldsymbol{\mu}_*]\,.\end{aligned} \tag{6.2}$$

The estimate of $\boldsymbol{\mu}$ is given by the mode of $f(\boldsymbol{x}\,|\,\boldsymbol{\mu}, \sigma^2)\cdot\pi(\boldsymbol{\mu}_*\,|\,\sigma^2_A, \sigma^2_P, \sigma^2_C, \sigma^2)$, the computation of which is easily performed by using the weighted least square method. The value of ABIC is evaluated exactly by

$$\mathrm{ABIC}=(IJ)\log\hat{\sigma}^2-\log|D_s'S^{-1}D_s|+\log|D_w'W^{-1}D_w+D_s'S^{-1}D_s|\,, \tag{6.3}$$

where $\hat{\sigma}^2=[(\boldsymbol{x}-D\hat{\boldsymbol{\mu}})'W^{-1}(\boldsymbol{x}-D\hat{\boldsymbol{\mu}})+\hat{\boldsymbol{\mu}}_*D_s'S^{-1}D_s\hat{\boldsymbol{\mu}}_*]/(IJ)$, and $D_w=D_*-(\mathbf{1}'W^{-1}\mathbf{1})^{-1}\mathbf{1}\mathbf{1}'W^{-1}D_*$. The optimal model which gives the smallest ABIC is selected.

7. Discussions

In order to get over the difficulties of the identification problem in cohort analysis, we have developed Bayesian cohort models, which can be classified by sampling model and type of cohort tables into the following four typologies:

Sampling Model	Type of Cohort Table	
	Standard	General
Binomial	BLS	BLG
Normal	BNS	BNG

where B stands for a Bayesian cohort model, L for Logit, N for Normal, S for a Standard table, and G for a General one.

As to the sampling models, the binomial one is for the data with a qualitative dependent variable, and the normal one for the data with a quantitative dependent variable. Therefore, the two sampling models are considered to cover the wide range of data we usually analyse. We can also suppose the other sampling models and easily develop the corresponding cohort models.

As to the type of cohort tables, there is essentially no difference between the standard one and the general one. In order to deal with the general cohort table, we have only to arrange the design matrix appropriately. In addition, our Bayesian procedure is so flexible that we can analyse the cohort tables with some missing observations by only eliminating the missing cell quantities from the dependent vector and the corresponding rows from the design matrix.

The assumption that the successive parameters change gradually is considered a kind of the principle of parsimony. Adopting the Bayesian modeling with this assumption, not only can we overcome the identification poblem in cohort analysis, but we may also analyse various kinds of cohort data as well.

Acknowledgement

The author would like to thank Professor Yosiyuki Sakamoto for his suggestions and encouragement, and the referees for their useful comments on the earlier version of this paper.

THE INSTITUTE OF STATISTICAL MATHEMATICS

Appendix A

Computational algorithm for the Bayesian logit cohort model

The following steps describe the computational algorithm for estimating the effect parameters of the Bayesian logit cohort model. Concrete procedure by APL program can be seen in Appendix C. As for the normal-type model, it is not necessary to perform an iteration process.

Step 0. Compute the initial value of $\hat{\boldsymbol{\mu}}$.
Set $\hat{\boldsymbol{n}}^{(0)}=\boldsymbol{n}$. Compute $\hat{\boldsymbol{p}}=[\hat{p}_l]=[\hat{n}_l/N_l]$ and $\hat{\boldsymbol{q}}=[\hat{q}_l]=[\log \hat{p}_l/(1-\hat{p}_l)]$. Obtain the diagonal matrix V whose diagonal element is $v_{ll}=[N_l\hat{p}_l(1-\hat{p}_l)]^{-1}$, and then compute D_v. Solve the equation $[D_v'V^{-1}D_v+D_s'S^{-1}D_s]\hat{\boldsymbol{\mu}}_*^{(1)}=D_v'\hat{\boldsymbol{q}}$ and compute $\hat{\mu}^{(1)}=[1'V^{-1}\hat{\boldsymbol{q}}-(1'V^{-1}D_v+1'S^{-1}D_s)\hat{\boldsymbol{\mu}}_*^{(1)}]/(1'V^{-1}1)$. Then set the initial value $\hat{\boldsymbol{\mu}}^{(1)}$ and compute $\hat{\boldsymbol{n}}^{(1)}$ through $\hat{\boldsymbol{q}}=D\hat{\boldsymbol{\mu}}^{(1)}$. Set $t=1$.

Step 1. Check the convergence.
If $\hat{\boldsymbol{n}}^{(t)}$ and $\hat{\boldsymbol{n}}^{(t-1)}$ are sufficiently close, complete the algorithm. Otherwise continue to the next step.

Step 2. Update the $\hat{\boldsymbol{\mu}}$.
Update V and D_v in the same way as in Step 0, and solve equation $[D_v'V^{-1}D_v+D_s'S^{-1}D_s]\Delta\boldsymbol{\mu}=D_v'[\boldsymbol{n}-\hat{\boldsymbol{n}}^{(t)}]-D_s'S^{-1}D_s\hat{\boldsymbol{\mu}}^{(t)}$ to determine the step size $\Delta\boldsymbol{\mu}$. Then update the estimate $\hat{\boldsymbol{\mu}}_*^{(t+1)}=\hat{\boldsymbol{\mu}}_*^{(t)}+\Delta\boldsymbol{\mu}$, compute $\hat{\mu}^{(t+1)}$, and set $\hat{\boldsymbol{\mu}}^{(t+1)}$. Increment t and return to Step 1. The determination of $\Delta\boldsymbol{\mu}$ is based on the equivalence of maximum likelihood and weighted least squares estimates in the exponential family (see Bradley [2]).

Appendix B

Derivation of ABIC for the Bayesian logit cohort model

Using the fact that $x_l=\log n_l/(N_l-n_l)$ has an approximate normal distribution with mean $q_l=\log p_l/(1-p_l)$ and variance $v_{ll}=[N_l\hat{p}_l(1-\hat{p}_l)]^{-1}$, we can replace f in (3.2) with the normal distribution:

$$\text{(B.1)}\qquad f\sim(2\pi)^{-L/2}|V|^{-1/2}\exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{q})'V^{-1}(\boldsymbol{x}-\boldsymbol{q})\right\}.$$

Therefore, we have

$$\text{(B.2)}\quad f\cdot\pi\sim C_1|V|^{-1/2}|D_s'S^{-1}D_s|^{1/2}\exp\left\{-\frac{1}{2}\begin{pmatrix}\boldsymbol{x}-D\boldsymbol{\mu}\\ D_s\boldsymbol{\mu}_*\end{pmatrix}'\begin{pmatrix}V & 0\\ 0 & S\end{pmatrix}^{-1}\begin{pmatrix}\boldsymbol{x}-D\boldsymbol{\mu}\\ D_s\boldsymbol{\mu}_*\end{pmatrix}\right\}$$

$$=C_1\cdot B\exp\left\{-\frac{1}{2}(\boldsymbol{y}-\tilde{D}\boldsymbol{\mu})'\tilde{\Sigma}^{-1}(\boldsymbol{y}-\tilde{D}\boldsymbol{\mu})\right\},$$

where

$$\tilde{\Sigma}=\begin{bmatrix}V&0\\0&S\end{bmatrix},\quad \tilde{D}=\begin{bmatrix}\mathbf{1}&D_*\\\mathbf{0}&D_s\end{bmatrix},\quad \boldsymbol{y}=\begin{bmatrix}\boldsymbol{x}\\0\end{bmatrix},\quad \boldsymbol{\mu}=\begin{bmatrix}\mu\\\boldsymbol{\mu}_*\end{bmatrix},$$

$$B=|V|^{-1/2}|D_s'S^{-1}D_s|^{1/2},$$

and C_1, C_2, $\cdots$ are properly chosen constants.

Applying the maximum likelihood estimate $\hat{\boldsymbol{\mu}}$ (taken for the weighted least square estimate) to the above equation, we can write

$$\text{(B.3)}\qquad f\cdot\pi\sim C_1\cdot B\exp\left[-\frac{1}{2}\{(\boldsymbol{y}-\tilde{D}\hat{\boldsymbol{\mu}})'\cdot\tilde{\Sigma}^{-1}(\boldsymbol{y}-\tilde{D}\hat{\boldsymbol{\mu}})+(\tilde{D}\boldsymbol{\mu}-\tilde{D}\hat{\boldsymbol{\mu}})'\cdot\tilde{\Sigma}^{-1}(\tilde{D}\boldsymbol{\mu}-\tilde{D}\hat{\boldsymbol{\mu}})\}\right],$$

and hence

$$\begin{aligned}\text{(B.4)}\qquad \int f\cdot\pi\,d\boldsymbol{\mu}_*&\sim C_1\cdot B\exp\left[-\frac{1}{2}\{(\boldsymbol{y}-\tilde{D}\hat{\boldsymbol{\mu}})'\tilde{\Sigma}^{-1}(\boldsymbol{y}-\tilde{D}\hat{\boldsymbol{\mu}})\}\right]\\&\quad\times\frac{1}{C_2}|(\tilde{D}'\tilde{\Sigma}^{-1}\tilde{D})^{-1}|^{1/2}\\&\quad\times\int C_2|(\tilde{D}'\tilde{\Sigma}^{-1}\tilde{D})^{-1}|^{-1/2}\exp\left\{-\frac{1}{2}(\boldsymbol{\mu}-\hat{\boldsymbol{\mu}})'\tilde{D}'\tilde{\Sigma}^{-1}\tilde{D}(\boldsymbol{\mu}-\hat{\boldsymbol{\mu}})\right\}d\boldsymbol{\mu}_*\\&=C_3\cdot B\exp\left\{-\frac{1}{2}(\boldsymbol{y}-\tilde{D}\hat{\boldsymbol{\mu}})'\tilde{\Sigma}^{-1}(\boldsymbol{y}-\tilde{D}\hat{\boldsymbol{\mu}})\right\}\\&\quad\times|\tilde{D}'\tilde{\Sigma}^{-1}\tilde{D}|^{-1/2}\cdot|\mathbf{1}'V^{-1}\mathbf{1}|^{1/2}\exp\left\{-\frac{(\mathbf{1}'V^{-1}\mathbf{1})}{2}(\mu-\hat{\mu})^2\right\}.\end{aligned}$$

Here we set $\mu=\hat{\mu}$ so as to maximize (B.4), that is, to minimize ABIC, and use $|\tilde{D}'\tilde{\Sigma}^{-1}\tilde{D}|=|\mathbf{1}'V^{-1}\mathbf{1}|\cdot|D_v'V^{-1}D_v+D_s'S^{-1}D_s|$. From (B.4), we have

$$\begin{aligned}\text{(B.5)}\qquad &\int f\cdot\pi\,d\boldsymbol{\mu}_*\\&\sim C_3\cdot B|D_v'V^{-1}D_v+D_s'S^{-1}D_s|^{-1/2}\exp\left\{-\frac{1}{2}(\boldsymbol{y}-\tilde{D}\hat{\boldsymbol{\mu}})'\tilde{\Sigma}^{-1}(\boldsymbol{y}-\tilde{D}\hat{\boldsymbol{\mu}})\right\}.\end{aligned}$$

According to the definition of ABIC in (4.2), we have

$$\begin{aligned}\text{(B.6)}\qquad \mathrm{ABIC}&=-2\log\int f\cdot\pi\,d\boldsymbol{\mu}_*\\&\sim-2\log B+\log|D_v'V^{-1}D_v+D_s'S^{-1}D_s|\\&\quad+(\boldsymbol{y}-\tilde{D}\hat{\boldsymbol{\mu}})'\tilde{\Sigma}^{-1}(\boldsymbol{y}-\tilde{D}\hat{\boldsymbol{\mu}})-2\log C_3\end{aligned}$$

$$=\log|V|-\log|D_s'S^{-1}D_s|+\log|D_v'V^{-1}D_v+D_s'S^{-1}D_s|$$
$$+(\boldsymbol{x}-D\hat{\boldsymbol{\mu}})'V^{-1}(\boldsymbol{x}-D\hat{\boldsymbol{\mu}})+\hat{\boldsymbol{\mu}}_*D_sS^{-1}D_s\hat{\boldsymbol{\mu}}_*+C_4\,.$$

Setting $\hat{p}_l=\hat{n}_l/N_l$ and $\hat{\boldsymbol{q}}=D\hat{\boldsymbol{\mu}}$, and taking the logarithm on the both sides of (B.1), we have the following correspondence

$$\text{(B.7)}\qquad \sum n_l\log n_l/\hat{n}_l+\sum(N_l-n_l)\log(N_l-n_l)/(N_l-\hat{n}_l)+C_5$$
$$\sim-\frac{1}{2}\log|V|-\frac{1}{2}(\boldsymbol{x}-D\hat{\boldsymbol{\mu}})'V^{-1}(\boldsymbol{x}-D\hat{\boldsymbol{\mu}})\,.$$

Applying this to (B.5), then we have

$$\text{(B.8)}\qquad \mathrm{ABIC}\sim 2\sum\{n_l\log n_l/\hat{n}_l+(N_l-n_l)\log(N_l-n_l)/(N_l-\hat{n}_l)\}$$
$$+\hat{\boldsymbol{\mu}}_*D_s'S^{-1}D_s\hat{\boldsymbol{\mu}}_*$$
$$-\log|D_s'S^{-1}D_s|+\log|D_v'V^{-1}D_v+D_s'S^{-1}D_s|+C_6\,.$$

When comparing ABICs for different combinations of hyper parameters, we can ignore the constant term in (B.8). Therefore, we have ABIC for the logit cohort model in (4.3).

Appendix C

APL program for solving the Bayesian logit cohort model

The following APL function '*BLOGIT*', which computes the estimates of the parameters and ABIC, is the main part of the functions for the Bayesian logit cohort model discussed in this paper.

The calling syntax is

$$O\leftarrow Z\ BLOGIT\ X\,,$$

where the first column of the right-hand argument array X is the vector of observed frequencies $\boldsymbol{n}$, the second column of X is the vector of sample sizes (N_{ij}), and from the third to the last columns of X give the design matrix D less its first column; the first column of the left-hand argument array Z is the vector of the diagonal elements of S, and from the second to the last columns of Z give the difference matrix D_s less its first column; the first element of the output vector O is ABIC, and from the second to the last elements of the vector O give the estimated parameter vector $\hat{\boldsymbol{\mu}}$.

This function needs two functions '*LOGDET*' and '*LRCHISQ*'. The former computes the natural logarithm of the determinant of a matrix A, and the latter the likelihood-ratio chi-square statistic for the observed frequency vector N and the corresponding expected frequency vector M.

```
 ∇ O←DS BLOGIT D;ITER;RL;AB;K;T;V;N;S;DME;DSD;MM;U0;U;Y
   ;W;M;G;E;DU;A;CONV
  ITER←(RL←0.00001),(AB←0.001),(K←10),CONV←0
  MM←N←D[;1]
  V←÷T←D[;2]
  D← 0 2 ↓D
  S←÷DS[;1]
  DS← 0 1 ↓DS
  U←(1↓⍴D)⍴U0←0
  Y←((W←V×M×Y)×⍟M÷Y←T-M←0.5⌈N⌊T-0.5),(⍴S)⍴0
  DSD←(⍉DS)+.×DS×⍉(⌽⍴DS)⍴S
L0:M←MM
  DME←(D-(⍴D)⍴E←(+⌿D×⍉(⌽⍴D)⍴W)×G←÷+/W)
  A←(((⍉DME)×(⌽⍴DME)⍴W)+.×DME)+DSD
  →CONV/L1
  U←U+DU←(Y+.×DME,[1] DS)⌹A
  U0←U0+((G×+/(⍴T)↑Y)-+/E×DU)
  Y←(N-MM←T÷1+*-U0+D+.×U),-S×DS+.×U
  W←V×MM×T-MM
  →L0,CONV←(0≥K←K-1)∨(∧/RL>M÷MM)∨(∧/AB>M←|M-MM)
L1:O←(N LRCHISQ MM)+((T-N) LRCHISQ T-MM)++/U×DSD+.×U
  O←(O+(-⍟G)+(1↓LOGDET A)-1↓LOGDET DSD),U0,U
 ∇
```

```
 ∇ T←LOGDET A;N;K;B;L;SIGN
  SIGN←1+T←K←N-N←1↑⍴A
L0:→(1=L←(|B)⍳⌈/|B←,A[;1])/L1
  SIGN←-SIGN,0⍴(A[1;]←B),0⍴(A[L;]←A[1;]),0⍴B←A[L;]
L1:→(N≤K←K+1,0⍴T←T+⍟|A[1;1],0⍴SIGN←SIGN××A[1;1])/L2
  →L0,⍴A←(1 1 ↓A)-A[B;1]∘.×A[1;B←1↓⍳1↑⍴A]÷A[1;1]
L2:T←SIGN,T
 ∇
```

```
 ∇ X2L←N LRCHISQ M
  X2L←2×+/((N>0)/N)×⍟(N>0)/N÷M
 ∇
```

References

[1] Akaike, H. (1980). Likelihood and the Bayes procedure, *Bayesian Statistics* (eds. J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith), Valencia, University Press.

[2] Bradley, E. L. (1973). The equivalence of maximum likelihood and weighted least squares estimates in the exponential family, *J. Amer. Statist. Ass.*, **68**, 199-200.

[3] Fienberg, S. E. and Mason, W. M. (1978). Identification and estimation of age-period-cohort models in the analysis of discrete archival data, *Sociological Methodology 1979* (ed. K. F. Schuessler), London, Jossey-Bass.

[4] Glenn, N. D. (1977). *Cohort Analysis*, Biverly Hills and London, Sage Publications.

[5] Ishiguro, M. and Akaike, H. (1981). A Bayesian approach to the trading-day adjustment of monthly data, *Time Series Analysis*, (eds. O. D. Anderson and M. R. Perryman), Amsterdam, North-Holland.

[6] Ishiguro, M. and Sakamoto, Y. (1984). A Bayesian approach to the probability density estimation, *Ann. Inst. Statist. Math.*, B, **36**, 523-538.

[7] Kashiwagi, N. (1982). A Bayes estimation procedure for fertilities in field experi-

ments, *Res. Memo.*, No. 220, Tokyo, the Institute of Statistical Mathematics.

[8] Knoke, D. and Burke, P. J. (1980). *Log-Linear Models*, Biverly Hills and London, Sage Publications.

[9] Mason, W. M. and Fienberg, S. E. (eds.) (1985). *Cohort Analysis in Social Research*, New York, Springer-Verlag.

[10] Mason, K. O., Mason, W. M., Winsborough, H. H. and Poole, W. K. (1973). Some methodological issues in cohort analysis of archival data, *American Sociological Review*, **38**, 242-258.

[11] Nakamura, T. (1982). A Bayesian cohort model for standard cohort table analysis, *Proc. Inst. Statist. Math.*, **29**, 77-97. (In Japanese)

[12] Pullum, T. W. (1978). Parameterizing age, period, and cohort effects: an application to U.S. delinquency rates, 1964-1973, *Sociological Methodology* 1978 (ed. K. F. Schuessler), San Francisco, Jossey-Bass.

[13] Ryder, N. B. (1965). The cohort as a concept in the study of social change, *American Sociological Review*, **30**, 843-861.

[14] Smith, M. D. (1979). Increasing in youth violence: age, period or cohort effect, Presented at the Meetings of the American Sociological Association, Boston (Cited by Knoke and Burke (1980)).