THE NATURAL VARIABILITY OF VITAL RATES AND ASSOCIATED STATISTICS

BY

DAVID R. BRILLINGER*

TECHNICAL REPORT NO. 44
JUNE 1985

*RESEARCH PARTIALLY SUPPORTED BY
NATIONAL SCIENCE FOUNDATION GRANT DMS-8316634

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA

The Natural Variability of Vital Rates and Associated Statistics

David R. Brillinger

Statistics Department, University of California, Berkeley,

California, 94720

SUMMARY

The first concern of this work is the development of approximations to the distributions of crude mortality rates, age-specific mortality rates, age-standardized rates, standardized mortality ratios and the like, for the case of an open population. It is found that assuming Poisson birth times and independent lifetimes implies that the number of deaths and the mid-year population have a bivariate Poisson distribution. The Lexis diagram is seen to make the result immediate. It is suggested that in a variety of cases, it will be satisfactory to approximate the distribution of the number of deaths given the population size, by a Poisson with mean proportional to the population size. It is further suggested that situations in which explanatory variables are present may be modeled via a doubly stochastic Poisson distribution for the number of deaths, with mean proportional to the population size and an exponential function of a linear combination of the explanatories. Such a model is fit to mortality data for Canadian females classified by age and year and a dynamic variant of the model is further fit to the time-series of total female deaths by year. The models with extra-Poisson variation are found to lead to substantially improved fits.

Key words: age-adjusted death rate, dynamic model, extra-Poisson variation, point process, Poisson regression, standardized mortality ratio, uncertainty estimation, vital statistics, weights

1. Introduction

Vital statistics are data on the fundamental events of human lives such as births, deaths, marriages and the like. They usually take the form of counts or rates and are often collected via censuses and legally required registrations. They are used for: summarization, comparison, forecasting, detection of change, hypothesis generation, surveillance and studying public health generally.

A continual presence is a wish to make comparisons; comparisons between regions, (comparisons between) time periods, (comparisons) between social groups, and so on. Now, in many circumstance the data are virtually complete so that it is a fact that say death ratesdiffer for two counties or two years or two races. What is more likely of concern then is; do two death rates differ by more than some level of natural fluctuations? The purpose of this study is the conceptualizationand formalization of the natural variability of vital statistics to support their use in comparisons and other analyses. The particular cases of mortality and of groups specifically delineated in age and time will be emphasized; however results for a broad variety of other cases should be apparent.

Quite a number of distinct vital statistics are in common use. These include counts, rates and ratios , the emphasis being on the latter two.

<u>Counts</u>. One sets down the total number of deaths in a given time period for a population of interest, perhaps separately by age, region or cause. Figure 1, middl gives this data for the population of Canadian females annually for the time period 1926 to 1982. An issue that arises, always, in mortality studies is whether the population whose deaths have been recorded is closed or

open. In the former case, the group of individuals whose deaths are recorded is unchanging. In the latter case, the group membership changes because of birthdays, emigration, immigration and the like. This is the circumstance for the Canadian data. This paper is concerned with open populations.

Rates. Rates are relative frequencies. For example the crude death rate is the number of deaths in a population of interest during a specified time period, divided by the number of person-years lived by the population during the time period. A complication that often arises is that person-years lived has to be estimated. In the case of annual rates, an estimate of the mid-year population is often used. Figure 1 gives the Canadian 1 June population estimate and the corresponding annual crude death rate. (The three figures display rising numbers of deaths and population members, but a falling death rate. The kink in the population series in 1949 resulted from Newfoundland's entering the country.) In the case that a death rate is given for a specified age group, it is referred to as an age-specific mortality rate. These rates are important because mortality experience usually varies substantially with age and a crude rate may not display important phenomena. Age-adjusted rates are an attempt to provide single rates that allow direct comparison of populations with differing age compositions. They are weighted combinations of age-specific rates. (For example, the weights may correspond to the composition of some standard population.)

Ratios. The standardized mortality ratio (SMR) may be mentioned. It is the ratio of observed total deaths to "expected" deaths using the rates of some standard population and the given person-years lived. It is often

used in making comparisons.

The purpose in setting down the above material has been to bring out the basic quantities involved in constructing vital statistics - counts and estimates of population size. These are the quantities whose variability will be fundamental. Discussion of vital statistics generally and details concernin particular cases may be found in: Chiang (1961), Keyfitz (1966), Benjamin (1968), Fleiss (1981), Benjamin and Pollard (1980), Inskip, Beral and Fraser (1983) for example.

The structure of this paper is the following: the next section sets the scene and presents some variability measures in common use; Section 3 sets down a conceptual model for the physical process of concern and shows how the Lexis diagram and the methodology of point processes may be used; Section 4 presents specific formulas for a number of cases of interest; the following section discusses the results obtained and describes a simplifying approximation; Section 6 turns to the (regression) case where measurements of explanatory are available and presents models and worked examples for cases where greater than Poisson variation is present; the final section draws some conclusions and indicates problems for further study.

2. Background

Discussion will focus on the case of an age-specific death rate for a given year. Let D_x denote the number of deaths in the year for the age group x to x+K/and let P_x denote the mid-year population for that age group. Then the age x death rate is (usually taken to be)

$$M_x = D_x/P_x \quad . \tag{2.1}$$

(In practice P_x has to be estimated, but D_x may be obtained from official records. For the moment however, P_x will be assumed available.)

In statistical studies, D_x is often assumed to be distributed as a binomial variate with parameter $n = P_x$ and its variance is estimated by $D_x(1 - M_x)$, (see for example Pollard (1971), Daw (1974), Tukey and Mosteller (1977) Section 11C.) Conceptually however, this assumption has to be viewed as an approximation for this case of an <u>open</u> population. Some individuals enter the population during the year, when they reach age x, others leave during the year, when they reach age x+K+1.The exposures of the individuals are not all the same and the realizations of the individual life histories are not identically distributed - as is required for the binomial. Further P_x is not the number of individuals in the study, rather it is an estimate of the average number alive aged x to x+K during the year.

In research to obtain a more valid variance estimate, Chiang (1961) created a hypothetical population of size $N_x = (P_x + K(1-a_x)D_x)/K$ and assumed D_x binomial with parameter $n = N_x$. (Here a_x is taken to represent the average number of years lived in the year by an individual who dies.

It is usually taken to be 1/2.) Chiang's estimate of the variance of D_x is then

$$D_x(1 - D_x/N_x) \tag{2.2}$$

In the case that the death rate is low, both Chiang's and the preceding variance estimate are approximately D_x. This last corresponds to Poisson variation. It will be argued in this paper that it is this Poisson estimate that should be employed generally, whether or not the death rate is low. Conditional variances are also presented below.

Death rates are often subjected to regression analyses when explanatory variables are available. The discussion of what is an appropriate variance estimate there becomes the question of what weights to employ in the regression analysis. Some references are: Fryer et al. (1979), Hogan et al. (1979), Pocock, Cook and Beresford (1981). This issue will be returned to later in the paper.

3. A Conceptual Model

An issue that arises with death counts and rates is - are these not facts or exact values and hence subject to no uncertainty? There is clearly a conceptual basis for treating times and lifetimes as random however. Among justifications that may be provided are: moments of birth and death appear unpredictable; there exists an immense biological variability; there exists substantial environmental diversity; there are epidemics; there are medical advances, accidents, violent deaths; periods of extreme weather occur; and researchers have constructed useful chance mechanisms for fertility and diease. At the same time it may be mentioned that there do exist near deterministic aspects, in particular babies are often induced and there do exist seasonal fluctuations. In the framework to be presented both times of birth and lifetimes will be assumed stochastic leading to natural variability of vital statistics.

Before developing sampling results, some notation and assumptions will be set down. It is convenient to display individual's life histories by slope 1 lines in the age time-of-death plane, i.e. via the Lexis diagram. (This technique is discussed and employed in Benjamin and Pollard (1980) for example.) In the diagram the axes have equal scales, lifelines begin at birth and end at death, for a set A of the plane the number of lines ending in A gives the number of deaths in A. Figure 2 provides an example of a Lexis diagram. Let N(A) denote the number of deaths (endpoints) in the region A. Then, for example, the crude death rate for 1980 may be represented by N(A)/N(B) with A and B the regions given in Figure 3.

In setting down definitions and developing results, it will be convenient to make use of the machinery of stochastic point processes. Cox and Lewis (1966) is one reference to this material. Briefly, a <u>linear point process</u> is a random scattering of points along the real line. Its realizations may be denoted by $\{\sigma_j\}$ with σ_j the coordinate value of the j-th point. (Shortly the σ_j will be taken to be the birth times of the population members.) An important parameter of a linear point process is its <u>intensity function</u>, $\beta(\cdot)$, given by

$$\text{Prob}\left\{\text{point in } (t,t+h)\right\} \sim \beta(t)h \tag{3.1}$$

for h small. Supposing I to be an interval, and M(I) to be the number of points in I, then

$$E\left\{M(I)\right\} = \int_I \beta(t)dt = B(I) \tag{3.2}$$

say. The (linear) <u>Poisson process</u> with intensity $\beta(\cdot)$ may now be defined by the requirement that for disjoint intervals $I_1,\cdots,I_K$ the counts $M(I_1),\cdots,M(I_K)$ are independent Poisson variates with means $B(I_1),\cdots,B(I_K)$ respectively, $K = 1,2,\cdots$.

A <u>planar point process</u> is a random scattering of points in the plane. Its <u>intensity function</u> $\lambda(\cdot)$ is given by

$$\text{Prob}\left\{\text{point in } (t,t+h)\times(x,x+h')\right\} \sim \lambda(t,x)hh' \tag{3.3}$$

for h, h' small. If A is a region of the plane and N(A) the number of points in A, then

$$E\left\{N(A)\right\} = \int_A \lambda(t,x)dtdx \tag{3.4}$$

The <u>planar Poisson process</u> with intensity $\lambda(\cdot)$ may now be defined by the requirement that for disjoint regions $A_1,\cdots,A_K$ the counts $N(A_1),\cdots,N(A_K)$ are independent Poisson variates with means $\Lambda(A_1),\cdots,\Lambda(A_K)$ respectively, $K = 1,2,\cdots$. Here

$$\Lambda(A) = E\{N(A)\} = \iint_A \lambda(t,x)\,dt\,dx \quad . \tag{3.5}$$

Properties of the Poisson process include: i) var $N(A) = E\{N(A)\} = \Lambda(A)$, ii) for A contained in a region B the distribution of $N(A)$ given $N(B) = n$ is binomial with parameters n and $\Lambda(A)/\Lambda(B)$.

Returning to the discussion of vital statistics; suppose that the times of birth of the population of concern are $\sigma_1, \sigma_2, \sigma_3,\cdots$. Let $M(I)$ denote the number of σ_j in the interval I . Supposing that $M(I)$ is a stochastic point process, its intensity function $\beta(t)$ will be referred to as the <u>birth intensity</u>. (We remark that, for example, $\beta(t)$ would be periodic were there a seasonal effect present.) Next, suppose that individuals live random lengths of time. Let X denote the lifetime of an individual born at time σ. The distribution of X is conveniently described by the <u>force of mortality</u>, $\mu(t,x)$, defined by

$$\text{Prob}\{x < X \leq x+h \mid X \geq x\} \sim \mu(t,x)h \tag{3.6}$$

with $t = \sigma+x$ and h small. For example, the probability that an individual born at time σ survives to age x is given by

$$\exp\left\{-\int_0^x \mu(\sigma+y,y)\,dy\right\} \tag{3.7}$$

The <u>death process</u> is defined to be the planar point process with points at the positions (date of death, age at death) , specifically supposing the j-th individual is born at time σ_j and dies aged x_j , then N(A) denotes the number of points (σ_j+x_j, x_j) in the region A. It corresponds to the endpoints of the lifelines of the Lexis diagram. Let $\lambda(t,x)$ denote the <u>death intensity</u> (i.e. the intensity of the point process N(.)), then

$$\Lambda(A) = \iint_A \lambda(t,x)\,dt\,dx = E\{N(A)\} \tag{3.8}$$

where

$$\lambda(t,x) = \beta(t-x)\ \mu(t,x)\ \exp\left\{-\int_0^x \mu(t-x+y,y)\ dy\right\} \quad . \tag{3.9}$$

This last follows by first principles, the three factors on the right having the interpretations; "born at t-x","die at t age x","survive to t" respectively.

A theorem describing the distribution of the death process may now be stated. Its proof is given in the Appendix.

<u>Theorem 1</u>. If a) the birth process $\{\sigma_j\}$ is Poisson with intensity $\beta(t)$, b) the lifetimes $\{X_j\}$ of the individuals are independent of each other, independent of the birth process and correspond to the force of mortality $\mu(t,x)$, then the death process N(.) is planar Poisson with intensity (3.9) .

This result may be used to derive the distributions of various vital statistics. It is particularly convenient because for the Poisson process counts corresponding to disjoint regions are statistically independent. It is worth remarking specifically that the resulting Poisson distributions

for death counts arise not from rarity (small numbers) rather from the assumed total randomness (Poisson) of the birth process and the assumed randomness of lifetimes.

The assumptions of Poisson births and independent lifetimes were essential to the derivation of the Poisson conclusion. In fact the birth process may be expected to show some clustering because of twin births, further lifetimes will not be completely independent because of the existence of multiple deaths in accidents. These phenomena may be expected to have small effects generally however.

The following result will be used to set down the distributions of various statistics arising. It follows directly by writing the regions involved in terms of disjoint subregions and the fact that for a Poisson process, counts for disjoint regions are statistically independent.

<u>Corollary</u>. Under the conditions of the theorem: a) $\{N(A),N(B)\}$ is distributed as $\{U+W,V+W\}$ where U, V, W are independent Poissons with means $\Lambda(A-AB)$, $\Lambda(B-AB)$, $\Lambda(AB)$; b) $N(A)/N(B)$ is distributed as $(U+W)/(V+W)$; c) $N(A)$ given $N(B)$ is distributed as $U+S$ where U is Poisson with mean $\Lambda(A-AB)$ and S is independently binomial with $n = N(B)$ and $\pi = \Lambda(AB)/\Lambda(B)$.

4. Some Examples

The preceding theorem and corollary will now be used to set down distributions for various vital statistics.

Example 1 : Crude Death Rate. Let D denote the number of deaths in a given year and P the corresponding mid-year population. Then the crude death rate is D/P . It may be represented as N(A)/N(B) with A, B regions of the Lexis diagram, for example Figure 3 applies to 1980.

Assuming complete randomness of births and independent lifetimes as required in the theorem, it follows that $\{D,P\}$ has a bivariate Poisson distribution. Specifically $\{D,P\}$ is distributed as $\{U+W, V+W\}$ of the theorem, with $\Lambda(\cdot)$ given by (3.8), (3.9) . The crude death rate, D/P, is therefore distributed as (U+W)/(V+W) . (Incidentally, this representation shows that there is a chance that the denominator of this ratio is 0 when the numerator is not.) The bivariate Poisson is discussed in Haight (1967).

On some occasions one is interested in conditional distributions. It follows from the corollary that the distribution of D given P is U+S with U Poisson and S independently binomial. In particular this gives

$$E\{D|P\} = \Lambda(A-AB) + P\frac{\Lambda(AB)}{\Lambda(B)} \tag{4.1}$$

$$\text{var}\{D|P\} = \Lambda(A-AB) + P\frac{\Lambda(AB)}{\Lambda(B)}\left(1 - \frac{\Lambda(AB)}{\Lambda(B)}\right) \quad . \tag{4.2}$$

Restating (4.1), the regression coefficient of D on P is $\Lambda(AB)/\Lambda(B)$. This is initially surprising because the region AB refers only to deaths occurring in the second half of 1980 to persons born before 1 July 1980. The constant term clearly plays an important role in this case of an open population. The relations (4.1), (4.2) may be used to guide regression analyses. The conditional variance of D, and indeed its distribution, is made up of a Poisson and

a binomial part. In the case that the expected number of deaths in A-AB is small, the distribution is approximately binomial and one is led to the traditional assumption of binomial variation. In the case of a closed population, A is contained in B and A-AB is empty and the binomial is exact.

<u>Example 2 : Age-specific Death Rates</u>. The age x death rate has the form $M_x = D_x/P_x = N(A_x)/N(B_x)$ where A_x is the set $(1980,1981)\times(x,x+5)$ and B_x is the set of (t,y) satisfying $1980.5 < t$ and $x \leq y - (t-1980.5) < x+5$. $N(A_x)$, $N(B_x)$ count respectively how many die aged x to x+5 in 1980 and how many were alive and aged x to x+5 on 1 July 1980.5. Because of the (planar) Poisson nature of the death process a variety of distributions are now apparent from the theorem and its corollary. The distribution of D_x is Poisson with mean

$$\Lambda(A_x) = \int_{1980}^{1981} \int_{x}^{x+5} \lambda(t,y)\,dt\,dy \quad . \tag{4.3}$$

Its variance may be estimated by D_x. The distribution of $\{D_x, P_x\}$ is bivariate Poisson. The distribution of D_x given P_x is not generally simple. An approximation to the distribution of M_x will be presented in the next section.

One simple result is that M_x statistics for disjoint age intervals are statistically independent.

<u>Example 3 : Age-standardized Rates</u>. These have the form

$$\sum_x w_x N(A_x)/N(B_x) \tag{4.4}$$

for some given weights w_x. The distribution may be described in terms of Poisson variates; however it is non elementary.

<u>Example 4 : Ratios</u>. This are generally based directly on counts. For example the SMR is given by

$$N(A)/\sum_x M_{xs}N(B_x) \tag{4.5}$$

with the M_{xs} the rates of a selected standard population. The distribution here is clearly messy, although it may be represented directly in terms of statistically independent Poisson variates. An approximation to its variance will be suggested in the next section.

5. Some Discussion.

The principal purpose of this paper is to provide a conceptual basis on which sampling uncertainties of various vital statistics might be derived. Assuming birth times in accordance with a Poisson process, and assuming independent lifetimes, it has been found that the points (time of death, age at death) are distributed in the Lexis diagram in accordance with a planar Poisson process. This means that counts corresponding to disjoint regions of the Lexis diagram are independent Poissons. As many vital statistics may be written as functions of such counts, an expression for their distribution has been constructed. The results obtained differ from those of Chiang (1961) - the results here are simpler. Chiang's results typically involve Poisson terms and correction terms, such as the $(1 - D_x/N_x)$ of expression (2.2) . An implication is that variances computed under the present framework will generally be larger. An extreme case of this is provided by the case of the rate for those aged 85 and over. Chiang (1961), page 281, estimates the variance of this by 0. Here it would be estimated by D_x/M_x^2 .

The exact distribution of an age-specific death rate was seen to involve the bivariate Poisson. This is generally an inconvenient distribution to work with. In the case that the coefficient of variation of the population size is small, as the following theorem shows, an approximation may be employed. A further advantage acruing in this situation is that the particular choice made for the denominator (person-years-lived) is not so crucial. The approximation is to replace the denominator by its expected value.

<u>Theorem 2</u>. Suppose that D is Poisson with mean λ and that P has mean Λ and

variance σ^2, then

$$\left|\mathrm{Prob}\left\{\left(\frac{D}{P} - \frac{\lambda}{\Lambda}\right)\Big/\frac{\sqrt{\lambda}}{\Lambda} \leq x\right\} - \mathrm{Prob}\left\{\left(\frac{D}{\Lambda} - \frac{\lambda}{\Lambda}\right)\Big/\frac{\sqrt{\lambda}}{\Lambda} \leq x\right\}\right|$$

$$\leq \quad 3\left(\frac{\sigma}{\Lambda}(\sqrt{\lambda} + x)\right)^{2/3} + \frac{1}{\sqrt{\lambda}} \qquad \text{for } \sqrt{\lambda} + x > 0 \quad .$$

The proof of this result is given in the Appendix. Basically one wants the coefficient of variation of P, σ/Λ, to be small.

This result leads one, for example, to estimate the variance of the crude death rate D/P by D/P^2; to estimate the variance of an age-specific death rate $M_x = D_x/P_x$ by D_x/P_x^2; to estimate the variance of an age-standardized rate, $\Sigma\, w_x M_x$, by $\Sigma\, w_x^2 D_x/P_x^2$; and to estimate the variance of a standardized mortality ratio, $D/\Sigma\, M_{xs}P_x$, by $D/(\Sigma\, M_{xs}P_x)^2$. This is to be contrasted with the formula resulting from Chiang (1961), expression (18), namely $\Sigma\, w_x^2 D_x(1 - D_x/N_x)/P_x^2$. This last is smaller, particularly when D_x/N_x is not small.

In some situations one may have a parametric model of interest. One may then be able to set down a likelihood function and proceed to compute say maximum likelihood estimates. In particular cases that likelihood may factor and the term involving the population size separate, leading one to make inferences conditionally on it. This happens, for example, in the case of a closed population.

Another situation in which things simplify is when the individuals' person-years-lived values are known. Hoem (1984) discusses this case and presents variance estimates.

The assumed variability of the birth process was basic to obtaining the Poisson conclusion. In some circumstances, eg. manpower studies, there will be steady (deterministic) recruitment to the population and the results derived here will be inappropriate.

6. Regression

In many studies of mortality, measured explanatory variables are available. The most common of these are age and (time) period. Others include: cause, sex, region. An individual's mortality may be expected to depend on various of these. The measurements may be included, in a quantitative manner, by setting down a functional form for the force of mortality or related parameters. In this section, the case of Poisson regression will first be mentioned, then the case of extra-Poisson variation will be studied.

6a. Poisson Regression. The conceptual model of mortality, presented in this paper, led to a Poisson distribution for the number of deaths. In the case that the population size is P and that a (vector-valued) explanatory variable $\underset{\sim}{x}$ is available, one might assume that the number of deaths, D, is Poisson with mean $P\exp\{\underset{\sim}{x}'\underset{\sim}{\beta}\}$, $\underset{\sim}{\beta}$ being a parameter to be estimated. For example, Frome (1983) sets down such a model for lung cancer deaths of British physicians taking P to be man-years at risk and as explanatories years of smoking and number of cigarettes per day. The Poisson model is found to fit well - a deviance of 51.47 based on 48 degrees of freedom obtained.

6b. Extra-Poisson Variation. As is often the case in ordinary regression analyses, it is to be expected that in many situations essential explanatory variables will not have been observed. Were they all available, a rate $\exp\{\underset{\sim}{x}'\underset{\sim}{\beta}\}$ might be appropriate. In the case of omitted explanatories, we are led to consider a rate

$$\exp\{\underset{\sim}{x}'\underset{\sim}{\beta}+\varepsilon\} \tag{6.1}$$

with ε normal, mean 0 and variance σ^2. The parameter σ^2 provides a

measure of the extra-Poisson variation. Following the work of Hinde (1982) and Brillinger and Preisler (1983), the maximum likelihood estimates of $\underset{\sim}{\beta}$ and σ here may be determined by a combination of numerical integration and the EM algorithm. Those papers provide glim macros and sessions illustrating the technique. That the error ε has been assumed normal is not crucial, rather the distribution of the error should be known up to a finite dimensional parameter.

Briefly the approach is as follows: let U denote a latent variate with density function $f(u|\beta)$ depending on a parameter β. (In the present case U is ε and β is σ.) Let Y be an observable variate with probability mass (or density) function, given $U = u$, $f(y|u,\alpha)$ depending on the parameter α. (In the present case Y is Poisson and α is $\underset{\sim}{\beta}$.) Then the marginal probability mass function of Y is

$$f(y|\alpha,\beta) = \int f(y|u,\alpha)f(u|\beta)du \tag{6.2}$$

Let $\theta = (\alpha,\beta)$ and

$$\psi(y|\theta) = \frac{\partial \log f(y|\theta)}{\partial \theta} \tag{6.3}$$

Supposing that observations $y_1,\cdots,y_n$ are available, the maximum likelihood equation for estimating θ is given by

$$\sum_{i=1}^{n} \psi(y_i|\hat{\theta}) = 0 \tag{6.4}$$

Elementary manipulations allow this last to be written

$$\sum_{i=1}^{n} \int \psi(u|\hat{\beta})f(y_i|u,\hat{\alpha})f(u|\hat{\beta})du/f(y_i|\hat{\alpha},\hat{\beta}) = 0 \tag{6.5}$$

The difficulty that arises in many cases, particularly the present one, is that the integration in (6.5) may not be carried out analytically. The approach is to carry it out numerically, replacing the probability element $f(u|\beta)du$ by a discrete approximation

$$f(u|\beta)du = \sum_{m=1}^{M} p_m \delta\{u - u_m\} \tag{6.6}$$

$\delta\{u\}$ denoting unit mass at $u = 0$. The u_m are nodes and the p_m are corresponding weights. This all leads to the approximate likelihood equations

$$\begin{aligned} \sum_{i=1}^{n} \sum_{m=1}^{M} \mathcal{H}(y_i|u_m,\hat{\alpha})w_m(y_i|\hat{\alpha},\hat{\beta}) &= 0 \\ \sum_{i=1}^{n} \sum_{m=1}^{M} \mathcal{H}(u_m|\hat{\beta})w_m(y_i|\hat{\alpha},\hat{\beta}) &= 0 \end{aligned} \tag{6.7}$$

where the w_m are weight functions given by

$$w_m(y|\alpha,\beta) = f(y|u_m,\alpha)p_m / \sum_{m=1}^{M} f(y|u_k,\alpha)p_k \tag{6.8}$$

These equations are conveniently solved iteratively.

The procedure will now be illustrated by two sets of computations. One set involves the fitting of a dynamic (time series) model to the historial data on Canadian female mortality given in Figures 1. However the first set refers to a data set having both age and period as explanatories. The computations made use of Gauss-Hermite integration with 11 nodes, see David and Rabinowitz (1975) for the formulas. The standard errors were estimated as in Brillinger and Preisler (1983).

The data of the first example are female deaths by age group for Canada during the period 1950 - 1972 and the corresponding age-specific death rates. The data are given in Tables 5 and 6 of Statistics Canada (1976). They are here displayed in Figures 4 and 5. In those figures, the diameter of the circle plotted at a given(age, year)position is proportional to the number of deaths in Figure 4 and the rate in Figure 5 . Examination of the figures shows: high death counts and rates for the 0-1 age group, with both falling as time passes. It shows death counts at the high age groups increasing, (the population size is steadily increasing, see Figure 1), but death rates falling.

Let D_{ij} denote the number of deaths in age group i for year j and let P_{ij} denote the corresponding (midyear) population. The model fit is one of D_{ij} given ε_{ij} being Poisson of rate

$$P_{ij}\exp\{\alpha_i + \beta_j + \varepsilon_{ij}\} \tag{6.9}$$

with the ε_{ij} independent normals of mean 0 and variance σ^2 . The α_i and β_j are age and period effects respectively. The deviance obtained for a pure Poisson fit ($\sigma^2 = 0$) was 4619. on 396 degrees of freedom. The deviance with the extra-Poisson variation was 1452. on 395 degrees of freedom - a substantial reduction for the inclusion of a single further parameter. It is to be expected that the deviance may be driven down substantially further by including further explanatories, for example a cohort effect; however the purpose of the present study was principally to illustrate that mortality data can be non-Poisson and that a direct procedure was available to handle the extra-variation.

Figures 6 and 7 provide the estimated age and period effects, $\hat{\alpha}_i$ and $\hat{\beta}_j$. (Actually the model was reparametrized to $\gamma + \alpha_i + \beta_j$, with $\alpha_1, \beta_1 = 0$ to avoid aliasing.) The age effects show a "bathtub" shape - corresponding to high mortality at the youngest ages, then a dip followed by a steady increase with age. On the other hand, the period effects evidence a steady decrease in mortality with time.

The fit of a model is conveniently studied by the standardized residuals, as well as the deviance. These are defined as $(D - \hat{\mu})/\hat{\Sigma}$, where $\hat{\mu}$, $\hat{\Sigma}$ are estimates of $E\{D\}$ and $\sqrt{\text{var } D}$ under the model being considered. Figure 8 is an estimate of the density of the standardized residuals under the Poisson model. (The estimate was computed via the procedure "density" of Becker and Chambers (1984).) The distribution is exceedingly broad. Figure 9 is the estimated density for the model (6.9). This figure provides further evidence of a substantial improvement in fit being provided by the model with extra-Poisson variation. Examination of the residuals themselves brought out the presence of a clear outlier in the published values (Table 6, Statistics Canada (1976)) - namely the value 2.9 for those 35-39 in 1951. The estimate of σ was .064 .

It is worth remarking that when the present extended model was fit to the Frome data of 6a above, there was no reduction in the deviance - it just fluctuated with round-off error.

Our second example involves a time series modelling of the data, on all Canadian female deaths during the time period 1926 to 1982, presented earlier in Figure 1 and taken from Table 1 of Statistics Canada (1976) and a supplement provided by D. Nagnur. The model fit is analagous to an autoregressive process of order 1. Let d_t denote the number of deaths in time period t and let p_t denote a corresponding measure of population size. Let m_t denote the latent death rate at time t and suppose that it evolves in accordance with

$$\log m_t = \alpha + \beta \log m_{t-1} + \varepsilon_t \qquad (6.10)$$

the ε_t being independent normal variates with mean 0, variance σ^2 and suppose further that given m_t and the past, d_t is distributed as Poisson with mean $p_t m_t$. A model of this sort may be expected to be of some use in forecasting.

The model (6.10) was fit by maximum likelihood as in the earlier example. The deviance for a Poisson model ($\sigma^2 = 0$) was 1844. based on 54 degrees of freedom. For the dynamic model it was 276.6 based on 53 degrees of freedom - a substantial improvement in fit. The estimates and their estimated standard errors were $\hat{\alpha} = .062$ (.007), $\hat{\beta} = .964$ (.003), $\hat{\sigma} = .0168$ (.0005) . the first two were highly correlated. Figure 10 is a plot of the (conditionally) standardized residuals versus time. It evidences a definite suggestion of the variability reducing with time.

7. Concluding Remarks

The goals of this paper have been to provide a conceptual basis for the description of the natural variability of certain vital statistics and to make use of that description in the analysis of two data sets. It was found that under two elementary assumptions (one re the birth process, the other re lifetimes) that basic counts of deaths were Poisson, with those corresponding to disjoint regions of the Lexis diagram independent. It was further demonstrated that sometimes, perhaps because of omitted explanatory variables, Poisson variability was insufficient. A general model involving extra variability was set down and fit to the two data sets. These data sets were found to evidence substantial variability beyond the Poisson.

A continuing issue in analyses of mortality rates, with measured explanatory variables, by linear regression has been: what are the appropriate weights for the observations. Different choices are made in Fryer et al. (1979), Hogan et al. (1979) and Pocock et al. (1982) for example. Employing a full likelihood analysis, as is proposed in this paper, is clearly an alternate way to address the issue. Noting that the present computations were in fact carried out by iteratively reweighted least squares makes the connection even more apparent.

Finally it is to be noted that this paper has taken the basic quantities to be analyzed ~~to~~ be simple counts and rates. Clearly other quantities, perhaps specific estimates of probabilities as in Hoem (1984) or subtle variants such as the Mosteller (1969) rate $D/(P + cD)$ will be of interest.

ACKNOWLEDGEMENTS

The author would like to thank J. W. Tukey for getting him interested in this problem. He would further like to thank C-L. Chiang, H. Preisler, J. Fryer and J. W. Tukey for comments on the work. Finally, he would like to thank D. Nagnur for providing the Canadian data analysed. The research was supported in part by the National Science Foundation Grant DMS-8316634. Most of the figures included were made via the S system of Becker and Chambers (1984).

REFERENCES

Becker, R. A. and Chambers, J. M. (1984). S. Belmont: Wadsworth.

Benjamin, B. (1968). Health and Vital Statistics. London: G. Allen and Unwin.

Benjamin, B. and Pollard, J. H. (1980). The Analysis of Mortality and Other Actuarial Statistics. London: Heinemann.

Brillinger, D. R. and Preisler, H. K. (1983). Maximum likelihood estimation in a latent variable problem. pp. 31-65 in Studies in Econometrics, Time Series, and Multivariate Statistics (Eds. T. Amemiya and S. Karlin). New York: Academic.

Chiang, C. L. (1961). Standard error of the age-adjusted death rate. Vital Statistics-Special Reports, Vol. 47, No. 9. Washington: Public Health Service.

Cox, D. R. and Lewis, P. A. W. (1966). The Statistical Analysis of Series of Events. London: Methuen.

Daley, D. J. and Vere-Jones, D. (1972). A summary of the theory of point processes. pp. 299-383 in Stochastic Point Processes (ed. P. A. W. Lewis). New York: J. Wiley.

Davis, P. J. and Rabinowitz, P. (1975). Methods of Numerical Integration. New York: Academic.

Doob, J. L. (1953). Stochastic Processes. New York: Wiley.

Fleiss, J. L. (1981). Statistical Methods for Rates and Proportions. New York: Wiley.

Frome, E. L. (1983). The analysis of rates using Poisson regression models. Biometrics 39, 665-674.

Fryer, J. G., Harding, R. A., Macdonald, M. D., Read, K. L. Q., Crocker, G. R. and Abernathy, J. (1979). Comparing the early mortality rates of the local authorities in England and Wales. J. R. Statist. Soc. A 142, 199-209.

Haight, F. A. (1967). Handbook of the Poisson Distribution. New York: Wiley.

Hinde, J. (1982). Compound Poisson regression models. pp. 109-121 in GLIM82 (Ed. R. Gilchrist). Lecture Notes in Statistics 14. New York: Springer.

Hoem, J. M. (1984). A flaw in actuarial exposed-to-risk theory. Scand. Actuarial J. 187-194.

Hogan, M. D., Chi, P-Y., Hoel, D. G. and Mitchell, T. J. (1979). Association between chloroform levels in finished drinking water supplies and various site-specific cancer mortality rates. J. Env. Path. and Tox. 2, 873-887.

Inskip, H., Beral, V. and Fraser, P. (1983). Methods for age-adjustment of rates. Statistics in Medicine 2, 455-466.

Keyfitz, N. (1966). Sampling variance of standardized mortality rates. Hum. Biol. 38, 309-317.

Mosteller, F. M. (1969). Estimation of death rates. pp. 234-235 in The National Halothane Study (Ed. J. P. Bunker). Bethesda: National Institutes of Health.

Mosteller, F. and Tukey, J. W. (1977). Data Analysis and Regression. Reading: Addison-Wesley.

Statistics Canada (1976). Vital Statistics: General Mortality 1950-1972. Ottawa: Statistics Canada.

Tsaregradskii, I. P. (1958). On uniform approximations of the binomial distribution by infinitely divisible laws. Theory Prob. Appl. 10, 472-479.

Vere-Jones, D. (1968). Some applications of probability generating functionals to the study of input-output streams. J. Roy. Statist. Soc. B 30, 321-333.

APPENDIX

The proof of Theorem 1 will proceed via the method of probability generating functionals. The pertinent methodology may be found in Vere-Jones (1968) and Daley and Vere-Jones (1972). In particular it should be noted that for a general stochastic point process, with points located at positions $\underset{\sim}{r}_j$ the p. g. fl. is defined to be

$$E\left\{\prod_j \xi(\underset{\sim}{r}_j)\right\}$$

for a general function $\xi(\cdot)$. The p. g. fl. characterizes a point process. The p. g. fl. of a Poisson process with intensity function $\nu(\cdot)$ is given by

$$\exp\left\{\int[\xi(\underset{\sim}{r}) - 1]\nu(\underset{\sim}{r})\, dr\right\} \quad .$$

<u>Proof of Theorem 1</u>. The ends of lifelines in the Lexis diagram occur at the positions (σ_j+X_j, X_j) with σ_j denoting the birth time and X_j the lifetime of the j-th individual. The p. g. fl. of the death process is therefore

$$E\left\{\prod_j \xi(\sigma_j+X_j, X_j)\right\} \quad .$$

Now

$$\underset{X}{E}\left\{\xi(t+X,X)\right\} = \int \xi(t+x,x)\, \mu(t+x,x)\, \exp\left\{-\int_0^x \mu(t+x-y,y)dy\right\}dx = \eta(t)$$

say, with X denoting the lifetime of an individual born at time t. It has been assumed that the σ_j correspond to a Poisson process of intensity $\beta(\cdot)$. Therefore

$$E\left\{\prod_j \eta(\sigma_j)\right\} = \exp\left\{\int[\eta(t) - 1]\beta(t)dt\right\} \quad .$$

Combining these last expressions one sees that the death process is (planar) Poisson with the indicated intensity.

Doob (1953), pp. 405 - 407, is an early reference to this type of result - that random translations of Poisson processes are themselves Poisson.

<u>Proof of Theorem 2</u>. Let

$$X = \left(\frac{D}{\Lambda} - \frac{\lambda}{\Lambda}\right) / \frac{\sqrt{\lambda}}{\Lambda}$$

and

$$\varepsilon = (\sqrt{\lambda} + x)\left(\frac{P}{\Lambda} - 1\right)$$

then elementary manipulations show the quantity to be bounded is

$$\begin{aligned} &\leq \text{Prob}\left\{x - |\varepsilon| < X \leq x + |\varepsilon|\right\} \\ &\leq \text{Prob}\left\{|\varepsilon| > \delta\right\} + \text{Prob}\left\{|X - x| \leq \delta\right\} \end{aligned} \tag{A.1}$$

for any $\delta > 0$. Now, by Tchebyceff's Inequality, the first probability here is

$$\leq \text{var}\,\varepsilon/\delta^2 = (\sqrt{\lambda} + x)^2\sigma^2/\delta^2\Lambda^2 \quad . \tag{A.2}$$

For D a Poisson variate of mean λ it is the case that

$$|\text{Prob}\{D \leq u\} - \text{Prob}\{D \leq v\}| \leq (u - v)c/\sqrt{\lambda} \tag{A.3}$$

for u,v integers $u > v$ with $c = \pi\sqrt{2\pi}/8$. This comes from Theorem 2 Tsaregradskii (1958) and from bounding the absolute value of the characteristic function of D by $\exp\{-2\lambda t^2/\pi^2\}$.

The second probability in (A.1) for $\sqrt{\lambda} + x > 0$ is $\leq$ Prob $\{v < D \leq u\}$ with $u - v \leq 2\delta\sqrt{\lambda} + 2$. The result of the theorem now follows by adding (A.3) and (A.2) and then choosing δ to give the smallest total.

FIGURE CAPTIONS

Figure 1. Top graph provides the estimated 1 June number of females in Canada for 1926 to 1982. Middle graph provides the year total number of female deaths for the same time period. Bottom graph is the ratio of the previous two, the crude death rate.

Figure 2. A Lexis diagram, with the sloping lines representing individual's lifetimes. The lines begin at the moment of birth and end at death. Those ending in the region A represent individuals dying in the corresponding age and time intervals.

Figure 3. Lexis diagrams representing the crude death rate, N(A)/N(B) for the year 1980 taking N(B) to be the midyear, i.e. 1980.5, population.

Figure 4. A circle diagram to represent the counts of females dying annually from 1950 to 1972 separated into 19 age groups (age intervals: 0-1, 1-4, 5-9, 10-14,..., 80-84, 85+). The radius of the circle plotted is proportional to the corresponding count.

Figure 5. A circle diagram of the age-specific death rates corresponding to Figure 4.

Figure 6. The estimated age effects for the data of Figure 4 and the model (6.9). The values plotted are defined so the age 0-1 value is 0.

Figure 7. The estimated period (year) effects for the data of Figure 4, defined so the 1950 value is 0.

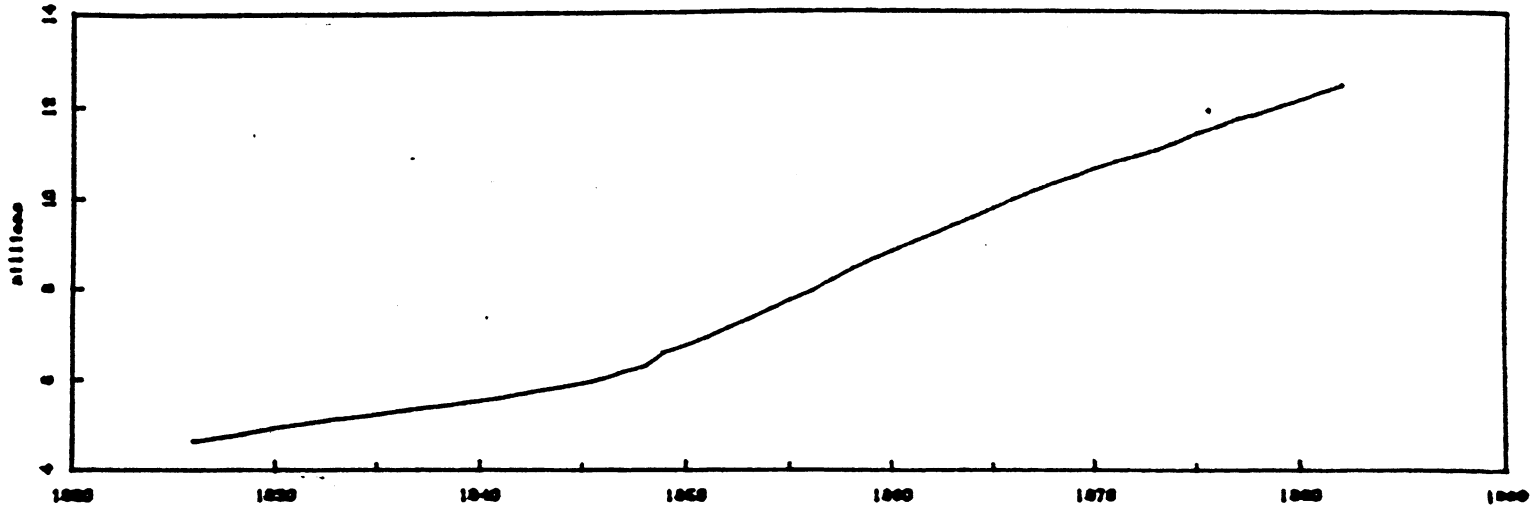
Figure 8. An estimate of the density function of the standardized residuals resulting from fitting a simple ($\sigma = 0$) Poisson model to the data of Figure 4.

Figure 9. An estimate of the density function of the standardized residuals resulting from fitting the model (6.9) to the data of Figure 4.

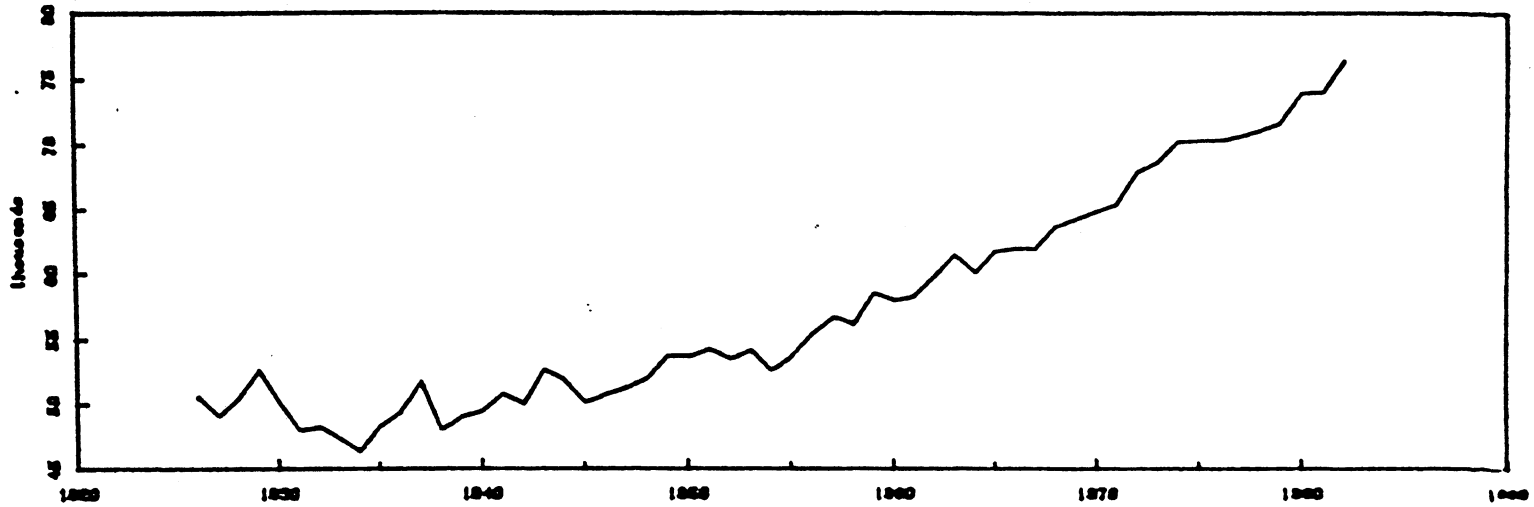
Figure 10. The standardized residuals, plotted versus year, resulting from fitting the dynamic model (6.10) to the total counts of Canadian female deaths during 1926 to 1982.

CANADIAN MORTALITY (F), 1926-1982

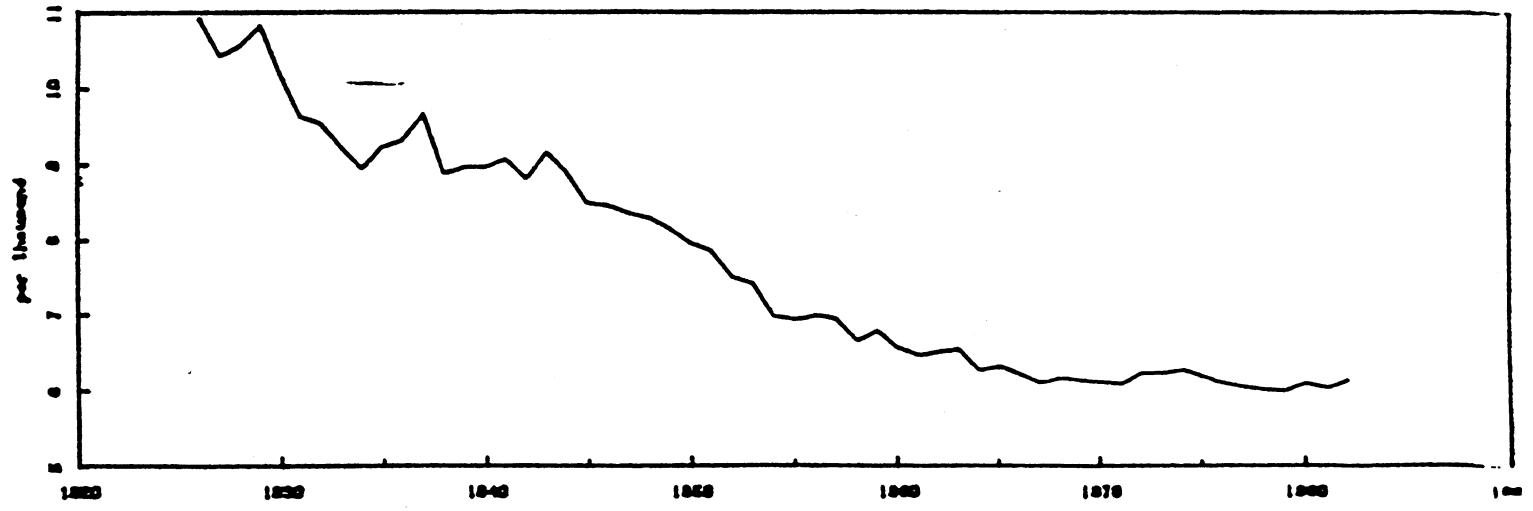
Population

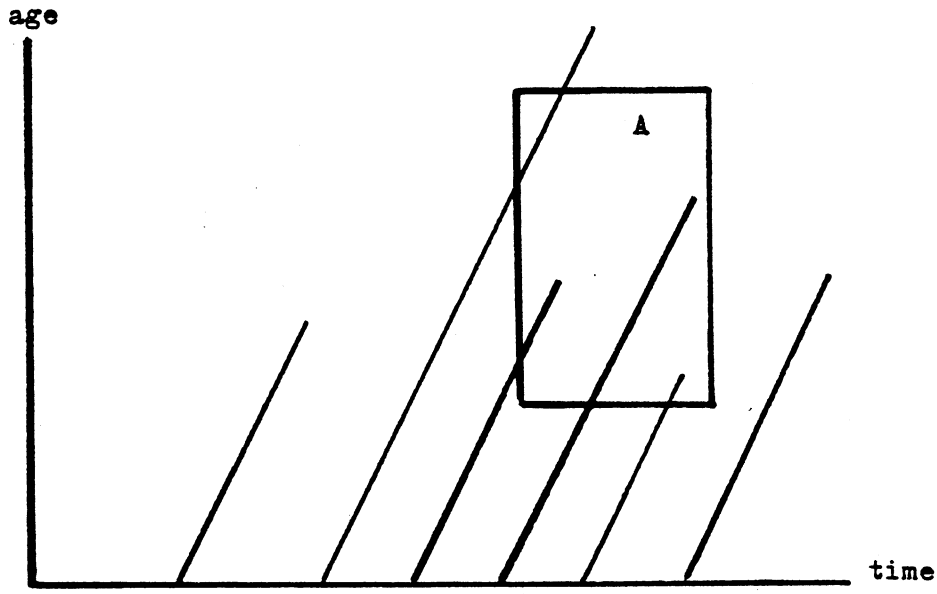


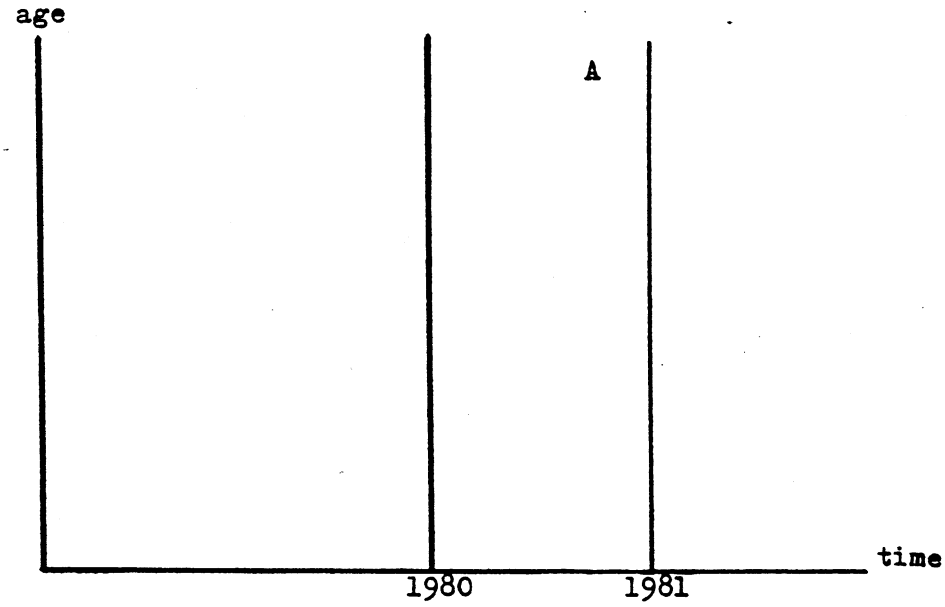
All Deaths

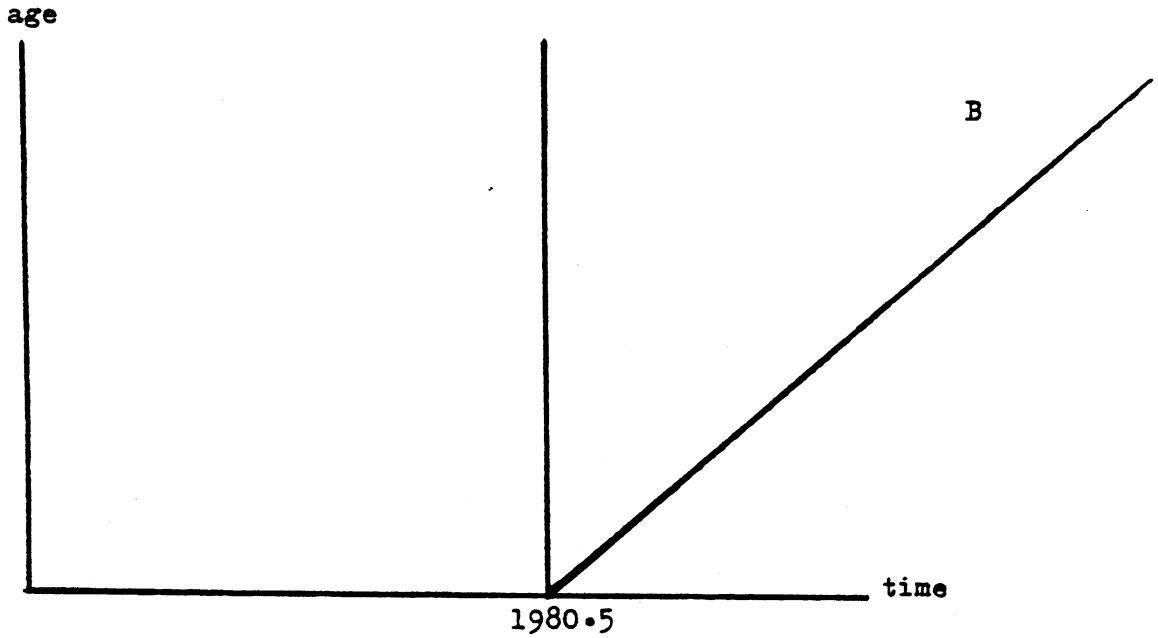


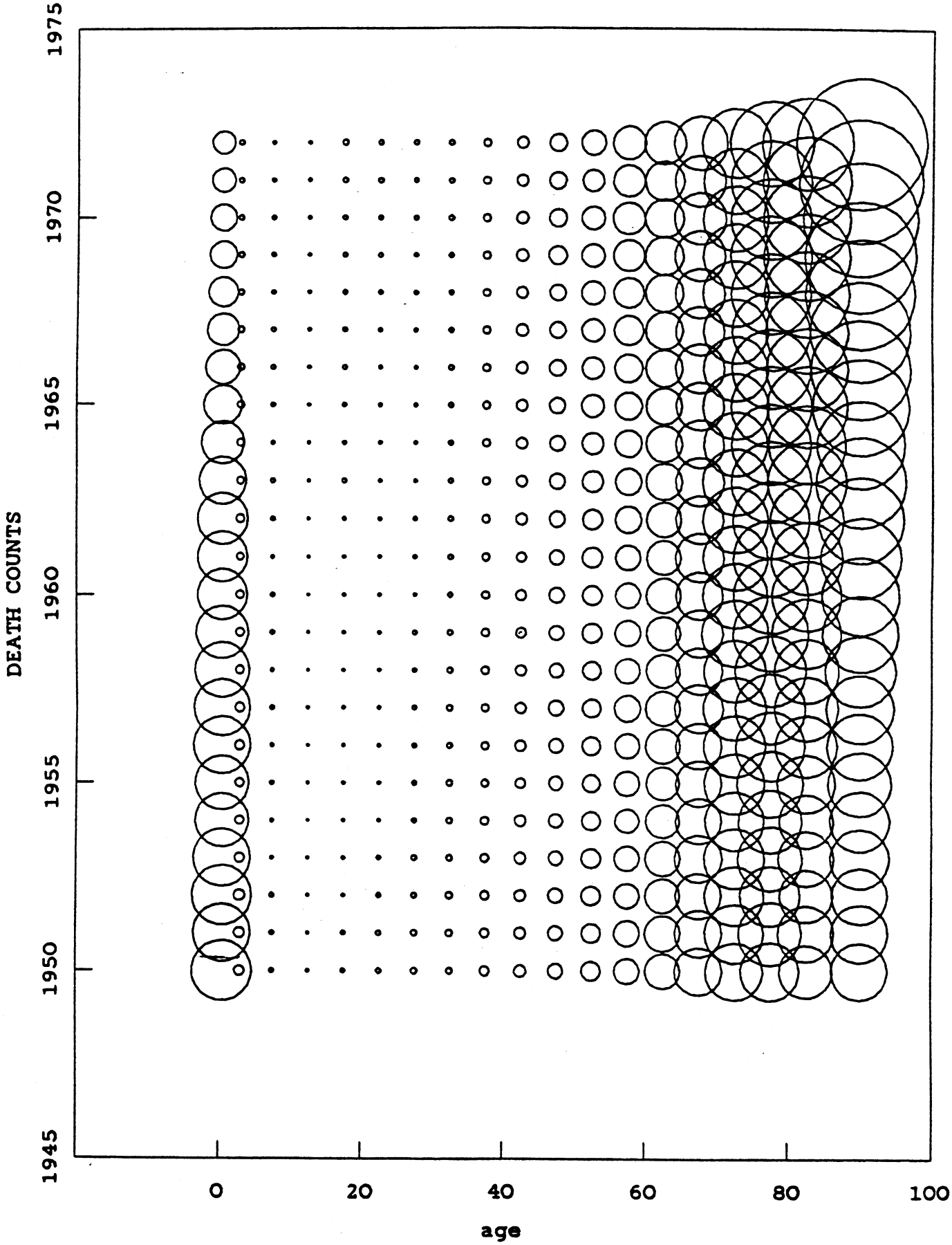
Crude Death Rate




age
A
time


age
A
time
1980
1981


age
B
time
1980.5


DEATH COUNTS
1945
1950
1955
1960
1965
1970
1975
0
20
40
60
80
100
age

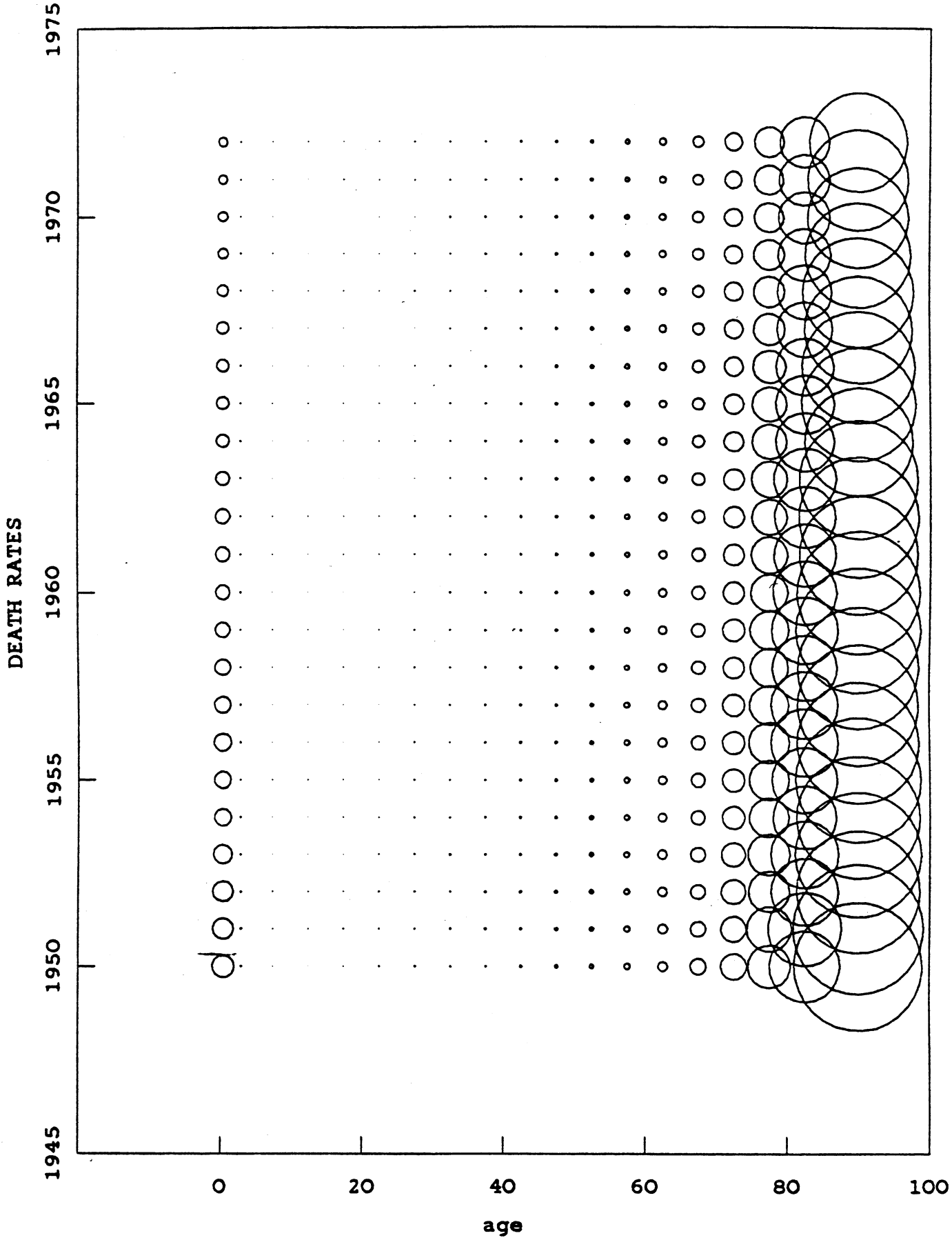

DEATH RATES
1945
1950
1955
1960
1965
1970
1975
0
20
40
60
80
100
age

Fig. 6

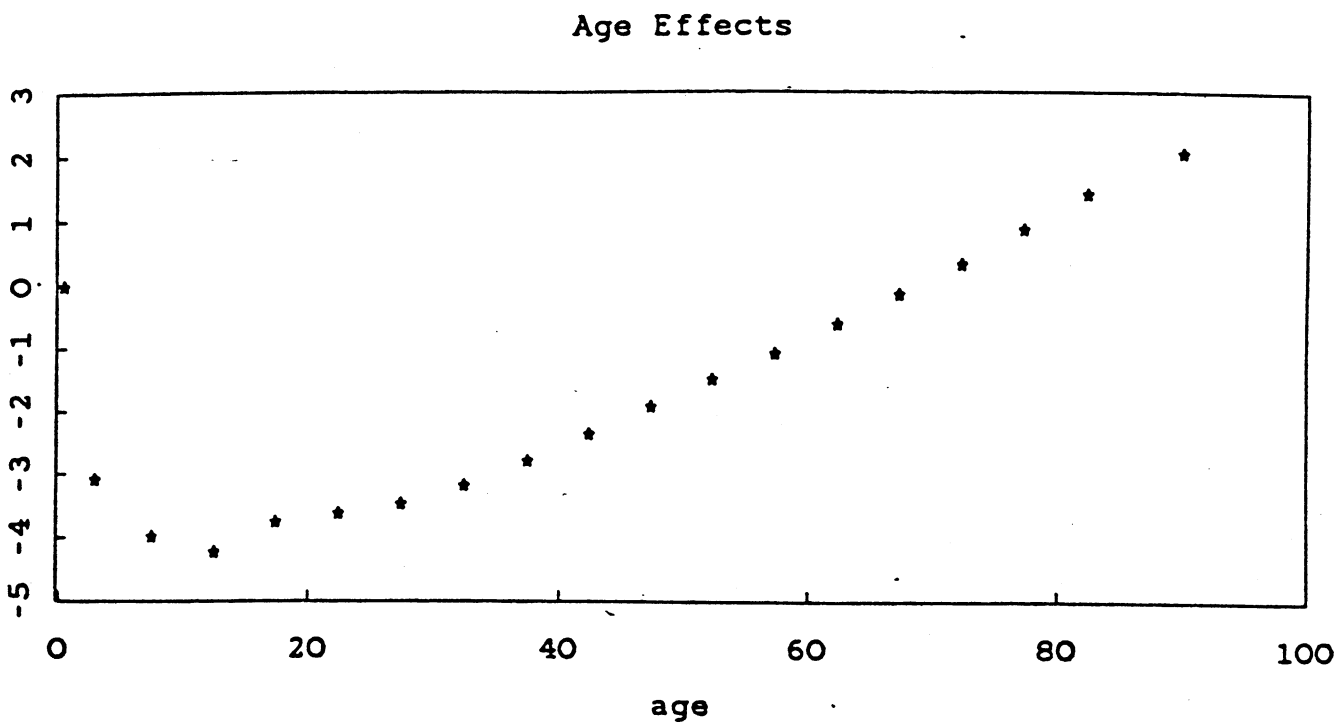
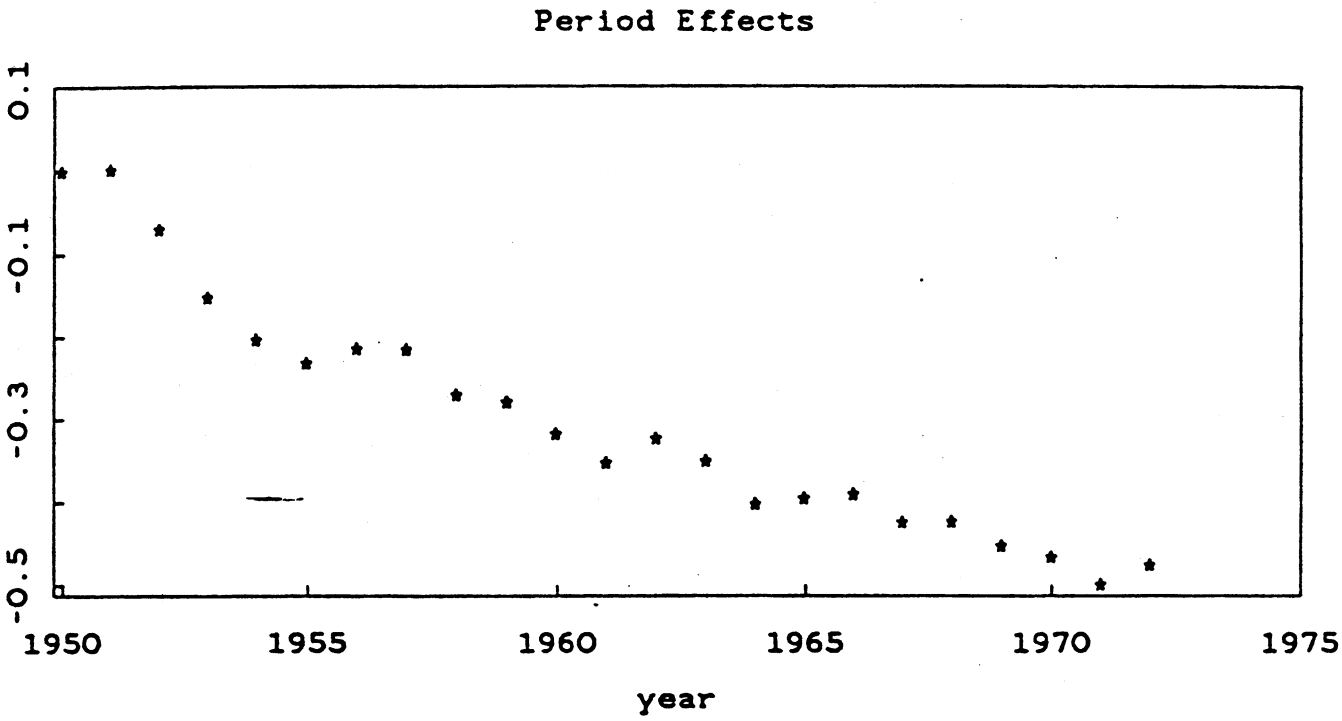

Age Effects
3
2
1
0
-1
-2
-3
-4
-5
0
20
40
60
80
100
age

Fig. 7


Period Effects
0.1
-0.1
-0.3
-0.5
1950
1955
1960
1965
1970
1975
year

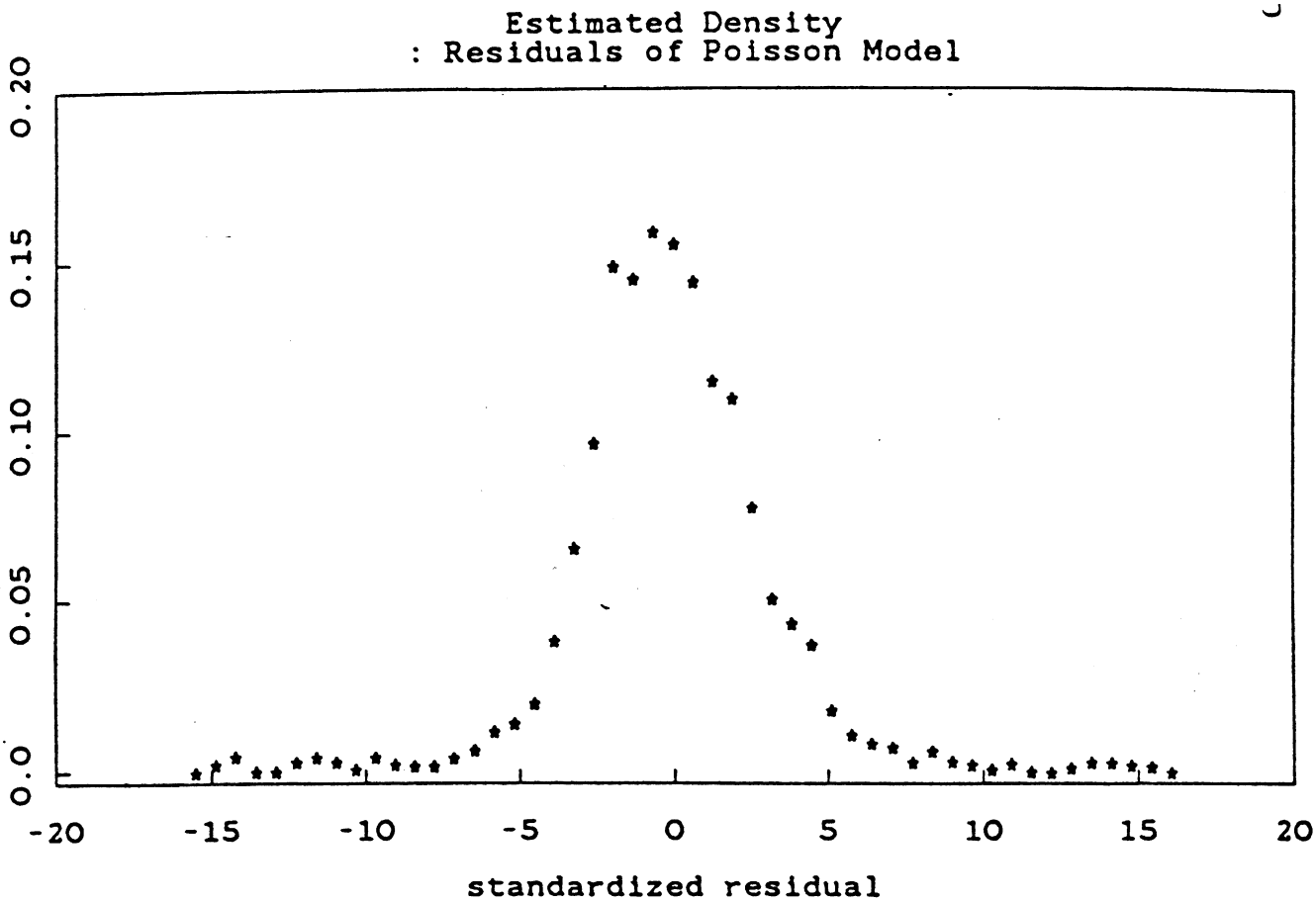

Estimated Density
: Residuals of Poisson Model
0.20
0.15
0.10
0.05
0.0
-20
-15
-10
-5
0
5
10
15
20
standardized residual

Fig. 9

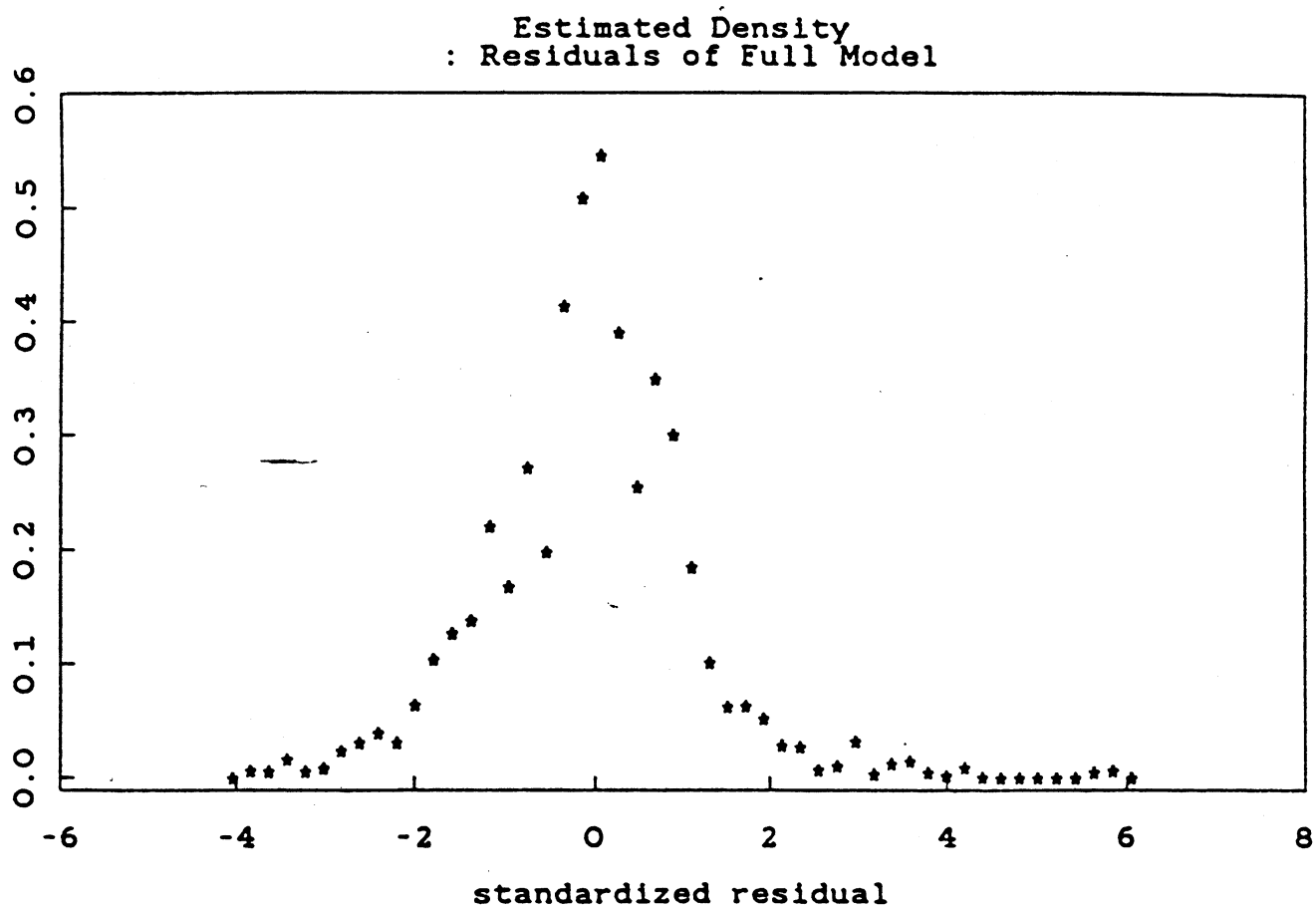

Estimated Density
: Residuals of Full Model
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-6
-4
-2
0
2
4
6
8
standardized residual

Fig. 5

Standardized Residuals