



Sampling Designs for Short Panel Data

Bianca L. De Stavola

Econometrica, Vol. 54, No. 2 (Mar., 1986), 415-424.

Stable URL:

<http://links.jstor.org/sici?sici=0012-9682%28198603%2954%3A2%3C415%3ASDFSPD%3E2.0.CO%3B2-7>

Econometrica is currently published by The Econometric Society.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/econosoc.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact jstor-info@umich.edu.

SAMPLING DESIGNS FOR SHORT PANEL DATA

BY BIANCA L. DE STAVOLA¹

Some aspects of the design and analysis of short panel data are considered. It is assumed that each individual may be at any instant in one of two states, for example employed and unemployed, and that individuals may switch from one to the other state. Alternative ways of observing individuals in discrete time are compared via the Fisher information matrix. The effect of the observational interval on the efficiency of the estimation is stressed and optimal time intervals for alternative sampling schemes are computed. A generalization to a model with covariates is outlined.

1. INTRODUCTION

PANEL DATA are becoming increasingly common in economic and social studies (see Heckman and Flinn [5], Lancaster [6], Lancaster and Nickell [7]). In this paper we show that it is possible to design equivalently efficient sampling schemes for the analysis of two-state Markov processes, so that survey costs may be controlled. If, for example, the aim of a study is the estimation of the proportion of time spent in unemployment, criteria are given to balance the amount of information to be recorded at each time point (e.g., "Are you unemployed at present?" or "How many unemployment spells did you experience over the last year? and Are you unemployed now?"), the dimension of the sample size, and the interval between observations in order to achieve a desired degree of efficiency.

Consider a population of individuals each of whom is at any instant in one of two states, labelled 0 and 1. For example, the states might represent unemployment and employment. We assume that individuals switch from time to time from one state to the other and, initially, take the individuals as homogenous: the dependence of the transition rates on explanatory variables is an important issue and will be considered in the last section. We assume also that the system is stationary and, at least in the initial analysis, that the times spent in state 0 or in state 1 are independent exponentially distributed random variables with parameters ρ_0 and ρ_1 , respectively. Thus the equilibrium probability of being in state 0 is $\rho_1/(\rho_0 + \rho_1)$.

Suppose that a random sample of n individuals is chosen. Ideally, we should monitor the individual transitions from one state to the other separately for each individual continuously over a period of time, say T . When, as in most empirical studies of social behavior, this is not feasible, observations are made at discrete time points. We assume that k observations are made for each individual in the sample, Δ units of time apart, where $(k-1)\Delta = T$ is the observed period. While in practice Δ is not necessarily constant, for simplicity it will be held so throughout this paper. Two questions arise: (a) what kind of information do we need at each

¹ I would like to express my gratitude to Professor D. R. Cox for his guidance throughout the preparation of this paper and to Professor S. Nickell for stimulating discussions. I would also like to thank a referee for helpful comments. The work, supported by Consiglio Nazionale delle Ricerche, Italy, was carried out at Imperial College, London.

point-in-time observation to obtain estimates with a desired degree of accuracy (for given n); and, in the same context (b) what is an optimal value for Δ .

We restrict our discussion to the case of small k . Among the possible ways of obtaining information we will discuss two in particular: (1) the recording of the states occupied by each individual at each sampling point; (2) as in (1), with the inclusion of some qualitative information about the process development over each Δ interval, e.g. the number or the sequence of changes of state.

In order to determine when survey designs more complicated than (1) are worth considering, we compare the estimating efficiencies achieved by these observational schemes with: (3) the continuous time recording of the individual processes over the whole period T .

In Section 2 the likelihood functions for (1)–(3) are computed; in Section 3 the asymptotic relative efficiencies (ARE's) are discussed in the light of different objectives and the time interval effect is explicitly considered. Section 4 generalizes the results to the case of heterogeneous individuals.

2. SCHEMES CONFIGURATION

The assumptions that the population is homogeneous and that the waiting times to switch into one state or the other are independently exponentially distributed imply that for each individual who is in state s at time t the hazard function is affected neither by the length of time already spent in that state nor by the previous sequence of transitions between the two states.

Denote by $X(t)$ the state occupied by individual i at time t . Then, for $k=2$, the sampling scheme (1) will record for each sampled individual one of four possible sequences: $C_{00} = (X(0)=0, X(\Delta)=0)$; $C_{01} = (X(0)=0, X(\Delta)=1)$; $C_{11} = (X(0)=1, X(\Delta)=1)$; $C_{10} = (X(0)=1, X(\Delta)=0)$. The associated probabilities are (Cox and Miller [3, p. 172]):

$$(2.1) \quad \begin{aligned} P(C_{00}) &= \frac{\rho_1}{\rho_0 + \rho_1} \left\{ \frac{\rho_0}{\rho_0 + \rho_1} e^{-(\rho_0 + \rho_1)\Delta} + \frac{\rho_1}{\rho_0 + \rho_1} \right\}, \\ P(C_{01}) &= \frac{\rho_1}{\rho_0 + \rho_1} \left\{ \frac{\rho_0}{\rho_0 + \rho_1} - \frac{\rho_0}{\rho_0 + \rho_1} e^{-(\rho_0 + \rho_1)\Delta} \right\}, \\ P(C_{11}) &= \frac{\rho_0}{\rho_0 + \rho_1} \left\{ \frac{\rho_1}{\rho_0 + \rho_1} e^{-(\rho_0 + \rho_1)\Delta} + \frac{\rho_0}{\rho_0 + \rho_1} \right\}, \\ P(C_{10}) &= \frac{\rho_0}{\rho_0 + \rho_1} \left\{ \frac{\rho_1}{\rho_0 + \rho_1} - \frac{\rho_1}{\rho_0 + \rho_1} e^{-(\rho_0 + \rho_1)\Delta} \right\}, \end{aligned}$$

where the equality of $P(C_{01})$ and $P(C_{10})$ derives from the time reversibility inherent in this Markov process. In each equation, the first factor is the equilibrium probability for the initial state and the second factor corresponds to the conditional probability of being or not being in the same state at the end of the Δ interval.

The extension to the case of 3 or more point-in-time observations is straightforward. When $k=3$ the probabilities associated to each outcome at time 2Δ ,

conditionally on $X(\Delta)$, are exactly equal to the conditional probabilities in the equations (2.1). Thus the full probabilities associated to the 8 possible outcomes are computed.

Similar results are obtained for the sampling scheme (2). As a special case, let m equal one if there has been at least one change of state in the time interval Δ , $m = 0$ otherwise. Then when $k = 2$ we define 6 possible outcomes: $E_{00,0} = (X(0) = 0, X(\Delta) = 0, m = 0)$; $E_{00,1} = (X(0) = 0, X(\Delta) = 0, m = 1)$; $E_{01,1} = (X(0) = 0; X(\Delta) = 1, m = 1)$. $E_{11,0}$; $E_{11,1}$; $E_{10,1}$ are similarly defined with state 1 in place of state 0 and vice-versa. The associated probabilities, factorizable as before in equilibrium and transition probability components, are:

$$\begin{aligned}
 P(E_{00,0}) &= \frac{\rho_1}{\rho_0 + \rho_1} \{e^{-\rho_0 \Delta}\}, \\
 P(E_{00,1}) &= P(C_{00}) - P(E_{00,0}) \\
 &= \frac{\rho_1}{\rho_0 + \rho_1} \left\{ \frac{\rho_1}{\rho_0 + \rho_1} - e^{-\rho_0 \Delta} + \frac{\rho_0}{\rho_0 + \rho_1} e^{-(\rho_0 + \rho_1) \Delta} \right\}, \\
 (2.2) \quad P(E_{01,1}) &= P(C_{01}) = \frac{\rho_1}{\rho_0 + \rho_1} \left\{ \frac{\rho_0}{\rho_0 + \rho_1} - \frac{\rho_0}{\rho_0 + \rho_1} e^{-(\rho_0 + \rho_1) \Delta} \right\}, \\
 P(E_{11,0}) &= \frac{\rho_0}{\rho_0 + \rho_1} \{e^{-\rho_1 \Delta}\}, \\
 P(E_{11,1}) &= P(C_{11}) - P(E_{11,0}) \\
 &= \frac{\rho_0}{\rho_0 + \rho_1} \left\{ \frac{\rho_0}{\rho_0 + \rho_1} - e^{-\rho_1 \Delta} + \frac{\rho_1}{\rho_0 + \rho_1} e^{-(\rho_0 + \rho_1) \Delta} \right\}, \\
 P(E_{10,1}) &= P(C_{10}) = \frac{\rho_0}{\rho_0 + \rho_1} \left\{ \frac{\rho_1}{\rho_0 + \rho_1} - \frac{\rho_1}{\rho_0 + \rho_1} e^{-(\rho_0 + \rho_1) \Delta} \right\}.
 \end{aligned}$$

Again note the identity $P(E_{01,1}) = P(E_{10,1})$. For $k = 3$ the same remarks as before are applicable here for computing the probabilities associated with the 18 possible outcomes. In general, for any value of k the computation consists of multiplying the equilibrium probabilities by a sequence of transition probabilities, the initial definition of which is given by the terms between $\{ \}$ in (2.1) and (2.2).

On the basis of these results the log-likelihood functions l_1 and l_2 for the observational schemes (1) and (2) are computed. For $k = 2$,

$$(2.3) \quad l_1 = \sum_{r=0}^1 \sum_{s=0}^1 n_{rs} \ln \{P(C_{rs})\},$$

where n_{rs} represents the number of observed transitions between state r and state s over the Δ interval. Indeed l_2 has a similar form. The log-likelihood for scheme (3) includes terms for the equilibrium probabilities of the individuals starting in state 0 and in state 1 and terms for the probabilities associated with each spell spent in a state. These derive from $\rho_0 e^{-\rho_0 t}$ and $\rho_1 e^{-\rho_1 t}$, the density functions for

completed spells in 0 of length t and in 1 of length t' . Incomplete spells contribute via $e^{-\rho_0 t}$, $e^{-\rho_1 t'}$ respectively. With a similar notation to (2.3) we obtain:

$$(2.4) \quad l_3 = \{(n_{11} + n_{10}) \ln \rho_0 + (n_{00} + n_{01}) \ln \rho_1 - n \ln (\rho_0 + \rho_1)\} \\ + \left\{ m_0 \ln \rho_0 - \rho_0 \sum_{i=1}^{m_0} t_i + m_1 \ln \rho_1 - \rho_1 \sum_{i=1}^{m_1} t'_i \right\},$$

where m_0 and m_1 represent the total number of completed spells spent by the n individuals in state 0 and state 1 and (t_1, \dots, t_{m_0}) , (t'_1, \dots, t'_{m_1}) are the lengths of the completed and censored spells. Note that $\sum t_i + \sum t'_i = nT$.

If the observed process is correctly specified as an alternating Poisson process, then the Fisher information matrices associated with (1), (2), and (3) are the expected values of the negative of the second derivatives of the log-likelihoods. The information matrices relative to the first and second scheme are easily derivable when regularity conditions hold (Cox and Hinkley [3, p. 107]). Scheme (1), for example, has information matrix,

$$I(1) = \{I_{ij}(1)\} = \left\{ \sum_{r,s} [P(C_{rs})]^{-1} \frac{\partial P(C_{rs})}{\partial \rho_i} \frac{\partial P(C_{rs})}{\partial \rho_j} \right\}.$$

The computation of the information matrix of scheme (3) involves instead the expectations of the number of transitions out of each state, as well as the equilibrium probabilities. The result is

$$I(3) = \begin{bmatrix} \frac{E(m_0) + E(n_{10} + n_{11})}{\rho_0^2} - \frac{n}{(\rho_0 + \rho_1)^2} & -\frac{n}{(\rho_0 + \rho_1)^2} \\ -\frac{n}{(\rho_0 + \rho_1)^2} & \frac{E(m_1) + E(n_{00} + n_{01})}{\rho_1^2} - \frac{n}{(\rho_0 + \rho_1)^2} \end{bmatrix}$$

where $E(m_0) = E(m_1) = \rho_0 \rho_1 \Delta / (\rho_0 + \rho_1)$ and $E(n_{01} + n_{00}) = n \rho_1 / (\rho_0 + \rho_1)$; $E(n_{10} + n_{11}) = n \rho_0 / (\rho_0 + \rho_1)$.

3. COMPARISON OF SAMPLING SCHEMES

3.1. Asymptotic Relative Efficiencies

In designing a survey we may be interested in some or all of the following objectives: (a) the estimation of the proportion of time spent in state 0, say $\lambda = \rho_1 / (\rho_0 + \rho_1)$; (b) the estimation of one of the rate parameters, say ρ_0 , or equivalently the corresponding mean sojourn time $1/\rho_0$; (c) the joint estimation of all the parameters in the complete model, in this case both ρ_0 and ρ_1 .

For each of these objectives we compute the A.R.E.'s of the schemes (1) and (2) with respect to (3), the formal definition being given in the Appendix. The factors $\rho_0 \Delta$ and $\rho_1 \Delta$ in the ARE contour pictures are to be interpreted as ratios between the time interval Δ and the mean sojourn time in state 0 and state 1 respectively, so that the interval (0.1, 1.5) corresponds to the cases of the means being between $\frac{2}{3}$ and 10 times Δ .

Figure 1 gives the ARE contour levels for the estimation of λ , the ARE's being defined as the ratio of the variances of the estimators. The dotted lines represent

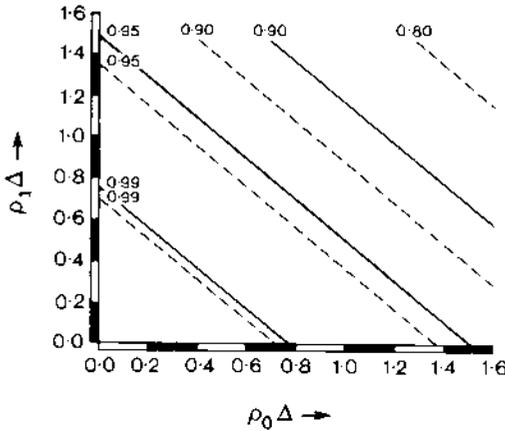


FIGURE 1—ARE's of estimates of λ based on scheme (1) vs. scheme (3) and based on scheme (2) vs. scheme (3).
 --- (1) vs. (3)
 — (2) vs. (3)

the ARE levels for the estimator derived from scheme (1) and the continuous lines the ARE levels for the estimator derived from scheme (2). Both decrease for increasing values of $\rho_0 \Delta$ and $\rho_1 \Delta$ or, equivalently, the ARE's decrease for decreasing mean sojourn times in state 0 and 1, for a given interval Δ . Since $\text{var}(\hat{\lambda})$ depends symmetrically on the two parameters (see Appendix), both ARE's behave symmetrically about the axes.

Similarly, Figure 2 gives the ARE contour levels for the estimates of ρ_0 . Here, however, the two schemes perform quite differently. Take $\Delta = 1$: the parameter

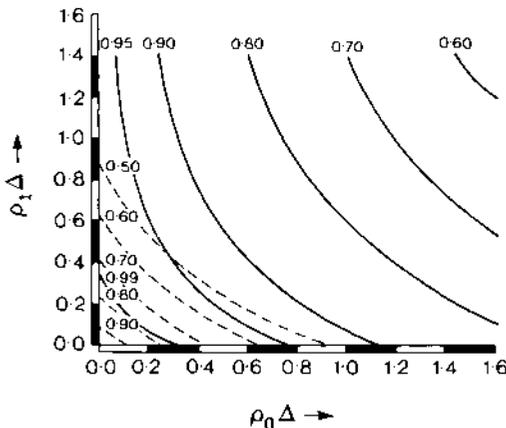


FIGURE 2—ARE's of estimates of ρ_0 based on scheme (1) vs. scheme (3) and based on scheme (2) vs. scheme (3).
 --- (1) vs. (3)
 — (2) vs. (3)

region for which scheme (1) gives an estimate of ρ_0 at least 70 per cent efficient covers only a small area of the (ρ_0, ρ_1) plane. The equivalent space defined by scheme (2) includes instead many (ρ_0, ρ_1) values. Furthermore scheme (2) allows ρ_1 , the nuisance parameter, to be large relative to ρ_0 and still allows the achievement of a given level of ARE. In other words, when the mean sojourn time in state 1 is considerably smaller than the mean sojourn time in state 0, scheme (2) increases its efficiency relatively to scheme (1).

Figure 3 gives the generalized ARE's. They are defined as ratios of the square roots of the asymptotic generalized variances. Again, the loss of efficiency due to collecting discrete time data instead of continuous time data increases steadily with the rates. Here, however, the symmetry about the axes is recovered.

When $k > 2$ the contour levels in Figure 1-3 are smoothly pushed upwards, showing the ARE improvement of the systematic sampling schemes when the total number of discrete observations increases, when T is fixed.

3.2. The Effect of the Time Interval

The comparison has so far been conducted assuming Δ fixed. When Δ varies, i.e. when the observations are, say $\Delta' = h\Delta$ units apart, the transition rates are reparameterized accordingly. Further, the joint information matrix is now h^2 times the original, h^2 being the square value of the Jacobian of the transformation (see Cox and Hinkley [3, p. 130]). For given values of the parameters it is therefore possible to find values of h which minimize the generalized variance or, otherwise, one of the variances.

Table I, for instance, gives the factors, labelled h_1 and h_2 , which minimize $\text{var}(\hat{\rho}_0)$ for scheme (1) and (2) and for a selection of parameter values. As one would expect, decreasing values of the mean sojourn times in the two states are

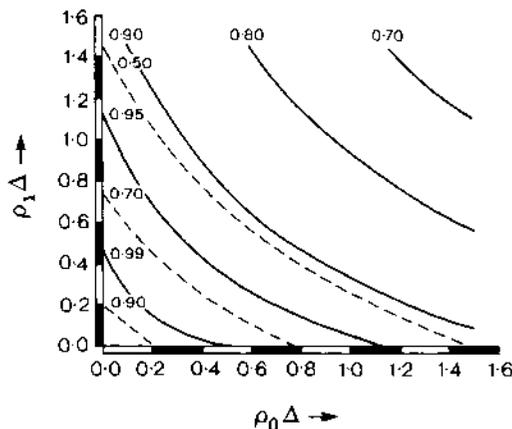


FIGURE 3—Generalized ARE's for estimates of (ρ_0, ρ_1) based on scheme (1) vs. scheme (3) and based on scheme (2) vs. scheme (3).

--- (1) vs. (3)
— (2) vs. (3)

TABLE I
FACTORS h_1 AND h_2 WHICH MINIMIZE $\text{VAR}(\hat{\rho}_0)$ FOR A SELECTION OF PARAMETER VALUES.^a

ρ_0	$1/\rho_0$	Balanced mean sojourn times		Unbalanced mean sojourn times ^b	
		h_1	h_2	h_1	h_2
.05	20.0	8.16	31.23	4.84	25.56
.10	10.0	4.08	15.62	2.42	12.78
.20	5.0	2.04	7.81	1.21	6.38
.40	2.5	1.02	3.91	.60	3.19
.60	1.7	.68	2.60	.40	2.13
.80	1.3	.51	1.95	.30	1.58
1.00	1.0	.41	1.56	.24	1.26
1.50	.7	.27	1.04	.16	.84

^a The term h_1 refers to scheme (1) and the term h_2 to scheme (2).

^b The values are computed for $\rho_1 = 2.5\rho_0$.

associated with smaller optimal time intervals, the rate of their decrement being exactly the same as that in the means. On the other hand, the values h_1 and h_2 are fairly different, with the latter allowing for much longer intervals between observations to achieve the smallest variances. This is even more noticeable when the mean sojourn times in the two states are not balanced.

Slightly different considerations apply to $\text{var}(\hat{\lambda})$. When the first scheme is considered (Cox [1, p. 89]),

$$\text{var}(\hat{\lambda}) = \frac{\rho_0\rho_1}{2(\rho_0 + \rho_1)^2} \{1 + \exp\{-(\rho_0 + \rho_1)\Delta\}\}.$$

The first term is exactly the binomial variance corresponding to two independent observations with $P(\text{success}) = \lambda$ while the second term collapses to 1 as $\Delta \rightarrow \infty$. Then the further apart are the observations, the smaller is $\text{var}(\hat{\lambda})$. The same is true for the second scheme. Of course these remarks refer to a given total number of observations: the cost in a real study may depend in a rather complex way on both the number of observations and the total period of observation.

This leads to some broad conclusions. When interest lies in overall indicators, like the proportion of time spent in one of the two states and no continuous time information is available, the best strategy is to observe the process either through scheme (1) or scheme (2) at time points sufficiently distant to be considered independent. If, on the contrary, interest lies in the structural parameters, observations which include some qualitative information lead to better estimates and an appropriate—but not necessarily accurate—choice of Δ' is important. This has been shown for a very simple case of additional qualitative information: any more informative scheme than (2) would improve the efficiency. Furthermore, if some a priori knowledge of the parameters size is available, we can achieve more efficient estimates of ρ_0 through the results in Table I or similarly, more efficient joint estimates of (ρ_0, ρ_1) , when a similar argument is applied to the generalized variance.

4. HETEROGENEOUS POPULATIONS

The results of the previous sections can be extended to the more interesting case of a nonhomogeneous population. Very often in fact, the transition rates between the two states depend on explanatory variables z , fixed for each individual. Denote by $\rho_0(z)$, $\rho_1(z)$ the hazard functions associated respectively with state 0 and state 1. The mean sojourn times, $[\rho_0(z)]^{-1}$ and $[\rho_1(z)]^{-1}$, therefore, are not generally the same for all the individual processes. To represent the effect of z on the transition densities, we consider the particular forms $\rho_0(z) = \alpha_0 e^{\beta_0 z}$ and $\rho_1(z) = \alpha_1 e^{\beta_1 z}$, where the parameters $\varrho = (\alpha_0, \alpha_1)$ define the baseline functions corresponding to a process with $z=0$ and the second terms represent the multiplicative effect of z in a log-linear form. For simplicity we restrict attention to the case of just one covariate, z , and compute the log-likelihood functions for the parameters $\theta = (\alpha_0, \alpha_1, \beta_0, \beta_1)$ associated with the three schemes, as before. In order to compare the schemes we consider the asymptotic covariance matrix of $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1)$ near $\beta = (0, 0)$. This is defined by the bottom right-hand corner of the inverse of the joint information matrix. Similarly to Section 3.1 we compare the efficiencies in estimating one or both the coefficients β , which are now the parameters of interest. Figure 4 gives the contour levels of estimates for β_0 when the model is overspecified, i.e. when z does not influence the transition densities. The same remarks as for Figure 2 are valid here.

The generalized ARE's for the vector $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1)$ are defined as ratios of the square roots of the asymptotic generalized variances of $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1)$; see the Appendix. Figure 5 shows the results. Again, they confirm the previous findings about the efficiency gain achieved when some additional information about the individual realizations is included in the model specification. In general, the ARE's are not affected by $\text{var}(z)$ or by its distribution.

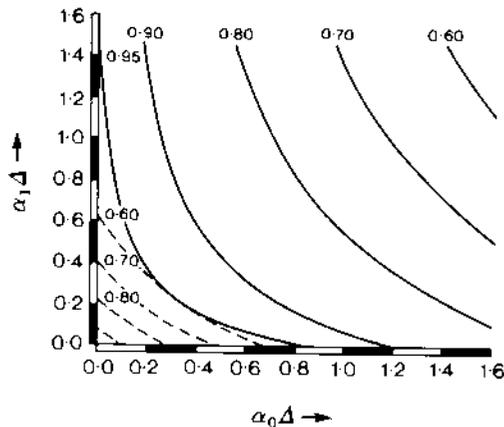


FIGURE 4—ARE's of estimates of β_0 when $(\beta_0, \beta_1) = (0, 0)$ based on scheme (1) vs. scheme (3) and based on scheme (2) vs. scheme (3), $z \sim N(0, 1)$.

--- (1) vs. (3)
— (2) vs. (3)

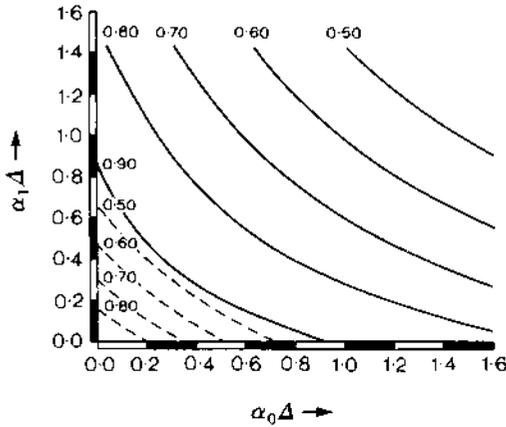


FIGURE 5—Generalized ARE's of estimates of (β_0, β_1) when $(\beta_0, \beta_1) = (0, 0)$ based on scheme (1) vs. scheme (3) and based on scheme (2) vs. scheme (3), $z \sim N(0, 1)$.
 --- (1) vs. (3)
 — (2) vs. (3)

The results presented here apply to alternating Markov processes in which transition rates do not vary with time. Nevertheless, when realizations are observed only over a short period of time and the process is locally stationary, the assumption of exponential sojourn time distributions in the two states may be suitable and the criteria described in Sections 3 and 4 would apply.

Medical Research Council Centre, Cambridge, England

Manuscript received May, 1984; final revision received September, 1985.

APPENDIX

DEFINITION OF THE ASYMPTOTIC RELATIVE EFFICIENCIES

We define here the ARE's on which the sampling schemes (1) and (2) are compared. The ARE's for the estimation of the parameter $\lambda = \rho_1 / (\rho_0 + \rho_1)$ are defined as

$$ARE_{\lambda}(1:3) = \text{var} \{ \hat{\lambda}(3) \} / \text{var} \{ \hat{\lambda}(1) \},$$

$$ARE_{\lambda}(2:3) = \text{var} \{ \hat{\lambda}(3) \} / \text{var} \{ \hat{\lambda}(2) \},$$

respectively for scheme (1) and scheme (2), $\hat{\lambda}(i)$ being the maximum likelihood estimate derived via the information collected by scheme (i), $i = 1, 2, 3$. In general,

$$\begin{aligned} \text{var}(\hat{\lambda}) &= \{ (\partial \rho / \partial \lambda)^T I(\partial \rho / \partial \lambda) \}^{-1} \\ &= \frac{\rho_0^2 \text{var}(\hat{\rho}_1) + \rho_1^2 \text{var}(\hat{\rho}_0) - 2\rho_0\rho_1 \text{cov}(\hat{\rho}_0, \hat{\rho}_1)}{(\rho_0 + \rho_1)^4}, \end{aligned}$$

where I is the Fisher information matrix for $\rho = (\rho_0, \rho_1)$. The ARE'S for $\hat{\lambda}$ are therefore defined in terms of the elements of the covariance matrices, I^{-1} when $\rho = \hat{\rho}$.

Similarly, when $I(i)$ is the Fisher information matrix relative to scheme (i) , $(i = 1, 2, 3)$, and $I^{jj}(i)$ the j th element of the diagonal of $I^{-1}(i)$,

$$\text{ARE}_{\rho_0}(1:3) = \text{var}\{\hat{\rho}_0(3)\}/\text{var}\{\hat{\rho}_0(1)\} = I^{11}(3)/I^{11}(1),$$

$$\text{ARE}_{\rho_0}(2:3) = \text{var}\{\hat{\rho}_0(3)\}/\text{var}\{\hat{\rho}_0(2)\} = I^{11}(3)/I^{11}(2).$$

The generalized ARE's are defined as ratios of the square roots of the information matrices determinants:

$$\text{Gen. ARE}_{\rho}(1:3) = |I(3)|^{-1/2} \cdot |I(1)|^{1/2},$$

$$\text{Gen. ARE}_{\rho}(2:3) = |I(3)|^{-1/2} \cdot |I(2)|^{1/2}.$$

Finally, in the heterogeneous population case, when I_{θ} denotes the information matrix for $\theta = (\alpha_0, \alpha_1, \beta_0, \beta_1)$ and I_{β} the information matrix for $\beta = (\beta_0, \beta_1)$,

$$\text{ARE}_{\rho_0}(1:3) = \text{var}\{\hat{\beta}_0(3)\}/\text{var}\{\hat{\beta}_0(1)\} = I_{\beta}^{33}(3)/I_{\beta}^{33}(1),$$

$$\text{ARE}_{\rho_0}(2:3) = \text{var}\{\hat{\beta}_0(3)\}/\text{var}\{\hat{\beta}_0(2)\} = I_{\beta}^{33}(3)/I_{\beta}^{33}(2),$$

and

$$\text{Gen. ARE}_{\beta}(1:3) = |I_{\beta}(3)|^{-1/2} \cdot |I_{\beta}(1)|^{1/2},$$

$$\text{Gen. ARE}_{\beta}(2:3) = |I_{\beta}(3)|^{-1/2} \cdot |I_{\beta}(2)|^{1/2}.$$

REFERENCES

- [1] COX, D. R.: *Renewal Theory*. London: Chapman and Hall, 1962.
- [2] ———: "Regression Models and Life Tables," *Journal of the Royal Statistical Society, B*, 34(1972), 187-202.
- [3] COX, D. R., AND D. HINKLEY: *Theoretical Statistics*. London: Chapman and Hall, 1974.
- [4] COX, D. R., AND H. MILLER: *The Theory of Stochastic Processes*. London: Methuen, 1965.
- [5] HECKMAN, J., AND C. FLINN: "New Methods for Analysing Structural Models of Labor Force Dynamics," *Journal of Econometrics*, 18(1982), 115-186.
- [6] LANCASTER, T.: "Econometric Methods for Duration of Unemployment," *Econometrica*, 47(1979), 939-956.
- [7] LANCASTER, T., AND S. NICKELL: "The Analysis of Re-Employment Probabilities," *Journal of the Royal Statistical Society, A*, 143(1980), 141-165.