Adaptive Sampling in Behavioral Surveys

Steven K. Thompson

ABSTRACT

Studies of populations such as drug users encounter difficulties because the members of the populations are rare, hidden, or hard to reach. Conventionally designed large-scale surveys detect relatively few members of the populations so that estimates of population characteristics have high uncertainty. Ethnographic studies, on the other hand, reach suitable numbers of individuals only through the use of link-tracing, chain referral, or snowball sampling procedures that often leave the investigators unable to make inferences from their sample to the hidden population as a whole. In adaptive sampling, the procedure for selecting people or other units to be in the sample depends on variables of interest observed during the survey, so the design adapts to the population as encountered. For example, when self-reported drug use is found among members of the sample, sampling effort may be increased in nearby areas. Types of adaptive sampling designs include ordinary sequential sampling, adaptive allocation in stratified sampling, adaptive cluster sampling, and optimal model-based designs. Graph sampling refers to situations with nodes (for example, people) connected by edges (such as social links or geographic proximity). An initial sample of nodes or edges is selected and edges are subsequently followed to bring other nodes into the sample. Graph sampling designs include network sampling, snowball sampling, link-tracing, chain referral, and adaptive cluster sampling. A graph sampling design is adaptive if the decision to include linked nodes depends on variables of interest observed on nodes already in the sample. Adjustment methods for nonsampling errors such as imperfect detection of drug users in the sample apply to adaptive as well as conventional designs.

INTRODUCTION

Surveys to estimate human behavioral characteristics such as drug use encounter a number of inherently difficult sampling and estimation problems. Among the factors making sampling and estimation difficult for such populations are the rarity and geographic unevenness of some of the populations of interest, the elusiveness or hiddenness of individuals in the population, and the variability of behaviors within and between subpopulations. In addition, the sensitive nature of the behaviors of interest gives rise to nonsampling errors including nonresponse and underreporting of stigmatized behaviors. Each of these issues arises in design and implementation of a large survey on drug use such as the National Household Survey on Drug Abuse as well as in ethnographic studies focusing on specific drug-using populations (cf., Lambert 1990; Turner et al. 1992*c*; Weppner 1977). Similar sampling and estimation problems arise in surveys of persons infected with rare diseases, populations defined by sexual orientation or behavior, sex workers and others involved in underground economic activities, homeless people, and other underrepresented groups. Some of the statistical problems arising with human populations have arisen also in environmental and biological surveys so that methods first developed for one area have subsequently been applied to another (Freedman 1991; Kalton and Anderson 1986; Thompson 1992; Wolter 1986, 1991).

Conventionally designed large-scale surveys detect relatively few members of rare or hidden populations so that estimates of population characteristics have high uncertainty. For example, the original impetus for the national Health and Social Life Survey (Laumann et al. 1994), a national probability sample survey of sexual behavior, was in large part provided by concern regarding the acquired immunodeficiency syndrome (AIDS) epidemic. Funding constraints limited the sample size to 3,432 people, and the number of people in the sample who reported having tested positive for the human immunodeficiency virus (HIV) was only 6.

Ethnographic studies and other studies focusing on the behaviors of people in rare and hidden populations find suitable numbers of individuals for their samples only through the use of link-tracing, chain referral, or snowball sampling procedures that usually leave the investigators unable to make inferences from their sample to the hidden population as a whole. For example, to investigate sex-for-crack exchanges in Philadelphia and Newark (French 1993), the sample of 100 crack cocaine users was obtained by the investigators going to known drug-dealing areas and talking with persons they believed were users or dealers, who then referred the investigators to other users. To analyze how some opiate addicts had overcome their addiction on their own, Biernacki (1986) used referral chains. Finding starting points for the chains was difficult, however, for the people of interest had little contact with treatment centers or social service agencies.

One approach to increasing the sample representation of members of a rare or hidden population while still obtaining unbiased estimates of population characteristics is through the use of adaptive sampling procedures. In adaptive sampling, the selection of people or other units to include in the sample adapts to observations made during the survey. In particular, whenever "interesting" values are observed, sampling intensity may be adaptively increased for neighboring or linked units. Examples of interesting values that could be specified include reported drug use, involvement in underground economic activities, high-risk sexual behaviors, or a positive HIV test result. In a spatial context, additional units may be added to the sample from the geographic vicinity of any unit in which an interesting value is observed. Linkages such as social contact or kinship could be used in place of spatial proximity. The condition for adaptively adding units can be based either on the variable of interest, such as self-reported drug use, or on an auxiliary variable such as tobacco use. For some populations, adaptive designs can produce gains in efficiency, relative to conventional designs, for estimating the population mean or total. In addition, adaptive sampling designs can substantially increase the yield of interesting units in the sample.

Graph sampling refers to situations with nodes connected by edges. In studies of hidden populations, the nodes may represent individual people and the edges represent social links or geographic proximity between people. An initial sample of nodes or edges is selected and edges are subsequently followed to bring other nodes into the sample. Graph sampling designs include network sampling, snowball sampling, link-tracing, chain referral, and adaptive cluster sampling. A graph sampling design is adaptive if the decision to include linked nodes depends on variables of interest observed in nodes already in the sample.

In this chapter, the use of adaptive sampling and graph sampling methods for studies of behavioral characteristics in rare and hidden populations is examined. Methods of adjusting for the nonsampling errors that arise in such studies are discussed for both adaptive and conventional designs.

ADAPTIVE SAMPLING STRATEGIES

For populations that are rare, unevenly distributed, hidden, or hard to reach, conventional sampling designs such as simple random sampling lead to estimates with high variances and potential biases. With sufficient previous knowledge of the population, precision can be increased through such devices as stratification, systematic designs, and use of auxiliary information in the design and estimation stages (Cochran 1977; Thompson 1992).

Often, however, the uneven patterns in the populations cannot be predicted before the survey. For example, patterns of drug use may change over time, epidemics progress through cycles, neighborhoods may change their compositions, and economic changes occur; similarly, natural populations of animals or fish may change unpredictably in spatial pattern. For such populations, adaptive sampling strategies can be useful.

Adaptive sampling designs are those in which the procedure for selecting units to include in the sample may depend on values of the variable of interest observed during the survey. For spatially clustered populations, additional observations may be added in the neighboring vicinity when- ever high abundance is encountered. Whenever an infected person appears in a survey sample of a rare, contagious disease, close contacts of that person might be added to the sample. In a drug use study, sampling intensity could be adaptively increased in the neighborhood of respondents with self-reported use.

Types of Adaptive Designs

Adaptive sampling designs include sequential stopping designs, adaptive allocation designs, optimal model-based strategies, and adaptive cluster sampling.

With sequential stopping designs, sampling continues until a given criterion, based for example on observed incidence or sample variance, is attained. The procedure may be based on sequentially observing the variable of interest and evaluating the criterion as each unit is selected, or it may be based on batches of units. Much of the statistical literature in sequential analysis (Chernoff 1972; Siegmund 1985; Wald 1947; Woodroofe 1982) concerns sequential stopping problems, in which the size of a random sample is determined sequentially from the observed values.

With adaptive allocation designs, an initial stratified sample is selected. Based on the observed values for the initially selected units, an additional stratified sample is selected with allocation of sample sizes to strata based on the initial observations. For example, after an initial stratified sample is obtained in a drug use epidemiological survey, the remaining sample sizes could be allocated to give larger sample sizes in strata showing high reported drug use in the initial sample. For adaptive allocation designs, the strata may be defined either spatially or through sociological variables. With adaptive allocation, the usual stratified estimate is not unbiased for the population total. Unbiased adaptive allocation strategies are described in Kremers (1987), Thompson and associates (1992), and Thompson and Seber (1996). Other adaptive allocation strategies are described in Solomon and Zacks (1970), Francis (1984), Gasaway and colleagues (1986), and Jolly and Hampton (1990).

Optimal model-based sampling strategies are often adaptive ones. Much of survey sampling practice is design based. That is, no assumptions are made about the population itself, and properties such as unbiasedness of estimates are calculated over all possible samples that might have been selected. In the model-based approach to sampling, a probabilistic model is assumed for the population. For example, the values of the variable of interest may be assumed to have a multivariate lognormal distribution, with units that are close to each other either geographically or socially having positive correlation. For many assumed population models, the theoretically optimal sampling strategy can be shown to be an adaptive one (Thompson 1988; Thompson and Seber 1996; Zacks 1969). However, the theoretically optimal strategies are not necessarily the most practical because they may require an unattainable amount of previous information about the population and tend to be computationally and implementationally complex (Solomon and Zacks 1970). Simpler designs that are model based and have some features of optimality have been applied to environmental sampling problems by Englund and Herari (in press) and by Geiger (1994).

In adaptive cluster sampling, an initial sample is selected based on a conventional sampling design such as simple random sampling, systematic sampling, cluster sampling, or stratified sampling. Whenever the variable of interest of a unit in the sample satisfies a specified condition, units in the neighborhood of that unit are added to the sample. If in turn any of the added units satisfies the condition, still more units are added, and so on. For example, neighboring units may be added whenever the variable of interest has a high or interesting value, such as reported drug use, incidence of a disease, sexual behaviors of interest to the study, high economic activity, high observed occurrence of homelessness, high abundance of an animal species, or high concentration of a pollutant. A variety of adaptive cluster sampling designs are described in Thompson (1990, 1991*a*, 1991*b*, 1992, 1993, 1994*a*, in press), Seber and Thompson (1993), Thompson and Seber (1996), and Munholland and Borkowski (1993). The use of adaptive cluster sampling to estimate the prevalence of insect infestation in forest trees is described by Roesch (1993). Adaptive cluster sampling of winter waterfowl populations is described by Smith and colleagues (1995). Adaptive cluster sampling for rare household characteristics is described by Danaher and King (1994).

The neighborhoods of adaptive cluster sampling may be defined spatially, as with the geographically neighboring sample sites of environmental and ecological surveys and with urban blocks and larger geographic regions in human surveys, but they also may be defined by social or institutional connections. For example, in a survey of a rare disease, the neighborhood of a person in the sample could be defined to include that person's siblings or close social contacts. Determining key social links of persons in the sample can present many methodological challenges. Wiebel (1990) reports good success in obtaining the identities of sexual partners of active intravenous drug users once the trust between outreach workers and subjects solidified. Additional effective methodologies for obtaining sensitive information on hidden populations are described in Adler (1990) and Goldstein and associates (1990). However, as pointed out below in the section on graph sampling, estimation in adaptive cluster sampling is based on the empirical or observable links, not on hidden or underlying links, so that unbiased estimation of a population total is possible even though some of the underlying links between sample respondents may remain hidden.

Typically, a survey is used to obtain estimates of more than one characteristic of interest. Generalizations of adaptive cluster sampling results to the multivariate case produce unbiased estimators of the population mean and total for each variable as well as unbiased estimators of variances and covariances (details are found in Thompson 1993). The results hold whether the condition for additional sampling depends on just one of the variables or on a function of all of them. The result giving conditions under which adaptive sampling produces more precise estimates than conventional sampling is also generalized to the multivariate case (Thompson

1993). Further, adaptive addition of units can be based on an easier-to-measure auxiliary variable or one that is less sensitive, such as tobacco use, rather than the variable of interest, such as illegal drug use.

One problem with conventional survey designs when applied to populations with rare characteristics of interest, such as heroin use or HIV infection, is that typically few cases with the characteristic show up in the sample. Adaptive designs, in addition to potentially increasing the precision of survey estimates, have the potential to increase the yield of the sample in terms of the characteristic of interest. For example, the number of self-reported drug users or people who have tested positive for HIV in the sample may be increased by adaptively either increasing the allocation to strata in which such people are encountered or following social links from such people as in adaptive cluster sampling. The objective is to obtain data on more individuals of the rare population in order to more effectively carry out analytic studies such as evaluations of drug use outcomes, outcomes of treatment programs, and identification of risk factors. Indeed, without adaptively following leads and links from initially encountered individuals, it may not be possible to penetrate a hidden population sufficiently for study (Adler 1990; Frank and Snijders 1994). Importantly, unbiased estimates of population totals and other parameters are still possible with such surveys even though the sample contains members of the target population in higher proportion than the population as a whole.

Unbiased Estimation With Adaptive Designs

Estimators that are unbiased with conventional sampling designs may be no longer unbiased with adaptive designs. For example, adaptive cluster sampling typically produces a sample with a higher than representative yield of the variable of interest—more birds or whales sighted, more persons reporting drug use, or more persons infected with the disease— than would occur with a random sample. With such a sample, the conventional expanded sample mean would tend to overestimate the total in the population. Fortunately, unbiased estimators are available for use with adaptive designs. The simplest of these estimators are design unbiased, meaning that the unbiasedness is based on the way the sample is selected and does not depend on any assumptions about the population itself.

Suppose the initial sample consists of a simple random sample of n units. For a unit in the initial sample whose y-value is observed to satisfy the specified condition, the units in its neighborhood are added to the sample. For any of those added units satisfying the condition, the neighboring units are added, and so on. For any of the added units not satisfying the condition, on the other hand, no neighboring units are added. Thus, in the final sample associated with each initial unit that satisfied the condition, there is a network of units satisfying the condition and a number of added edge units that do not satisfy the condition. Unbiased estimation in adaptive cluster sampling must deal with the fact that selection or inclusion probabilities cannot be determined from the sample data for every unit in the sample. Even so, simple unbiased estimates can be computed.

The simplest of the unbiased estimates is obtained by averaging the *y*-values within networks (but excluding edge units). Let w_i be the average of the *y*-values for the network associated with the *i*-th unit of the initial sample. Any unit not satisfying the condition is considered a network of size one. An unbiased estimate of the population mean is given by
An unbiased estimator that is only slightly more complicated to compute but that in practice tends to be more efficient than that shown above is obtained by computing for each network intersected by the initial sample the probability a of that network being intersected. Edge units are again ignored. Suppose that networks have been intersected by the initial sample and let y^* denote the total of the *y*-values in the - th network. The unbiased estimator is where is the number of units in the population.
An unbiased estimate of the population total is obtained by multiplying the estimate of the mean by . Unbiased estimates are also available for adaptive cluster sampling with other initial designs such as cluster, systematic, and stratified sampling. Unbiased estimates of variances are also readily computed. The efficiency of the above estimators can be improved using the Rao-Blackwell method, so that edge units receive some weight in the estimates, but the improved estimates are more complicated to compute. Full details on estimation with adaptive cluster sampling are given in Thompson (1992) and Thompson and Seber (1996).

Efficiency of Adaptive Sampling

For some populations, particularly those that are rare and clustered, adaptive sampling strategies have been found to produce remarkable increases in precision or efficiency compared to conventional sampling designs of equivalent sample size. In addition, adaptive designs can significantly increase the yield of interesting observations in the sample. Efficiency comparisons for specific populations are given in Thompson (1990, 1991*a*, 1991*b*, 1992, 1994*a*), Roesch (1993), Thompson and associates (1992), Francis (1984), and Smith and colleagues (1995). Factors influencing the relative efficiency of adaptive cluster sampling to simple random sampling are described in Thompson (1994*a*) and summarized below. The efficiency of an adaptive cluster sampling design compared to a conventional design for household surveys of rare characteristics was estimated using a trial survey by Danaher and King (1994).

The relative efficiency of adaptive cluster sampling to simple random sampling depends on characteristics of the population, the design, and the cost of sampling. Any of the following characteristics tend to increase the efficiency of adaptive cluster sampling relative to conventional random sampling: (1) the within-network variance is a high proportion of the total population variance (i.e., the population is clustered or aggregated with high variability within aggregations); (2) the population is rare; (3) the expected final sample size with adaptive cluster sampling is not much larger than the initial sample size; (4) the cost of observing units in clusters or networks is less than the cost of observing the same number selected at random; (5) the cost of observing units not satisfying the condition is less than the cost of observing units satisfying the condition; (6) the condition for extra sampling may be based on an auxiliary variable that is easy to measure; and (7) an efficient estimator or Rao-Blackwell improved estimator is used with the adaptive cluster sampling.

Because the final sample size depends on what is observed during the survey, practical measures are needed to ensure that the final sample size does not exceed the time or funding resources available for the survey. Ideally, a good choice of the criterion for extra sampling (as described above) limits the adaptively added units to a relatively small proportion of the total. Further, if the population has been stratified, then the criterion in any stratum can be changed adaptively based on the time spent or observations made in previous strata, without affecting the unbiasedness of the estimates. Because of the design unbiased of the procedure within any stratum, the average value of the estimate over all possible samples equals the population mean for that stratum even though a previous stratum may have influenced the choice of the adaptive condition to be used. Thus, if time is running short halfway through the survey, the criterion can be made more stringent or adaptive sampling dispensed with completely for the remaining strata. If the stratification has not been done at the design stage, a pragmatic approach is to use poststratification at the estimation stage to approximate the same result. Thus, if adaptive sampling is discontinued part way through the survey, estimates can be poststratified with adaptive cluster sampling estimators used for that portion of the population in which the sampling was adaptive and the conventional estimator used in that portion in which the sampling was conventional. Other methods for limiting sample size include adding the extra units only for the top few values of the initial sample (Thompson, in press) and stopping sampling as soon as a specified total sample size has been reached (Brown 1994).

GRAPH-SAMPLING METHODS

Sampling methods such as network sampling, snowball sampling, chain referral sampling, adaptive cluster sampling, and other link-tracing designs in which investigators use links between people to find other people to include in the sample are examples of survey sampling in graphs. A directed graph consists of a set of nodes such as people or other units, and a set of edges linking some nodes to others. For two people (nodes), the links (edges) could be provided by physical proximity such as living on the same block, by hereditary relationship such as siblinghood, or by a social relationship. The edges can be directional so that two nodes i and j can be linked from i to j, from j to i, in both directions, or in neither direction. For example, individual i might provide investigators with the name of individual j, while user j either did not know or would not reveal individual j.

Associated with the i-th node is a variable of interest, y_i. For example, with nodes representing individual people, the variable of interest could be an indicator of cocaine use or dollar amount spent on heroin. The basic problem in graph sampling is to select a sample of nodes or edges by some means and then estimate some population quantity, such as the total of the y-values, of the nodes or edges. The population quantities of interest could be number of cocaine users in the population, dollar amount spent on heroin, or average number of partners with whom needles are shared.

A graph sampling design is adaptive if decisions on whether to follow links depend on the observed y-values in the sample. For example, if an individual in the sample is asked to name sexual partners only if the individual reports intravenous drug use, the survey is adaptive, whereas it is not adaptive if every person sampled is asked to name sexual partners. The inherent links in the population, such as the sexual partners each individual would name if asked (regardless of intravenous drug use status), correspond to the neighborhood connections of adaptive cluster sampling in the spatial setting. The links that are followed, connecting groups of intravenous drug users, determine the networks of units that satisfy the condition in adaptive cluster sampling. By the previously cited results on multivariate adaptive cluster sampling, in which the adaptive condition could be based on an auxiliary variable rather than the variable of interest, it might be sensible to base the links on information that is less sensitive than drug use or sexual partners. For example, instead of being asked to name other cocaine users, individuals could be asked to name the people with whom they spend the most time.

Network sampling was introduced by Birnbaum and Sirken (1965) to estimate the number of people with a rare disease when a random sample of medical centers was selected. A person with the disease who had been treated at more than one center would have a higher probability of being included in the sample than a person who had been treated at only one center, so the configuration of such linkages needed to be taken into account to allow for unbiased estimation of prevalence. Subsequent uses of network sampling included surveys in which each person in the sample would be asked to report not only on themselves but also on persons, such as siblings, linked to them. A variety of linking rules and sampling designs have been investigated (Czaja et al. 1986; Faulkenberry and Garoui 1991; Kalton and Anderson 1986; Levy 1977; Nathan 1976; Sirken 1970, 1972*a*, 1972*b*; Sirken and Levy 1974; Sudman et al. 1988; Thompson 1992).

The term "snowball sampling" has been applied to a variety of graph sampling procedures (cf., Thompson 1992). In one type (Kalton and Anderson 1986), members of a rare population in an initial sample are asked to identify other members of the population, those so identified are asked to identify others, and so on, for the purpose of obtaining a nonprobability sample or constructing a frame from which to sample. In another type (Goodman 1961), individuals in the sample are asked to identify a fixed number of other individuals, who in turn are asked to identify other individuals, for a fixed number of stages, for the purpose of estimating the number of mutual relationships or social circles in the population. Uses of snowball sampling for surveying drug users and other hidden populations are reviewed by van Meter (1990), who notes the difficulty of putting estimation on a sound statistical basis with such surveys without either assuming a specific stochastic process giving the original sample or the prohibitive requirement of knowing the linkage structure for the entire population. However, recent approaches to estimation with snowball and other graph samples (Frank 1977, 1979; Frank and Snijders 1994; Snijders 1992; Snijders et al. 1995; Spreen 1992; Spreen and Zwaagstra 1994) appear to be very promising. In addition, the estimation methods of network sampling and adaptive cluster sampling apply to a number of graph sampling situations.

In network sampling, the links generally are symmetric, or at least the links between units in the sample and those outside the sample are known. Further, the addition of linked units to the sample does not depend on observed values of the variable of interest. Thus, with network sampling it is possible to calculate the selection or inclusion probability of any person in the sample from the sample data. With that information, unbiased estimates of the population total or mean can be obtained, including the Horvitz-Thompson estimator and the multiplicity estimator (Birnbaum and Sirken 1965).

With the snowball sampling procedures described by Frank (1977, 1979), Frank and Snijders (1994), and Snijders (1992), and with adaptive cluster sampling, estimation is complicated by the fact that the selection or inclusion probabilities cannot generally be calculated from the sample data for every unit in the sample. This results from the asymmetry of some of the directional links, so that for some units (people) in the sample the investigators do not know how many other units would potentially have directed investigators to that unit. In graph sampling terminology, the in-degree of that unit is unknown (Frank 1977). An even more fundamental estimation difficulty for many snowball samples as obtained in practice is the lack of a well-defined probability sampling procedure for obtaining the initial sample.

The snowball sampling procedure described by Frank and Snijders (1994) for estimating the number of people in a hidden population illustrates both the possibilities and the difficulties. The design was used to estimate the number of cocaine users in a town in The Netherlands. The estimates obtained were in fact consistent with current police and social agency estimates. The initial sample of cocaine users was obtained not from a designed probability sample but from police and social service encounters. In the first wave, users in the initial sample were questioned for the names of other cocaine users, and the names edited to eliminate duplicates. In general, the survey procedure involves additional names from a second wave provided by the first wave people and so on, but in this particular study only the first wave was carried out. Estimates were then obtained based on a variety of assumptions. Because the initial sample was not obtained from a deliberate probability sampling procedure, estimation was based on the assumption that it arose as a Bernoulli procedure; that is, it is assumed that each individual in the hidden population had the same unknown probability of being included in the initial sample and inclusion was independent between individuals. For the model-based estimators, the additional assumption was made that directional links between individuals were independent and identically distributed Bernoulli random variables, so that for example whether individual A knows individual B is independent of whether individual B knows individual A. A design-based (subject to the Bernoulli assumption) estimator was also used, based on expanding the number of individuals added in the first wave by dividing that number by the proportion of individuals in the initial sample who were linked to any other individuals in the initial sample. One potential difficulty with such an estimator is the possibility that the proportion of initial sample units so linked would be zero, in which case that estimator could not be calculated.

The design unbiased estimates of adaptive cluster sampling could be used with a graph sampling or snowball procedure such as the one described above provided that the probability basis of the initial sample could be determined and that the addition of linked units was completed through all waves. The ideal would be to have a probability survey sample at the first stage and then follow through using the links provided by self-reported users in the survey sample. With such a survey, nonusers would be sampled along with users. For example, the variable of interest could be cocaine use, with $y_i = 1$ if the i-th individual is a user and $y_i = 0$ otherwise. The total of the y-values is then the total number of users in the population. The estimate would also need to be adjusted for inaccurate reporting as described in the next section. For estimation purposes, a network would consist of a connected set of people, so that for any two people in the network it is possible to get from one to the other through the directed links. Any unit connected to the initial sample only by an asymmetric link—individual j who was reported by person i, but who when questioned similarly does not reveal individual i—would be treated as edge units in the estimation.

Interestingly, since estimation deals only with empirical or observable links and not with any underlying or true links, the unbiasedness of the adaptive cluster sampling estimates of the population total or mean is not affected by misreporting of links. For example, suppose in a survey of intravenous drug users links from each respondent consist of the names of sexual partners given by the respondent to investigators. Suppose that individual i has actual sexual partners j

and *k* but reveals only *k* and not *j* to the investigators. The investigators then add individual *k* to the sample. If individual *k* in turn gives the name of *i* to investigators, then *i* and *k* are in the same network. Individual *j* remains outside the sample unless independently selected as part of the initial sample or of another network. Even if *j* gives the name of *i* to investigators, the two will not be in the same network because of the asymmetric link. Thus, for the design and estimation, the links from an individual are defined to be the names that person would give if included in the sample and asked, not necessarily the actual estimators. Although misreporting of the links would not bias the estimate of the population total, it would bias estimates of interest to epidemiologists regarding the actual pattern of sexual contacts in the population.

The use of link-tracing, chain referral, snowball, networking, or other graph-related sampling methods pervades the field of behavioral and ethnographic studies. Examples include studies of cocaine use and associated sexual behaviors (French 1993; Inciardi 1993), marijuana and cocaine dealing (Adler 1985), marijuana use (True and True 1977), heroin use (Agar 1977; Soloway and Walters 1977), street drug culture (Preble and Miller 1977), opiate addiction in women (Rosenbaum and Murphy 1990), recovery from addiction (Biernacki 1986), prostitution and drug use (McNamara 1994), pickpockets (Inciardi 1977), sexual behavior of selected ethnic or age-defined groups (Sterk-Elifson 1994; Thompson 1994*b*), and sexually transmitted diseases (Bailey and Aunger 1995). Further, even behavioral surveys producing a standard probability sample of individuals often seek estimates of network or graph-related characteristics of the population, such as the sexual networks relevant to sexually transmitted infections (Laumann et al. 1994).
Because of the prevalence of graph-related methods, adaptive and otherwise, for obtaining samples from rare and hidden populations, improvements in design and estimation methodologies for such studies are highly desirable. Some investigators who use snowball and other link-tracing designs to obtain a suitably large sample for study do not try to make inference from the sample to the population from which it comes but instead include in their study description a disclaimer that statistical inference is impossible or questionable. The disclaimer misses the point, however, because characteristics of the sample are commonly summarized as means or percentages, and the interpretation of the meaning of such means and percentages requires consideration of how the sample was obtained. Quite possibly, a more meaningful sample mean or percentage could be obtained using a weighted average with the weights reflecting the distinct networks in

the sample. When a random initial sample is not possible to obtain, Snijders (1992) suggests drawing respondents as much as possible from independent sources to satisfy the assumptions of estimation methods as closely as possible.

NONSAMPLING ERRORS IN CONVENTIONAL AND ADAPTIVE SAMPLING

Because of the sensitivity of issues related to behavioral characteristics such as drug use, such surveys potentially involve prominent nonsampling errors related to incomplete candor in self-reporting, imperfect detectability of persons in high-risk groups, and other special factors in addition to the usual nonsampling errors associated with frame development, nonresponse, measurement, and data recording (Biemer et al. 1991; Lessler and Kalsbeek 1992). Sources of nonsampling variability in surveys of drug use and other sensitive behaviors include untruthful or incorrect self-reporting (Gfroerer et al. 1992; Rouse et al. 1985), inconsistent answers to survey questions (Cox et al. 1992), misinterpretations of questions by respondents (Forsyth et al. 1992), and item and unit nonresponse (Rubin 1987; Witt et al. 1992; Caspar 1992, Graham et al. 1994). Underreporting, self-selection bias, and other sources of nonsampling errors in sexual behavior surveys are reviewed in Clement (1990) and Berk and colleagues (1995). The role of nonsampling errors in general in surveys of sensitive topics is discussed in Turner and associates (1992*a*). Obtaining the best possible estimates from surveys involves, in addition to using a good sampling design, developing methods for reducing nonsampling errors and using methods to assess and adjust for nonsampling errors that do occur. Adjusting for nonsampling errors can be illustrated by modeling incomplete self-reporting in a drug use survey as a problem in detectability. Assume a sample of households and people within households is selected according to the survey design, but not all drug users in the sample are detected by the self-reporting. Suppose for illustration that independent studies comparing self-reported use to bioassay results indicate that only half of users report use, so that the rate of detection is 50 percent. Then a simple form of adjusted estimate, whether of prevalence of use or of another variable such as amount spent on drugs during a 2-week period, is obtained by taking the naive estimate from the survey and dividing by one-half, so that the adjusted estimate is twice the initial estimate. The effects of such adjustments on survey estimates are analyzed in Thompson and Seber (1994). When the detection rates differ for different subpopulations or for different kinds of individuals, separate adjustments can be made for each observation based on the individual detectability rate that applies (Thompson and Seber 1994). Imperfect detectability in surveys has been estimated with such techniques as double sampling, distance sampling, and capture-recapture methods, as well as studies comparing self-reported and bioassay results.

With any conventional sampling design, nonsampling errors affect only the values recorded for units in the sample, while with an adaptive sampling design nonsampling errors may additionally affect what sample is selected. Potentially, the problem looks much more complicated for the adaptive design, but a conditioning argument given every possible sequence shows that adjustment and analysis methods are as straightforward for adaptive as for conventional designs (Thompson and Seber 1994).

The adjustment methods for imperfect detectability are required to produce realistic estimates of population characteristics, while the analysis is required to evaluate the effect of each source of sampling and nonsampling error and to determine the most effective means of improving estimates and reducing overall mean square error. In surveys of self-reported drug use, estimates of prevalence can be adjusted to account for the estimated proportion of users reporting no use, but the variance of the resulting estimates includes the following three important terms. The first is sampling variance due to the difference of one sample of households from another under the design. The second is a detectability error due to drug users reporting that they do not use. The third term is associated with the uncertainty in estimating the proportion of users who report no use; such estimates typically come from comparative and criterion studies. To reduce the uncertainty associated with the first component involves improving sampling design, increasing sample sizes, and using more efficient estimation methods. Reduction of the second term requires interview methods that increase the accuracy of self-reporting among users. Such methods include question wording and interview mode of administration, such as face-to-face, computer, or telephone (Gfroerer and Hughes 1992; Schober et al. 1992; Turner et al. 1992*a*, 1992*b*). Reduction of the third term is achieved with larger, more specific, and more effective comparative and criterion studies, such as comparisons between self-reporting and the results of bioassays such as hair or urine tests.

ACKNOWLEDGMENT

Research on which this article is based was supported by the National Science Foundation.

REFERENCES

Adler, P. Ethnographic research on hidden populations: penetrating the drug world. In: Lambert, E.Y., ed. *The Collection and Interpretation of Data from Hidden Populations*. National Institute on Drug Abuse Research Monograph 98. DHHS Pub. No. (ADM)90-1678. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1990.

Adler, P. *Wheeling and Dealing, an Ethnography of an Upper-Level Drug Dealing and Smuggling Community*. New York: Columbia University Press, 1985.

Agar, M.H. Ethnography in the streets and in the joint. In: Weppner, R.S., ed. *Street Ethnography: Selected Studies of Crime and Drug Use in Natural Settings*. Beverly Hills, CA: Sage, 1977. pp. 143-155.

Bailey, R.C., and Aunger, R.V. Sexuality, infertility, and sexually transmitted disease among farmers and foragers in Central Africa. In: Abramson, P.R., and Pinkerton, S.D., eds. *Sexual Nature, Sexual Culture*. Chicago: University of Chicago Press, 1995. pp. 195-222.

Berk, R.; Abramson, P.R.; and Okami, P. Sexual activities as told in surveys. In: Abramson, P.R., and Pinkerton, S.D., eds. *Sexual Nature, Sexual Culture*. Chicago: University of Chicago Press, 1995. pp. 371-386.

Biemer, P.P.; Groves, R.M.; Lyberg, L.E.; Mathiowetz, N.A.; and Sudman, S. *Measurement Errors in Surveys*. New York: Wiley, 1991.

Biernacki, P. *Pathways from Heroin Addiction*. Philadelphia: Temple University Press, 1986.

Birnbaum, Z.W., and Sirken, M.G. *Design of Sample Surveys To Estimate the Prevalence of Rare Diseases: Three Unbiased Estimates*. National Center for Health Statistics, Vital and Health Statistics, Series 2, No. 11. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1965.

Brown, J.A. The application of adaptive cluster sampling to ecological studies. In: Fletcher, D.J., and Manly, B.F.J., eds. *Statistics in Ecology and Environmental Monitoring*. Otago Conference Series No. 2. Dunedin, New Zealand: University of Otago Press, 1994. pp. 86-97.

Caspar, R. Followup of nonrespondents in 1990. In: Turner, C.F.; Lessler, J.T.; and Gfroerer, J.C., eds. *Survey Measurement of Drug Use: Methodological Studies*. DHHS Pub. No. (ADM)92-1929. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1992. pp. 155-173.

Chernoff, H. *Sequential Analysis and Optimal Design*. Philadelphia: SIAM, 1972.

Clement, U. Surveys of heterosexual behavior. (Translated by U. Dous.) *Ann Rev Sex Res* 1:45-74, 1990.

Cochran, W.G. *Sampling Techniques*. 3d ed. New York: Wiley, 1977.

Cox, B.; Witt, M.; Traccarella, M.; and Perez-Michael, A. Inconsistent reporting of drug use in 1988. In: Turner, C.F.; Lessler, J.T.; and Gfroerer, J.C., eds. *Survey Measurement of Drug Use: Methodological Studies*. DHHS Pub. No. (ADM)92-1929. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1992. pp. 109-154.

Czaja, R.F.; Snowdon, C.B.; and Casady, R.J. Reporting bias and sampling errors in a survey of a rare population using multiplicity counting rules. *J Am Stat Assoc* 81:411-419, 1986.

Danaher, P.J., and King, M. Estimating rare household characteristics using adaptive sampling. *N Z Stat* 29:14-23, 1994.

Englund, E.J., and Herari, N. Phased sampling for soil remediation. *Environm Ecolog Stat*, in press.

Faulkenberry, G.D., and Garoui, A. Estimating a population total using an area frame. *J Am Stat Assoc* 86:445-449, 1991.

Forsyth, B.; Lessler, J.; and Hubbard, M. Cognitive evaluation of the questionnaire. In: Turner, C.F.; Lessler, J.T.; and Gfroerer, J.C., eds. *Survey Measurement of Drug Use: Methodological Studies*. DHHS Pub. No. (ADM) 92-1929. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1992. pp. 13-52.

Francis, R.I.C.C. An adaptive strategy for stratified random trawl surveys. *N Z J Marine Freshwater Res* 18:59-71, 1984.

Frank, O. Survey sampling in graphs. *J Stat Plan Inference* 1:235-264, 1977.

Frank, O. Estimation of population totals by use of snowball samples. In: Holland, P.W., and Leinhardt, S., eds. *Perspectives on Social Network Research*. New York: Academic Press, 1979.

Frank, O., and Snijders, T. Estimating the size of hidden populations using snowball sampling. *J Official Stat* 10:53-67, 1994.

Freedman, D.A. Adjusting the 1990 census. *Science* 252:1233-1236, 1991.

French, J.F. Pipe dreams: Crack and the life in Philadelphia and Newark. In: Ratner, M.S., ed. *Crack Pipe as Pimp*. New York: Lexington, 1993. pp. 205-232.

Gasaway, W.C.; DuBois, S.D.; Reed, D.J.; and Harbo, S.J. "Estimating Moose Population Parameters from Aerial Surveys." Biological Papers of the University of Alaska (Institute of Arctic Biology) Number 22. Fairbanks, AK: University of Alaska, 1986.

Geiger, H.J. A Bayesian approach for estimating hatchery contribution in a series of salmon fisheries. *Alaska Fishery Res Bull* 1:66-75, 1994.

Gfroerer, J., and Hughes, A. Collecting data on illicit drug use by phone. In: Turner, C.F.; Lessler, J.T.; and Gfroerer, J.C., eds. *Survey Measurement of Drug Use: Methodological Studies*. DHHS Pub. No. (ADM)92-1929. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1992. pp. 277-295.

Gfroerer, J.; Gustin, J.; and Turner, C. Introduction. In: Turner, C.F.; Lessler, J.T.; and Gfroerer, J.C., eds. *Survey Measurement of Drug Use: Methodological Studies*. DHHS Pub. No. (ADM)92-1929. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1992. pp. 3-10.

Goldstein, P.J.; Spunt, B.J.; Miller, T.; and Bellucci, P. Ethnographic field stations. In: Lambert, E.Y., ed. *The Collection and Interpretation of Data from Hidden Populations*. National Institute on Drug Abuse Research Monograph 98. DHHS Pub. No. (ADM)90-1678. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1990. pp. 80-95.

Goodman, L.A. Snowball sampling. *Ann Math Stat* 32:148-170, 1961.

Graham, J.W.; Hofer, S.M.; and Piccinin, A.M. Analysis with missing data in drug prevention research. In: Collins, L.M., and Seitz, L.A., eds. *Advances in Data Analysis for Prevention Intervention Research*. National Institute of Drug Abuse Research Monograph 142. NIH Pub. No. 94-3599. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1994. pp. 13-63.

Inciardi, J.A. In search of the class cannon: A field study of professional pickpockets. In: Weppner, R.S., ed. *Street Ethnography: Selected Studies of Crime and Drug Use in Natural Settings.* Beverly Hills, CA: Sage, 1977. pp. 55-77.

Inciardi, J.A. Kingrats, chicken heads, slow necks, freaks, and blood suckers: A glimpse at the Miami sex-for-crack market. In: Ratner, M.S., ed. *Crack Pipe as Pimp.* New York: Lexington, 1993. pp. 37-67.

Jolly, G.M., and Hampton, I. A stratified random transect design for acoustic surveys of fish stocks. *Can J Fisheries Aquatic Sci* 47:1282-1291, 1990.
Kalton, G., and Anderson, D.W. Sampling rare populations. *J Royal Stat Soc Ser A* 149:65-82, 1986.
Kremers, W.K. "Adaptive Sampling to Account for Unknown Variability among Strata." Preprint No. 128. Institut für Mathematik, Universitat Augsburg, Federal Republic of Germany, 1987.
Lambert, E.Y., ed. *The Collection and Interpretation of Data from Hidden Populations*. National Institute on Drug Abuse Research Monograph 98. DHHS Pub. No. (ADM)90-1678. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1990.
Laumann, E.O.; Gagnon, J.H.; Michael, R.T.; and Michaels, S. *The Social Organization of Sexuality, Sexual Practices in the United States*. Chicago: University of Chicago Press, 1994.
Lessler, J.T., and Kalsbeek, W.D. *Nonsampling Error in Surveys*. New York: Wiley, 1992.
Levy, P.S. Optimum allocation in stratified random network sampling for estimating the prevalence of attributes in rare populations. *J Am Stat Assoc* 72:758-763, 1977.
McNamara, R.P. *The Times Square Hustler: Male Prostitution in New York City*. Westport, CT: Praeger, 1994.
Munholland, P.L., and Borkowski, J.J. "Adaptive Latin Square Sampling + 1 Designs." Technical Report No. 3-23-93. Bozeman, MT: Department of Mathematical Sciences, Montana State University, 1993.
Nathan, G. An empirical study of response and sampling errors for multiplicity estimates with different counting rules. *J Am Stat Assoc* 71:808-815, 1976.
Preble, E., and Miller, T. Methadone, wine, and welfare. In: Weppner, R.S., ed. *Street Ethnography: Selected Studies of Crime and Drug Use in Natural Settings*. Beverly Hills, CA: Sage, 1977. pp. 229-248.
Roesch, F.A., Jr. Adaptive cluster sampling for forest inventories. *Forest Sci* 39:655-669, 1993.

Rosenbaum, M., and Murphy, S. Women and addiction: Process, treatment, and outcome. In: Lambert, E.Y., ed. *The Collection and Interpretation of Data from Hidden Populations*. National Institute on Drug Abuse Research Monograph 98. DHHS Pub. No. (ADM)90-1678. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1990.

Rouse, B.A.; Kozel, N.J.; Louise G.; and Richards, L.G., eds. *Self-Report Methods of Estimating Drug Use, Meeting Current Challenges to Validity*. National Institute on Drug Abuse Research Monograph 57. DHHS Pub. No. (ADM)85-1402. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985.

Rubin, D.B. *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley, 1987.

Schober, S.; Fe Caces, M.; Pergamit, M.; and Branden, L. Effect of mode of administration on reporting of drug use in the national longitudinal survey. In: Turner, C.F.; Lessler, J.T.; and Gfroerer, J.C., eds. *Survey Measurement of Drug Use: Methodological Studies*. DHHS Pub. No. (ADM)92-1929. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1992. pp. 267-276.

Seber, G.A.F., and Thompson, S.K. Environmental adaptive sampling. In: Patil, G.P., and Rao, C.R., eds. *Handbook of Statistics.* Vol. 12, *Environmental Statistics*. New York: North Holland/Elsevier Science Publishers, 1993. pp. 201-220.

Siegmund, D. *Sequential Analysis: Tests and Confidence Intervals*. New York: Springer, 1985.

Sirken, M.G. Household surveys with multiplicity. *J Am Stat Assoc* 63:257-266, 1970.

Sirken, M.G. Stratified sample surveys with multiplicity. *J Am Stat Assoc* 67:224-227, 1972*a*.

Sirken, M.G. Variance components of multiplicity estimators. *Biometrics* 28:869-873, 1972*b*.

Sirken, M.G., and Levy, P.S. Multiplicity estimation of proportions based on ratios of random variables. *J Am Stat Assoc* 69:68-73, 1974.

Smith, D.R.; Conroy, M.J.; and Brakhage, D.H. Efficiency of adaptive cluster sampling for estimating density of wintering waterfowl. *Biometrics* 51:777-788, 1995.

Snijders, T.A.B. Estimation on the basis of snowball samples: How to weight. *Bull Methodologie Sociologique* 36:59-70, 1992.

Snijders, T.; Spreen, M.; and Zwaagstra, R. The use of multilevel modeling for analysing personal networks: Networks of cocaine users in an urban area. *J Quant Anthropol* 5:85-105, 1995.

Solomon H., and Zacks, S. Optimal design of sampling from finite populations: A critical review and indication of new research areas. *J Am Stat Assoc* 65:653-677, 1970.

Soloway, I., and Walters, J. Workin' the corner: The ethics and legality of ethnographic fieldwork among active heroin addicts. In: Weppner, R.S., ed. *Street Ethnography: Selected Studies of Crime and Drug Use in Natural Settings*. Beverly Hills, CA: Sage, 1977. pp. 159-178.

Spreen, M. Rare populations, hidden populations, and link-tracing designs; what and why? *Bull Methodologie Sociologique* 36:34-58, 1992.

Spreen, M., and Zwaagstra, R. Personal network sampling, outdegree analysis and multilevel analysis: Introducing the network concept in studies of hidden populations. *Int Sociol* 9:475-491, 1994.

Sterk-Elifson, C. Sexuality among African-American women. In: Rossi, A.S., ed. *Sexuality Across the Life Course*. Chicago: University of Chicago Press, 1994. pp. 99-126.

Sudman, S.; Sirken, M.G.; and Cowan, C.D. Sampling rare and elusive populations. *Science* 240:991-996, 1988.

Thompson, S. Changing lives, changing genres: Teenage girls' narratives about sex and romance, 1978-1986. In: Rossi, A.S., ed. *Sexuality Across the Life Course*. Chicago: University of Chicago Press, 1994*b*. pp. 209-232.

Thompson, S.K. Adaptive sampling. *Proceedings of the Section on Survey Research Methods of the American Statistical Association*. Arlington, VA: American Statistical Association, 1988. pp. 784-786.

Thompson, S.K. Adaptive cluster sampling. *J Am Stat Assoc* 85:1050-1059, 1990.

Thompson, S.K. Adaptive cluster sampling: Designs with primary and secondary units. *Biometrics* 47:1103-1115, 1991*a*.

Thompson, S.K. Stratified adaptive cluster sampling. *Biometrika* 78:389-397, 1991*b*.

Thompson, S.K. *Sampling*. New York: John Wiley and Sons, 1992.

Thompson, S.K. Multivariate aspects of adaptive cluster sampling. In: Patil, G.P., and Rao, C.R., eds. *Multivariate Environmental Statistics*. New York: North Holland/Elsevier Science Publishers, 1993. pp. 561-572.

Thompson, S.K. "Factors Influencing the Efficiency of Adaptive Cluster Sampling." Technical Report 94-0301. University Park, PA: Center for Statistical Ecology and Environmental Statistics, Department of Statistics, Pennsylvania State University, 1994*a*.

Thompson, S.K. Adaptive cluster sampling based on order statistics. *Environmetrics*, in press.

Thompson, S.K., and Seber, G.A.F. Detectability in conventional and adaptive sampling. *Biometrics* 50:712-724, 1994.

Thompson, S.K., and Seber, G.A.F. *Adaptive Sampling*. New York: Wiley, 1996.

Thompson, S.K.; Ramsey, F.L.; and Seber, G.A.F. An adaptive procedure for sampling animal populations. *Biometrics* 48:1195-1199, 1992.

True, W.R., and True, J.H. Network analysis as a methodological approach to the study of drug use in a Latin city. In: Weppner, R.S., ed. *Street Ethnography: Selected Studies of Crime and Drug Use in Natural Settings*. Beverly Hills, CA: Sage, 1977. pp. 125-141.

Turner, C.F.; Lessler, J.T.; and Devore, J. Effects of mode of administration and wording on reporting of drug use. In: Turner, C.F.; Lessler, J.T.; and Gfroerer, J.C., eds. *Survey Measurement of Drug Use: Methodological Studies*. DHHS Pub. No. (ADM)92-1929. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1992*a*. pp. 177-220.

Turner, C.F.; Lessler, J.T.; George, B.; Hubbard, M.; and Witt, M. Effects of mode of administration and wording on data quality. In: Turner, C.F.; Lessler, J.T.; and Gfroerer, J.C., eds. *Survey Measurement of Drug Use: Methodological Studies*. DHHS Pub. No. (ADM)92-1929. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1992*b*. pp. 221-244.

Turner, C.F.; Lessler, J.T.; and Gfroerer, J.C. Future directions for research and practice. In: Turner, C.F.; Lessler, J.T.; and Gfroerer, J.C., eds. *Survey Measurement of Drug Use: Methodological Studies*. National Institute of Drug Abuse. DHHS Pub. No. (ADM)92-1929. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1992*c*. pp. 299-306.

Van Meter, K.M. Methodological and design issues: Techniques for assessing the representatives of snowball samples. In: Lambert, E.Y., ed. *The Collection and Interpretation of Data from Hidden Populations*. National Institute on Drug Abuse Research Monograph 98. DHHS Pub. No. (ADM)90-1678. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1990.

Wald, A. *Sequential Analysis*. New York: Wiley, 1947.

Weppner, R.S. *Street Ethnography: Selected Studies of Crime and Drug Use in Natural Settings*. Beverly Hills, CA: Sage, 1977.
Wiebel, W.W. Identifying and gaining access to hidden populations. In: Lambert, E.Y., ed. *The Collection and Interpretation of Data from Hidden Populations*. National Institute on Drug Abuse Research Monograph 98. DHHS Pub. No. (ADM)90-1678. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1990.
Witt, M.; Pantula, J.; Folsom, R.; and Cox, B. Item nonresponse in 1988. In: Turner, C.F.; Lessler, J.T.; and Gfroerer, J.C., eds. *Survey Measurement of Drug Use: Methodological Studies*. DHHS Pub. No. (ADM)92-1929. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1992. pp. 85-108.
Wolter, K.M. Some coverage error models for census data. *J Am Stat Assoc* 81:338-346, 1986.
Wolter, K.M. Accounting for America's uncounted and miscounted. *Science* 253:12-15, 1991.
Woodroofe, M. *Nonlinear Renewal Theory in Sequential Analysis*. Philadelphia: SIAM, 1982.
Zacks, S. Bayes sequential designs of fixed size samples from finite populations. *J Am Stat Assoc* 64:1342-1349, 1969.

AUTHOR

Steven K. Thompson, Ph.D.
Associate Professor
Department of Statistics
Pennsylvania State University
326 Classroom Building
University Park, PA 16802

Click here to go to page 320