J1
802:
S19
pt. 2
c. 2

PROPERTY OF LIBRARY
OFFICE OF STATE BUDGET

SCHS
State Center For Health Statistics

STATISTICAL PRIMER

Department of Human Resources • Division of Health Services • P.O. Box 2091 • Raleigh, NC 27602-2091 • 919/733-4728

No. 6 July 1984

SAMPLING AND MEASUREMENT ERROR
PART 2: ERRORS IN VITAL RATES

by

Paul A. Buescher

Introduction

Part 1 of this statistical primer on sampling and measurement error (May 1984) dealt primarily with the problems of sampling errors. In addition, some general concepts were presented concerning random vs. nonrandom errors and sampling vs. nonsampling errors. The reader may want to refer to Part 1 for background in these areas. This second part addresses the problem of errors in measures based on a "complete count," and particularly the problem of errors in vital rates. Formulas are presented for calculating random errors of measurement in terms of the standard error of simple rates, and the standard error of the difference between two rates. The general concept of a standard error is explained in Part 1 of this statistical primer.

Random Measurement Error In a Complete Enumeration

The discussion in Part 1 considered random error associated with measurements based on a sample, but the basic concept of random variability in surveys also applies to measurement error in general. Even a measure based on a "complete count" of the population has a random error component that can be assessed in terms of standard errors. Such a measure may also contain a very large nonrandom error or bias, due to such factors as a poor measurement instrument or inadequate field procedures, and it is of course very important to minimize this type of error. But this paper addresses only the problem of random error. A measure based on a complete enumeration that is observed at a point in time may be considered as one possible outcome of a chance process, in addition to the problem of random errors due to imperfect measurements.

Let us use the example of random error in a death rate (number of deaths in a given year per 1000 population) that is based on complete registration of deaths. The rate measured in a given year can be considered as a sample of one of a large number of possible measurements, all of which cluster in a normal distribution around the "true" death rate of the population. In fact, the larger the numerator of the measured death rate, the smaller the standard error associated with this "sampling distribution," and therefore the better this death rate will estimate the true or underlying rate of the population. This idea of an "underlying" rate is an abstract concept, since the death rate observed in one year did actually occur, but it is this underlying rate that health policies should seek to address rather than annual rates which may fluctuate dramatically. Since death occurring to an individual is subject to chance, the death rate for the population may vary from one time to another even if the force of mortality or underlying death rate remains constant.

The concepts presented below apply to the standard error of any simple or "crude" rate (death rate, birth rate, marriage rate, etc.), which is probably the most frequently used type of health measure. Standard errors may also be estimated for adjusted rates and other more complex measures, but description of these is beyond the scope of the present paper (*1*).

Any rate with a small number of events in the numerator will be unstable, with possibly large random fluctuations from year to year that do not comprise a significant trend. It has been shown that events of a rare nature follow a Poisson probability distribution and that the following approximation may be used to estimate the **standard error of a simple rate**. In the case of a death rate, where D is the number of deaths in the numerator and P