EXHIBIT N

Sampling of Populations

Methods and Applications

Fourth Edition

Paul S. Levy
RTI International
Statistical Research Division
Research Triangle Park, North Carolina

Stanley Lemeshow
The Ohio State University
College of Public Health
Columbus, Ohio

WILEY

A JOHN WILEY & SONS, INC., PUBLICATION

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data:

Levy, Paul S.
Sampling of populations : methods and applications / Paul S. Levy, Stanley Lemeshow. — 4th ed.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-470-04007-2 (cloth)
1. Population—Statistical methods. 2. Sampling (Statistics) I. Lemeshow, Stanley. II. Title.
HB849.49.L48 2008
304.601'51952—dc21 2008004934

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

Table 2.9 Data for the Burn Area Estimates

Student	Mean (%)	Variance $(\%)^2$
Dave	37	64
Don	42	9
Virginia	50	9

where z is the standard normal deviate.

Don's burn measurements are usually high, but it is not often that he would be more than 10 percentage points away from the true value. For Don

$$\Pr(X > 47) + \Pr(X < 27) = \Pr\left(z > \frac{47 - 42}{\sqrt{9}} = 1.67\right) + \Pr(z < -5.00) = .05.$$

Finally, as reflected by her large MSE, the probability that Virginia misses the target burn area by more than 10 percentage points is quite high:

$$\Pr(X > 47) + \Pr(X > 27) = \Pr\left(z > \frac{47 - 50}{\sqrt{9}} = -1\right) + \Pr(z < -7.67) = .84.$$

We have seen in this example that, when evaluating particular estimates, it is important to examine both bias and variance. Both entities play important roles in determining the size of the mean square error.

2.4.3 Validity, Reliability, and Accuracy

In earlier sections, we spoke of the desirability of using sample designs that yield reliable and valid estimates. However, we have never defined just what the terms "reliable" and "valid" mean in terms of characteristics of estimates.

We now have developed enough concepts and notation concerning estimates to be able to define these two terms as well as a third term, the "accuracy" of an estimate, which we will see is derived from the validity and reliability.

The *reliability* of an estimated population characteristic refers to how reproducible the estimator is over repetitions of the process yielding the estimator. If we assume that there is no measurement error in the survey, then the reliability of an estimator can be stated in terms of its sampling variance or, equivalently, its standard error. The smaller the standard error of an estimator, the greater is its reliability.

The *validity* of an estimated population characteristic refers to how the mean of the estimator over repetitions of the process yielding the estimate, differs from the true value of the parameter being estimated. Again, if we assume that there is no measurement error, the validity of an estimator can be evaluated by examining the bias of the estimator. The smaller the bias, the greater is the validity.