EXHIBIT O

☆ ☆ ☆ ☆ ☆ ☆

Sixth Edition

STATISTICAL METHODS

GEORGE W. SNEDECOR

Professor Emeritus of Statistics
and Former Director, Statistical Laboratory
Iowa State University

WILLIAM G. COCHRAN

Professor of Statistics
Harvard University

THE IOWA STATE UNIVERSITY PRESS AMES, IOWA, U.S.A.

GEORGE W. SNEDECOR is professor emeritus of statistics, Iowa State University, where he taught from 1913 to 1958 and where he was for fourteen years director of the statistical laboratory. His writings include a body of scientific journal articles, research bulletins, and books, including *Correlation and Machine Calculation* (with H. A. Wallace), *Calculation and Interpretation of Analysis of Variance and Covariance*, and *Statistical Methods*. He holds a master of science degree from the University of Michigan, and honorary doctor of science degrees from North Carolina State University and Iowa State University. He is a member of the International Statistical Institute, past president of the American Statistical Association, and an honorary Fellow of the British Royal Statistical Society. He has served also as consultant, Human Factors Division, U.S. Navy Electronics Laboratory, San Diego, California, where he now lives.

WILLIAM G. COCHRAN is professor of statistics, Harvard University. He has served formerly on the faculties of Johns Hopkins University, North Carolina State University, and Iowa State University. He holds master of arts degrees from Glasgow University and Cambridge University and an honorary master of arts degree from Harvard University. He is past president of the American Statistical Association, the Institute of Mathematical Statistics, and the Biometric Society. His writings include many research papers in the professional journals of his field; *Sampling Techniques*, 2nd ed., 1963; and *Experimental Designs* (with Gertrude M. Cox), 2nd ed., 1957.

Printed in the U.S.A.

First edition, 1937
Second edition, 1938
Third edition, 1940

Fourth edition, 1946
Second printing, 1946
Third printing, 1948
Fourth printing, 1950
Fifth printing, 1953
Sixth Printing, 1955
Seventh printing, 1955

Fifth edition, 1956
Second printing, 1957
Third printing, 1959
Fourth printing, 1961
Fifth printing, 1962
Sixth printing, 1964
Seventh printing, 1965
Eighth printing, 1966

Sixth edition, 1967
Second printing, 1968
Third printing, 1969
Fourth printing, 1971

International Standard Book Number: 0–8138–1560–6
Library of Congress Catalog Card Number: 67–21577

Preface

In preparing the sixth edition we have kept in mind the two purposes this book has served during the past thirty years. Prior editions have been used extensively both as texts for introductory courses in statistics and as reference sources of statistical techniques helpful to research workers in the interpretation of their data.

As a text, the book contains ample material for a course extending throughout the academic year. For a one-term course, a suggested list of topics is given on the page preceding the Table of Contents. As in past editions, the mathematical level required involves little more than elementary algebra. Dependence on mathematical symbols has been kept to a minimum. We realize, however, that it is hard for the reader to use a formula with full confidence until he has been given proof of the formula or its derivation. Consequently, we have tried to help the reader's understanding of important formulas either by giving an algebraic proof where this is feasible or by explaining on common-sense grounds the roles played by different parts of the formula.

This edition retains also one of the characteristic features of the book—the extensive use of experimental sampling to familiarize the reader with the basic sampling distributions that underlie modern statistical practice. Indeed, with the advent of electronic computers, experimental sampling in its own right has become much more widely recognized as a research weapon for solving problems beyond the current skills of the mathematician.

Some changes have been made in the structure of the chapters, mainly at the suggestion of teachers who have used the book as a text. The former chapter 8 (Large Sample Methods) has disappeared, the retained material being placed in earlier chapters. The new chapter 8 opens with an introduction to probability, followed by the binomial and Poisson distributions (formerly in chapter 16). The discussion of multiple regression (chapter 13) now precedes that of covariance and multiple covariance (chapter 14).

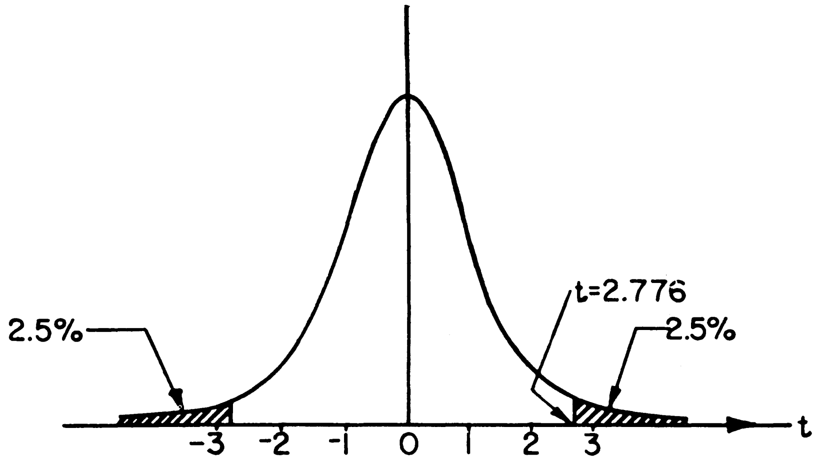

FIG. 2.15.1—Distribution of t with 4 degrees of freedom. The shaded areas comprise 5% of the total area. The distribution is more peaked in the center and has higher tails than the normal.

as "*Student's*" *t-distribution*, this result was discovered by W. S. Gosset in 1908 (7) and perfected by R. A. Fisher in 1926 (8). This distribution has revolutionized the statistics of small samples. In the next chapter you will be asked to verify the distribution by the same kind of sampling process you used for chi-square; indeed, it was by such sampling that Gosset first learned about it.

The quantity t is given by the equation,

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}}$$

That is, t is the deviation of the estimated mean from that of the population, measured in terms of $s/\sqrt{n}$ as the unit. We do not know μ though we may have some hypothesis about it. Without μ, t cannot be calculated; but its sampling distribution has been worked out.

The denominator, $s/\sqrt{n}$, is a useful quantity estimating $\sigma/\sqrt{n}$, the *standard error* of $\bar{X}$.

The distribution of t is laid out in table A 4, p. 549. In large samples it is practically normal with $\mu = 0$ and $\sigma = 1$. It is only for samples of less than 30 that the distinction becomes obvious.

Like the normal, the t-distribution is symmetrical about the mean. This allows the probability in the table to be stated as that of a larger absolute value, sign ignored. For a sample of size 5, with 4 degrees of freeedom, figure 2.15.1 shows such values of t in the shaded areas; 2.5% of them are in one tail and 2.5% in the other. Effectively, the table shows the two halves of the figure superimposed, giving the sum of the shaded areas (probabilities) in both.

EXAMPLE 2.15.1—In the vitamin C sampling of table 2.8.1, $s_{\bar{X}} = 3.98/\sqrt{17} = 0.965$ mg./100 gm. Set up the hypothesis that $\mu = 17.954$ mg./100 gm. Calculate t. Ans. 2.12.

EXAMPLE 2.15.2—For the vitamin C sample, degrees of freedom $= 17 - 1 = 16$, the denominator of the fraction giving s^2. From table A 4, find the probability of a value of t larger in absolute value than 2.12. Ans. 0.05. This means that, among random samples of $n = 17$ from normal populations, 5% of them are expected to have t-values below -2.12 or above 2.12.

EXAMPLE 2.15.3—If samples of $n = 17$ are randomly drawn from a normal population and have t calculated for each, what is the probability that t will fall between -2.12 and $+2.12$? Ans. 0.95.

EXAMPLE 2.15.4—If random samples of $n = 17$ are drawn from a normal population, what is the probability of t greater than 2.12? Ans. 0.025.

EXAMPLE 2.15.5—What size of sample would have $|t| > 2$ in 5% of all random samples from normal populations? Ans. 61. (Note the symbol for "absolute value," that is, ignoring signs.)

EXAMPLE 2.15.6—Among very large samples ($d.f. = \infty$), what value of t would be exceeded in 2.5% of them? Ans. 1.96.

2.16—Confidence limits for μ based on the t-distribution. With σ known, the 95% limits for μ were given by the relations

$$\bar{X} - 1.96\sigma/\sqrt{n} \leq \mu \leq \bar{X} + 1.96\sigma/\sqrt{n}$$

When σ is replaced by s, the only change needed is to replace the number 1.96 by a quantity which we call $t_{0.05}$. To find $t_{0.05}$, read table A 4 in the column headed 0.050 and find the value of t for the *number of degrees of freedom in s*. When the *d.f.* are infinite, $t_{0.05} = 1.960$. With 40 *d.f.*, $t_{0.05}$ has increased to 2.021, with 20 *d.f.* it has become 2.086, and it continues to increase steadily as the number of *d.f.* decline.

The inequalities giving the 95% confidence limits then become

$$\bar{X} - t_{0.05}s/\sqrt{n} \leq \mu \leq \bar{X} + t_{0.05}s/\sqrt{n}$$

As illustration, recall the vitamin C determinations in table 2.8.1; $n = 17$, $\bar{X} = 20$ and $s_{\bar{X}} = 0.965$ mg./100 gm. To get the 95% confidence interval (interval estimate):

1. Enter the table with $d.f. = 17 - 1 = 16$ and in the column headed 0.05 take the entry, $t_{0.05} = 2.12$.
2. Calculate the quantity,

$$t_{0.05}s_{\bar{X}} = (2.12)(0.965) = 2.05 \text{ mg./100 gm.}$$

3. The confidence interval is from

$$20 - 2.05 = 17.95 \text{ to } 20 + 2.05 = 22.05 \text{ mg./100 gm.}$$

If you say that μ lies inside the interval from 17.95 to 22.05 mg./100 gm., you will be right unless a 1-in-20 chance has occurred in the sampling.

The point and 95% interval estimate of μ may be summarized this way: 20 ± 2.05 mg./100 gm.