Statistics for Chemistry

Hubert L. Youmans

Statistics for Chemistry

Statistics for Chemistry

Hubert L. Youmans
Western Carolina University

Charles E. Merrill Publishing Company
A Bell & Howell Company
Columbus, Ohio

Published by
Charles E. Merrill Publishing Co.
A Bell & Howell Company
Columbus, Ohio 43216

International Standard Book Number: 0-675-09042-3
Library of Congress Catalog Number: 72-86906
1 2 3 4 5 6 7 8 9 10—73 74 75 76 77 78

Printed in the United States of America

Preface

This manual resulted from efforts to make simple data treatment and experimental design methods palatable to undergraduates in analytical chemistry. Most of the topics are treated in varying degrees in standard analytical texts. The methods usually, however, are sketchy or excessively theoretical.

In a two-course undergraduate sequence in quantitative analysis, one can use all the methods of the manual in laboratory work. But with increasing emphasis on quantitation in freshman chemistry, the manual should find application as an adjunct there. In fact, I judge it as about the minimum level any chemistry major should master and, to attain this level, more than the efforts of the analytical teacher are required.

Sir Ronald Fisher may be considered the father of small-sample statistics. If he wasn't the father, then he was the midwife. His book, *Statistical Methods for Research Workers*, was published in 1925. It introduced researchers to statistics as a tool in experimental design and interpretation. Today, the methods of statistics are increasingly used in laboratory work, but to many workers the estimated population standard deviation still represents the zenith of statistics. Chemists may not be slow to take a chance, but they seem slow to use it. But electronic computers are changing that.

The manual is meant to be one small step man may take into the world of using chance. The material is somewhat cookbookish. It was meant to be. But once a person becomes familiar with ideas and terms covered in it, he can progress to an increasing understanding of statistics and experimental design. This can be by self-study or by attending courses.

Statistics is the sort of tool that a person should become more proficient in by use. But one can't take a series of courses in everything and I think a series in statistics is precluded for most chemists. One can, however, continuously build on a small foundation. There are many excellent books for this, as indicated in Chapter 1. After a certain proficiency is attained, primary journal articles are more useful than books. *Analytical Chemistry* reviews these biannually; the reviews are good secondary sources to learn of new and useful experimental designs.

Tabulation of data at the various stages of accumulation and assimilation makes statistical methods easier to use. For this reason, we include

blank tables at the end of many chapters. Although tabulations of data occur in the explanatory examples, blank tables make it easier for students to apply the methods to their laboratory problems. These, and what we hope is a simple introduction to a complex field, seem to make the manual worth publishing.

Chance apparently is ubiquitous to all universes this side of heaven. As a result, the manual should be useful to others than undergraduate chemistry majors.

A certain indebtedness to those from whom I may appear to have plagiarized is hereby admitted. But why publish something if you don't want it used?

I appreciate, also, the typing of Mrs. Judy Baker, Mrs. Karen Barbee, and Mrs. Kay Pool. And I thank Mr. Joseph Hartswicke for being my computer consultant.

Hubert L. Youmans

Contents

Tables for Statistical Calculations

Tables and Graphs for Functional Computation

CHAPTER 1

Introduction

Purpose

To introduce small-sample statistics as a useful tool for systematizing experimental design and interpretation, to suggest this manual as a practical introduction to statistics in chemistry, and to present a selected bibliography of applied statistics for further study.

Statistics is a field of specialization, in a manner similar to analytical chemistry. Statistics is an area of specialization in mathematics; analytical chemistry is a specialized area of chemistry. As analytical chemistry has many sub-areas useful under different circumstances, so has statistics. The methods of this manual make up part of what is called small-sample statistics.

Small-sample statistics resulted from the efforts of R. A. Fisher and others to systematize experimental design and interpretation. It is the application of statistical methods to the small samples of data usually obtainable by experimentation. This manual consists of some simple methods of general applicability to chemistry. They should serve as a basis for much experimentation and as background for learning more specialized or complex procedures.

The following books have been used extensively in writing this manual:

Langley, R., *Practical Statistics Simply Explained*, Revised Edition, 1971, Dover Publications, Inc., New York.

Fisher, R. A., *Statistical Methods for Research Workers*, 13th Edition, 1963, Oliver and Boyd, Ltd., Edinburgh.

Fisher, R. A., *The Design of Experiments*, 8th Edition, 1966, Hafner Publishing Company, New York.

Wortham, A. W. and T. E. Smith, *Practical Statistics in Experimental Design*, 1959, Dallas Publishing House, Dallas.

Snedecor, G. W. and W. G. Cochran, *Statistical Methods*, 6th Edition, 1967, The Iowa State University Press, Ames, Iowa.

Davies, O. L., editor, *The Design and Analysis of Industrial Experiments*, 2nd edition, 1967, Hafner Publishing Company, New York.

Appropriate journal articles will be cited throughout the text.

An attempt is made not to limit the statistical methods to specific chemical problems. Types of uses, along with specific examples, are suggested; some of these are used to illustrate the methods. *Calculation tables* are included to aid the novice.

Each statistical treatment is presented as a single experiment. Judicious choice of experimental conditions, however, permits several treatments of the same set of data.

The manual makes no attempt to cover statistical theory. Several monographs are available to the interested student. The emphasis here is on *application*. Only that *theory* necessary for understanding the purpose and execution of the experiments is included.

The *relation of statistics to chemistry* is similar to the relation of analytical chemistry to other areas of science. A biologist may fruitfully use analytical methods without an understanding of them comparable to an analytical chemist's. There is no shame in using a cookbook, if it is intelligently done. One may be a competent technician in many fields where time and inclination preclude thorough theoretical knowledge.

But a chemist should learn enough statistics to solve his simpler statistical problems. For some problems, he will have to consult with a statistician. He then should know enough statistics to criticize and control the solution to these problems. Otherwise, he may end up with a beautiful statistical study that does not solve his chemical problem—for his statistician probably won't know much chemistry.

CHAPTER 2

Small-Sample Statistics in Chemistry

Purpose

To give an overview of the role of statistics in chemistry. Some ideas, purposes, and methods of small-sample statistics are briefly discussed. Statistical designs have as their basis attempts to characterize systems from relatively small numbers of measurements on the systems. Such incomplete data collections do not permit infallible conclusions, but statistics makes possible experimental design and interpretation improvements that include estimates of error in system characterizations.

Small-sample statistics is a mathematical tool that may be used to describe a large set of data from a small sample of that data. The complete set of data, called a *population*, may or may not be known. It may be physically impossible to obtain all the data or the population may simply be too large to conveniently measure. Some of the data may not actually exist; these may be hypothetical or potential.

Statistics has many *uses*. In analytical chemistry, sampling and analytical errors are efficiently studied by statistical methods. Analytical methods may be compared. The effects of the variables in a chemical process may be evaluated. Optimum operating conditions for a process can be determined.

Experimental statistics is used to test a scientific hypothesis. Most *statistical tests* are based on the so-called *null hypothesis*. That is, the hypothesis is used to determine if a difference exists among the things tested. The hypothesis in some way states that no difference or a known difference exists among the things tested.

A *null hypothesis* can be disproved, but not proved, with a small number of data. For instance, we may hypothesize that two analytical procedures give statistically equivalent results. Statistical treatment of the results will then permit us to determine with a certain probability if this is a false statement. If we fail to disprove the null hypothesis, we must accept it as true. If our test disproves the null hypothesis, we must accept the *alternate hypothesis* that the procedures do not give statistically equivalent results.

Statistics cannot say absolutely that a difference exists among things tested. What it can do is estimate the probability that a difference exists. The *probability* of a given measurement is the number of data of that size divided by the total number of data in the population. *Percent probability* is this number multiplied by 100. If the population is continuous, the probability must be given for the difference between two measurements, rather than for a single measurement. This is because the probability of an event in a continuous distribution is zero.

If in 100 reference samples 18 contained 3.82% molybdenum, the probability that a randomly selected sample from this finite population of samples would contain 3.82% Mo is $18 \div 100$ or 0.18.

The number of measurements in a continuous distribution is represented by the area under the distribution curve. A point on this curve has height but not width, so there can be only zero area beneath it. Then the probability of the point is zero divided by the area beneath the curve. This gives zero for the probability of the point. But there is area beneath the part of a curve connecting two points, hence the distance between the two points has a probability greater than zero.

The statistics that we shall use is based on a *normal or Gaussian curve*, which is continuous. Thus a range of values around some average is obtained for a statistic.

In estimating a *population parameter or statistic*, such as the arithmetic mean, we shall obtain a range of values with some measure that the range contains the absolutely correct value. An analysis may give some such result as $12.55 \pm 0.05\%$ with a 95% probability. This may not sound like a satisfactory state of affairs, but it's a more accurate and truthful evaluation than we could get from just a single number, such as the midpoint of the range.

Statistics increases the *efficiency* of evaluating data. It will help us extract the maximum amount of information from our experiments. It also will help reduce the amount of experimentation. This involves experimental design.

Data collected without design are much more difficult to treat statistically than are data gathered with a design amenable to statistical evalua-

tion. Yet the paramount purpose of our experimenting is chemical and must not be sacrificed for statistics. We should use statistics only as it serves our reasons for experimenting. Our best *experimental designs*, however, will harmonize the requirements of chemistry and statistics.

When one sets out to gather data from experiments, he always plans to arrange the results in some manner. This is required for presentation and assimilation. If the collecting and arranging has been haphazard in the least, statistics can usually provide designs that produce an equal amount of information for less effort. *Statistical designs* will also frequently produce information not readily obtained otherwise. But statistics cannot create information; it can only extract that inherent in the data.

Wisely used statistics forces *improvements in experimentation*. To choose an experimental design, one must consider what he is seeking and how best to arrive at it. This necessity of carefully identifying the problem is frequently avoided, if at all possible. Problems are often created or changed to use and justify poorly conceived research. This is the backward approach. Statistics helps us to force identification of problems before experimentation takes place.

A *good experimental design* yields the desired information with a minimum of experimental effort. A good design should correctly formulate the questions to be answered by experimenting. It should take into account the accuracy required and the experimental difficulties that may be encountered. The kind, number, spacing and interrelations of the various observations should be correctly chosen.

Another advantage of using statistical designs is that all possible *sources and magnitudes of errors* must be considered. Decisions must be made about how errors are to be treated or their effects nullified. Statistics requires a careful consideration of the number of measurements to be made and the magnitude of the variables, such as sample size.

Because experimental results are often interpreted by statistical tests of significance, *levels of significance* must be chosen. They must always be chosen before the data are collected, in order to avoid bias. For instance, if we compare two analytical methods, we may wish to make the level "significant." We therefore determine from probability tables the value of some statistic, such as Student's t, that has a probability 0.05 of being exceeded if all results by the two methods belong to the same population. We do this before experimenting. If the tabular statistic is subsequently exceeded in the experiments, our null hypothesis of equivalency is said to be disproved. The difference between the two methods is "significant." The "highly significant" level has a probability 0.01 of being exceeded.

In the above test, we assume that if the calculated statistic exceeds the

tabular value, it does not belong on the same Gaussian curve as the tabulated statistic, even though 5% or 1% of the values on the *reference curve* exceed the tabular value.

There is obviously a chance that the assumption is not justified. We thus could make a *wrong decision* about the methods, based on statistics. But the probability is not 0.05 or 0.01. It is usually considerably less, but generally is unknown. The 0.05 and 0.01 simply represent probabilities that certain numbers will be exceeded in a single population.

There are *two ways of making wrong decisions.* Two pieces of data may belong to the same population but be outside the chosen limits of a probability level. They would be incorrectly said to belong to different populations. Or two pieces of data may fall within the chosen limits, but belong to different populations. They would be incorrectly said to belong to the same population.

The best *defenses against such errors* are large samples and thorough randomization of data. One should also remember that statistics implies there is something tentative about every accepted hypothesis. This should be the first axiom of the experimentalist.

The statistical methods of this manual are useful in *disciplines other than chemistry.* The methods for comparing measurements of things don't necessarily change with the nature of the things measured. There are relations in nature.

CHAPTER 3

Significant Figures and Rules for Their Use

Purpose

To differentiate between accuracy and precision and to give rules for using numbers so that properly warranted accuracy or precision will be implied by measured data.

Measured numbers are *approximate numbers*. Weighings, volumetric measurements and the like always have error associated with them. Significant figures are used to indicate the accuracy of measured numbers.

Accuracy tells how correctly a measurement was made. *Precision* shows the repeatability of a measurement. Frequently, accuracy is unknown or at best poorly estimated. In such cases, significant figures may only reflect precision. This may be unavoidable. Sometimes it is desirable. One should never, however, equate precision with accuracy. This section is largely concerned with expressing accuracy. The rules are equally applicable, though, to calculations with numbers of unknown accuracy but known precision.

Numbers may be classified as exact or measured. An *exact number* has no error in it. Exact numbers are either theoretical or counted numbers. If we wish to convert a decimal fraction to percent, we multiply by 100, a theoretical number. There is no error in the 100. It is by definition exactly 100. Other examples of theoretical numbers are:

1000, to convert grams to milligrams;
2, to double a sample size;
60, to convert minutes to seconds.

An example of a counted number would be the number of replicates of an analysis. This is usually so small as to preclude error.

Fractions may be exact or measured quantities. To differentiate between the two, express exact fractions as common fractions and measured fractions as decimals. Examples are:

$\frac{2}{3}$ of a counted dozen samples,
0.100 of a measured volume.

Counted numbers are exact because of their size and nature. A student can count the number of flasks in his laboratory desk. A chemist can count the new compounds that he has made.

Counted numbers shade over into *measured numbers*. For instance, a chemistry professor can count the number of students in class on a given day. Frequently, however, the registrar can't give the number of students enrolled in all chemistry courses. Students may have withdrawn recently from courses, the computer section may be behind in its work—many factors may force the registrar into giving an approximate number.

We normally think of the United States census as a count of this country's people. Yet Avogadro's number is more accurately known than is the population of New York City. Both are measured numbers because of inaccuracies of counting.

A measured number has *limited accuracy* because of the limitations in measurement technique and instruments. On a triple beam balance we may weigh a sample to the nearest 0.1 g. The same sample can be weighed on an analytical balance to 0.0001 g.

A measured number should *indicate the accuracy* of the measurement. If we report a sample weight as 1.8264 g., the actual weight should be between 1.82635 and 1.82645 g. In other words, the inaccuracy of the number should be in the last digit to the right. A measured number thus implies an error of up to $\pm\frac{1}{2}$ of 1 in the last digit to the right; 1.8264 g. implies that the true value is in the range 1.8264 ± 0.00005 g. No probability of accuracy is associated with this.

Accuracy is frequently confused with precision. Although good precision is desirable, it can falsely suggest good accuracy. A *constant error* causes poor accuracy; it affects precision not at all.

The accuracy of measured numbers is indicated by significant figures. The term "significant" as related to a number is generally used to indicate a confidence level, discussed later. One normally speaks of significant figures, that is, digits within a number. In a number where the error in the number is in the right-hand digit:

all non-zero digits are significant,
zeros between other digits are significant,
terminal zeros are significant,
zeros preceding the first non-zero digit are not significant.