

A Statistical Manual for Chemists

Second Edition

Edward L. Bauer

*Winthrop Laboratories
Rensselaer, New York*



Academic Press 1971 New York and London

A Subsidiary of Harcourt Brace Jovanovich, Publishers

COPYRIGHT © 1971, BY ACADEMIC PRESS, INC.

ALL RIGHTS RESERVED

**NO PART OF THIS BOOK MAY BE REPRODUCED IN ANY FORM,
BY PHOTOSTAT, MICROFILM, RETRIEVAL SYSTEM, OR ANY
OTHER MEANS, WITHOUT WRITTEN PERMISSION FROM
THE PUBLISHERS.**

ACADEMIC PRESS, INC.

111 Fifth Avenue, New York, New York 10003

United Kingdom Edition published by

ACADEMIC PRESS, INC. (LONDON) LTD.

24/28 Oval Road, London NW1

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 73-154404

Second Printing, May 1973

Third Printing, November 1973

PRINTED IN THE UNITED STATES OF AMERICA

Preface to the Second Edition

One of the cherished dreams of mankind is to be able to relive one's life, to have a second chance. Such a Faustian experience is granted to a few lucky souls, among whom are the authors of second editions. Given the chance to "do it over again," some bold persons would live an entirely different life, but most would probably try to keep the best of the old, risking the new and unknown only when the venture stood a good chance of producing an improvement.

This is what I have done. I still believe there is a need for a simple book on statistics for the working chemist, hence the purpose of the book (as set forth in the preface to the first edition) has remained the same.

In this edition, major changes have been made in every chapter except Chapter I, and a chapter on control charts has been added.

Many of these changes were the result of letters received from readers of the first edition. I am very grateful to these people (most of them chemists) for their ideas and suggestions. I hope I have answered their questions and have in this way made this a more useful book.

Preface to the First Edition

This book has been written for chemists who perform experiments, make measurements, and interpret data. Conversations with my colleagues have convinced me that too few chemists are taking advantage of the help statistical tools can give them: (1) maximum economy in experimentation, (2) maximum information from measurement data, (3) maximum accuracy and precision from test results. There seems to be a feeling that statistics are either too complicated to learn or too time consuming to use. My purpose has been to provide techniques which are both simple and fast and which will enable the chemist to analyze his own data.

The book is intentionally elementary in content and method. It is not meant to be a complete text on statistical techniques, but rather a manual for the working chemist. Throughout the book, use is made of rapid methods of calculation requiring only addition, subtraction, and the ability to use a slide rule. Discussions of statistical theory have been kept to a minimum because I believe many chemists are awed by a page of integrations, just as many mathematicians are abashed by a page of structural formulas.

In order to keep the manual simple and understandable to the neophyte, it was necessary to omit some very useful but sophisticated techniques. It is my hope that this manual will serve as a statistical primer; having mastered the fundamentals, the reader will be prepared to graduate to the more complex techniques as the need arises. At the end of each chapter the reader will find references, some

of which are not cited in the text. These were included as a suggested step for further study.

I am grateful to Professor E. S. Pearson and *Biometrika* Trustees for permission to publish certain tables which appear in the Appendix.

I am also grateful to my secretary, Miss Pat Ibarreche, who so unflinchingly undertook the task of transcribing my original notes; to the Winthrop Laboratories' Librarian, Miss Ethel Center, for her tireless forages into the jungle of statistical literature for obscure references; and to Helen, my wife, for typing the manuscript, and for her criticism and coffee—both hot, strong, black, and without sugar.

List of Symbols

The reader often finds the terminology of a strange subject somewhat confusing. This is particularly true of statistical terminology where various authors use different symbols. The following glossary lists and defines the symbols used in this book, except in Chapter 6 which uses control chart symbols.

SYMBOL	DEFINITION
a	The intercept of a regression line
A	A factor that, when multiplied by the range, gives the confidence limits of the average
AD	Average deviation
α (alpha)	The risk of making a Type-I error
β (beta)	The risk of making a Type-II error
b	The slope of a regression line
c_1, c_2	Factors convert range to an unbiased estimate of the square root of the variance
CL	Confidence Limits of an average
d	A factor that converts average range to standard deviation
df	Degrees of freedom
f	Equivalent degrees of freedom
F	The critical value of the variance ratio test
I	A factor that when multiplied by the range gives the tolerance intervals of individual measurements
k	The number of groups in a series of observations

SYMBOL	DEFINITION
L	The critical value for the one-sample substitute t test based on range
M	The critical value for the two-sample substitute t test based on range
μ (mu)	The true mean of the population
n	Sample size or the number of observations in a group
N	kn —the total number of observations
p_1	The fraction defective of an acceptable lot
p_2	The fraction defective of an unacceptable lot
q	The critical value for the Studentized range
R	The range—the difference between the largest and the smallest of a group of observations
\bar{R}	The average range
s	The standard deviation of a sample
s_d	The standard deviation of a difference
$s_{\bar{x}}$	The standard deviation of an average = s/\sqrt{n}
σ (sigma)	The standard deviation of the population
σ^2	Variance of the population
Σ	Summation
t	The critical value of the t test
V	The variance of a sample
X	An observation, and the independent variable in a regression
\bar{X}	The average of a number of observations
$\bar{\bar{X}}$	The grand average
Y	The dependent variable in a regression

Contents

<i>Preface to the Second Edition</i>	ix
<i>Preface to the First Edition</i>	xi
<i>List of Symbols</i>	xiii
1 FUNDAMENTALS	
1.1 Introduction	1
1.2 Experimental Error	2
1.3 The Average	2
1.4 The Normal Distribution	3
1.5 The <i>t</i> Distribution	5
1.6 Accuracy and Precision	6
1.7 The Average Deviation	7
1.8 The Variance and Standard Deviation	8
1.9 The Range	10
References	12
2 THE AVERAGE	
2.1 Replication	13
2.2 Confidence Limits	14
2.3 Degree of Confidence	15
2.4 Illustration of Confidence Limits	15
2.5 Calculation of Confidence Limits	16
2.6 Confidence Limits of Large Groups of Data by Range	19
2.7 Tolerance Limits	21
2.8 Invalid Measurements	22
2.9 Derivations and Proofs	23
References	26
3 EXPERIMENTAL DESIGN AND THE ANALYSIS OF VARIANCE	
3.1 Experimental Design	27
3.2 Nomenclature of Statistically Designed Experiments	29

vi *Contents*

3.3	Tests of Significance	30
3.4	The Analysis of Variance	31
3.5	Block Design: One-Way Classification	32
3.6	Block Design: Two-Way Classification	36
3.7	Models of ANOVA	42
3.8	Components of Variance	42
3.9	Expected Mean Square (EMS) Components	44
3.10	Latin Square	47
3.11	Factorial Experiments	49
3.12	Nested Factorial Experiment	55
4	THE COMPARISON OF TWO AVERAGES	
4.1	The t Test	61
4.2	Uses of a t Test	62
4.3	Substitute t Tests	67
4.4	Uses of Substitute t Tests	68
5	ANALYSIS OF VARIANCE BY RANGE	
5.1	Introduction	71
5.2	Block Design: One-Way Classification	72
5.3	Block Design: Two-Way Classification	75
5.4	Interaction	77
5.5	The Latin Square Design	80
5.6	Factorial Experiments	84
	References	93
6	CONTROL CHARTS	
6.1	Introduction	95
6.2	Nomenclature	95
6.3	Theory of Control Charts	96
6.4	Control Limits	97
6.5	The Chart for Averages	98
6.6	The Chart for Ranges (or Standard Deviations)	98
6.7	Subgroups	98
6.8	Calculation of Control Limits	99
6.9	Significance of Control Limits	100
6.10	Runs	100
6.11	Making a Control Chart	101
6.12	Lack of Control	106
	References	108

7 CORRELATED VARIABLES

7.1	Linear Regression	109
7.2	A Laboratory Use of Regression	110
7.3	Shortcut Methods	117
7.4	Shortcut Method When X_n Does Not Equal nX_1	119
7.5	Colorimetric Analysis	120
7.6	Confidence Limits for X	126
7.7	Nonlinear Functions	127
	Reference	130

8 SAMPLING

8.1	The Sample and the Population	131
8.2	The Theory of Sampling	133
8.3	Sample Size	135
8.4	Attribute Sampling	136
8.5	Sampling by Variables	137
8.6	Use of Components of Variance	139
8.7	Variables Plan Based upon Normal Distribution	141
	References	144

9 CONTROL OF ROUTINE ANALYSIS

9.1	Problems of the Routine Analyst	145
9.2	Test for Outliers	146
9.3	Precision of the Analyses	147
9.4	Difference between Analysts	149
9.5	Accuracy	149
9.6	Precision of Optical Rotation Measurements	151
9.7	Precision of Colorimetric Analysis	152
9.8	Reduced Sample Size	153
9.9	Compliance with Specifications	156
9.10	Control Charts in the Analytical Laboratory	157
9.11	Interlaboratory Studies	158
	Reference	163

APPENDIX	165
----------	-----

Subject Index	191
---------------	-----

1

Fundamentals

1.1 INTRODUCTION

We use numbers in two ways—to enumerate objects and to designate the magnitude of measurements. If we were to count the number of words in the first sentence, we would find there are sixteen. No matter who counted them, when they were counted, or how they were counted, we would still get exactly sixteen words. This is an example of numbers used to enumerate objects. It is an absolute value—it does not change with time or method of measurement.

As a rule, chemists are not as interested in enumeration data as they are in measurement data. The information important to the chemist comes not from counting objects, but from weighing, reading burets, measuring volumes, and reading instruments. All of these operations involve measurements, and all measurements involve a region of uncertainty.

For example, consider the results obtained in reading the absorptivity of a spectrophotometric analysis. It is standard technique to make readings at about 0.43 absorbance to achieve minimum error. In this region, the absorbency scale of a well-known spectrophotometer is graduated so that there are 0.01 units between scale markings. The analyst must interpolate between 0.43 and 0.44, and the best he can hope to do is estimate one tenth of the least count of the instrument, or 0.001 units. This introduces a doubtful value into each reading. For instance, if the true absorbance of a solution is 0.435,

and he reads the instrument scale as 0.436, he is making a relative error of 0.23 %; if he reads 0.433 or 0.437, the relative error is 0.46 %.

1.2 EXPERIMENTAL ERROR

Texts on analytic chemistry classify errors as determinate and indeterminate. Determinate errors are defined as those that can be avoided once they are recognized. This type of error is caused by such factors as:

- (1) Improper calibration of glassware or instruments, or improper standardization of reagents.
- (2) Personal errors, such as the tendency of an analyst to misjudge a color change.
- (3) Prejudice.
- (4) A constant error in method.

Determinate errors introduce a bias into the measurements. For example, if the analyst stands to one side of the hairline on the scale, his readings will all be high or low because of parallax.

Indeterminate errors cannot be eliminated. They exist by the very nature of measurement data. For example, the slight errors in interpolation are indeterminate. The analyst does not know their magnitude, or whether they are positive or negative. It is these indeterminate errors which we call "experimental error." They affect the precision of all chemical work, and we attempt to contain them in as narrow a zone as possible.

The most common way to minimize experimental error is to make a series of measurements on the same object and report the average.

1.3 THE AVERAGE

The average is the sum of the measurements divided by the number of measurements:

$$\bar{X} = (X_1 + X_2 + X_3 + \cdots + X_n)/n. \quad (1.1)$$

Two facts are evident from Eq. (1.1):

- (1) The average is a measure of the central tendency; the sum of the deviations from it is zero.
- (2) Since it comprises a number of different observations, it cannot be an absolute value.

The reliability of the average depends upon the range of values from which we obtain it: 16 is the average of 0 and 32; it is also the average of 15 and 17. As the average of 0 and 32 we could put little or no reliance upon its validity. As the average of 15 and 17, we can be reasonably certain that it is a good estimate of a true value. In neither case, can we feel as certain of the validity of the average as we can that there are sixteen words in the first sentence of this chapter.

When a chemist calculates an average, he is using statistics. He is intuitively making use of the laws of probability by taking advantage of the fact that he will make small errors more frequently than he will make large errors, and that in the long run, the plus-and-minus errors will cancel each other, leaving the average as a good estimate of the true value.

A rigorous definition of mathematical probability that would satisfy all statisticians would be too difficult and involved for this book. For our purpose, it is sufficient to describe, rather than define, it. Mathematical probability may be described as *expected frequency in the long run*.

To the statistician, the "long run" means a large number of data, variously distributed.

1.4 THE NORMAL DISTRIBUTION

Suppose the analyst makes a very large number of absorbance measurements (say 1000) on the same solution, and plots the magnitude of the measurement against the frequency of its occurrence. He would find the measurements distributed in a bell-shaped manner, with most of the measurements in the center and an equal number

distributed with decreasing frequency on either side of the center. The distribution of such data can be described by a curve like Fig. 1.1, curve *a*. This is the Gaussian, or normal distribution curve. It is the theoretical distribution of the relative frequency of a large number of observations made on the same object. It is, therefore, a description of the expected frequency (or probability).

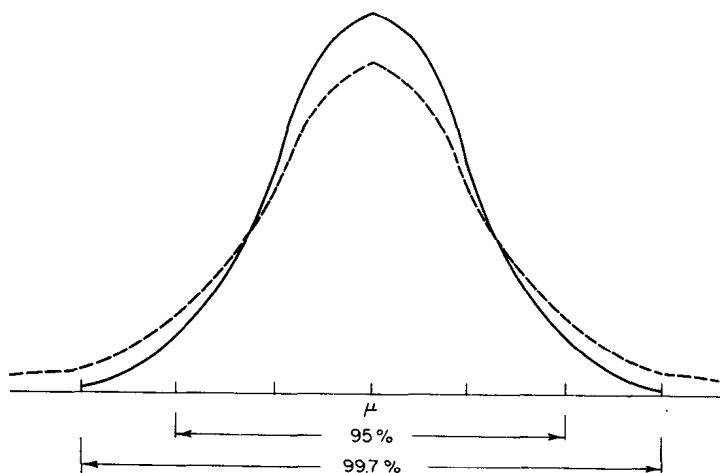


Figure 1.1. Curve *a* (—): Normal distribution curve. Curve *b* (---): Student distribution curve.

The curve has two properties that make it valuable to anyone using statistics:

(1) It can be completely described by the average (μ), which fixes the location of the center of the curve with reference to the x axis, and by the standard deviation (σ), which describes the spread of the data along the x axis.

$$\sigma = [\Sigma(X - \mu)^2/n]^{1/2}. \quad (1.2)$$

(2) The distribution of the frequency of the data has been thoroughly established. For example, 95 % of the individual measure-

ments will lie within $\mu - 1.96\sigma$ and $\mu + 1.96\sigma$, and a spread of $\mu \pm 3\sigma$ will include 99.7% of the measurements.

This means that the analyst who is making 1000 absorbance readings would expect 50 readings to be outside $\mu \pm 1.96\sigma$ and only 3 readings to be outside the limits $\mu \pm 3\sigma$.

From a practical viewpoint, the converse is important. Suppose the analyst is making readings of a solution whose absorbance is 0.435, and he knows $\sigma = 0.005$. A reading outside of the range 0.425–0.445 would happen only 5 times in 100, and hence must be suspect. This concept is the basis of tests of significance.

1.5 THE *t* DISTRIBUTION

The theory of the normal distribution was developed from large amounts of data, and does not necessarily apply to small numbers of observations. In the laboratory, we cannot afford to make a very large number of observations; as a result, statistical tests based on the normal distribution could lead the laboratory worker to draw false conclusions. This fact was recognized by W. S. Gosset, an Irish chemist. In 1908, he published a paper under the pseudonym, "Student" entitled "The Probable Error of a Mean" (1). Partly by means of theoretical considerations, and partly by drawing small random samples, he derived the theoretical distribution of the average of small samples drawn from a normal distribution.

If we do not use large samples, we cannot know the true standard deviation σ or the true population mean μ . However, we can replace σ by the sample standard deviation (s). When we do this, we must use a new distribution, which is independent of σ . This is the concept introduced by Gosset that has become known as "Student's *t*"

$$t = (\bar{X} - \mu)/s_{\bar{X}}. \quad (1.3)$$

Student demonstrated that the distribution of *t* is dependent only on the sample size (*n*). Figure 1.1, curve *b* (dashed line), shows the relationship of the *t* distribution to the normal distribution. The *t* curve is flatter than the normal curve, but approaches it as the sample

size increases, becoming equal to the normal curve as n approaches infinity. For practical purposes, we usually use the normal distribution for sample sizes greater than 30.

It is necessary to understand the concept of the t distribution because it is the foundation upon which all tests of significance involving the comparison of two averages from small samples are based.

1.6 ACCURACY AND PRECISION

Accuracy may be defined as the correctness of a measurement. if

μ = the true value,

X = the value obtained experimentally,

E = the error,

then

$$\mu = X \pm E.$$

In chemical work, μ is often unknown, and therefore must be estimated from $X \pm E$. If E is zero, $\mu = X$, and the measurement is accurate.

Precision is a measure of the reproducibility of the measurements. The terms "accuracy" and "precision" are sometimes used interchangeably. They are not necessarily synonymous, as Fig. 1.2

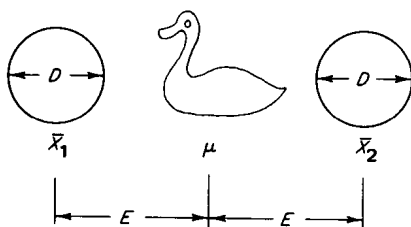


Figure 1.2

demonstrates. A hunter fired both barrels of a shotgun at a duck, with the results illustrated. Both barrels shot precisely, but the aim was not accurate—the duck flew away.

The true value is μ , the duck. The averages of the bursts with maximum distribution D are \bar{X}_1 and \bar{X}_2 , a measure of the precision. The distance from \bar{X}_1 or \bar{X}_2 to μ , is E , a measure of the accuracy. It is only when E is small compared with D that accuracy and precision are the same.

If D were 10 ft, the hunter would miss the duck by $E - D/2 = 5$ ft. If, however, E were 5 ft, he would have a duck dinner.

Most statistical techniques measure precision rather than accuracy. However, statistical techniques are essential to the measurement of accuracy, because precision must be known before accuracy can be evaluated. A chemist cannot say a method is “accurate within the limits of experimental error” if he has no knowledge of the magnitude of the experimental error.

There are three common ways of evaluating the precision: (i) the average deviation, (ii) the variance, (iii) the range.

1.7 THE AVERAGE DEVIATION

If we sum all the X 's in Fig. 1.3, regardless of whether they are

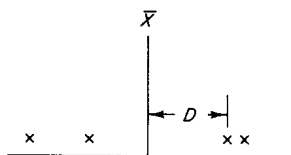


Figure 1.3

positive or negative, we will obtain the total deviation. Dividing this by the number of deviations will give the average deviation (AD).

$$AD = \Sigma(X - \bar{X})/n. \quad (1.4)$$