



ANALYTICAL CHEMISTRY BY OPEN LEARNING

Measurement, Statistics and Computation

Measurement, Statistics and Computation

Analytical Chemistry by Open Learning

Authors:

DAVID MCCORMICK

Manchester Polytechnic

ALAN ROACH

Paisley College of Technology

Editor:

NORMAN B. CHAPMAN

on behalf of ACOL

**Published on behalf of. ACOL, Thames Polytechnic,
London**

**by
JOHN WILEY & SONS**
Chichester • New York • Brisbane • Toronto • Singapore

© Crown Copyright, 1987

Published by permission of the Controller of
Her Majesty's Stationery Office

All rights reserved.

No part of this book may be reproduced by any means, or
transmitted, or translated into a machine language without the
written permission of the publisher.

Library of Congress Cataloging in Publication Data:

McCormick, David.

Measurement, statistics and computation:

(Analytical Chemistry by Open Learning)

1. Chemistry, Analytic—Statistical methods—Programmed instruction.
 2. Chemistry, Analytic—Measurement—Programmed instruction.
 3. Chemistry, Analytic—Data processing—Programmed instruction.
 4. Chemistry, Analytic—Programmed instruction.
- Roach, Alan. II. Chapman, N. B. (Norman Bellamy), 1916— III. ACOL
(Project) IV. Title. V. Series: Analytical Chemistry by Open Learning
(Series)

QD75.4.S8M4 1987 543 87-10470

ISBN 0 471 91366 9

ISBN 0 471 91367 7 (pbk.)

British Library Cataloguing in Publication Data:

McCormick, David.

Measurement, Statistics and computation.—(Analytical Chemistry).

1. Chemistry, Analytic—Statistical methods

I. Title II. Roach, Alan III. Chapman, Norman B.

IV. ACOL V. Series

519.5'024541 QD75.4.S8

ISBN 0 471 91366 9

ISBN 0 471 91367 7 Pbk.

Printed and bound in Great Britain

Analytical Chemistry

This series of texts is a result of an initiative by the Committee of Heads of Polytechnic Chemistry Departments in the United Kingdom. A project team based at Thames Polytechnic using funds available from the Manpower Services Commission 'Open Tech' Project has organised and managed the development of the material suitable for use by 'Distance Learners'. The contents of the various units have been identified, planned and written almost exclusively by groups of polytechnic staff, who are both expert in the subject area and are currently teaching in analytical chemistry.

The texts are for those interested in the basics of analytical chemistry and instrumental techniques who wish to study in a more flexible way than traditional institute attendance or to augment such attendance. A series of these units may be used by those undertaking courses leading to BTEC (levels IV and V), Royal Society of Chemistry (Certificates of Applied Chemistry) or other qualifications. The level is thus that of Senior Technician.

It is emphasised however that whilst the theoretical aspects of analytical chemistry can be studied in this way there is no substitute for the laboratory to learn the associated practical skills. In the U.K. there are nominated Polytechnics, Colleges and other Institutions who offer tutorial and practical support to achieve the practical objectives identified within each text. It is expected that many institutions worldwide will also provide such support.

The project will continue at Thames Polytechnic to support these 'Open Learning Texts', to continually refresh and update the material and to extend its coverage.

Further information about nominated support centres, the material or open learning techniques may be obtained from the project office at Thames Polytechnic, ACOL, Wellington St., Woolwich, London, SE18 6PF.

How to Use an Open Learning Text

Open learning texts are designed as a convenient and flexible way of studying for people who, for a variety of reasons cannot use conventional education courses. You will learn from this text the principles of one subject in Analytical Chemistry, but only by putting this knowledge into practice, under professional supervision, will you gain a full understanding of the analytical techniques described.

To achieve the full benefit from an open learning text you need to plan your place and time of study.

- Find the most suitable place to study where you can work without disturbance.
- If you have a tutor supervising your study discuss with him, or her, the date by which you should have completed this text.
- Some people study perfectly well in irregular bursts, however most students find that setting aside a certain number of hours each day is the most satisfactory method. It is for you to decide which pattern of study suits you best.
- If you decide to study for several hours at once, take short breaks of five or ten minutes every half hour or so. You will find that this method maintains a higher overall level of concentration.

Before you begin a detailed reading of the text, familiarise yourself with the general layout of the material. Have a look at the course contents list at the front of the book and flip through the pages to get a general impression of the way the subject is dealt with. You will find that there is space on the pages to make comments alongside the

text as you study—your own notes for highlighting points that you feel are particularly important. Indicate in the margin the points you would like to discuss further with a tutor or fellow student. When you come to revise, these personal study notes will be very useful.

- When you find a paragraph in the text marked with a symbol such as is shown here, this is where you get involved. At this point you are directed to do things: draw graphs, answer questions, perform calculations, etc. Do make an attempt at these activities. If necessary cover the succeeding response with a piece of paper until you are ready to read on. This is an opportunity for you to learn by participating in the subject and although the text continues by discussing your response, there is no better way to learn than by working things out for yourself.

We have introduced self assessment questions (SAQ) at appropriate places in the text. These SAQs provide for you a way of finding out if you understand what you have just been studying. There is space on the page for your answer and for any comments you want to add after reading the author's response. You will find the author's response to each SAQ at the end of the text. Compare what you have written with the response provided and read the discussion and advice.

At intervals in the text you will find a Summary and List of Objectives. The Summary will emphasise the important points covered by the material you have just read and the Objectives will give you a checklist of tasks you should then be able to achieve.

You can revise the Unit, perhaps for a formal examination, by re-reading the Summary and the Objectives, and by working through some of the SAQs. This should quickly alert you to areas of the text that need further study.

At the end of the book you will find for reference lists of commonly used scientific symbols and values, units of measurement and also a periodic table.

Study Guide

Consider the scope of the activities in which analytical chemists the world around are engaged. The starting-point is usually that someone – we shall call him the client – is confronted with a situation. It may be that the fish in a river are dying, that a chemical product (perhaps a drug) is no longer being manufactured to its usual specification, or that there is a need to discover whether or not a fire was the result of arson. The client decides, sooner rather than later, we hope, that the situation would benefit from the involvement of an analytical chemist. In discussion, analytical problems emerge, and in the light of all the circumstances of the case, procedures and methods suggest themselves for the solution of the problem.

The relationship between the client and the chemist is clear. The client needs to make rational decisions. To do this, trustworthy information is needed. Analytical chemists are not the only people involved in providing help to such clients. However, when the chemical aspects of a situation have been identified, the information given by the analytical chemist is of paramount importance. In choosing the procedures and methods for the analysis, the chemist will have taken into account matters such as the required sensitivity and accuracy of the analysis. The chemist must also be concerned with questions such as whether the analysis is being carried out on samples which are representative of the whole. It may be that meaningful information can be provided only if several samples are analysed. It may be good enough to carry out one determination on each sample; alternatively, the client's requirement with regard to the uncertainty in a measurement may be such that several determinations on each sample are needed. In making decisions about such matters, the chemist relies heavily on that branch of mathematics known as statistics. The major purpose of this Unit is to introduce you to the application of statistical ideas in the context of analytical chemistry.

There are, however, *three* words in our Unit title – Measurement, Statistics, and Computation. It is not our aim to introduce you to a formal and rigorous exposition of statistical theory in the abstract. Above all, analytical chemists are practical people. The great weight of our expertise and experience is directed to the competent handling of our sample, and then to subjecting it to a carefully selected procedure for making *measurements*. Reaching that stage is the hard job, but it is also important to ensure that we can extract the maximum benefit from our labours. This Unit is thus fundamentally about drawing quantitative conclusions from experimental measurements, assessing the value of our results, and suggesting at times additional work which might usefully be undertaken.

The methods of *statistics* give us the tools we need for this task. In introducing you to these methods our aims will be threefold; that you should:

- (a) be able to apply a given test competently;
- (b) understand clearly the issues addressed by a test;
- (c) develop, through practice, a 'feeling' for the significance of the conclusions you reach.

We shall not be concerned with formal derivations, and we shall keep mathematical manipulations to a minimum. Nor do we intend to give an exhaustive survey of methods. The Unit concentrates on helping you to obtain a good practical grasp of a limited, but useful, range of methods.

The third word in our title is *computation*. The microcomputer revolution is completely transforming our field of work, as it is many others. Analytical instruments are more and more commonly interfaced with computers. Quality control measurements are becoming increasingly computer-automated. The quantity of analytical data which we can expect to have to handle is increasing substantially. Also, you will find, statistical manipulation can involve a lot of rather tedious arithmetic, as well as a certain amount of intensive activity scouring through tables. Much checking and re-checking is required, to make sure that no slips in arithmetic have occurred. This work lends itself admirably to a computer approach. Once our

data is inside our computer, we can, with relative ease, perform a variety of different tests, investigate the effects of rejecting suspect measurements, or of including others, etc. So to study this Unit you need access to a computer. We make no assumptions about *which* computer, as there are very many different models of microcomputer, any of which will suffice. If you have your own machine at home you have a head start. Otherwise you should arrange to obtain access to a machine at work or at college. (An interactive terminal linked to a mainframe computer, will, of course, also fill the bill).

If you have had no previous experience of computing, the last paragraph may have come as a shock to you. But it needn't. There are no two ways about it, computers are becoming more and more significant to our work. This Unit gives you an opportunity to make a beginning. Our approach is to use the computer to extend rather effortlessly your experience in applying the methods of statistics to the analysis of the results of measurement. At the same time, we use our need to learn about statistics as a vehicle to help you to become familiar with using a computer. It may be that you already have had considerable experience with computers. Equally, it may be that you have had none. To cope with both possibilities, our practical introduction to using a machine is given in an Appendix. The time which you will need to devote to studying that depends on your previous experience.

Now just what previous knowledge do we expect of you, if you are to undertake our foray into statistics and computation? The answer is – very little! If you have a basis in mathematics up to GCE O-level or equivalent, that will be quite adequate. The processes we have to use are essentially arithmetical. The only other material which you may need at hand for occasional reference is the user's guide of the micro-computer you use. There are many available text books covering the area of interest to us, but the time to refer you to those is when you have completed the Unit.

This does not necessarily mean that you will find the work easy. Understanding what exactly it is that you are doing in statistics requires clear and careful thinking. Great pains must be taken to develop a thorough insight into a number of concepts which on first acquaintance may be found quite confusing. For this reason we shall start

slowly by discussing relatively familiar terms such as *accuracy*, *precision* and *errors*; after this, we shall introduce statistical ideas such as *sample*, *population*, *confidence interval*. Later in the Unit we shall discuss matters such as *correlation* and *regression*. By then, we believe, you will be quite competent in the use of statistics in analytical chemistry.

If you find the going easy to begin with, please don't try to cut corners. Attempt all the exercises: that is most important. The danger is that you may *think* you understand concepts which you may have only grasped superficially. Working through examples, and checking your responses, will help to ensure that this does not happen. Statistics can be misused, and they often are, hence the famous maxim about the three categories of lies – 'lies, damped lies and statistics'. If you understand statistical methods properly, and if you apply them properly, they will never deceive you. On the contrary, statistical methods provide the *only* basis for making a sure-footed evaluation of a set of measurements. As for the computing, you really should find that you can take it in your stride. If you are a beginner you must be prepared to work through the Appendix rather thoroughly. You should then find it quite straightforward to tackle the computation in the main body of the Unit. You will quite soon appreciate what a useful tool the computer can be!

Because of the nature of the ground we are covering, the general style of this Unit is a little different from most of the others in the Analytical Chemistry series. You will find relatively fewer 'normal' SAQs, but that does not mean that you will be left as a passive reader. On the contrary, much of the time the argument is developed by discussing specific examples. If you want to make a success of your study it is important that *you* work through the details, and, *inter alia*, check our arithmetic. This work supplements the tasks explicitly set for you in the text, in SAQ's, and in computer exercises.

So to business! First tackle our Appendix, which is reprinted from the *Microprocessor Applications* volume in this series. When you have dealt with that you will be ready to start on the main body of the Unit. It is important to study that in the order in which it is presented. As you start, we wish you good luck!

Bibliography

There are very many books which deal with topics which are of interest in this Unit, and these approach their subject matter from a great diversity of points of view. Other sources are, we feel, best consulted after you have worked your way through to the end of the present Unit. At this point, in Section 8.7, we are in a position to comment rather more fully about the available literature. You will find that we there highlight the following representative texts.

1. J. C. Miller and J. N. Miller, *Statistics for Analytical Chemistry*, Ellis Horwood, 1984.
2. C. Chatfield, *Statistics for Technology*, Chapman and Hall, 1983.
3. D. Cooke, A. H. Craven and G. M. Clarke, *Basic Statistical Computing*, Edward Arnold, 1982.
4. I. M. Kolthoff and P. J. Elving, *Treatise on Analytical Chemistry*, Part 1, Volume 1, John Wiley, 1975.
5. O. L. Davis and P. L. Goldsmith, *Statistical Methods in Research and Production*, Longman, 1984.
6. R. E. Walpole and R. H. Myers, *Probability and Statistics for Engineers and Scientists*, Collier McMillan, 1985.
7. E. L. Grant and R. S. Leavenworth, *Statistical Quality Control*, McGraw-Hill, 1980.
8. C. Liteanu and I. Rica, *Statistical Theory and Methodology of Trace Analysis*, Ellis Horwood, 1980.
9. M. R. Spiegel and R. W. Boxer, Schaums's Outline Series: *Theory and Problems of Statistics*, McGraw-Hill, 1972.

Acknowledgements

Figures 6.1a, 6.b, 6.6c and 6.6d are taken from the *British Regional Heart Survey*, as published in the *British Medical Journal*, May 1980.

Contents

Study Guide	xv
Bibliography	xix
Acknowledgements	xx
1. Accuracy and Precision	1
1.1 Replicate Measurements and the Arithmetic Mean	1
1.2 Accuracy	3
1.3 Replicate Measurements and Precision	7
1.4 Doing the Chores by Computer	13
1.5 Comparing Accuracy and Precision	16
1.6 Errors	23
1.7 Systematic Errors	33
1.8 The Structure of Error	40
1.9 Summary	46
2. Probability and the Distribution of Error	50
2.1 Probability	51
2.2 Probability and Random Error	65
2.3 The Normal Distribution	78
2.4 The Normal Curve and Random Error	92
2.5 Finding the Areas by Computer	98
2.6 A Taste of Things to Come	106
3. Samples, Estimation, and Hypothesis Testing	110
3.1 Population, Random Samples, and Sample Size	111
3.2 Estimating a Population Mean from a Sample Mean	127
3.3 Estimating a Population Standard Deviation from a Sample Standard Deviation	139
3.4 The t -distribution	148
3.5 Hypothesis Testing	181
3.6 One- and Two-tail Tests	198

3.7 Taking Stock	202
4. Comparison of Means and of Standard Deviations	208
4.1 Hypothesis Testing – a Modified Procedure	209
4.2 Comparison of Two Sample Means: Confidence Intervals	218
4.3 Comparison of Two Sample Means: Hypothesis Testing	229
4.4 Paired Differences	239
4.5 Comparing Two Variances	245
4.6 Completion of Program STAT	267
4.7 The Analysis of Variance: an Introduction	303
4.8 Summary	325
5. The Elementary Statistics of Calibration	330
5.1 Fitting a Straight Line to Calibration Data	333
5.2 Making Estimates by Using a Regression Line	343
5.3 Dangers of Extrapolation	358
5.4 Programming the Regression Calculation	361
5.5 Fitting Non-linear Calibration Curves	371
5.6 Summary	382
6. Correlation	387
6.1 The Scatter Diagram – Linear Regression Revisited	390
6.2 Explained and Unexplained Variation, and the Coefficient of Correlation	393
6.3 Testing the Significance of Evidence of Correlation	397
6.4 Programming the Correlation Calculation	401
6.5 Measuring ‘Goodness-of-fit’ for a Straight-line Relationship	405
6.6 Further Comments on Correlation Studies	406
6.7 Conclusions	418
7. Statistics in Quality Control	422
7.1 Batch Inspection and Process Control	423
7.2 Shewhart Control Charts	427
7.3 Cusum Charts	437
7.4 Variation in a Process Standard-deviation	452

7.5 Taking Stock	461
8. Accuracy and Precision in Handling Results	465
8.1 Significant Figures	466
8.2 The Propagation of Random Error	478
8.3 The Propagation of Systematic Error	493
8.4 The Dilemma of Discordant Results	499
8.5 The Limit of Detection	508
8.6 Statistics in the Life of an Analytical Chemist	513
8.7 Overview and Bibliography	516
Self Assessment Questions and Responses	523
Units of Measurement	644
Appendix	650
Appendix Self Assessment Questions and Responses	727

1. Accuracy and Precision

The symbol for the Society of Analytical Chemistry (SAC) contains two words – 'Accuracy and Precision'. To those not engaged in scientific or technological measurement it might be thought that the SAC has made use of two words where one would have been sufficient. To the layman the words accuracy and precision are interchangeable. To the analytical chemist, however, the two words have quite different meanings. It is the purpose of this Part of the course both to define and to give quantitative meaning to these terms. We shall also investigate and classify various categories of error which occur in experimental measurements. Let us make a start.

1.1. REPLICATE MEASUREMENTS AND THE ARITHMETIC MEAN

All chemists are aware that when we carry out determinations we often do more than one measurement, ie we do *replicate measurements*. Practically every student of chemistry first experiences this in the context of titrimetric analysis. With this in mind, we shall use analytical results obtained by the titrimetric method as the basis for the ensuing discussion.

We start with the averaging of a series of replicate measurements. The most commonly used averaging procedure is to calculate the *arithmetic mean*. This is done by adding together individual values, x_1, x_2, \dots, x_n for a series of measurements and dividing this sum by the number of measurements, n .

In mathematical notation we give the arithmetic mean \bar{x} , in equation 1.1.

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} = \frac{\sum_{i=1}^{i=n} x_i}{n} \quad (1.1)$$

Let us put this relationship to use. Suppose that we had to determine the alkali concentration of a solution and that in order to do so we carried out five replicate titrations with standard acid. Let us assume that the values obtained were: 20.02 cm³, 20.00 cm³, 20.02 cm³, 20.04 cm³, 20.02 cm³. Now, almost without thinking, we would use these values to obtain the arithmetic mean of the titres (from now on called the mean titre), ie

$$\bar{x} = \frac{\sum_{i=1}^{i=n} x_i}{n} = \frac{100.10}{5} = 20.02 \text{ cm}^3$$

This value would then be used in our calculations of the alkali concentration. For example, if the acid-base reaction had a 1:1 stoichiometry, we would use

$$M_{\text{alkali}} = \frac{M_{\text{acid}} V_{\text{acid}}}{V_{\text{alkali}}}$$

The question now arises. Why did we use the mean value in our calculation? The simple answer is that we believe *the mean value to be the best estimate of the experimental titre*. This is not the same as saying that the mean value is closest to the true value, ie that it gives the best result. It may well have been that the true alkali concentration would have been obtained from a titre of 20.10 cm³. This being so, we would have obtained a result closer to the truth by using the highest experimental titre (20.04 cm³), rather than the mean titre, (20.02 cm³). However, we were in no position to pick and choose from our titres: we were dealing with an unknown. All that we had to work on was our experimental results. Since there was variation in these, we chose to work with the mean because we believed it reflected better the results of our measurements. Without wishing to labour the point, we repeat that the mean is used because it is taken to be the best estimate of our experimental measurements.