

Notes on Statistics and Data Quality for Analytical Chemists

Michael Thompson • Philip J Lowthian

$$b_w = \frac{\sum_i w_i (x_i - \bar{x}_w)(y_i - \bar{y}_w)}{\sum_i w_i (x_i - \bar{x}_w)^2}$$



Imperial College Press

Notes on Statistics and Data Quality for Analytical Chemists

Michael Thompson and **Philip J. Lowthian**
Birkbeck University of London, UK



Published by

Imperial College Press
57 Shelton Street
Covent Garden
London WC2H 9HE

Distributed by

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

NOTES ON STATISTICS AND DATA QUALITY FOR ANALYTICAL CHEMISTS

Copyright © 2011 by Imperial College Press

All rights reserved. This book, or parts thereof, may not be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the Publisher.

For photocopying of material in this volume, please pay a copying fee through the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA. In this case permission to photocopy is not required from the publisher.

ISBN-13 978-1-84816-616-5

ISBN-10 1-84816-616-8

ISBN-13 978-1-84816-617-2(pbk)

ISBN-10 1-84816-617-6(pbk)

Typeset by Stallion Press

Email: enquiries@stallionpress.com

Printed in Singapore by B & Jo Enterprise Pte Ltd

**Notes on
Statistics and Data Quality
for Analytical Chemists**

Preface

This book is based on the experience of teaching statistics for many years to analytical chemists and the specific difficulties that they encounter. Many analytical chemists find statistics difficult and burdensome. That is a perception that we hope to dispel. Statistics is straightforward and it adds a fascinating extra dimension to the science of chemical measurement. In fact it is hardly possible to conceive of a measurement science such as analytical chemistry that does not have statistics as both its conceptual foundation and its everyday tool. Measurement results necessarily have uncertainty and statistics shows us how to make valid inferences in the face of this uncertainty. But, over the years, it has become apparent to us that statistics is much more interesting when it makes full use of the computer revolution.

It would be hard to overstate the effect that easily-available computing has had on the practice of statistics. It is now possible to undertake, often in milliseconds and with perfect accuracy, calculations that previously would have been impracticably longwinded and error-prone. A simple example is the calculation of the probabilities associated with density functions. Moreover, we can now produce in seconds several accurate graphical images of our datasets and select the most informative. These capabilities have transformed the applicability of both standard statistical methods and more recent computer-intensive methods. Textbooks for the most part have not caught up with this revolution and, to an unnecessary degree, are still stressing pencil-and-paper methods of calculation. Of course, a small number of pencil-and-paper examples of some elementary examples can assist learning, but they are too prone to mistakes for 'real-life' application. Many textbooks place a heavy stress on the mathematical basis of statistics. We regard this as inappropriate in an applied text. Analytical chemists do not need too many details of statistical theory, so we have kept these to a minimum. Drivers don't need to know exactly how every part of a car works in order to drive competently.

With ease of computation, there is, of course, a concomitant danger that people are tempted to use one of the many excellent computer statistics packages (or perhaps one of the not-so-excellent ones) without understanding what the output means or whether an appropriate method has been used. Analytical chemists have to guard against that serious shortcoming by exercising a proper scientific attitude. There are several ways of developing that faculty in relation to statistics. A key practice is the habitual careful consideration of the data before any statistics is undertaken. A visual appraisal of a graphical image is of paramount importance here, and the book is profusely illustrated with them: almost every dataset discussed or analysed is depicted. Another essential is developing an understanding of the exact meaning of the results of statistical operations. Finally, practitioners need the experience of both guided and unsupervised consideration of many examples of relevant datasets. Drivers don't need too many details of how the car works, but they do need the Highway Code, a road map, some driving lessons and as much practice as they can get.

The book is divided into quite short sections, each dealing with a single topic, headed by a 'Key points' box. Most sections are terminated by 'Notes and further reading' with useful references for those wishing to pursue topics in more detail. The sections are as far as possible self-contained, but are extensively cross-referenced. The book can therefore be used either in a systematic way by reading the sections sequentially, or as a quick reference by going directly to the topic of interest. Every statistical method and application covered has at least one example where the results are analysed in detail. This enables readers to emulate this analysis on their own examples. The statistical results on these examples have been cross-checked by at least two different statistics packages. All of the datasets used in examples are available for download, so that readers can compare the output of their own favourite statistical package with that shown in the book and thus verify that they are entering data correctly.

Statistics is a huge subject, and a problem with writing a book such as this is knowing where to stop. We have concentrated on providing a selection of techniques and applications that will satisfy the needs of most analytical chemists most of the time, and we make no apology for omitting any mention of the numerous other interesting methods and applications. Statisticians may be surprised at the relative emphasis placed on different topics. We have simply used heavier weighting on the topics that experience has told us that analytical chemists have most difficulty with. The book is cast in two parts, a technique-based approach followed by an application-based

approach. This engenders a certain amount of overlap and duplication. Analytical chemists are thereby encouraged to create their own overview of the subject and see for themselves the relationship between tasks and techniques.

Statisticians will also notice that we use the '*p*-value' approach to significance testing. This was adopted after careful consideration of the needs of analytical chemists. It greatly improves the transparency of significance testing so long as the exact meaning of the *p*-value is borne in mind, and we stress that meaning repeatedly. The alternative approaches tend to cause more difficulty. Experience has shown that analytical chemists find the somewhat convoluted logic of using statistical tables confusing and hard to remember. The confidence interval approach is simple but almost universally misunderstood among non-statisticians. Both of these methods also have the disadvantage of engendering the idea that the significance test can validly dichotomise reality — once you have set a level of confidence the test tells you 'yes' or 'no'. This tempts practitioners to use statistics to replace judgement rather than to assist it.

Finally, here are ten basic rules for analytical chemists undertaking a statistical analysis.

1. If you can, plan the experiment or data collection before you start the practical work. Make sure that it will have sufficient statistical power for your needs. Ensure that the data collection is randomised appropriately, so that any inferences drawn will be valid.
2. Make sure that you know how to enter the data correctly into your statistical software. After you have entered it, print it out for checking.
3. Examine the data as one or more graphical displays. This will often tell you all that you need to know. In addition it will tell you if your dataset is unlikely to conform to the statistical model that underlies the statistical test that you are proposing to use. Important features to look out for are: suspect data; lack of fit to linear calibrations; and uneven variance in regression and analysis of variance.
4. Select the correct statistical test, e.g., one-tailed or two-tailed, one sample, two sample or paired.
5. Make sure that you know exactly what the statistical output means, especially the *p*-value associated with a test of significance.
6. Be careful how you express the outcome in words. Avoid attributing probabilities to hypotheses (unless you are making a Bayesian analysis — not within the scope of this book).

7. Report the magnitude of an effect as well as its significance. Distinguish between effects that are statistically significant and those that are practically important.
8. After a regression always make plots of the residuals against the predictor variable. This will give you valuable information about lack of fit and uneven variance. It is sometimes useful to make other plots of the residuals, e.g., as a time series to detect drift in the measurement process.
9. If in doubt, ask a statistician.
10. Have fun!

Michael Thompson & Philip J Lowthian
Birkbeck University of London, UK
May 2010

* * * * *

Data files used in the book can be downloaded from http://www.icpress.co.uk/chemistry/p739.html

Contents

<i>Preface</i>	v
Part 1: Statistics	1
1. Preliminaries	3
1.1 Measurement Variation	3
1.2 Conditions of Measurement and the Dispersion (Spread) of Results	4
1.3 Variation in Objects	5
1.4 Data Displays	6
1.5 Statistics	6
1.6 Levels of Probability	7
1.7 An Example — Ethanol in Blood — a One-Tailed Test .	8
1.8 What Exactly Does the Probability Mean?	9
1.9 Another Example — Accuracy of a Nitrogen Analyser — a Two-Tailed Test	10
1.10 Null Hypotheses and Alternative Hypotheses	12
1.11 Statements about Statistical Inferences	12
2. Thinking about Probabilities and Distributions	15
2.1 The Properties of the Normal Curve	15
2.2 Probabilities Relating to Means of n Results	17
2.3 Probabilities from Data	19
2.4 Probability and Statistical Inference	20
2.5 Pre-computer Statistics	21
2.6 Confidence Intervals	22
3. Simple Tests of Significance	25
3.1 One-Sample Test — Example 1: Mercury in Fish	25
3.2 One-Sample Test — Example 2: Alumina (Al_2O_3) in Cement	27

3.3	Comparing Two Independent Datasets — Method	28
3.4	Comparing Means of Two Datasets with Equal Variances	30
3.5	The Variance Ratio Test or <i>F</i> -Test	31
3.6	Two-Sample Two-Tailed Test — An Example	32
3.7	Two-Sample One-Tailed Test — An Example	34
3.8	Paired Data	36
3.9	Paired Data — One-Tailed Test	38
3.10	Potential Problems with Paired Data	40
4.	Analysis of Variance (ANOVA) and Its Applications	43
4.1	Introduction — The Comparison of Several Means	43
4.2	The Calculations of One-Way ANOVA	45
4.3	Example Calculations with One-Way ANOVA	47
4.4	Applications of ANOVA: Example 1 — Catalysts for the Kjeldahl Method	49
4.5	Applications of ANOVA: Example 2 — Homogeneity Testing	51
4.6	ANOVA Application 3 — The Collaborative Trial	54
4.7	ANOVA Application 4 — Sampling and Analytical Variance	57
4.8	More Elaborate ANOVA — Nested Designs	60
4.9	Two-Way ANOVA — Crossed Designs	63
4.10	Cochran Test	65
4.11	Ruggedness Tests	66
5.	Regression and Calibration	71
5.1	Regression	71
5.2	How Regression Works	73
5.3	Calibration: Example 1	76
5.4	The Use of Residuals	78
5.5	Suspect Patterns in Residuals	80
5.6	Effect of Outliers and Leverage Points on Regression	82
5.7	Variances of the Regression Coefficients: Testing the Intercept and Slope for Significance	83
5.8	Regression and ANOVA	85
5.9	Correlation	87

5.10	A Statistically-Sound Test for Lack of Fit	90
5.11	Example Data/Calibration for Manganese	91
5.12	A Regression Approach to Bias Between Methods	94
5.13	Comparison of Analytical Methods: Example	97
6.	Regression — More Complex Aspects	101
6.1	Evaluation Limits — How Precise is an Estimated x -value?	101
6.2	Reducing the Confidence Interval Around an Estimated Value of Concentration	104
6.3	Polynomial Regression	105
6.4	Polynomial Calibration — Example	108
6.5	Multiple Regression	109
6.6	Multiple Regression — An Environmental Example	111
6.7	Weighted Regression	116
6.8	Example of Weighted Regression — Calibration for ^{239}Pu by ICP-MS	119
6.9	Non-linear Regression	122
6.10	Example of Regression with Transformed Variables	124
7.	Additional Statistical Topics	127
7.1	Control Charts	127
7.2	Suspect Results and Outliers	130
7.3	Dixon's Test for Outliers	132
7.4	The Grubbs Test	133
7.5	Robust Statistics — MAD Method	135
7.6	Robust Statistics — Huber's H15 Method	137
7.7	Lognormal Distributions	139
7.8	Rounding	141
7.9	Non-parametric Statistics	142
7.10	Testing for Specific Distributions — the Kolmogorov-Smirnov One-Sample Test	144
7.11	Statistical Power and the Planning of Experiments	146
Part 2: Data Quality in Analytical Measurement		151
8.	Quality in Chemical Measurement	153
8.1	Quality — An Overview	153
8.2	Uncertainty	154

8.3	Why Uncertainty is Important	156
8.4	Estimating Uncertainty by Modelling the Analytical System	158
8.5	The Propagation of Uncertainty	160
8.6	Estimating Uncertainty by Replication	162
8.7	Traceability	163
8.8	Fitness for Purpose	165
9.	Statistical Methods Involved in Validation	167
9.1	Precision of Analytical Methods	167
9.2	Experimental Conditions for Observing Precision	169
9.3	External Calibration	171
9.4	Example — A Complete Regression Analysis of a Calibration	172
9.5	Calibration by Standard Additions	175
9.6	Detection Limits	177
9.7	Collaborative Trials — Overview	180
9.8	The Collaborative Trial — Outlier Removal	182
9.9	Collaborative Trials — Summarising the Results as a Function of Concentration	184
9.10	Comparing two Analytical Methods by Using Paired Results	186
10.	Internal Quality Control	189
10.1	Repeatability Precision and the Analytical Run	189
10.2	Examples of Within-Run Quality Control	192
10.3	Internal Quality Control (IQC) and Run-to-Run Precision	195
10.4	Setting Up a Control Chart	197
10.5	Internal Quality Control — Example	198
10.6	Multivariate Internal Quality Control	201
11.	Proficiency Testing	205
11.1	Proficiency Tests — Purpose and Organisation	205
11.2	Scoring in Proficiency Tests	207
11.3	Setting the Value of the Assigned Value x_A in Proficiency Tests	208
11.4	Calculating a Participant Consensus	210

11.5	Setting the Value of the 'Target Value' σ_p in Proficiency Tests	215
11.6	Using Information from Proficiency Test Scores	216
11.7	Occasional Use of Certified Reference Materials in Quality Assurance	219
12.	Sampling in Chemical Measurement	221
12.1	Traditional and Modern Approaches to Sampling Uncertainty	221
12.2	Sampling Uncertainty in Context	223
12.3	Random and Systematic Sampling	226
12.4	Random Replication of Sampling	227
12.5	Sampling Bias	229
12.6	Sampling Precision	231
12.7	Precision of the Estimated Value of σ_s	234
12.8	Quality Control of Sampling	236
	Index	239

PART 1

Statistics



Chapter 1

Preliminaries

This chapter sets the scene for statistical thinking, covering variation in measurement results and the properties of objects, and its graphical representation. The basis for statistical inference derived from analytical data is associated with the probability of obtaining the observed results under the assumption of appropriate hypotheses.

1.1 Measurement Variation

Key points

- Variation is inherent in results of measurements.
- We must avoid excessive rounding to draw valid conclusions about the magnitude of variation.

The results of replicated measurements vary. If we measure the same thing repeatedly, we get a different result each time. For example, if we measured the proportion of sodium in a finely powdered rock, we might get results such as 2.335, 2.281, 2.327, 2.308, 2.311, 2.264, 2.299, 2.295 per cent mass fraction (%). This variation is not the outcome of carelessness, but simply caused by the uncontrolled variation in the activities that comprise the measurement, which is often a complex multistage procedure in chemical measurement. Sometimes it may appear that results of repeated measurements are identical, but this is always a false impression, brought about by a limited digit resolution available in the instruments used to make the measurement or by excessive rounding by the person recording or reporting the data. If the above data are rounded to two significant figures, they all turn into an identical 2.3%, which tells us nothing about the magnitude of