PRACTICAL MATHEMATICS FOR CHEMISTS

F. H. C. KELLY
D.Sc.(Tas.), M.Sc.(Melb.), A.Melb.T.C.,
F.R.A.C.I., A.Aust.I.M.M.
Reader in Applied Chemistry, University of Tasmania

LONDON
BUTTERWORTHS
1963

PREFACE

This booklet is intended to be neither a thesis nor a treatise, but extracts from the records of the endeavours of a chemist who has plodded slowly through problems both in the industrial plant and in the chemical laboratory. These extracts have been put together in a manner which it is hoped will be of help to his younger fellows, and which they might master at a sufficiently early stage in their studies to encourage them to proceed to the problems of greater complexity which will be the common lot of future generations. Mathematics is of no practical value to a chemist if used only as a paintbrush to describe with elegant formulae the phenomena which he observes, no matter how beautiful or aesthetic the picture may be. It must be used as a spade to unearth the hidden treasures of chemical behaviour, remembering always that a chemical reaction will take place because of the chemical nature of the reacting substances and not because of any picturesque mathematical concept that may go with it. If a man is to be master of the chemical reaction he must employ quantitative terms and suitable mathematical equipment, use it for the elucidation of complexities and not for the confusion of his colleagues, make it an aid to clarity in his own thinking and a means for the effective communication to others of his own observations.

It is hoped that it will not only provide an elementary guide for the growing chemist, but may also be found to be a useful pocket book carried for ready reference to aspects of the subject which many find difficult to carry within their minds beyond examination day, but which should be drawn upon on numerous occasions when the real problems of life may be encountered. The effective study of the operation of some item of processing equipment calls for a diagnostic procedure as important to the study as any employed by a medico for the diagnosis of ailments of the human body, a procedure which is not by any means incomprehensible in its complexity, but one requiring patient application of well established techniques. Here are some elements of mathematics which can be part of such studies aiding in their clarification, clearing confusion and perhaps even pointing a way to better or more efficient operation.

CONTENTS

1

STATISTICAL TREATMENT OF DATA

Statistics may be usefully employed by chemists and chemical engineers for the objective study of laboratory or plant data, if the study is combined with a common sense appreciation of the problems involved. It is important to first become acquainted with the statistical methods made available by mathematicians, and then from this foundation to proceed to the design of experiments and analysis of data, whether the experiments be laboratory, pilot plant, or in an industrial plant. Statistical methods are also useful when specifying the precision of a laboratory technique, either to compare one technique with another or to determine the influence of a particular technique in a series of operations. There is also value in being able to objectively fit a curve to data gathered when carrying out an experiment.

THE COLLECTION AND PRESENTATION OF DATA

Information suitable for use in any statistical treatment may be collected by personal investigation, from reports, or from routine data recorded during the normal operations of an industrial plant. No statistical analysis is any better than the data being examined; it does not improve the inaccuracy of poor data but describes it in objective terms. Every endeavour should be made to use reliable measuring equipment and effective methods for sampling.

Statistical Presentation

Information accumulated during the course of experiments or during the normal operation of a plant needs to be recorded accurately and completely, both for study by the original investigator and also for future study by an investigator who may wish to re-appraise the findings in the light of later knowledge or a change of circumstances. All records and comments on observations should be logged in suitable books and not on loose paper, although using a primary and secondary book is often helpful.

This information also needs to be analysed, summarized and effectively presented, so that a bird's eye view may be obtained to save the time and effort of others who may have to make a decision on the results of the investigation.

Two general methods of presentation are available; tabular, which summarizes results in cold figures, and graphical, which presents a picture to the mind which may more effectively express trends or situations to which attention is drawn, and which some people may more readily memorize. For a graph illustrating a trend an equation may be derived. For diagrams illustrating the distribution of data we may calculate average values and the degree and type of distribution or dispersion. Relationships of similar character may have their degree of similarity expressed as a degree of correlation. This type of treatment may be termed a statistical analysis—for which a system of terminology has grown up, at times more confusing than the unanalysed tabulated or graphically illustrated data.

Statistical Control

In common industrial terminology a process spoken of as being 'under control' generally utilizes automatic control techniques. Strictly speaking this is too restricted a use of the phrase, and from the point of view of statistics any particular variable is under control as long as the variable is being measured. Thus the temperature fluctuations in a chemical reactor are considered to be under statistical control if a thermometer is inserted in the reactor and a record kept of temperature observations, irrespective of whether or not changes in the reading of the thermometer have any consequent influence on the temperature of the reactor.

The first stage of statistical control is, therefore, to make a series of measurements or record a series of values. The second stage is to study the variations in such measurements. In statistics these variations are referred to as 'errors', but it must be remembered that in the terminology of statistics the word 'error' does not mean 'mistake'. The third stage is to study the character of the variations, that is whether they are random or biassed. Random variations are those which appear to occur without any detectable specific influence and which are studied according to the rules of probability of behaviour. If some specific influence is exercising a control on variations, this is a bias. The effect may be identified by statistical examination of information, and when it is established the cause may be energetically sought. From this point we can continue studies in the fields of trends and correlation, precision, interdependence or wherever else the particular problem may lead. In the case of chemical behaviour we employ mathematical weapons to elucidate the mechanism, which not only provides us with an understanding of the real reasons for the observed behaviour, but also allows us to exercise more specific control, and generally to extend our influence and minimize our effort. At all times, however, it is important to remember that observations correctly made are more important than the mathematical analysis or personal interpretation. Chemistry is a science built up on the foundations of experiment and experimental observations into which any mathematical superstructure must be correctly integrated. Analysis is followed by synthesis, and the fundamental chemistry must be kept in mind at all times to avoid fallacy and absurdity.

Sampling

Information collected about a chemical reaction, whether carried out in the laboratory or on factory scale, is only as useful as the sample on which the measurement is made. In the laboratory it is generally not particularly difficult to obtain a sample which is truly representative of the whole, whereas this is seldom the case on a factory scale. We can, in fact, spend a great deal of the expensive time of an analyst in carrying out an analysis in duplicate or triplicate which does not increase the precision of our information, because of the difficulties in effectively sampling the system. Statistical analysis of results should be able to tell us whether a succession of samples is good, bad or indifferent, and which of such comments should also be applied to the combined effect of the analytical procedures. Common sense and practical experience are the best prescriptions when seeking improvements in sampling techniques.

Effective sampling is based on the theory of probability, and two laws need to be borne in mind.

First Law of Sampling—This is also referred to as the law of statistical regularity, which states that the characteristics of a small group or quantity (the sample) chosen at *random* from a larger group or quantity (the system) will not be different from the characteristics of the larger group or quantity. As the size of the sample increases, the reliability will increase. Generally speaking the reliability of the sample is proportional to the square root of its size. We may measure the size of a sample in terms of weight or volume or number of articles, but a sample which is four times the size of another should be twice as reliable from the point of view of sampling technique.

Second Law of Sampling—There is likely to be less change in a large bulk of material or group of items than in a small bulk of material or group of items. In sampling it is essential that every portion of the bulk of material or every item in the group under consideration shall have the same chance of being selected. To achieve this effectively, it is essential to have adequate mixing and make selection a truly random procedure.

In chemical plant and process analysis by far the larger number of sampling problems involve material in bulk—in storage, in a process vessel, or in some type of handling or transfer operation. The material may be homogeneous in character *i.e.* a single phase of solid, liquid or gas; or it may be heterogeneous, involving a mixture of two or more phases. Systems which contain two or more solid phases (*e.g.* gold in quartz) can be just as difficult to sample effectively as systems made up of mixtures of solid and liquid, or liquid and gas, or solid and gas.

In the case of systems made up only of solid phases, or in which the solid phase is more than about 60 per cent by volume of the total, bulk mixing as the term is commonly understood must be preceded by break-down in particle size. The size to which material should be broken down for effective sampling is best studied from the quantitative statistical point of view in association with a study of the variation of results, and it will later be considered in this context. Generally speaking, however, the larger the largest particle the greater the weight of the sample will need to be. A better criterion might be the size of the largest particle of the substance for which the bulk is being analysed. From either point of view we might quantify our approach by specifying a constant for the ratio of weight of largest particle to total weight of sample, and use this constant in association with our study of variations in analytical results.

Stratification is one of the most difficult features of heterogeneous systems from the point of view of sampling, and, in fact, we can use sampling to study the degree and extent of stratification. This may be met in the case of an ingot of an alloy which has cooled slowly, or when discrete particles are suspended in a gas as a smoke or mist. Stratification must always be suspected in heterogeneous systems and, if it cannot be overcome by effective mixing, it must be measured and allowed for by a planned sampling procedure.

When sampling bulk material of large particle size, such as lumps of coal or ore or sticks of sugar cane, it is essential to reduce the size of the material before a small enough sample for chemical analysis can be truly representative of the whole. In the case of lumps of material, these may be pulverized by suitable equipment and the size of the sample reduced progressively. It may finally be necessary, for chemical reasons, to reduce the particle size so that all may pass a 200 mesh sieve. Such a procedure would be most costly if applied to a bulk of perhaps a hundred pounds of sample of which only a few grams are needed for analysis. Specifications for sampling procedure to be applied to coal and mineral ores have been set out by the U.S. Bureau of Mines and other bodies, and in fact procedures have been specified for most commercial materials where a buyer and seller relationship exists. These specifications have usually been developed over a period of years mainly as the result of experiment and common sense, and very often the law of the land requires that the specified procedures be followed. Procedures specified for a particular situation should be carefully studied, and if not available then those developed for a related situation, combined with common sense and adequate statistical analysis, will provide an effective approach. An important objective of statistical examination should be to minimize expenditure of time and effort on the sampling and analytical procedures, in order to obtain the degree of precision which the situation warrants. If circumstances demand a higher degree of precision then the steps required to meet this demand should be clearly indicated, and it will be for the management to determine the extent to which the costs of meeting the demand are warranted.

Random selection is often as difficult to achieve as adequate mixing. The human mind has a tendency to exercise a preference according to some physical characteristic which influences the senses. Often it is 'eye-appeal' which is influenced, with a tendency to select large lumps, damaged containers, lumps of quartz showing free gold or some other unintentional preference. A conscientious sampler may even overcorrect himself if he recognizes an influence of this character. To overcome this type of influence, complete mechanization of the sampling procedure is desirable, or it may be necessary to draw a sketch plan of the heap of material and systematically lay out locations for sampling. Sometimes it may be necessary to select a number of items, in which case an individual may have a personal bias for particular numbers. This may be overcome by consulting a table of random numbers prepared by mathematicians for this purpose.

If one does not have a set of random numbers available it may be more convenient to draw marbles from a jar, toss coins, throw dice or draw cards from a pack. A small regular decagon may be cut and mounted on a spindle to be spun, and with each segment suitably marked it is possible to obtain at random a digit or succession of digits according to the size of the number required.

In *Table* 1.01 is listed a set of 600 numbers selected by the author at random. This is only a small set compared with some which have been prepared for more general mathematical purposes running to 10 or 20 times as many numbers. However for most requirements of chemists the smaller table is sufficient and more convenient. It is not unusual for an industrial chemist to be required to produce random numbers for sampling or related purposes. One may be prepared for such occasions by carrying either a pocket selector of the mechanical type or the set of 600 random numbers printed on a card small enough to be conveniently carried.

Table 1.01. Random Numbers

573	955	458	122	039	679	334	978
112	037	476	478	278	077	096	279
710	043	220	977	671	019	354	913
680	180	617	834	449	980	209	116
591	516	421	385	955	708	269	717
389	374	235	336	703	495	608	895
431	908	261	883	888	426	685	728
570	600	872	457	448	134	714	586
564	559	287	181	044	156	734	501
147	449	619	738	697	830	269	627
074	633	464	088	765	433	897	249
946	917	653	092	521	214	696	084
992	491	649	994	565	315	265	393
229	981	217	206	186	175	338	130
355	751	216	282	266	569	267	735
628	438	205	751	823	140	318	932
875	014	812	437	666	673	604	022
499	061	409	551	003	270	208	835
438	746	295	455	345	713	565	841
119	001	209	972	088	354	874	639
319	176	401	222	580	302	507	072
825	662	960	211	709	280	382	584
535	828	458	653	351	064	393	865
937	169	141	097	127	064	768	971
459	556	020	713	790	204	925	430

In *Table 1.01* the digits are arranged as shown merely for convenience of reading. No individual number carries any more weight than any other number, each digit being an independent sample from a group in which the digits 0 to 9 are equally likely to appear. Taking the table as a whole each digit appears on 60 occasions, which is very seldom attainable in a group of this magnitude simply by random selection, but which has been attained in this case by a process of successive random selection; it could also have been obtained by placing 600 suitably numbered marbles in a barrel and extracting them as in a lottery. Actually any table of lottery numbers frequently published in the daily newspaper could be used as a table of random numbers. Also it is an interesting exercise just to analyse a table of random numbers or lottery results from the purely statistical point of view to observe the frequency with which particular digits appear in samples of varying size. This exercise, however, would best be left until after Chapters 3 and 4 have been completed. For the moment it is sufficient to indicate that a random number may be selected from the table by placing, with one's eyes closed, a pointer on the table and selecting the nearest digit. The subsequent digits are then random appearances. If numbers involving two or three digits are required then each digit may be selected in turn or taken as a group of two or three. Alternatively, the first two randomly selected digits may be used to specify a row and column from which to take the actual digits used, and there are other ways of using the table. It is important, however, to decide *before* starting the selection process as to the details of procedure which will be followed, otherwise personal bias can creep in.

PLANNING INDUSTRIAL EXPERIMENTS

Experiments carried out in an industrial plant covering one or more unit operations and/or unit processes are generally expensive, involving extra manpower and also frequently resulting in dislocation of the main processing operations. It is of paramount importance, therefore, to plan the experiments with the utmost care, so that the maximum amount of information may be gained from the minimum expenditure of effort. As a general rule the following procedure has been found to provide a very effective approach:

(1) Write down the flow sheet covering the whole of the

stage being studied and, if necessary, also associated stages. The flow sheet may have been well established for many years, but there are occasions when the custom has become so well established that its weaknesses are not revealed until some investigator makes a fresh approach, applying a systematic analysis and avoiding the pitfall of accepting the apparently obvious as necessarily being the most advantageous technique.

(2) Prepare a materials balance.

(3) Prepare an energy balance (this may or may not be necessary depending upon the scope of the investigation).

(4) Study equilibrium conditions operating. These will generally be conditions of dynamic equilibrium, but may also include conditions of static equilibrium, *e.g.* solubility, pH, vapour pressure and kindred determinations. In the case of a continuous operation involving chemical reactions or physical rate processes, such as rate of solution or rate of crystallization, a condition of dynamic equilibrium will generally be established.

(5) Study the rate processes involved. These include not only chemical reactions but also physical processes such as fluid flow and heat transfer, as well as processes of mass transfer such as absorption, crystallization, solution, *etc.*

(6) Make a statistical examination of data to establish their real significance in relation to the problem under examination.

In preparing the programme of investigation we may identify items for which fundamental information is required, for example solubility or vapour pressure. The information on pure substances may be available in the literature or from other sources. Failing this, it may be necessary to carry out laboratory tests to obtain it, and this is referred to as the collection of 'background information', *i.e.* information of a fundamental character but specifically related to the problem in hand.

The employment of 'factorial design' techniques can often be of distinct advantage in minimizing the number of plant tests which need to be carried out for a particular investigation. For example, if we have to study the influence of three variables such as time, temperature and concentration on a particular result, we first prepare a diagram of a type similar to that shown in *Figure 1.01*. We then postulate the two end conditions for each variable we propose to

	Conc. 1			*Conc. 2*			*Conc. 3*		
Time	*temp. 1*	t_2	t_3	t_1	t_2	t_3	t_1	t_2	t_3
T_1	1	2	3	10	11	12	19	20	21
T_2	4	5	6	13	14	15	22	23	24
T_3	7	8	9	16	17	18	25	26	27

Figure 1.01.

study, and fix the intermediate condition half way between. If we wish to study the system at additional levels, then the steps between each level should be equal in magnitude. This arrangement will appreciably simplify statistical examination of the data.

If we proceed to fill in all the blank pigeon holes in the above diagram, it is evident that 27 results will be needed. But by virtue of the systematic manner in which we have gone about the investigation, we will have as much information as we would have got from 81 experiments if we had tested completely the condition at each particular time, temperature and concentration separately (this number of experiments is made up by failing to recognize the similarities in such combinations as $T_1\,C_2\,t_3$, $t_3\,C_2\,T_1$ and $C_2\,t_3\,T_1$).

If any of the relationships are simple and linear in character we can reduce the number of levels at which the experiments are conducted. For example, if the result is directly proportional to the time, then it would be sufficient to carry out our experiments at only two levels of time and thereby reduce the number from 27 to 18.

On the other hand, even if all inter-relationships are non-linear in character we can still reduce the number of experiments and at the same time gather a lot of very useful and often sufficient information. For example, three points are sufficient to derive the constants in a parabolic type relationship. The squares 1, 5, 9, 16, 14, 12, 19, 23 and 27 will cover three points in each of the three inter-relationships, but we can reduce the number of experiments from 9 to 7 and still obtain the knowledge we are seeking.

Thus: 1, 2 and 3 will show the *effect of temperature* change at constant time and concentration.

1, 10 and 19 will show the *effect of concentration* change at constant time and temperature.

1, 4 and 7 will show the *effect* of change in *time* at constant temperature and concentration.

There are other combinations of conditions which should produce the same result, and the alternative combinations should be studied in association with the particular problem, as the system may well be more easily set to one particular combination of conditions rather than another.

If we know the exact character of the inter-relationships *i.e.* whether they be linear, logarithmic, exponential, parabolic, *etc.*, then we can reduce the number of experiments to 4, *viz.* 1 and 3, 1 and 19, and 1 with 7. It is not necessary to know more than the character of the relationship as most can be reduced to a straight line *e.g.* by log–log or log–linear plot. These will be discussed in more detail in association with a study of curves and curve-fitting.

EXERCISES

1. Prepare plans and specifications for the study of individual unit operations such as crushing, sedimentation, filtration, multiple effect evaporation, crystallization, distillation, extraction.
2. Give similar consideration to reactor systems in which unit processes are occurring such as neutralization, precipitation, nitration, sulphonation, oxidation, reduction.
3. Examine combinations of two or more unit operations and/or unit processes.