

Notes on

The application of statistics to paper testing

The following is a list of members of the Paper Industry Standards Committee, who have had the preparation of this standard under consideration:—

Chairman: Mr. L. G. Cottrall

Vice-Chairman: Mr. G. F. Underhay

Mr. B. G. E. Heywood, M.C.		<i>Board of Trade</i>
Mr. A. Kirk	}	<i>British Federation of Master</i>
Mr. L. G. White		<i>Printers</i>
Dr. N. R. Hood		<i>British Paper and Board Industry</i>
		<i>Research Association</i>
Dr. C. Binns	}	
Mr. R. W. H. Bray		
Mr. F. D. Broadbent		
Dr. W. Burt		
Mr. E. E. Capon		
Mr. L. G. Cottrall		
Dr. Julius Grant		
Mr. A. Gemmell		
Dr. H. Ainsworth Harrison		<i>British Paper and Board Makers'</i>
Mr. S. H. Hough		<i>Association</i>
Mr. H. G. Paul		
Dr. H. F. Rance		
Mr. S. F. Smith		
Mr. F. H. Llewellyn Thomas		
Mr. G. F. Underhay		
Mr. A. C. Vincent		
Mr. W. Whiteley, J.P.		<i>British Railways, The British</i>
Mr. J. O'Neill		<i>Transport Commission</i>
Mr. G. Newman	}	<i>Coated Paper and Board Makers'</i>
Mr. K. Timberlake		<i>Association</i>
Mr. H. L. Morbey		<i>Envelope Makers' and Manufacturing</i>
		<i>Stationers' Association</i>
Mr. L. G. Smith		<i>H.M. Stationery Office</i>
Mr. E. L. Hill		<i>Ministry of Supply</i>
Mr. J. E. Pickering	}	<i>National Association of Paper</i>
Mr. C. J. Thorne, O.B.E.		<i>Merchants</i>
Dr. V. G. W. Harrison		<i>Printing, Packaging and Allied</i>
		<i>Trades Research Association</i>
Mr. F. Bolam	}	<i>Technical Section of the British</i>
Dr. H. F. Rance		<i>Paper & Board Makers' Association</i>
Mr. G. Thompson		<i>Co-opted. (Chairman of Technical</i>
		<i>Committee on Methods of Test)</i>

This British Standard, having been approved by the Paper Industry Standards Committee and endorsed by the Chairman of the Chemical Divisional Council, was published under the authority of the General Council on 27 May 1958

© BSI 03-2000

The following BSI references relate to the work on this standard:—
Committee reference PAC/11
Draft for comment CX(PAC)8756

ISBN 0 580 35830 5

The following is a list of members of the Technical Committee on Methods of Test for Paper, Board and Pulp: the Committee actually responsible for the preparation of the standard:—

Chairman: Mr. G. Thompson

Vice-Chairman: Mr. C. A. Chester

Mr. A. Kirk		<i>British Federation of Master Printers</i>
Mr. W. E. Bennett		<i>British Paper and Board Industry Research Association</i>
Mr. A. Baker	}	<i>British Paper and Board Makers' Association</i>
Mr. L. G. Cottrall		
Mr. L. G. S. Hebbs		<i>British Woodpulp Association</i>
Mr. R. Allum		<i>Crown Agents for Oversea Governments and Administrations</i>
Mr. E. Halson, M.B.E.		<i>H.M. Stationery Office</i>
Mr. K. MacIntyre		<i>Ministry of Supply</i>
Dr. C. Binns	}	<i>Technical Section of the British Paper and Board Makers' Association</i>
Mr. F. Bolam		
Mr. F. Bridge		
Mr. C. A. Chester		
Mr. T. H. Farebrother		
Mr. G. F. Glover		
Mr. I. F. Hendry		
Mr. L. F. Hopkins		
Dr. F. L. Hudson		
Mr. L. A. Lawrence		
Mr. D. McNeill		
Dr. H. F. Rance		
Mr. G. Thompson		
Mr. A. F. Tout		
Mr. H. G. Tyler		
Mr. G. F. Underhay		
Dr. V. G. W. Harrison	}	<i>Printing, Packaging and Allied Trades Research Association</i>
Mr. R. R. Coupe		
Mr. J. Meighan		<i>Society of British Soap Makers</i>

Amendments issued since publication

Amd. No.	Date	Comments

Contents

	Page
Foreword	ii
Introduction	1
Section 1. A practical problem from paper testing	2
Section 2. Application of statistical tests to the example	4
Section 3. Introduction to statistical methods	6
1 Distribution of results	6
2 Sampling from a population	6
3 Standard deviation (σ)	7
4 Coefficient of variation (v)	7
5 Estimation of the mean and the standard deviation from a sample	7
6 Precision of the mean: Confidence limits	8
7 Number of tests required for a given precision of the mean	9
8 Significance of difference between means of two tests	9
9 Student's t test	10
Section 4. Example of calculation of Standard deviation, precision of mean and application to t tests, together with recommended computing procedures	14
Section 5. Applications to paper testing	16
References	25
Appendix A Rejection of abnormal results	18
Appendix B Note on the use of range	19
Appendix C Coefficient of variation for various paper tests	19
Appendix D Table for t for various numbers of degrees of freedom v	21
Appendix E Formulae for easy reference	21
Appendix F Glossary of statistical terms and symbols	22
Figure 1	6
Figure 2	8
Figure 3	10
Figure 4	11
Figure 5	12

Foreword

This standard makes reference to the following British Standards:

BS 600, *Application of statistical methods to industrial standardization and quality control*.

BS 2846, *The reduction and presentation of experimental results*.

BS *Sampling paper for testing*¹⁾.

This British Standard is based upon British PBMA Method S1:1955, prepared by the Statistics Committee of the Technical Section of the British Paper and Board Makers' Association, to which Committee grateful acknowledgment is made for the considerable amount of work involved in its preparation.

The standard, which is one of a series for paper testing, is published under the authority of the Paper Industry Standards Committee. The other standards so far published in this series are:—

BS 2644, *Sizing properties of paper — Method of testing the degree of water resistance*.

BS 2699, *The absorbency of blotting paper — Determination of ink absorbency time*.

BS 2916, *Absorbency test for bibulous paper*.

BS 2922, *Methods for testing the strength of wet paper*.

BS 2923, *The printing opacity of paper*.

BS 2924, *Methods for determining the pH value of aqueous extracts of paper*.

BS 2925, *Methods for determining the air permeability and air resistance of paper*.

This standard has two purposes. One is to arouse the interest of persons engaged on paper testing in elementary statistical methods. The other is to provide an easy source of reference to the formulae required for the treatment of data, so as to satisfy the general reporting clauses included in most of the paper testing methods. To that extent it is complementary to the British Standards enumerated above.

It is hoped that this explanation will make the layout of this standard more intelligible to the general reader, for the dual purpose has inevitably resulted in the juxtaposition of elementary and more advanced concepts.

A British Standard does not purport to include all the necessary provisions of a contract. Users of British Standards are responsible for their correct application.

Compliance with a British Standard does not of itself confer immunity from legal obligations.

Summary of pages

This document comprises a front cover, an inside front cover, pages i to iv, pages 1 to 24, an inside back cover and a back cover.

This standard has been updated (see copyright date) and may have had amendments incorporated. This will be indicated in the amendment table on the inside front cover.

¹⁾ In course of preparation.

Introduction

The problems of sampling and the reliability of results obtained in paper testing are clearly cases for the application of statistical methods, and in recent years a number of papers have been published indicating the growing interest in statistical methods in the paper trade.

The results of paper tests are usually subject to appreciable variation, due not only to experimental error, but also to variation in the raw materials and to the operation of many unknown or random factors in the manufacturing process. The interpretation of test results therefore requires some care and the chief value of statistical methods in this connection lies in their ability to replace purely subjective judgments of the data by objective criteria such as the t test, which is described in the text. Statistical analysis of the data cannot add to the accuracy of an experiment by turning an uncertain result into a certainty, but it can enable the conclusions to be expressed more precisely in terms of a definite probability.

Although statistical methods have their foundation in mathematical theories of probability, a knowledge of such theories is not required for the successful application of these methods. Only straightforward arithmetical processes are involved and, in many cases, the calculation does not require the use of a machine. One or two shortcut methods to simplify calculations are given.

The purpose of the present notes is to give only a brief outline of the subject, but with sufficient detail to enable the reader to carry out the simpler statistical calculations on his test data, such as those for ascertaining the precision of test means and the significance of differences in test means. For the more advanced techniques, such as that of analysis of variance, which can be of great utility for the comparison of instruments and for the analysis of causes of variation in test readings, reference is made to appropriate textbooks.

A secondary purpose of these notes is to secure greater uniformity in the presentation of test results and in the use of statistical terms and symbols (see Glossary, Appendix F).

In the opinion of the Committee responsible for the preparation of these notes, the problems of sampling depend too much on particular conditions (such as of manufacture and storage) to be capable of statistical treatment here, although some consideration has been given to these matters in the preparation of BS ²⁾ *“Sampling paper for testing.”* Problems of sampling are being considered by the Statistics Committee of the Technical Section of the British Paper and Board Makers' Association and will be the subject of a future publication.

²⁾ In course of preparation.

Pending completion, see British PBMA Method PT2:1951, published by the Technical Section of the British Paper and Board Makers' Association.

Section 1. A practical problem from paper testing

Imagine that a paper merchant is accustomed to supplying a certain kraft paper A for bag making. He knows from experience over a long period of time that this material is satisfactory, but recently he has been offered quantities of papers B, C and D at attractive prices. As a safeguard against complaints, he decides to test the papers for strength before attempting to pass them on to his clients. He carries out burst tests, ten for each paper, under comparable conditions and gets the following results:

Burst tests (lb/sq. in.)			
A	B	C	D
59	59	57	63.5
56	59	67.5	60.5
55	61.5	62	53
62.5	59	54	52.5
54	56.5	70	43
61.5	66.5	57	62.5
58.5	64	65.5	61
54.5	62.5	65	57
52.5	56	60	57.5
60	59	66	58

The purpose of these tests is to show up differences of strength of the four papers, and probably the first thing the merchant will do with the figures is to run his eye up and down the columns, This may be sufficient to answer the questions, but, if the figures are difficult to compare, he will calculate the means or average values of the sets of ten results.

	A	B	C	D
Mean	57.4	60.3	62.4	56.9

In doing this, although he perhaps neither knows nor cares, he has applied statistical methods of analysis to the sets of figures. He has deduced a *derived quantity* or *statistic* which gives information in a succinct form about the readings.

What do these mean values tell him?

One person may conclude from the four means that papers B and C are stronger than A and paper D is weaker. Another may be more cautious. He knows that any test on paper is subject to considerable variability — that is why he has done ten tests instead of one — and there is therefore some uncertainty about the mean values.

He may conclude that, since the test average of C is 8 per cent higher than that of A, C is appreciably stronger, and B possibly stronger than A. The result for D is about 1 per cent weaker and, although he may have doubts about the significance of this result, he will perhaps reject paper D, just to be on the safe side.

The mean value is not enough

The variability of the readings has to be considered. In the unrealistic case where all ten readings on paper A happened to be 57.4 and those on B 60.3 it could be confidently asserted that B is stronger than A. In practice, some variability is to be expected, and the greater this is the more doubtful does the conclusion become. As the variability increases further, it becomes obvious that there is no reliable evidence of a difference of strength between the two papers.

Ways of assessing variability

A rough impression can be obtained by running over the figures by eye, but conclusions drawn from this procedure are subjective. If the variability has to be expressed in a report the *highest and lowest values* can be used, or the *range*, but again the interpretation of these statements is open to doubt. The *mean deviation* (see Glossary, Appendix F) is a better expression of variability but it is of limited usefulness.

The most satisfactory way to express variability is to calculate the *standard deviation*, also called the *root mean square deviation*. This is the most *efficient* way of expressing the variability, because it remains more constant than the others in repeated sets of figures. The range, for example, may vary relatively more than the standard deviation from one set of readings to another because it depends only on the two extreme values in each set. The standard deviation is more difficult to calculate than the range or mean deviation, but this is more than compensated by its versatility. It is fundamental to statistical analysis.

The *coefficient of variation* is closely related. It is the standard deviation expressed as a percentage of the mean value. It often has the useful property of typifying a material. If a series of kraft papers were made from the same stock the bursting strength and standard deviation should increase with basis weight, but the coefficient of variation should be relatively constant.

The sample itself is another variable

Only a small part of the consignment of paper A was tested and this sample has been assumed to be representative of the paper as a whole.

If another sample were to be tested, it is likely that a different value of burst would be obtained and this might lead to different conclusions. When a large series of similar samples is tested, however, a definite pattern begins to arise. Different values of the mean will occur with different frequencies which will approach steady values as more and more samples are tested. On the basis of this phenomenon, statements can be made about a *single sample* which take care of the variability of the samples.

Why use statistics in paper testing?

In certain simple situations, conclusions can be reached without any reasonable doubt. It should be recognized, however, that personal judgment often plays a large part in arriving at a decision. This may be tempered by experience but, unless statistical methods are used, at least one guess is involved.

Statistics eliminates the guesses. Instead of a forthright statement that the bursting strength of paper A is 57.4 lb/sq. in. (which may be proved wrong by the next sample tested), a conclusion firmly based on statistics might be, "The odds are 19 to 1 that the true bursting strength of paper A lies between 55.0 and 59.8 lb/sq. in."

The question whether it is worth while to employ statistics in a particular problem depends entirely upon the importance of the answer. If the most objective conclusion is sought, for scientific or commercial reasons, then statistical methods should be considered.

In the following Section 2, some statistical tests will be applied to our example. The methods used are developed in Section 3 and those who are unfamiliar with these methods should preferably read Section 3 before proceeding to the application.

Section 2. Application of statistical tests to the example

The figures in the example in Section 1 can now be treated by the statistical methods developed in Section 3.

The standard deviation is calculated by the formula in Section 3.5.

The least significant difference is calculated by the formula in Section 3.9 b). The pooled standard deviation for the pairs being tested is worked out and the value of t used is 2.10 for a 1 in 20 level of significance and 10 readings in each column [See Section 3.9 a)].

	A	B	C	D
Mean	57.4	60.3	62.4	56.9
Observed difference from sample A	—	2.9	5.0	− 0.5
Standard deviation	3.42	3.30	5.25	6.10
Least significant difference from sample A	—	3.2	4.2	4.6

Consider first papers A and C. If in fact the two papers were of equal strength, then there would be only a 1 in 20 chance of drawing samples whose means differ by 4.2 or more. Since the observed difference in this case is 5.0, which exceeds this significant value, it is fairly safe to conclude that the papers are not of equal strength, and that C is stronger than A.

The differences between A and B and between A and D are less than the significant value and one must conclude that, on the basis of the tests done, there is no evidence that either paper B or D differs in bursting strength from paper A.

The difference between the means of A and B is about 3 lb and the merchant still has the feeling, in spite of the conclusion of the ten tests, that a true difference of strength is being concealed, which he would be able to detect if a large number of tests were carried out.

In pursuit of this idea, he does ten more bursts each on second samples of papers A and B and gets the following readings:

- A 54, 62.5, 60, 53, 57, 59.5, 59.5, 57.5, 51.5, 54.
B 60, 59, 62, 60, 55, 60, 55, 60, 57.5, 65.

If these are combined with those of A and B above, the mean values for twenty tests become:

- A — 57.1 and B — 59.8.

Because of the larger number of tests, the least significant difference will be smaller and becomes 2.1, compared with the 3.2 previously obtained.

The actual difference between the means is 2.7 and thus we can conclude, on the basis of twenty bursts on each paper, that paper B is stronger than A.

It will be noticed that in the example it was concluded that there was no reliable evidence of any difference between A and B; this conclusion was modified when twenty bursts were carried out. A difference of about 3 lb escaped detection when only ten tests were done. If the merchant had been only interested in detecting differences of 5 lb ten tests would have been sufficient. For a 2-lb difference, more than twenty tests would have been required and in general the number of tests must be related to the minimum size of difference in which one is interested. In fact, to catch small fish, a fine net must be used.

Sample D contains one very low reading. In Appendix 1, a simple rule is given for rejecting abnormal results. The rejection criterion r_{11} in this case takes the value 0.487 which is above the limit of 0.477 for 10 tests. The extreme value 43 is therefore rejectable. Whether it should be rejected will depend upon considerations discussed in Appendix A.

From the treatment given in this section to a merchant's typical problem, it will be seen that the application of statistical methods enables one to draw objective conclusions as opposed to those which rely on the intuition and experience of the tester. This can be claimed as one of the principal advantages of using statistical methods.

Other more advanced techniques may be used with advantage in paper problems, but are beyond the scope of these notes, which should be regarded as only an introduction to the subject. The methods of analysis of variance, for example, are particularly useful in experimental work and, for details of such methods, reference should be made to the textbooks. One way in which the analysis of variance can be of use is in investigating the causes of variation in test results, which may be attributable to the testing instrument itself, to its method of operating or to irregularity in the paper sample. In a suitably designed experiment, such methods of analysis will enable these separate factors to be distinguished with the minimum of effort.