1.4 Measurement

The process of measurement can be defined as a quantitative comparison between a predefined standard and the object being measured. This definition shows that there are two parts to the measuring process: the comparison, which is the process that is usually thought of as measurement, and the predefined standard, which is the part that is easily overlooked. When an object is weighed in the laboratory on a single pan balance it gives a reading of the mass of the object and so the balance is the local standard. However, what is actually taking place is a comparison of the mass of the object with that of the international standard kilogram. The validity of the measurement relies on there being a clear link between the balance that is in use and the international standard. In other words the balance needs to be calibrated with standard masses that have themselves been calibrated against other masses that in turn have been calibrated against the international standard. This link needs to be documented at each calibration step as to when it was carried out and to what limits of accuracy it has been made so that the calibration of a given instrument can be traced back to the international standards. It is also important that this calibration is carried out at regular intervals as instrumental readings can change over time because of wear of mechanical parts and ageing of electronic circuits. Besides regular calibration being good laboratory practice it is specifically demanded by ISO 9000.

The actual process of measurement is always subject to errors which can be defined as the difference between the measured value and the 'true' value. However, the 'true' value of any parameter can never be known because the value can only be obtained through measurement and any measurement can only be an estimation of the value, subject to unknown errors.

The term **precision** as used by metrologists [9] means the same as repeatability. It is defined as the quality that characterises the ability of a measuring instrument to give the same value of the quantity measured. In other words it is an indication of how well identically performed measurements agree with each other. A measurement of a property may return a value of 2.9347, which because of the number of figures after the decimal point may impress with its precision. If the measurement is then immediately repeated by the operator on the same object a value of 2.8962 may be obtained, when it will be seen that the number of figures represents a spurious precision and that the actual precision is much less. The precision of any measurement can only be obtained by making a number of identical measurements and estimating the dispersion of the results about the mean. The standard deviation or coefficient of variation of a set of results is used as a measure of this. A single measurement is always of an unknown precision, although in general the precision of particular test procedures is known through repeated testing in a single laboratory.

However, a result may be very precise in that every time the measurement is made the same number is obtained but it may vary from the 'true' value due to systematic errors.

Accuracy may be defined as conformity with or nearness to the 'true' value of the quantity being measured. This can only be obtained by calibration of the measuring system against the appropriate standards at suitable intervals.

Sensitivity is defined as the least change in the measured quantity that will cause an observable change in the instrument reading. The sensitivity of a measuring instrument can be increased by amplifying the output or by using a magnifying lens to read the scale. Without an accompanying increase in the accuracy of the calibration and a reduction in sources of variation this may mean no more than an amplification of the errors as well.

1.4.1 Statistical terms

Most measurements in textile testing consist of a set of repeat measurements that have been made on a number of identical individuals constituting a sample taken from the bulk of the material. Certain statistical measures are used to describe the average of the results and their spread. A short guide to the terms employed is given below. For a more comprehensive explanation a textbook of statistics should be consulted [10–12].

Arithmetic mean or average

The arithmetic mean is the measure most commonly chosen to represent the central value of a sample. It is obtained by adding together the individual values of the variable x and dividing the sum by the number of individuals n. It is represented by the symbol $\bar{x}$:

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \ldots + x_n}{n}$$

Standard deviation

The standard deviation is the most widely used measure of the dispersion or spread of results about the mean value. The symbol σ is used for the standard deviation of the universe (population) containing all the possible measurements that could be made of the variable in question. The symbol s is generally used for the estimated value of the standard deviation from a sample which has been taken from the universe

$$s = \sqrt{\left[\frac{\sum (x - \bar{x})^2}{n - 1}\right]}$$

The units that the standard deviation is measured in are the same as those of the mean.

Coefficient of variation

A standard deviation of 1 for a property that has a mean value of 10 is far more significant than a standard deviation of 1 for a property with a mean value of 100. Because of this the coefficient of variation (CV) is often used as a measure of dispersion: it is the standard deviation expressed as a percentage of the mean. Therefore in the above example the first result would have a CV of 10% and the second result would have a CV of only 1%.

$$\text{CV} = \frac{\text{standard deviation} \times 100}{\text{mean}} = \frac{s}{\bar{x}} \times 100\%$$

Standard error of the mean

The standard error of the mean is a measure of the reliability of the mean value obtained from a sample of a particular size. It is the standard deviation of the means that would be obtained if repeated samples of the given size were measured:

$$\text{Standard error of mean} = \frac{\sigma}{\sqrt{n}}$$

where σ is standard deviation of the parent universe. In the case of a sample the standard error of the mean has to be estimated by using the standard deviation of the sample s in place of σ.

The standard error can be used to place confidence limits on the mean that has been measured. For example there is a 95% probability that the population mean lies within $\pm$(1.96 $\times$ standard error) of the measured mean value. This relationship only holds when the standard error has been calculated from the standard deviation of the parent universe σ or when the sample is large. For small samples where s has been used to calculate the standard error, the value of 1.96 should be replaced by the appropriate value of t obtained from statistical tables.

1.4.2 Determination of number of tests

In any test the number of individuals to be tested will depend on the variability of the material and the accuracy required from the measurement [1, 13]. If the material is repeatedly sampled at random and the test performed on n selected items each time, in 95% of cases the mean value which is calculated will be within $\pm(2C/\sqrt{n})\%$ of the population mean, where C is the coefficient of variation of the property being tested and n is the number of test specimens. The values of $(2C/\sqrt{n})\%$ are the confidence limits of error. For many standard tests the coefficient of variation is known approximately so that the number of tests necessary to achieve given confidence limits of error can then be calculated. For instance if the coefficient of variation for a yarn strength test is 10% and the number of tests carried out n is 5, there is a 95% chance that the mean value will lie within $\pm 8.9\%$ of the population mean. If the number of tests is increased to 10, then there is the same chance that the mean will lie within $\pm 6.3\%$ of the population mean, and if the number of tests is increased to 50, then it is likely that the mean will lie within $\pm 2.8\%$ of the population mean.

The use of the coefficient of variation in the above formula assumes that the error, in the form of the standard deviation, is proportional to the mean value. For example in the above case of yarn strength if the mean value was 10 N then the standard deviation would be 1 N, whereas if the mean value was 100 N then the standard deviation would be 10 N. With some measurements the error is relatively independent of the magnitude of the mean. If this is the case then the actual standard deviation should be used instead of the coefficient of variation so that in 95% of cases the measured mean value will lie within $\pm 2S/\sqrt{n}$ of the population mean, where S is the standard deviation.

1.4.3 Use of computers

The incorporation of computers and microprocessors has brought great changes to the instrumentation used for testing textiles. Their use falls into two main categories: recording and calculation of results and automation of the test procedure. Both of these uses may be found in the most advanced instruments.

Recording of results

In these applications the computer is usually connected via an analogue to digital converter to an existing instrument from where it collects the data that would previously have been written down on paper by the operator. The advantages of such an installation are as follows:

1 In the case of a graphical output the whole of the curve is recorded numerically so that results such as maxima, areas under the curve and slopes can be calculated directly without having to be read from a graph. This allows a more consistent measurement of features such as slopes which would previously have been measured by placing a rule on the graph by eye. However, it is important in such applications to be clear what criteria the computer is using to select turning points in the curve and at what point the slope is being measured. It is useful to have visual checks on these points in case the computer is making the wrong choice.
2 The ability to adjust the zero level for the instrument automatically. This can be done, for instance, by taking the quiescent output as being the zero level and subtracting this from all other readings.
3 The ability to perform all the intermediate calculations together with any statistical calculations in the case of multiple tests.
4 The ability to give a final neatly printed report which may be given directly to a customer.

It is important, however, to be aware of the fact that the precision of the basic instrument is unchanged and it depends on, among other things, the preparation and loading of the sample into the instrument by the operator and the setting of any instrumental parameters such as speed or range.

Automation of the test procedure

In such applications use is made of electronic processing power to control various aspects of the test rather than just to record the results. This means that steps such as setting ranges, speeds, tensions and zeroing the instrument can all be carried out without the intervention of an operator. The settings are usually derived from sample data entered at the keyboard. In the case of yarn-testing instruments the automation can be carried as far as loading the specimen. This enables the machinery to be presented with a number of yarn packages and left to carry out the required number of tests on each package.

The automation of steps in the test procedure enables an improvement to be made in the repeatability of test results owing to the reduction in operator intervention and a closer standardisation of the test conditions. The precision of the instrument is then dependent on the quality of the sensors and the correctness of the sample data given to the machine. The accuracy of the results is, however, still dependent on the calibration of the instrument. This is a point that is easily overlooked in instruments with digital outputs as the numbers have lost their immediate connection with the physical world. If the machine fails in some way but is still giving a numerical output, the figures may still be accepted as being correct.

To be generally acceptable automated instruments have to be able to carry out the test to the appropriate standard or have to be able to demonstrate identical results to those that have been obtained with the standard test method.

It is still possible even with advanced automation for results to be incorrect for such simple reasons as wrong identification of samples or failure to condition samples in the correct testing atmosphere.

1.4.4 Types of error

Errors fall into two types.

Bias or systematic error

With this type of error the measurements are consistently higher or lower than they should be. For instance if a balance is not zeroed before use then all readings taken from it will have the same small amount added to or subtracted from them. This type of error cannot be detected by any statistical examination of the readings. Systematic errors can only be eliminated by careful design of the tests, proper calibration and correct operation of the instruments.

Precision or random error

This type of error is present when repeated measurements of the same quantity give rise to differing values which are scattered at random around some central value. In such cases the error can be estimated by statistical methods.

1.4.5 Sources of error

Errors of both types can arise from a number of causes:

1 **Instrument reproducibility**: even when an instrument is correctly calibrated, mechanical defects can influence the readings unless they are taken in exactly the same fashion as the calibration values. Mechanical defects such as slackness, friction and backlash can cause measurements to vary. These effects can depend on the direction that the mechanism is moving so that the error may be different when the reading is increasing from that when it is decreasing. Electrical and electronic instruments can suffer from drift of settings over a period of time owing to an increase in temperature of components.
2 **Operator skill**: a great many tests are based on personal manipulation of the apparatus and visual reading of a resultant indication. An op-

erator may be called on to prepare a sample, load it into the instrument, adjust readings such as zero and maximum, and to take a reading from a scale. Each manipulation, adjustment and reading involves an uncertainty which can depend on the skill and the conscientiousness of the individual operator. The ideal in instrument and test method design is to reduce the amount and scope of operator intervention.

3 **Fineness of scale division**: a fundamental limit is set to the precision of a measurement by the instrument scale which is necessarily subdivided at finite intervals. It carries with it an immediate implication of a minimum uncertainty of one half of the finest scale division. In the case of a digital scale the last digit of the display sets the limit to the precision in a similar manner as it has by its nature to be a whole digit. The final digit implies that it is plus or minus half of what would be the next digit. However, digital scales usually read to more figures than the equivalent analogue scale.

4 **External factors**: these may come from sources outside the actual instrument such as line voltage fluctuations, vibration of instrument supports, air currents, ambient temperature and humidity fluctuations and such diverse factors as variation in the sunlight intensity through windows.

The above uncertainties in the measurement of textile properties derive from the measurement process. In addition to these uncertainties, textile materials also exhibit variation in properties throughout their bulk. These can be quite considerable in magnitude, particularly in the case of yarns and fibres. This variability, in a similar manner to the errors described above, falls into two types: systematic, as is the case when the properties of a fabric vary from the edge to the centre, and random, when the variability has no pattern. The effect of this is to add to the errors from the measurement itself to give a larger overall error from which it is difficult to separate out the variability of the material from the experimental error. Therefore, because of all the above sources of variation, the appropriate statistical analysis of results has a great importance in textile testing.

1.4.6 Repeatability and reproducibility

The true accuracy of a test method can only be gauged by repeated testing of identical material both within the same laboratory and between different testing laboratories that possess the same type of equipment. International round trials [14–17] are organised by sending out sets of test samples, all produced from the same batches of material, to participating laboratories and asking them to test the samples in a prescribed manner. The results are then correlated and the within (repeatability) and between (reproducibility) laboratory variations calculated. The variation between laboratories is always greater than the variation found within a single laboratory.

BS 5532 [18] defines repeatability and reproducibility as follows.

Repeatability

1 *Qualitatively*: the closeness of agreement between successive results obtained with the same method on identical test material, under the same conditions (same operator, same apparatus, same laboratory and short intervals of time).
2 *Quantitatively*: the value below which the absolute difference between two single test results obtained in the above conditions may be expected to lie with a specified probability. In the absence of other indication, the probability is 95%.

Reproducibility

1 *Qualitatively*: the closeness of agreement between individual results obtained with the same method on identical test material but under different conditions (different operators, different apparatus, different laboratories and/or different times).
2 *Quantitatively*: the value below which the absolute difference between two single test results on identical material obtained by operators in different laboratories, using the standardised test method may be expected to lie with a specified probability. In the absence of other indication, the probability is 95%.

Errors involved

In order to understand the difference between repeatability and reproducibility the error in the test result can be considered to be due to two components [14]:

1 A random error (standard deviation σ_r) which occurs even when the same operator is using the same apparatus in the same laboratory. The variance of this σ_r^2 is called the within-laboratory variance and is assumed to have the same value for all laboratories.
2 An error (standard deviation σ_L) due to the difference that occurs when another operator carries out the test in a different laboratory using a different piece of identical apparatus. The variance of this σ_L^2 is called the between-laboratory variance.

The total error in a result that combines several sources of error can be obtained by adding together their variances. The numerical values for repeatability and the reproducibility are then given by substituting a value of 2 for n in the above equation for confidence limits:

$$\text{Repeatability} = \frac{2}{\sqrt{2}}\sigma_r$$

$$\text{Reproducibility} = \frac{2}{\sqrt{2}}(\sigma_L^2 + \sigma_r^2)^{1/2}$$

1.4.7 Significant figures

The numerical expression of the magnitude of a measurement may contain some figures that are doubtful. This can arise either from an estimation between the scale divisions by the operator or, in the case of a digital readout, from the uncertainty in the choice of the last figure by the machine. For instance in the case of a measurement of length by a rule that is graduated in millimetres, the rule might show that the length is definitely between 221 mm and 222 mm. Estimation by the person making the measurement might put the value at 221.6 mm. The figures 221 are exact but the final digit (6) is doubtful because it is only estimated. However, all four figures are regarded as significant because they convey meaningful information. This can be seen if it is imagined that the true value is actually 221.7 mm; the error would then be 0.1 mm but if the figures had been taken to the nearest whole millimetre (222) the error would have been 0.3 mm.

Significant figures, therefore, include all the exact figures followed by one doubtful one. Any zeros before the figures are not included in the number of significant figures and zeros after the figures are included only if they are considered to be exact regardless of the position of the decimal point. Zeros that are only there to position the decimal point are not regarded as significant; for example, 540,000 has only two significant figures. If it is necessary to express the fact that some of the zeros are significant, it is better to write the number as, for instance, 5.40×10^5. Zeros after a decimal point should be included only if they are significant. For instance the value 3.0 has two significant figures.

Unless otherwise indicated the uncertainty in any written measurement of a continuous variable is taken to be plus or minus half a step of the last significant figure. For instance 25.4 mm is taken to mean 25.4 ± 0.05 mm but 25.40 mm would be taken to mean 25.40 ± 0.005 mm.

The number of significant figures written down only concerns the reading of figures from instruments. It is an entirely separate issue from how many of the figures are meaningful which can only be decided from repeat tests as described above.

Rounding off

When further calculations are carried out on measured values the number of figures in a result may increase but in general the number that are significant does not increase. Retaining these figures in the final result gives a misleading impression of the precision of the result. The discarding of any figures beyond the significant ones is known as rounding off.

The convention for rounding off is that the last figure to remain is left unchanged if the amount to be discarded is less than 0.5, but it is increased by one if the amount to be discarded is greater than 0.5. For example: 6.854 would be rounded to 6.85 to three significant figures or 6.9 to two significant figures. Note that the rounding up or down is done only in one stage, not firstly to 6.85 and then to 6.8.

If the amount to be discarded is exactly 0.5 of a step then the rounding is to the nearest even figure in the last place, the idea being that this gives a random choice with as many results being rounded up as are rounded down. For example 6.85 would be rounded down to 6.8 whereas 6.95 would be rounded up to 7.0.

When carrying out calculations involving results with different numbers of significant figures, the number of figures in the result is governed by the contribution with the largest error. For example in addition or subtraction:

$$2.71 + 11.814 = 14.52$$
$$6.4 + 123.625 + 5.7165 = 135.7$$
$$2000 + 2{,}400{,}000 = 2{,}400{,}000$$

In each case the result is governed by the number whose concluding figure is the furthest to the left. In the case of simple multiplication or division the result should not in general be credited with more significant figures than appear in the term with the smallest number of significant figures. For example:

$$63.26 \times 0.0217 = 1.37$$
$$0.356 \times 0.6149 = 0.219$$

In case of doubt the mathematical operations can be carried out on the results for the implicit range of values.

Any rounding off must be carried out only on the final result after all the calculations have been made.

General reading

The WIRA Textile Data Book, WIRA, Leeds, 1973.

Morton W E and Hearle J W S, *Physical Properties of Textile Fibres*, 3rd edn., Textile Institute, Manchester, 1993.

Massey B S, *Measures in Science and Engineering*, Ellis Horwood, Chichester, 1986.

References

1. BS 2545 Methods of fibre sampling for testing.
2. IWTO-26-74 Glossary of terms relating to sampling.
3. BS 4784 Determination of commercial mass of consignments of textiles Part 2 Methods of obtaining laboratory samples.
4. ASTM D 2258 Sampling yarn for testing.
5. Lunney H W M, 'The pressure coring of wool bales', *Wool Sci Rev*, 1978 **55** 2.
6. ASTM D 1060 Core sampling of raw wool in packages for determination of percentage of clean wool fiber present.
7. Anon, 'The measurement of wool fibre length', *Wool Sci Rev*, 1952 **9** 15.
8. BS 2010 Method for determination of the linear density of yarns from packages.
9. ASTME, *Handbook of Industrial Metrology*, Prentice Hall, New Jersey, 1967.
10. Leaf, G A V, *Practical Statistics for the Textile Industry: Part 1*, Textile Institute, Manchester, 1984.
11. Leaf, G A V, *Practical statistics for the textile industry: Part 2*, Textile Institute, Manchester, 1987.
12. Murphy T, Norris K P and Tippett L H C, *Statistical Methods for Textile Technologists*, Textile Institute, Manchester, 1960.
13. ASTM D 2905 Statement on number of specimens for textiles.
14. Ly N G and Denby E F, 'A CSIRO Inter-laboratory trial of the KES-F for measuring fabric properties', *J Text Inst*, 1988 **79** 198.
15. Clulow E E, 'Comparison of methods of measuring wettability of fabrics and inter-laboratory trials on the method chosen', *J Text Inst*, 1969 **60** 14–28.
16. Lord J, 'Draft method for determination of dimensional changes during washing: interlaboratory trials', *Tex Ins Ind*, 1972 **10** 82.
17. Smith P J, 'Proposed new ISO domestic washing and drying procedures for textiles a background paper', *Tex Ins Ind*, 1977 **15** 329.
18. BS 5532 Part 1 Glossary of terms relating to probability and general terms relating to statistics.