SAWTRI TECHNICAL REPORT

No. 361

Skewness and Kurtosis as Further Measures of Wool Fibre Distribution

by
E. Gee

SOUTH AFRICAN
WOOL AND TEXTILE RESEARCH
INSTITUTE OF THE CSIR

P. O. BOX 1124
PORT ELIZABETH
REPUBLIC OF SOUTH AFRICA

ISBN 0 7988 1077 7

SKEWNESS AND KURTOSIS AS FURTHER MEASURES OF WOOL FIBRE DISTRIBUTION

by E. GEE

ABSTRACT

The mean and standard deviation are useful descriptions of fibre diameter and length distributions but are not necessarily of high accuracy. The use of the higher moments of the observations to give a measure of skewness and kurtosis are described and their use to fit Pearson curves to the distributions are indicated. From these curves better descriptions of the peak of the distribution and of more extreme portions (tails) can be made.

INTRODUCTION

A widely used technique for characterising and comparing distributions (of any factor) is to assume normality and to use the calculated mean and standard deviation (σ). In many instances, especially when the CV is low, departure from normality hardly invalidates their use as adequate descriptions of the distributions. Fibre length and fibre diameter, however, show widely spread distributions. For example coefficients of variation (CV) of up to 50 *per cent* can be found. The extreme members of these populations can significantly affect the processing of the fibres, hence the need to measure the upper 5 *per cent* of length distributions[1], and so the mean and σ may not give an adequate description. By obtaining a measure of the length of the tails (skewness) and of the peakiness or flat-tops (kurtosis) a better description and interpretation may result.

Many workers have considered the distribution of diameter and of fibre length[2,3]. Some have considered the log form of the distribution, especially for diameter. It is reasonable to expect that this will provide a better fit for the positively skewed diameter distributions by effectively compressing the higher members. The technique, however, is not effective for negatively skewed distributions. Further, although logs may be relatively easy to manipulate they are not as easily assessed subjectively as are the absolute values of diameter and length.

Wegener and co workers considered a different function describe the distribution of fibre length which involved the sum of three exponentials each raised to different function of length. The attraction of this technique is that bimodal distributions can be represented.

Characterisation of distributions by use of the higher moments is not a new technique. The failure of the normal curve to adequately describe certain distributions found by actuaries led to the work of K. Pearson[4-6] and others.

Pearson was able to describe many different distributions by using a series of curves derived from one concept. Interestingly, the normal curve forms a special case in the series. These curves have become known as Pearson curves. E.S. Pearson[7, 8] and co workers used these to describe the warp strength distribution of a duck cloth.

Although the calculations are lengthy the continuing phenomenal growth of computing power enables the application of these curve fitting techniques to be made relatively easily. The calculations illustrated by Elderton[9] if done on a mechanical calculator are prohibitive. The advent of mini-programmable calculators has reduced the work to a mere entering of the observations, all the required answers appearing automatically.

THEORY

Distributions

A set of objects such as the yields of greasy wool samples, the strengths of yarn samples, the lengths of the individual fibres in a sample are never consitent in that every member has exactly the same value. When distributions are measured the variation is often a reflection of the accuracy of the measurement and can be adequately described by Gaussian or normal probability statistics, i.e. mean and standard deviation. In the case of the lengths or the diameters of wool fibres in a sample, be it is testing sample or a consignment for production, the fibres in the sample do have widely different values — they exist not as artifacts of measurement but as real different objects. The amount and nature of their variation and the influence on the processing performance are most important. For instance a sample in which all the fibres were 60 mm in length could process differently from another sample which had equal proportions of 20 mm, 60 mm and 100 mm fibres and hence the same 60 mm average.

The pattern of variation of length can be found by taking a representative sample of fibres, (about 1 000 fibres) and measuring the length of each fibre. The practical test is usually arranged so that each measurement is accurate to 5 mm i.e. a fibre which is longer than 52,5 mm but shorter than 57,5 mm is counted as a 55 mm long fibre. On completion, the number of fibres in each length category is counted and the results are arranged in a table, e.g.

5 fibres at 5 mm .....180 fibres at 50 mm ..... 2 fibres at 90 mm . Table I shows a typical example where 1602 fibres were measured into categories from 5 mm to 120 mm .

For ease of subsequent calulation these length categories have been divided by five and labelled 1 to 24.

By brief inspection, the information conveyed by this table is that this particular sample contained fibres as short as 5 mm and as long as 120 mm. The majority were in the region of 55 mm to 85 mm. More detailed inspection would be directed to the distribution in the 5 to 55 mm region and the 90 to 120 mm region. This picture is revealed more easily by constructing the histogram which is illustrated by the stepped pattern in Fig. 1. Examination and comparisons of

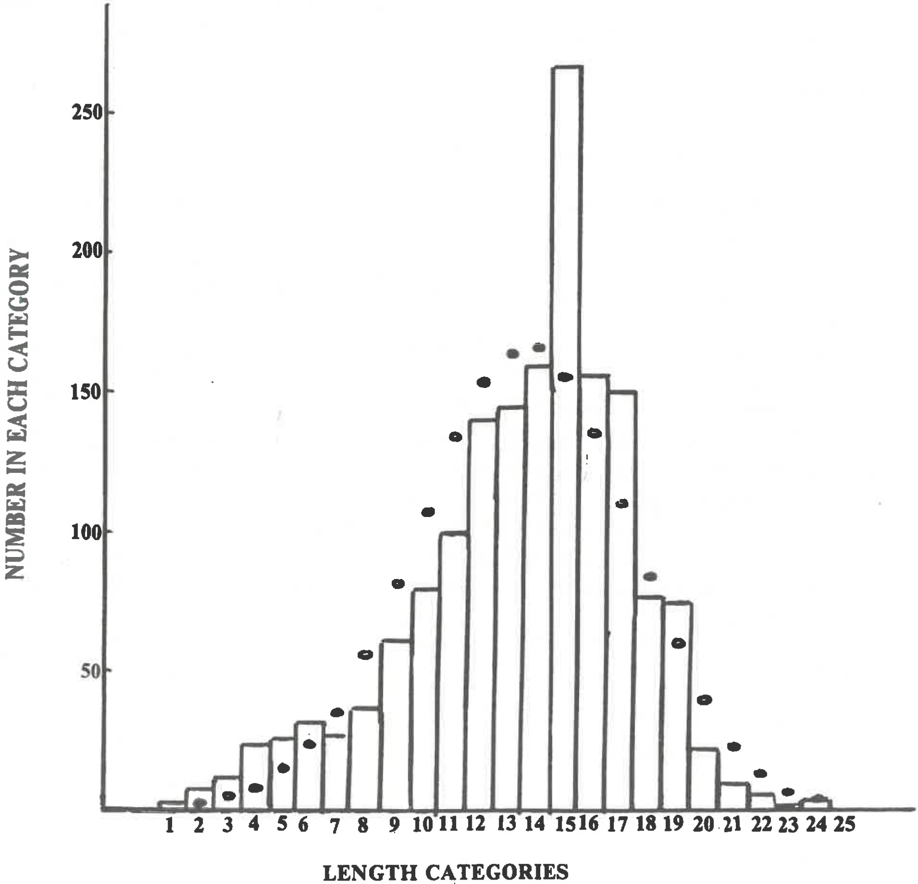

Fig. 1 Histogram and derived normal curve

histograms are instructive but only allow qualitative descriptions to be made. If the histogram can be represented by a few simple numbers then quantitative descriptions and calculated comparisons can be made.

To this end it is often assumed that the distribution obtained can be adequately represented by a normal or Gaussian distribution. Such a curve derived from the data is illustrated by the circles in Fig. 1. That the circles do not follow the histogram exactly can be due to either the inadequate sampling of the consignment or the normal curve not being really a faithful representation of the distribution. If the normal curve was sufficiently representative of the distribution then the known properties of such a curve could be used with advantage, and the well-understood properties of probability statistics could be applied.

Normal or Gaussian Distribution Curve

If the millions of fibres in one consignment did conform to a normal distribution then:

1. the average or mean (μ) would be the best representative value for the fibres.
2. the lengths of 68,2 *per cent* of the fibres would be within plus or minus one standard deviation (σ) of the mean and 95,4 *per cent* would be within plus or minus two standard deviations of the mean.

If the whole consignment or population is not measured, and whole populations rarely are, the procedure to calculate the mean and standard deviation can still be followed. The values obtained (now labelled $\bar{x}$ and s) will then only be the best estimates of the true values.

If the distribution was not normal then, although the same calculations can be made, the value obtained for the standard deviation will not give the 68,2 *per cent* range.

The shape of the normal curve and the properties of the normal equation are such that:

$$\sigma^2 = \frac{\Sigma(x - \bar{x})^2}{n - 1} \quad \text{.......(1)}$$

where x is an observation

$\bar{x}$ is the mean or average

n is the number of observations

and Σ means the sum of all such terms.

The practical calculation uses a different form of this expression, namely:

$$s^2 = \frac{\Sigma x^2 - (\Sigma x)^2/n}{n - 1} \quad \text{.......(2)}$$

$$= \frac{\Sigma x^2}{n} - \bar{x}^2$$

and can be described in words as 'the average of the squares less the square of the average'.

The characteristic $\bar{x}$ and s for the data in Table I are obtained from columns 2 and 3. Columns 4 and 5 are calculated and the totals of columns 3, 4 and 5 are the values for n, Σx and Σx^2 which are required in equation (2) to calculate s. The mean $\bar{x}$ is given by $\Sigma x/n$.

Thus

$$\bar{x} = 21736/1602 = 13{,}568$$

$$s^2 = \frac{317882 - \dfrac{(21736)^2}{1602}}{1601} = 14{,}345$$

This calculation procedure can be applied to any set of numbers, be they representative of a normal distribution, a rectangular distribution, zero distribution or any form of distribution. Consequently before $\bar{x}$ and s are used their adequacy or usefulness should be checked. The chi-squared test can be used here to check how representative of the distribution is this normal curve.

Chi-squared test

The test is made by calculating the chi-squared value from the data, and assessing its value. If it is too large when compared with standard table values then the hypothesis that the distribution is normal, is false. Chi-squared is given by the following equation:

$$\chi^2 = \Sigma \frac{(o - e)^2}{e}$$

where o = an observed value

e = the value expected from the hypothesis.

Before this can be used the expected values must be obtained by calculating the ordinate Z for each x value using $\bar{x}$ and s as the normal curve parameters. A table of the Normal probability function can be used for this purpose. These calculated Z values are scaled by use of the ratio (total number of observations) / (sum of the Z values) and are the required expected values. Column 8 of Table I lists these values while column 3 lists the observed values. The discrepancy between 1592 and 1602 is due to rounding errors. For the data in Table I chi-squared is calculated from columns 3 and 8 but omitting those values in column 3 which are less than 5 as these extremes can cause inaccuracies. Thus $[(8-2)^2 + (11-4)^2 + \text{-----}(5-14)^2]$ divided by 1580 gives a chi-squared value of 12,10. Of the 24 categories three have been omitted hence there are 20 degrees of freedom with which the significance of the chi-squared value can be checked. Statistical tables show that a value of 12,10 has about 90 *per cent* probability of arising by chance from a sample selected from a normal population. As it is usual to accept only a 5 *per cent* probability, i.e. to be able to make the statement that we are 95 *per cent* sure that the distribution is not normal, then we must conclude that this data can be represented by a normal distribution.

Pearson and Hartley[10] comment that this test, while it has certain advantages, may prove somewhat insensitive to real departures from normality, partly because the test only takes into account the magnitude of the differences between observed and expected frequencies and not their sign and arrangements

A more instructive procedure is to calculate the third and fourth moments from the data and hence estimate the significance of the skewness and kurtosis values.

Skew Curves

Although population distributions which are normal or Gaussian predominate and are amenable to manipulation permitting sound judgement of effects to be made, some related shapes can give useful descriptions. One such shape is the skew-curve. Where the normal curve is a precisely symmetrical bell-shaped curve which is dependent only on scale, a skew-curve is similar but not symmetrical. It has an extension to one side or the other. It has too many high (or low) values, and is said to have a tail (positive or negative). Such skewness could arise or could change during the conversion of raw wool into tops. Breakage of fibres during processing could result in a few unbroken long fibres giving a high tail or positive skewness in the distribution or short broken fibres could produce a distribution having a low tail, or negative skewness.

If both are present then a peaked curve could result. This curve is typified by the major portion of the distribution being close to the mean and the minor portion extended well out on either side, i.e. shape peak and long tails. This curve has significant kurtosis and is named a leptokurtic distribution. The reverse of this, namely a flatter top with sharp shoulders and shorter tails is called a platykurtic distribution.

A measure of skewness is obtained from the second and third moments while kurtosis is assessed by using the second and fourth moments.

Moments of Observations

For any set of data which represent a population with individual values x_1 x_2 etc up to x_n, there being n data values, the moments are defined as follows

1st moment, $m_1 = \frac{1}{n} \cdot \Sigma_1^n x_i \equiv \tilde{x}$ or mean value

2nd moment, $m_2 = \frac{1}{n} \cdot \Sigma_1^n (x_i - \tilde{x})^2 \equiv \sigma^2$ or variance

3rd moment, $m_3 = \frac{1}{n} \cdot \Sigma_1^n (x_i - \tilde{x})^3$

4th moment, $m_4 = \frac{1}{n} \cdot \Sigma_1^n (x_i - \tilde{x})^4$

Moment ratios of interest are:

Skewness: $\sqrt{\beta_1} = [m_3^2/m_2^3]^{1/2}$ (Note there are other definitions of skewness)

Kurtosis: $\beta_2 = m_4/m_2^2$

The departure of $\sqrt{\beta_1}$ values from the 'normal' value of zero is an indication of skewness in the distribution, while departure of β_2 values from the 'normal' value of 3 is an indication of kurtosis, values less than 3 reflecting platykurtosis.

The practical forms of the moments are:

$$m_3 = \frac{1}{n} \cdot \Sigma x^3 - \frac{3}{n} \cdot \tilde{x} \cdot \Sigma x^2 + 2(\tilde{x})^3$$

$$m_4 = \frac{1}{n} \cdot \Sigma x^4 - \frac{4}{n} \cdot \tilde{x} \cdot \Sigma x^3 + \frac{6}{n} \cdot (\tilde{x})^2 \cdot \Sigma x^2 - 3(\tilde{x})^4$$

The data in Table I, using the values on the last line of the table for the appropriate terms in these formulae give

$m_3 = -35,6006$ $m_4 = 704,00$

By using the previously calculated value for s^2 (or m_2) we obtain

$\sqrt{\beta_1} = -0,656$ $\beta_2 = 3,425$

Statistical tables[11] indicate that for 1 600 observations the distribution is skewed (95 *per cent* level of significance) if $\sqrt{\beta_1} > 0,100$ and shows a significant degree of kurtosis if $2,81 > \beta_2 > 3,21$.

It would be concluded that the distribution of the data of Table I shows significant negative skewness and is leptokurtic, i.e. it has an excess of short fibres and is peaked. One implication of peakiness is that, ignoring the tails, the majority of the fibres occupy a narrower range of lengths than would be indicated by the calculated standard deviation. The reverse holds for a platykurtic distribution. The influence of skewness on the practical use of $\tilde{x}$ and s is that the majority of fibres will occupy a somewhat narrower range than indicated by the standard deviation and that the middle of this limited range (the Mode or most frequently occurring value) will have a slightly lower value of x than $\tilde{x}$ for a positive skewness and slightly higher for negative skewness.

Calculation of position of Mode

The Pearson system of frequency curves ably demonstrated by Elderton and Johnson[9] has been used to characterise distributions. For curves of type I, IV and VI which cover the general range of possibilities, the position of the mode is given by

$$\text{Mean - Mode} = \frac{1}{2} \cdot \frac{m_3}{m_2} \cdot \frac{r + 2}{r - 2}$$

where m_2 and m_3 are the second and third moments about the mean after Sheppard's corrections have been applied for the grouped data

and $r = 6 (\beta_2 - \beta_1 - 1) / (6 + 3 \beta_1 - 2 \beta_2)$

The mode is above the mean for a negatively skewed distribution and below for positive.

Applying these formula to the data of Table I a value for r of 26,0 was obtained which positioned the mode at 1,456 x 5 or 7,28 mm above the mean.

Thus the mean fibre length of this data is 67,8 mm while the mode is at 75,1 mm.

Normal Curve Superimposed on the histogram

As mentioned earlier the histogram in Fig 1 illustrates the distribution. The distributions shown in Fig 2a and b have been produced by a computer and show the length distribution of the fibres in the grease and top states, respectively. The abscissa units are 5 mm . These histograms areshown by x while the circles delineate the derived assumed normal distribution adjusted to scale for the size of the measured population. The position of these circles was calculated as follows.

From the mean value of the data, $\tilde{x}$, and its standard deviation s, the ordinate (Z) or number in the group x was calculated using:

$$Z = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x - \tilde{x}}{s}\right)^2}$$

Alternatively, values of Z can be found from Tables of Z(x) in terms of the standardised normal deviate x, for each length group considered.

The sum of all these Z values and its ratio to the number of fibres measured was calculated. This is the ratio by which every calculated Z value must be multiplied in order to correctly scale the displayed normal distribution.

PRACTICAL APPLICATION

Skewness and Kurtosis of various wool samples

Values for skewness and kurtosis of the distributions of diameters and lengths have been obtained from data which have been collected for various projects. These are used here to illustrate their value as descriptive terms.

Data Set I contained five lots of wool with the mean diameters ranging systematically from about 20 μm to about 25 μm and which were used in a project to study the processing properties of blends[12]. Various 50/50 blends by mass were prepared. Their mean fibre lengths were fairly constant at 90 mm to 100 mm . Tables II and III show the properties of the distributions of diameter and of length of fibres for the greasy staple form. Significant skewness can be assumed when the value of β_1 exceeds about 0,1. Kurtosis is not present when β_2 lies between about 2,8 and 3,2.

For the diameter distribution each main component and each blend thus shows positive skewness and is high peaked or leptokurtic. The implication is that the most commonly occurring diameter (or mode) will be a little lower than

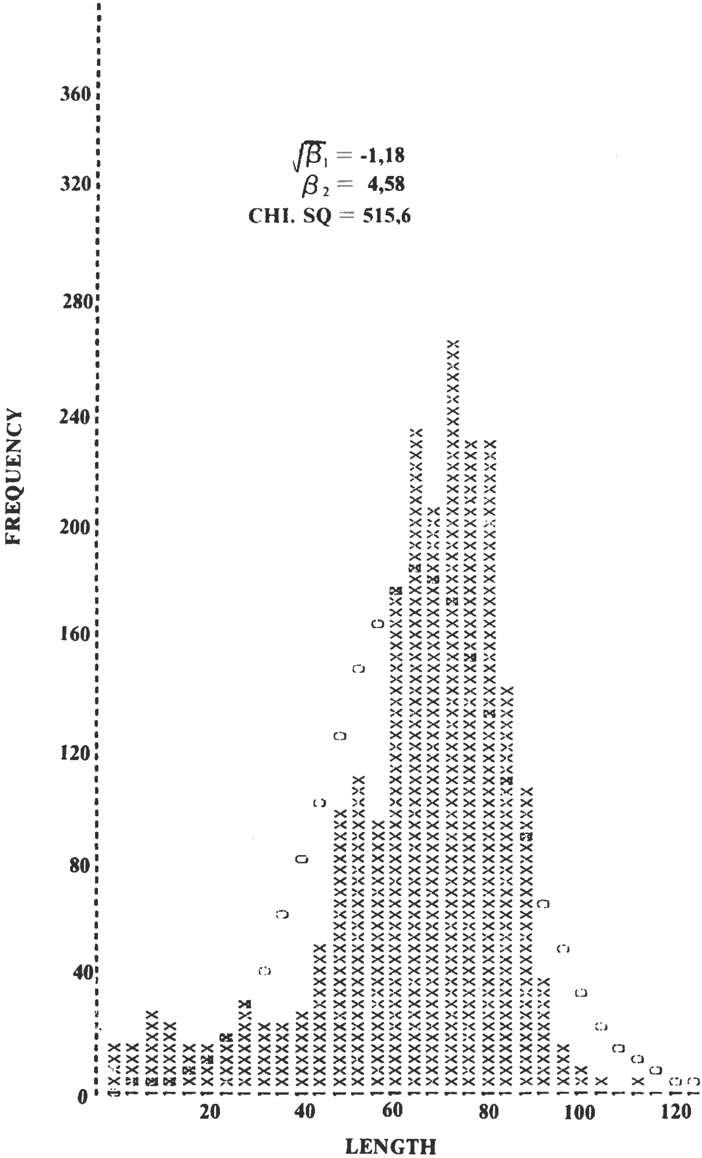

Fig. 2a Fibre length

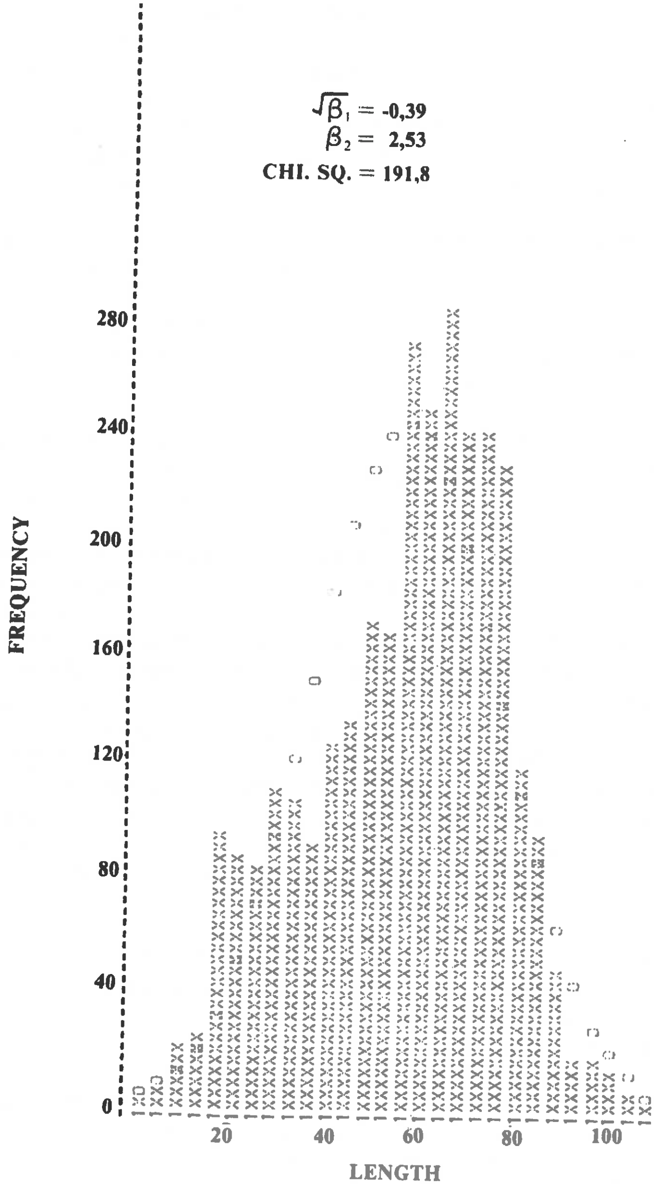

Fig. 2b Fibre length

the calculated mean and that the bulk of the distribution is spread over a narrower range than that indicated by the standard deviation or CV.

The distributions of length show that all these samples and blends have negative skewness or tails and have peaked distributions, i.e. the most commonly occurring lengths or mode will be longer than the calculated mean.

The leptokurtic distribution of fibre length in the greasy staple sample suggests that perhaps all the fibres in the sample were the same length reflecting a constant rate of growth over the skirted fleece and that the observed distribution of lengths is a consequence of several other factors. Among those could be variation in how close to the skin the wool was shorn, variable loss of tip by natural weathering, and fibre breakage during sample preparation.

If these factors are relevant then caution must be exercised when interpreting data by assuming normal distribution. This is not to deny the usefulness of assumed normal distributions; it is more of a refinement.

Data set II contained six wools of nominally constant diameter and of varying lengths (50 mm to 100 mm) together with various blends[1]. Tables IV and V show the diameter distribution in the grease or staple state and in the top state, while the length distributions for the greasy wool, gilled sliver and top respectively are given in Tables VI - VIII.

Again the diameter results show positive tails and peaked distributions. The mean values for each set of data are given in Table IX and shows that processing to the top has not affected the diameter distribution but has affected the length distributions. The fibre lengths of the staples show mainly negative tails, and when processed into tops these tails have been reduced and occasionally show positive tails. Similarly while most of the "staple" lengths show peaked distributions, their corresponding tops are less peaked and often change to flat topped distributions. The mean values of Table IX show this very clearly. Figs 2a and b illustrate a change in fibre length distribution.

These figures also show that the mean fibre length, which is the peak of the illustrated normal curve does not coincide with the peak of the measured distribution. For the negatively skewed distributions the mean fibre lengths is lower. This suggests that a mean fibre length is not the best representation of this distribution.

If a distribution of fibre lengths is Gaussian or Normal then the most commonly occurring length would be represented by the mean fibre length. If the opposing technological effects of the longer fibres and the shorter fibres are balanced then the mean fibre length would be a useful characteristic of the distribution. If, however, the distribution is skew then it is no longer the best characteristic feature. The mode, the most frequently occurring length could be more useful, especially if the effects of the tails are negligible or if a few long fibres can carry a greater number of short fibres (lengths usually have a negatively skewed distribution). If this latter proviso does not hold then the median or middle length value should be used. The median value is the length of the middle fibre when all are arranged in ascending order of length.

The position of the mode has been calculated for the above sets of data. For data set II, fibre length, the average position of the mode was 4,3 mm, 2,4 mm and 1.4 mm longer than the mean for the greasy staples, gilled sliver and tops respectively. Data set I, diameters in greasy staple form, showed the mode to be 1,5 μm below the mean while for data set II, also diameter, the mode was below the mean by a similar amount.

To illustrate the position of the mode by use of Pearson's curves two samples have been selected. Figs. 3a and b show the histogram, the derived normal curve (from calculated mean value and standard deviation) and the best fit Pearson curve of type I and IV. The samples, from data set two, were PP 67 and 66 respectively for the diameter distribution in the greasy state.

The distribution of area under the curve expressed as the percentage number of fibres shorter than any given diameter (or length) has been calculated from the curves of Figs. 3a and b. These are displayed in Fig 4 where the cummulative areas are plotted.

The area under the normal curve is divided 50:50 about the mean, that is there are an equal number of fibres shorter than the mean as there are longer. For the Pearson curves the division about the mode is shown to be about 34:66 and 41:59 for the two examples given.

Characteristic data for the curves of fig 3a and b are:

Normal: $$y = 199{,}5\ e \uparrow \left(-\tfrac{1}{2} \cdot \left[\frac{x - 10.695}{2.595} \right]^2 \right)$$

Pearson Type I:

$$y = 226{,}04 \left[1 + \frac{x}{3.169} \right]^{1,966} \cdot \left[1 - \frac{x}{82.055} \right]^{50,889}$$

for the origin at the mode and the Mode is 1,438 units below the mean.

Normal: $$y = 209{,}0\ e \uparrow \left(-\tfrac{1}{2} \cdot \left[\frac{x - 11{,}279}{2{,}519} \right]^2 \right)$$

Pearson Type IV:

$$y = 0{,}000818 \left[1 + \left(\frac{x}{6{,}610} + 1{,}290 \right)^2 \right]^{-10,809} \cdot e^{25,310 \tan^{-1} \left(\frac{x}{6{,}610} + 1{,}290 \right)}$$

for the origin at the mean and the mode is 0,789 units below the mean.
one unit of diameter = 2 μm

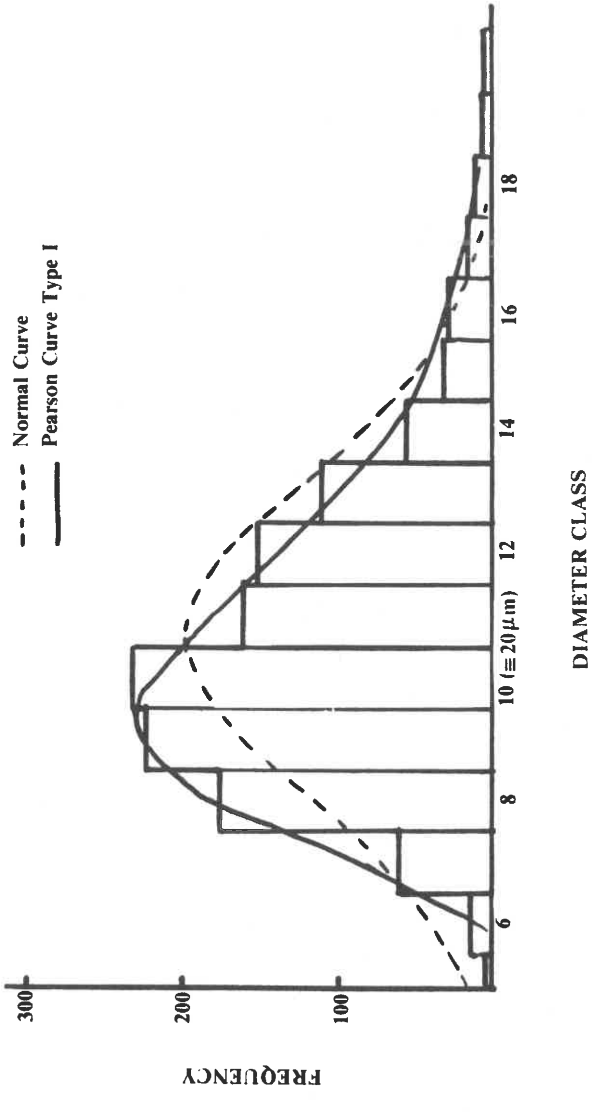

Fig. 3a Diameter distribution of sample PP 67 (grease wool)

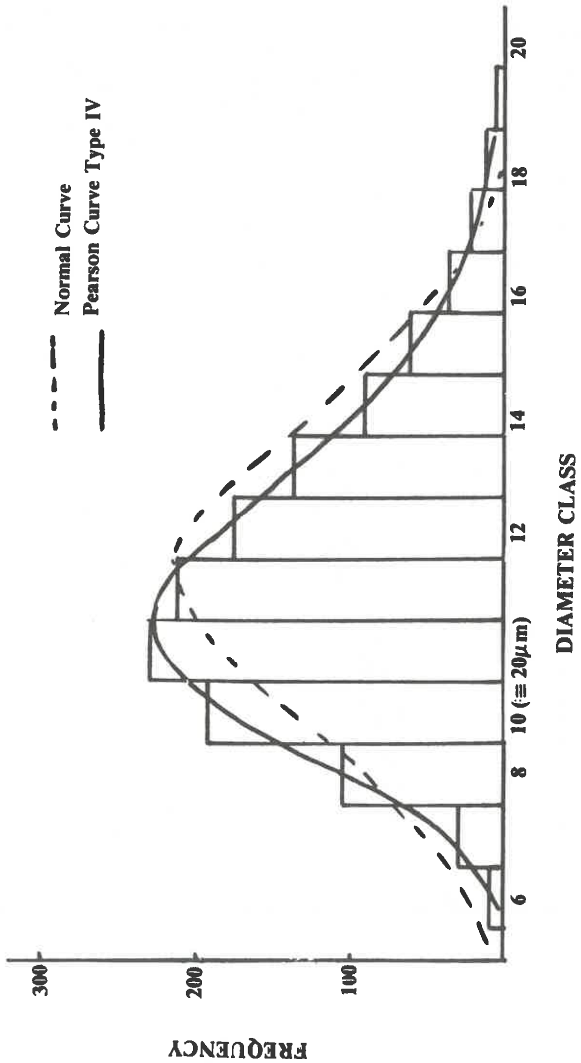

Fig. 3b Diameter distribution for sample PP 66 (grease wool)

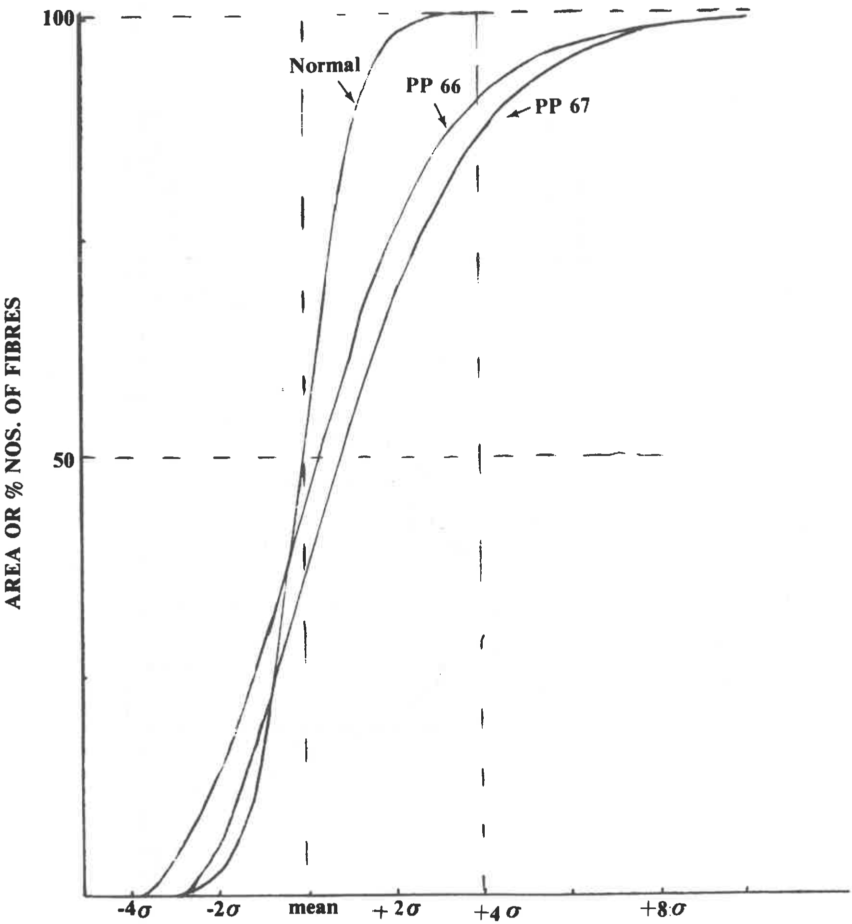

Fig. 4 Cumulative distributions for figures 3a and b

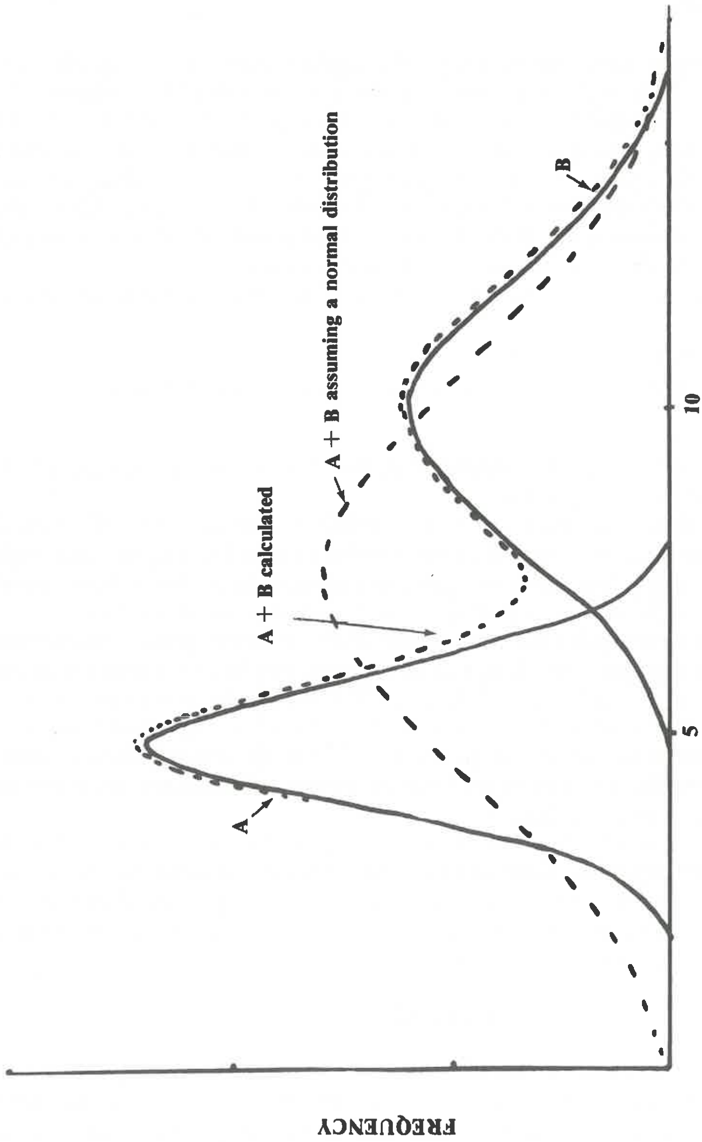

Fig. 5 Effect of a blend of two normal populations

Population blends

These results direct attention to the consequence of mixing different populations. Fig. 5 illustrates a possible situation where a population having a μ and σ of 10 and 2 is added in equal parts by numbers to one of 5 and 1. The resulting population is very different from a normal distribution. It is doubled peaked (bi-modal) and shows a minimum point where the assumed normal distribution shows a maximum. The calculated standard deviation is 2,96 about a mean of 7,5. Consideration of the 95 *per cent* range shows the inadequacy of the standard deviation as a measure of the distribution.

Thus the range within which 95 *per cent* of the fibres would be found are:

Actual range: 3,4 to 11,3
Implied range: 1,6 to 13,4 (from calculated standard deviation)

The actual 95 *per cent* range is 7,9 units while the 'normal' calculation gives 11,8 units which is 50 *per cent* greater.

The calculated skewness of 0,43 implies a positive tail reflecting a shortage of members at the low end of the distribution and an excess at the high end. The significantly low value of 2,07 for kurtosis, implying a flat-topped distribution, is a reflection of the double-peak and sharper shoulders.

Table X was compiled to illustrate the effect of merging two populations in equal parts by number, one chosen to have μ and σ of 10 and 2 with the other having in turn values of μ from 5 to 10 and CV from 20 *per cent* to 40 *per cent*. This shows that the variations of skewness of the blend is large, being negative (-0,43) for one extreme value and positive (0,34 for the other extreme value. Similarly the kurtosis shows changes from an extreme flat-topped distribution to a pronounced peaky distribution.

Tables XI and XII show the results for equal blends by mass for length distributions and diameter distributions respectively. Instead of using the number in each class interval the mass was used, i.e. for lengths the class interval contained "numbers in class times class length" and for diameters "the number in class times the square of the class diameter".

SUMMARY

The mean value and the standard deviation which characterise a normal distribution give useful measures of fibre length and of fibre diameter distributions. Calculation of the third and fourth moments of any distributions has shown that neither length nor diameter has a normal distribution, significant skewness (tails) and kurtosis (peaky or flat-topped varieties) are present. Diameter distributions tend to have positive tails and are peaked while lengths

tend to have negative tails and are also peaked.

Several blends of two components, in equal parts by mass were examined. For a range of different sample diameters little change in the pattern of skewness and kurtosis was observed while for a range of different sample lengths a progressive reduction of the negative tail and of kurtosis from a peaky state to a flat-topped state took place as the primary components of the blend were spaced further apart. By comparing the results from staples, gilled slivers and tops, the influence of processing was clearly shown. The general effect was a reduction of the negative tail (or increase in the positive tail) and a flattening of a peaked distribution.

The dependence of the standard deviation, skewness and kurtosis of an equal part (by number and by mass for length and for diameter) blend of two normal population whose relative means and deviations spanned a given range has been examined. As can be expected a near normal blend only results from the addition of two populations having the same mean. For different means and standard deviations the resultant population can be skewed in either direction and can be peaked or flat-topped.

The average modal length was 4,3 mm longer than the mean fibre length for one set of data while the modal diameter was 1,5 μm below the mean.

The 95 *per cent* range, obtained from the calculated standard deviation of a blend can differ significantly from the actual 95 *per cent* range, e.g. 8 mm to 67 mm compared with 17 mm to 57 mm respectively.

ACKNOWLEDGEMENTS

The author thanks the S.A. Wool Board for permission to publish this report, Dr D.W.F. Turpie for allowing the author to use some of his results and Mrs K. Belligan who helped with the computer programmes.

REFERENCES

1. Turpie, D.W.F. Processing and Characteristics of S.A. Wools Part XII. *SAWTRI Techn. Rep.* No. 354 (1977)
2. Palmer, R.C., Calculation of the Coefficient of Variance of fibre diameter in a batch of sorted wool by synthesis of variance. J. Text. Inst. **40**, T623 (1949).
3. Wegener W., Rosemann W., and Hoth E.G., Zeropoint equalization for the fibre length distribution in Combed Wool. Melliand Textilber. **46**, (8) 796 (1965).
4. Pearson, K., Phil Trans A **186** 343 (1895).
5. Pearson, K., Phil Trans A **197**, 445 (1901).
6. Pearson, K., Biometrika Vol 1 390 (1901).

7. Pearson, E.S. and Welch, B.L., Notes on some Statistical Problems raised in Mr. Bayes's Paper J. Royal Stats. Soc. Suppl. **4** 94 (1937).
8. Pearson, E.S. and Hartley H.O., Biometrika Tables for Statisticians Vol. 1 Cambridge Univ. Press (1956).
9. Elderton W.P. and Johnson N.L., Systems of Frequency Curves Cambridge Univ. Press (1969).
10. Ref. 8 p.61
11. Ref. 8 Tables 34 B and C.
12. Turpie, D.W.F., Processing and Characteristics of S.A. Wools, Part X *SAWTRI Techn. Rep.* No. 303 (1976).

TABLE I

LENGTHS OF SINGLE FIBRE (mm)

Length of Category		Nos of fibres n	ny	ny^2	ny^3	ny^4	Expected Value for n for normal curve
mm	y						
5	1	1	1	1	1	1	0
10	2	8	16	32	64	128	2
15	3	11	33	88	287	891	4
20	4	23	92	368	1 472	5 888	8
25	5	25	125	625	3 125	15 625	14
30	6	32	192	1 152	6 912	41 472	24
35	7	26	182	1 274	8 918	62 426	38
40	8	36	288	2 304	18 432	147 456	58
45	9	60	540	4 860	43 740	393 660	82
50	10	79	790	7 900	79 000	790 000	108
55	11	99	1 089	11 979	13 169	449 459	134
60	12	139	1 668	20 016	240 192	2 882 304	154
65	15	144	1 872	24 336	316 368	4 112 784	164
70	14	158	2 212	30 968	433 552	6 097 728	164
75	15	265	3 875	59 625	894 375	13 415 625	156
80	16	155	2 480	39 680	634 880	10 158 080	136
85	17	150	2 550	43 350	736 590	12 528 150	110
90	18	76	1 368	24 624	443 232	7 978 176	84
95	19	74	1 406	26 714	507 566	9 643 754	60
100	20	22	440	8 800	176 000	3 520 000	40
105	21	9	189	3 969	83 349	1 750 329	24
110	22	5	110	2 420	53 240	1 171 280	14
115	23	2	46	1 058	24 334	559 682	8
120	24	3	72	1 728	41 472	995 328	4
TOTALS		1 602 n	21 736 Σx	317 882 Σx^2	4 879 240 Σx^3	77 692 226 Σx^4	1 592

TABLE II
FIBRE DIAMETER OF GREASE WOOL — DATA SET I

Samples	Mean Diameter	Standard Deviation	Skewness	Kurtosis	No. of Fibres Measured
	μm	μm			
PP 49	20,0	4,9	0,94	4,72	1349
50	21,2	4,9	0,76	3,92	1469
51	22,5	5,1	0,80	5,14	1362
52	24,7	6,2	0,63	3,26	1326
53	25,4	6,2	0,53	3,36	1412
54	20,9	5,1	0,69	3,71	1306
55	22,1	5,0	0,92	4,51	1263
56	23,5	5,8	0,66	3,56	1332
57	24,8	6,2	0,79	3,90	1343
58	21,5	4,8	0,74	4,29	1272
59	22,8	5,7	0,77	3,75	1350
60	23,9	6,2	0,95	4,13	1337
61	22,0	5,7	0,72	3,77	1332
62	23,1	6,5	0,88	3,98	1373
63	22,8	6,5	0,82	4,43	1338

TABLE III
FIBRE LENGTHS OF GREASE WOOL — DATA SET I

Samples	Mean Lengths mm	Standard Deviation	Skewness	Kurtosis	No. of Fibres Measured
PP 49	103	3,8	-0,39	3,27	1356
50	104	3,7	-0,97	3,58	1332
51	92	2,1	-1,25	6,06	1684
52	91	2,7	-1,05	4,86	1317
53	89	2,6	-0,86	5,18	1249
54	97	31	-1,13	4,37	1377
55	87	30	-1,29	4,00	1328
56	92	27	-1,03	4,82	1560
57	95	28	-0,82	4,90	1596
58	90	30	-0,97	3,80	1879
59	101	30	-1,32	4,79	1698
60	88	24	-0,95	5,14	1810
61	98	28	-1,08	0,77	1547
62	96	30	-0,75	4,52	1964
63	101	31	-0,18	4,59	1646

TABLE IV

FIBRE DIAMETER OF GREASE WOOL — DATA SET II

Standard	Mean Values µm	Standard Deviation µm	Skewness	Kurtosis	No. of Fibres Measured
PP 64	22,0	5,7	0,83	3,68	1292
65	21,0	5,0	1,24	5,87	1278
66	21,6	4,7	0,78	4,24	1307
67	21,3	5,1	1,18	5,04	1288
68	22,6	5,0	0,83	4,15	1292
69	21,6	5,1	1,02	5,58	1319
70	2,16	5,1	0,83	3,96	1356
71	21,5	4,8	1,05	4,66	1312
72	21,3	4,9	0,97	4,98	1310
73	21,7	5,0	0,74	4,00	1337
74	22,2	5,3	0,83	3,90	1472
75	20,9	4,9	1,08	4,93	1325
76	22,9	4,8	0,54	3,74	1349
77	21,6	5,2	0,97	4,59	1381
78	21,9	5,1	1,01	4,87	1391
79	22,3	5,3	0,78	3,91	1357
80	22,0	5,5	1,05	4,97	1298

TABLE V

FIBRE DIAMETERS OF TOPS-DATA SET II

Sample	Mean Diameter μm	Standard Deviation μm	Skewness	Kurtosis	No. of Fibres Measured
PP 64	22,4	5,4	0,99	4,82	1274
65	21,0	4,8	1,19	6,38	1254
66	22,6	5,0	0,87	4,32	1306
67	21,4	5,2	1,01	4,53	1287
68	22,7	5,4	0,76	4,15	1296
69	21,7	5,1	0,51	3,23	1285
70	21,7	5,1	0,74	3,85	1276
71	22,0	5,2	1,39	7,02	1345
72	22,4	5,3	0,98	4,50	1288
73	22,1	5,5	0,92	4,21	1388
74	22,7	5,3	0,72	4,30	1261
75	21,9	5,3	1,20	5,66	1395
76	22,0	5,1	0,62	3,40	1311
77	22,3	5,4	0,75	3,68	1320
78	22,3	5,5	1,18	5,62	1286
79	22,5	5,5	0,83	3,92	1344
80	21,7	5,5	0,73	4,05	1341

TABLE VI

FIBRE LENGTHS OF GREASE WOOL — DATA SET II

Sample	Mean Lengths	Standard Deviation	Skewness	Kurtosis	No. of Fibres Measured
	mm	mm			
PP 64	103	42	-0,60	2,76	2174
65	87	28	-0,57	4,44	2673
66	83	24	-1,18	4,58	2280
67	70	19	-0,92	5,10	3251
68	6,0	18	-0,73	4,12	2822
69	52	17	-0,63	3,60	3827
70	90	31	-0,37	3,62	2943
71	84	26	-0,90	4,36	3202
72	76	22	-0,70	4,25	2722
73	66	20	-0,85	4,06	2317
74	95	37	-0,42	3,32	2516
75	85	26	-0,76	3,98	2638
76	78	25	-0,89	4,09	2477
77	89	36	-0,12	3,08	2176
78	83	23	-0,92	4,40	2201
79	89	39	-0,07	2,64	2443
80	74	39	0,58	2,87	5614

TABLE VII

FIBRE LENGTH GILLED SLIVER — DATA SET II

Sample	Mean Length mm	Standard Deviation mm	Skewness	Kurtosis	No. of Fibres Measured
PP 64	77	39	0,03	2,34	2033
65	69	31	-0,28	2,49	2346
66	70	27	-0,58	2,63	1804
67	61	21	-1,13	3,84	1541
68	57	24	-0,68	2,81	2125
70	71	32	-0,10	2,86	2362
71	69	30	-0,38	2,51	2347
72	65	26	-0,46	3,04	2516
73	61	21	-0,72	3,41	2751
74	77	34	-0,28	2,67	2056
75	69	27	-0,46	3,12	3213
76	61	27	-0,26	2,50	3331
77	66	31	0,15	3,16	3328
78	62	27	-0,20	2,81	3263
79	61	33	0,35	3,02	4025
80	54	32	0,90	3,98	3650

TABLE VIII

FIBRE LENGTH OF TOPS — DATA SET II

Sample	Mean Length mm	Standard Deviation mm	Skewness	Kurtosis	No. of Fibres Measured
PP 64	85	35	-0,02	2,39	1374
65	75	27	-0,32	2,88	1273
66	72	25	-0,39	2,53	1570
67	65	19	-0,66	3,44	1602
68	58	20	-0,42	2,86	1868
69	51	16	-0,55	3,04	1323
70	80	30	-0,15	2,69	2005
71	75	26	-0,33	2,71	1387
72	69	22	-0,43	3,12	1423
73	61	20	-0,61	3,19	1660
74	79	30	-0,04	2,67	1723
75	74	24	-0,34	3,07	1478
76	67	24	-0,19	2,76	1445
77	75	27	0,28	3,32	1291
78	66	25	-0,06	2,88	1522
79	73	31	0,31	3,18	1304
80	66	32	0,84	3,43	1470

TABLE IX

MEAN VALUES OF SAMPLES FOR DATA SETS I AND II

Sample	Mean	Standard Deviation	Skewness	Kurtosis
Fibre Length Values PP64-80 (mm)				
Greasy Staple	75	2,8	-0,58	3,84
Gilled Sliver	59	2,8	-0,27	2,94
Top	70	2,6	-0.18	2,95
Fibre Diameter Values PP64-80 (μ m)				
Greasy Staple	22,1	5,3	0,91	4,57
Top	22,1	5,3	0,85	4,56
Fibre Length Values PP49-63 (mm)				
Greasy Staples	95	2,9	-0,94	4,58
Fibre Diameter Values PP49-63 (μ m)				
Greasy Staples	22,8	5,7	0,77	3,50

TABLE X

PROPERTIES OF EQUAL BLENDS BY NUMBERS OF μ_a = 10, CV_a = 20% WITH VARIOUS μ_b AND CV_b VALUES

μ_b-Mean	Parameters of blend \ CV_b	20	25	30	35	40
5	μ	7,5				
	σ	2,96	3,01	3,06	3,12	3,19
	$\sqrt{\beta_1}$	0,43	0,34	0,23	0,12	0,02
	β_2	2,07	2,10	2,14	2,14	2,22
6	μ	8,0				
	σ	2,59	2,67	2,76	2,86	2,96
	$\sqrt{\beta_1}$	0,44	0,28	0,11	-0,04	-0,17
	β_2	2,40	2,41	2,45	2,50	2,54
7	μ	8,5				
	σ	2,27	2,40	2,54	2,68	-0,32
	$\sqrt{\beta_1}$	0,38	0,15	-0,05	-0,21	-0,32
	β_2	0,74	2,72	2,75	2,81	2,87
8	μ	9,0				
	σ	2,07	2,24	2,42	2,62	2,81
	$\sqrt{\beta_1}$	0,24	0,00	-0,18	-0,29	-0,34
	β_2	2,98	2,92	2,99	3,10	3,20
9	μ	9,5				
	σ	1,97	2,19	2,43	2,67	2,91
	$\sqrt{\beta_1}$	0,07	-0,08	-0,17	0,20	-0,20
	β_2	3,02	3,03	3,20	3,38	3,49
10	μ	10,00				
	σ	2,00	2,26	2,54	2,82	3,08
	$\sqrt{\beta_1}$	0,00	0,00	0,00	0,00	0,00
	β_2	3,00	3,38	3,55	3,55	3,61

TABLE XI
PROPERTIES OF EQUAL BLENDS BY WEIGHT (LENGTH) OF $\mu_a = 10$, $CV_a = 20\%$ WITH VARIOUS μ_b and CV_b VALUES

μ_b = mean	Parameters of blend \ CV_b	20	25	30	35	40
5	μ	6,67				
	σ	2,75	2,82	2,90	2,98	3,08
	$\sqrt{\beta_1}$	0,93	0,78	0,62	0,47	0,34
	β_2	2,88	2,76	2,65	2,55	2,46
6	μ	7,5				
	σ	2,48	2,58	2,70	2,82	2,95
	$\sqrt{\beta_1}$	0,72	0,51	0,30	0,13	-0,01
	β_2	2,86	2,62	2,62	2,56	2,51
7	μ	8,24				
	σ	2,23	2,37	2,53	2,70	2,87
	$\sqrt{\beta_1}$	0,50	0,24	0,02	-0,14	-0,24
	β_2	2,95	2,81	2,75	2,75	2,76
8	μ	8,89				
	σ	2,05	2,23	2,44	2,66	2,87
	$\sqrt{\beta_1}$	0,27	0,02	-0,16	-0,30	
	β_2	3,03	2,92	2,96	3,04	3,12
9	μ	9,47				
	σ	1,96	2,19	2,44	2,70	2,94
	$\sqrt{\beta_1}$	0,08	-0,07	-0,16	0,19	-0,19
	β_2	3,02	3,03	3,19	3,35	3,44
10	μ	10,0				
	σ	2,00	2,26	2,54	2,82	3,08
	$\sqrt{\beta_1}$	0	0	0	0	0
	β_2	3,00	3,14	3,38	3,55	3,61

$\mu_a = 10$ corresponds to a mean length of 50 mm

TABLE XII
PROPERTIES OF EQUAL BLENDS BY WEIGHT (DIAMETER) OF $\mu_a = 10$, $CV_a = 20\%$ WITH VARIOUS μ_b AND CV_b VALUES

μ_b-mean	Parameters of blend \ CV_b	15	20	25	30	35
	μ	8,77				
	σ	1,84	2,02	2,23	2,45	2,68
	$\sqrt{\beta_1}$	0,66	0,30	0,04	-0,14	-0,24
	β_2	3,56	3,10	2,93	2,94	3,01
9	μ	9,44				
	σ	1,74	1,96	2,20	2,45	2,71
	$\sqrt{\beta_1}$	0,31	0,08	-0,07	-0,16	-0,19
	β_2	3,41	3,02	3,02	3,18	3,34
10	μ	10,00				
	σ	1,77	2,00	2,26	2,53	2,80
	$\sqrt{\beta_1}$	0	0	0	0	0
	β_2	3,26	3,00	0,14	3,39	3,58
11	μ	10,45				10,44
	σ	1,91	2,15	2,41	2,68	2,93
	$\sqrt{\beta_1}$	- 0,13	0,06	0,18	0,22	0,21
	β_2	3,07	3,02	3,25	3,50	3,65
	μ	10,84				10,78
	σ	2,15	2,38	2,64	2,86	3,06
	$\sqrt{\beta_1}$	- 0,08	0,21	0,37	0,41	0,38
	β_2	2,90	3,05	3,30	3,48	3,59
	μ	11,13				11,06
	σ	2,45	2,66	2,86	3,03	3,19
	$\sqrt{\beta_1}$	0,08	0,36	0,51	0,53	0,49
	β_2	2,77	3,04	3,24	3,38	3,47

μ_a = corresponds to a mean diameter of 20μm

ISBN 0 7988 1077 7

Published by
The South African Wool and Textile Research Institute
P.O. Box 1124, Port Elizabeth, South Africa,
and printed in the Republic of South Africa
by P.U.D. Repro (Pty) Ltd., P.O. Box 44, Despatch