ELSEVIER

Dyes and Pigments 41 (1999) 111–126

DYES and PIGMENTS

More quality in textile finishing: reproducibility of processes in laboratory and production

R. Brossmann [a], T.V. Chambers [b], M. Krayer [c], M. Kunze [d,*], A. Lange [e], A.T. Leaver [f], H.-P. Oesch [g]

[a] *Im Kerberich 17, D-51519, Odenthal, Germany*
[b] *DyStar L. P., 9844 Southern Pine Boulevard, Charlotte, NC 28273, USA*
[c] *Ciba-Geigy AG, Leiter Qualitätsmanagement, K 1370 B. 501, Postfach CH-4002, Basel, Switzerland*
[d] *DyStar Textilfarben GmbH & Co., Deutschland KG, TO-QMI, Postfach 100 761, D-60007, Frankfurt, Germany*
[e] *BASF AG, ETM/SF-G 100, Postfach D-67056, Ludwigshafen, Germany*
[f] *Zeneca Colours, P.O. Box 42, Hexagon House, Blackley, Manchester, UK*
[g] *Clairiant (Schweiz) AG, Qualitätssicherung, Rothausstr. 61, CH-4132, Muttenz 1, Switzerland*

Received 3 August 1998; accepted 2 September 1998

Abstract

A trend to greater precision can be observed in production processes and technical testing procedures. Precision means that the process runs within narrow limits and that these limits are maintained with high probability. Even dyeing or printing processes in the laboratory or production plant are subject to greater or lesser fluctuations when the same processes are repeated. Quality Management according to the ISO 9000-9004 standard requires that, among other things, specifications, reproducibility levels and measurement uncertainties be defined or determined. Fundamental concepts of statistics are treated such as mean value, standard deviation, variation coefficient, confidence interval and measurement uncertainty and statistical process control (SPC). An attempt is made to reply to the question "How precisely can the colour strength of a dyestuff be determined?". Narrower specifications can only be realistically fulfilled when the measurement uncertainty factor can be correspondingly reduced, which can often only be achieved by additional, cost-intensive measures. Value is attached to the simplest possible presentation.

Keywords: Reproducibility; Specification; Laboratory; Production; Quality management

1. Introduction

The trend towards greater accuracy is a characteristic of modern industry that can be found in all production processes and studies associated with technical evaluations. The avoidance of errors takes precedence over subsequent correction. The knowledge and systematic examination of the process parameters which are relevant to quality in production and testing are a pre-requisite for achieving this objective.

This also applies to the textile finishing industry. Not only are efficient technical plants required, but also an extensive flow of information between suppliers and customers is required to achieve better reproducibility of manufacturing processes and thus a better uniform product quality. All necessary information must be mutually intelligible and

* Corresponding author.

0143-7208/99/$—see front matter
PII: S0143-7208(98)00074-6

clear to both sides working together as partners, in order to avoid costly disputes and associated losses. Only then can the individual stages of the textile finishing chain operate according to the key-lock principle with little expenditure on preliminary tests by the finisher.

Important elements of the information needed are the descriptions of the quality of both product and process, in the form of specifications. Moreover, quantitative details regarding the accuracy of the results are needed in order to interpret statements critically, to be able to evaluate their true values, and to avoid drawing the wrong conclusions.

In the case of the production process, accuracy means that the process takes place within narrow limits and that it is capable of adhering to these limits with a high degree of probability. Aids for assessing this "ability" are, for instance, "statistical process control" (SPC) and "processability analysis".

Dyeing processes in the laboratory or in the plant are also subject to a greater or lesser extent to variation when the same operating sequences are repeated. Furthermore, the resultant degree of variation in the quality of a manufactured product, and the tolerances resulting therefrom, are determined largely by experience and are only rarely defined on a statistical basis. However, this is a necessary requirement for optimization of the production sequences and therefore it will be demonstrated below, with an example of a laboratory dyeing process, how their accuracy can be described by means of statistical statements. The same procedure can also be applied to dyeing or printing processes in large-scale production, so that experiences of reproducibility on a large-scale can be collected and starting points for process optimization can be found.

Importance has been given to the use of the simplest possible descriptions and restricted to basic relationships, so as to show that important statements are possible, even by use of simple methods. The article is thus quite deliberately intended for readers who are not familiar with statistics. In this case, simplifications, whose clarification is dispensed with in the interest of clarity, are inevitably tolerated. The article is therefore to be thought of, above all, as a stimulus for more extensive study, and refers to more advanced text on statistics [1,2].

2. Quality management in accordance with ISO 9000-9004

For every customer, the certainty and confidence that a supplier is in a position to supply the contractual product in the quality demanded, are elementary requirements. The customer will accordingly want proof, as happens, for instance, as a result of adopting a quality management system run in accordance with standards ISO 9000-9004.

Equally, the supplier has a vital interest in being able to comply with agreements reached and to supply at the lowest possible costs, and in this, a quality management system also provides active support.

These ISO Standards define the Quality Management standard and require that, amongst other things, specifications, reproducibility and measuring uncertainties are determined and/or established. However, regarding numerical values, for instance in the form of minimum demands, the standards themselves make no statement. Thus, for instance, laying down specifications is the task of the interested parties, that is to say, of customer and supplier, according to the quality requirements. Often this does not take place in direct consultation, of course, but largely through feedback of market events and, influenced by the state of the art, the requirements become universally established.

The division of labour which occurs over several finishing stages is customary in textile production and calls for reliable quality of each of the prior stages. A useful tool for achieving this objective is a functioning quality management system that describes, and makes clear, the processing sequences and organisational responsibilities.

An important element of such a system is the testing institute for monitoring and control of processing sequences and testing of products. In ISO 9001, under 4.10.1 for testing, it is required that:

> ""The supplier must draw up and maintain process instructions for the test activities, in order that the stipulated quality requirement of the product will be met".

and, under 4.11.1, that

> ""The means of testing must be done in a way that ensures that the testing uncertainty is known and that it is consistent with the requirement in question".

The means of testing is understood to be the entire test, namely the method of carrying it out, including the testing device employed. Thus, in principle, each result of a test must be supplemented by statements regarding its reliability in the form of "testing uncertainty" (see Section 3.2), which is only possible based on statistical evaluations.

3. Practical statistical principles

The basic elements of the statistics needed for carrying out and evaluating statistical studies are formulated below.

3.1. Mean value, standard deviation, variation coefficient

Every physical or chemical process is subject to variation, even if it is repeated in a manner that is apparently always the same. These variations are often discernible only when a suitably sensitive test method is used. The following example should clarify this:

> "The length of a room is measured with a ruler with cm divisions. To start with, several repeated measurements give the same result. If, however, a metre rule with millimetre graduations, is used, different results are obtained when several measurements are taken. These variations in the test value can be caused by, for example, temperature differences or slightly inaccurate positioning of the ruler. Influences of this kind can never be totally avoided, even in the most careful work."

Over a large number of these length measurements, only a few will lie really close to the "true" value of the room length, which is not known

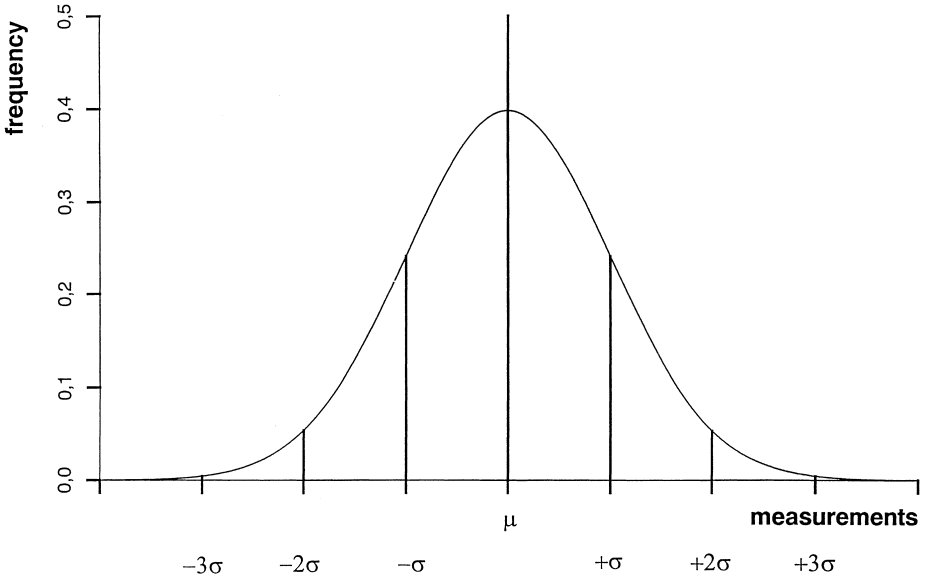

Fig. 1. Bell-shaped curve (normal distribution).

initially. If all results, consisting of many individual measurements are entered into a diagram with the measured values as abscissa and the number (frequency) of the measurements at the respective measurement value as ordinate, the following characteristic distribution is obtained in the form of a bell-shape curve.

A probability distribution function of this form is described as a normal (Gaussian) distribution. It is characterized by the maximum value at μ and points of inflections at $\mu - \sigma$ and $\mu + \sigma$.

Since the exact mean value μ can be determined only for an infinite number of measurements, that is to say, only theoretically (and can therefore never be determined as an exact value) in practice, with a limited number of measurements, instead of μ, the *arithmetical mean value* $\bar{x}$ is used. It is calculated according to the formula

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

that is to say, the sum of the n individual values is divided by the number n of the measurements.

Here, n is the number of measured values, x_i the value of the ith measurement.

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + \cdots + x_i.$$

A measure of the variation in the values around the mean value is the standard deviation σ. Since theoretically this, too, can be determined only from an infinite number of measured values, in practice the approximate value s, which is calculated as follows, is used:

$$s = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^{n} (x_i - \bar{x})^2}$$

$$= \sqrt{\frac{1}{n-1} \cdot \left[\sum_{i=1}^{n} x_i^2 - \frac{1}{n} \left(\sum_{i=1}^{n} x_i \right)^2 \right]}.$$

As the formula shows, the standard deviation is calculated from the differences between the individual values and the mean value. That is to say, it is proportional to the mean value.

Frequently, however, when accuracy is being considered, data expressed as a percentage is required, so that it is often advisable to use a value based on the mean value $\bar{x}$, which is called *variation coefficient* v, instead of the standard deviation (s):

$$v = \frac{s}{\bar{x}} \cdot 100(\%)$$

It is a measure of the relative variation of a distribution and is a measure of the mean variation of individual values.

The variation coefficient (v) characterizes the accuracy of measurements in a simple way, and makes a comparison of the different processes possible.

An important conclusion from these mathematical interrelationships is that, with a limited number of measured values, one can estimate the maximum variation that may occur between the mean value (the true value) and the measured value, or how far removed from the measured value the "true" value can lie; that is to say, how inaccurate the determined value is. A process with small standard deviation(s) or small variation coefficient (v), shows only a small variation of the measured values from mean value $\bar{x}$. The above will be explained using an example of a practical laboratory task, which can be used as a model case for the application of statistical methods:

3.1.1. How accurately can the tinctorial strength of a dyestuff be determined?

To determine tinctorial strength, test dyeings are made of a sample and standard dyestuff respectively and they are assessed by colorimetric evaluation [3]. The dependence of the dyeing on the substrate and its pre-treatment, the dyeing process, dyestuff concentration, liquor-to-goods ratio etc. is not considered here, but all of these parameters should be kept constant and the individual tests should be carried out under similar conditions. The variation and/or the measuring uncertainty of the determined tinctorial strength should then be established.

To work accurately, it is of course necessary to carry out all individual stages carefully and, if possible, to monitor all variables and keep them constant. This starts with the consideration of how accurately the dyestuff should be weighed – are scales with a weighing accuracy of 0.1 g or 10 mg adequate? – and includes the choice of substrate, definition of the dyeing programme (temperature profile), technical measuring evaluation, etc.

For the following discussion, which is simply a record of the actual situation in the determination of the accuracy of the end result, it actually has no significance. No knowledge whatsoever as to how greater accuracy can be achieved is gained by it either. However, with statistical evaluation, a tool is available with which improvements can be objectively measured and their success can be determined and quantified.

Statistical evaluation is the first step in a diagnosis, which, by systematic study of the individual parameters (e.g., substrate, liquor ratio, etc.), it is possible to plan improvements. Their realization is then an optimization process with which detailed problems of a technical and physico-chemical nature, and any organizational ones, can be solved.

This somewhat lengthy path is naturally not needed for all variables, but, with simple interrelationships, a direct quantitative estimation can lead to the destination. For example, an error of 0.1% in the weighed sample of dyestuff will bring about the same percentage error in the end result of the tinctorial strength. Looked at closely, however, one cannot manage here either, in the case of more precise statements, without statistics, as the following should make clear: what does the error in the weighed sample amounting to 0.1% mean? It can mean that with a weighing accuracy of 1 mg (with a weighed sample of 1 g) the error amounts to a maximum of 0.1% if it is read off correctly. This is a clear statement which does not need any statistical interpretation. But the situation is really different if it is not technically ensured, absolutely for certain, that the reading is always taken correctly from the scales – for example, by means of direct data transfer. If it is realistically assumed that reading-off errors are made when readings are taken manually, this, in turn, can be determined only statistically, for instance by determining the standard deviation.

Since the determination of tinctorial strength calls for two weighed samples, namely of test sample and standard, the weighing error also has to be taken into account twice. It would be wrong, though, simply to double it, for, according to the

Table 1

	Tinctorial strength characteristic: Weighted sum (K/S)		Tinctorial strength (%)
	Serial no.	Sample	Standard	
	1	68.9	70.3	98.01
	2	69.8	69.1	101.01
	3	69.9	69.6	100.43
	4	68.3	69.9	97.71
	5	69.1	70.5	98.01
	6	70.2	69.6	100.86
	7	69.5	68.9	100.87
	8	69.0	71.5	96.50
	9	70.9	70.2	101.00
	10	69.1	68.7	100.58
	11	68.7	71.8	95.68
	12	69.4	69.7	99.57
Number: n =		12		
Mean value: $\bar{x}$ =		69.40	69.98	99.19
standard deviation s:				1.91.
variation coefficient v:				1.93%.

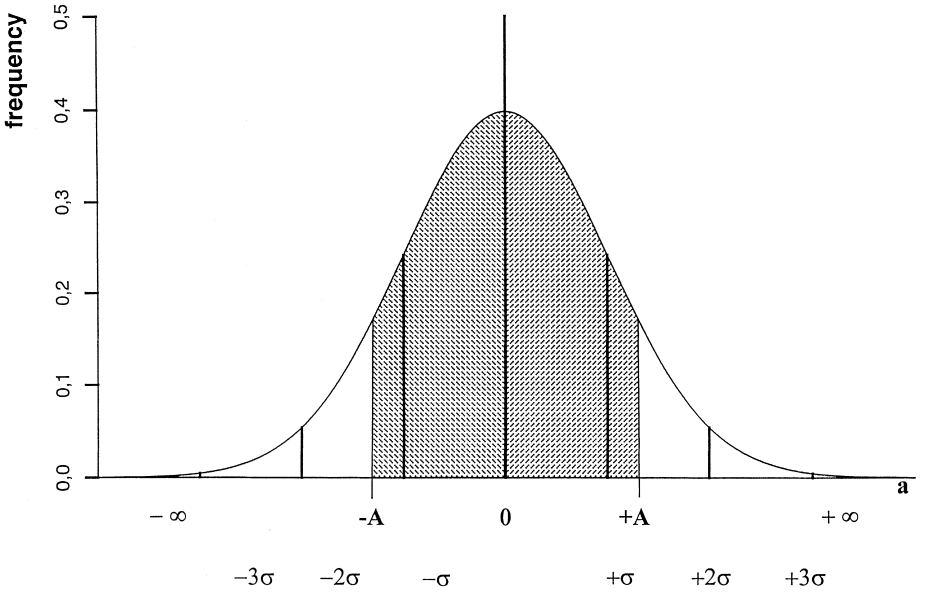

Fig. 2. Bell-shaped curve (normal distribution)(hatched areas $-A \ldots + A$).

Law of Propagation of Errors (see Section 3.3), it has to be calculated with the factor $\sqrt{2}$.

The test can be carried out in a different way, as explained herewith. First of all, it should be carried out in such a way that dyeings are always prepared in pairs, under conditions kept constant, and the dyeings of each pair are assessed against each other.

The data of such a test with 12 pairs are compiled in Table 1, which shows:

Tinctorial strength characteristic weighted sum (K/S) (absolute value for standard and test sample resulting from the colour measurement, see [3]).

Tinctorial strength of test sample (standard = 100%) (calculated as quotient of the tinctorial strength characteristic test sample/standard · 100).

Number n.

Mean value $\bar{x}$ of the tinctorial strength characteristic and the tinctorial strength.

Standard deviation s and variation coefficient v for the tinctorial strength.

Since, in this example, the tinctorial strength is expressed in % and its value is approximately 100, s and v are in close numerical agreement.

How are these values to be interpreted?

For this purpose the normal distribution shown in Fig. 1 should be viewed somewhat more meticulously (Fig. 2). The curve is symmetrical around 0 (corresponding to the mean value μ) and the abscissa a extends from $-\infty$ to $+\infty$.

This represents a probability distribution function (frequency density) and is standardized in such a way that the entire area under the curve has

Table 2

Range standard deviation	Probability P (%)	
$\pm 1s$	68	(*wrong in 1 of 3 cases*)
$\pm 2s$	95.5	(*wrong in 1 of 20 cases*)
$\pm 3s$	99.7	(*wrong in 1 of 300 cases*)
$\pm 1.9600s$	95	
$\pm 2.5758s$	99	(*wrong in 1 of 100 cases*)

Table 3

Probability P (%)	Uncertainty range U	Range limits for the first measured value from Table 1 tinctorial strength (%)
95.00%	1.9600 s (=3.74)	94.27 ... 98.01 ... 101.75
95.45%	2 s (=3.82)	94.19 ... 98.01 ... 101.83
99.00%	2.5758 s (=4.92)	93.09 ... 98.01 ... 102.93
99.73%	3 s (=5.73)	92.28 ... 98.01 ... 103.74

the value 1. In other words, all possible measuring occurrences (=100%) are included in the range $-\infty$ to $+\infty$. If only the range $-A$ to $+A$ is considered, the area enclosed (hatched) indicates the probability with which the measured result lies within this range.

This probability can be calculated, and the numerical values are available for the standardized curve in the form of tables (see, for example, [1,2]).

The value of the interval $a = \pm 1s$ (one standard deviation) amounts to about 0.68. The following statement can thus be made: the probability that a measured value is within the range $\pm 1s$ about the true value is 0.68 or 68%. Or, to put it another way: the true value lies in the range $\pm 1s$ around the measured value found, with a probability of 68%.

This probability is not very high, for, after all, the true value lies within the $\pm 1s$ range only in about 2 cases out of 3, in other words, *in one case out of 3 the statement probably is wrong*. Therefore, in the interest of greater certainty, the range is usually extended further. In the case of $\pm 2s$, the probability amounts to about 95%. This is a value very often taken as a basis in natural science and technology, and it conveys the fact that an incorrect statement can be expected only in one out of 20 cases. A few more frequently used "pairs of values" are compiled in Table 2.

Based on the above example (Table 1), the following statement is now possible for the 1st measurement (tinctorial strength = 98.01%): the true value of the tinctorial strength lies with a probability P within the uncertainty range U of the measured value in accordance with the following (Table 3):

3.2. Confidence range, measuring uncertainty,

If one does not restrict to a *single* measurement, but carries out n multiple measurements and calculates the mean value $\bar{x}$ from all n determinations, the uncertainty range of the mean value is reduced by the factor $1/\sqrt{n}$. That is to say, with an increasing number of measured values, the uncertainty of the mean value decreases by the inverse of the root of the number of measurements. For example, when $n=4$, the measuring uncertainty of the mean value compared with an individual measurement is halved. Thus, with the multiple test, one has a simple though costly means of improving accuracy (precision).

The range within which the true value lies with a given probability is described as the *confidence range for the expected value*. Instead of "given probability", as a rule the term "confidence level" is used and is expressed as P. The limits of the confidence range ("confidence limits") are given by

$$G = \bar{x} \pm \frac{t(n; P) \cdot s}{\sqrt{n}}.$$

Table 4
Student distribution

t- distribution Number n of the individual values	$P=95\%$ t
2	12.71
3	4.30
4	3.18
5	2.78
6	2.57
8	2.37
10	2.26
12	2.20
13	2.18
20	2.09
30	2.05
50	2.04
80	2.01
100	1.98
>200	1.96

Here the factor $t(n; P)$ is the value of the t distribution (Student distribution), which is a function of the number n of the measurements and of the selected probability P and which is tabulated for n and P (Table 4; see, for example [4]).

Factor t can be understood as the correction factor for taking the uncertainty into account when determining the standard deviation s in the case of small values of n. As n gets larger, the student distribution approaches the normal distribution and merges with it ($n > 200$; see Table 4).

In the example of Table 1, 12 individual tests (dyeings in pairs) were carried out for which a mean value of tinctorial strength of 99.19% and a standard deviation of 1.91% were calculated. Assuming that the standard deviation is known from previous trails, then the true value of the tinctorial strength lies with a probability of 95% within a range of $99.19\% \pm(1.96 \cdot 1.91/\sqrt{12}) = 99.19\% \pm 1.08\%$, thus at an interval reduced by a factor of about 0.29 ($=1 / \sqrt{12}$) compared with the individual measurement.

In addition to the mean value $\bar{x}$, the magnitude of an interval around $\bar{x}$, within which the true value of the measured variable presumably lies, also forms part of the complete data of the measured result of several individual measurements. The difference between the upper and lower limit and the mean value $\bar{x}$ is expressed as the *measuring uncertainty u* (definition in [4,5]). It corresponds to the value of the confidence range and is calculated according to the similar formula:

$$\text{measuring uncertainty } u = \frac{\pm t(n; P) \cdot s}{\sqrt{n}}.$$

However, it is often the case that the standard deviation(s) of a test is already precisely known, for example from previous studies. The value for s can then be included, and the measuring uncertainty is reduced quite considerably, in particular for small n, as the shape of the curve of factor t shows (Table 4). If s known, having been determined from many (∞) individual determinations, the formula for the measuring uncertainty u is:

measuring uncertainty

$$u = \frac{\pm t(\infty; P) \cdot s}{\sqrt{n}} \text{ (with known standard deviation).}$$

In the case of known standard deviation s, one profits from previous data, in that by determining s, uncertainty can be eliminated with a smaller number of tests and one thereby gains accuracy, as the numerical examples in Table 5 below show. The comments under 3.1 and in the first paragraph of 3.2 apply in each case to the known standard deviation s. The values given in Tables 2 and 3 correspond to a $t(n;P)$ factor for $n = \infty$.

From Table 1 compared with the total number of measurements (case 1) only the first 4 measurements (case 2) of the test considered as a further example. The following numerical values result for a confidence level of 95%. At ± 2.66, the measuring uncertainty of case 2, with 4 measurements, has more than doubled compared with case 1, despite a smaller standard difference. Case 3, also with the *first 4 measurement values*, but with a standard deviation s assumed to be known, is included as an additional example. It demonstrates how useful it is to use the standard deviation s from available data.

In the case of numerical examples in Table 5, above, and Table 6 (see below) the fact that various results which are expected to be the same differ more or less clearly from each other, will be

Table 5

	Case 1 ($n=12$) s is unknown tinctorial strength (%)	Case 2 ($n=4$) s is unknown tinctorial strength (%)	Case 3 ($n=4$) s is known tinctorial strength (%)
Mean value $\bar{x}$	99.19	99.29	99.29
Standard deviation s	1.91	1.67	1.91
Measuring uncertainty u	± 1.21	± 2.66	± 1.87

Table 6

	Tinctorial strength characteristic		Tinctorial strength (%)	
	Sample	Standard		Calculated from the mean values of the tinctorial strength characteristic sample/standard
Mean value	69.40	69.98	99.19	
				99.17
Standard deviation	0.71	0.95	1.91	
Variation coefficient (%)	$v_P = 1.03$	$v_S = 1.36$	1.93	Calculated from the individual values
			$v_{FS} = 1.71$	Calculated according to the Law of Propagation of Errors

puzzling. These are, for instance, the mean values of the tinctorial strength and the values of the standard deviation. The tinctorial strength values calculated for identical conditions according to Table 1 also differ slightly from each other when the mean value is calculated on the one hand from the 12 individual values of the tinctorial strength (=99.19%), and on the other hand is calculated as a quotient of the mean values of the tinctorial strength characteristics for test sample and standard (99.17%). A further example is the calculation of the variation coefficient according to the Law of Propagation of Errors (see below). It is true of all these cases that they are probability statements of the kind that specify a range within which the "measuring occurrence" is to be expected and/or within which the true value will lie. The true value is always unknown here, and different statements are therefore not incorrect as long as they do not contradict each other.

The applicability of these statistical laws is tied to conditions (amongst other things, the presence of a "normal distribution"), *which are usually fulfilled, but should not be disregarded altogether*. An "outlier" test (for example according to Dixon or Grubbs), is therefore advisable to check whether widely differing individual values should be regarded as "outliers", which could falsify the result and thus be deleted. For carrying out the test, reference is made to literature Refs [1,2,6].

3.3. Propagation of errors

Further evaluation of the test data by yet another variant is possible, as follows: first of all, the mean value of the tinctorial strength characteristic is calculated for all dyeings of the test sample, and then the mean value of the standard. The relative tinctorial strength is calculated as the quotient of the two mean values. The result (Table 6) is practically the same as in the case of the pairwise evaluations.

A statement about the variation of the result thus determined is provided by the *Law of Propagation of Errors* [1].

If an end results is calculated as the sum (total), difference, product or quotient from data (e.g. measured results) which is present as a mean value with known standard deviation, the following simple equations apply to the standard deviation and/or variation coefficient (the "error") of the end result:-

Sum total and differences:
The standard deviation of the result is calculated as a square root of the sum total of the squares of the individual standard deviations.

Products and quotients:
The variation coefficient of the result is calculated as a square root of the sum total of the squares of the individual variation coefficients.

Therefore, the following applies for the example of the tinctorial strength calculation:-

$$v_{FS} = \sqrt{v_P^2 + v_S^2},$$

where v_P is the variation coefficient of the measured values of the test sample, v_S the variation

coefficient of the measured values of the standard, v_{FS} the variation coefficient for tinctorial strength.

The values calculated from the data in Table 1 are compiled in Table 6. Here, it turns out that v_{FS} calculated according to the above formula yields a somewhat lower value (1.71%) than from the individual values of the tinctorial strength (1.93%).

3.4. Terms: Precision, repeat conditions

In connection with the description of measured results, a large number of often similar terms are used whose precise definition is necessary and is laid down in standards. A comprehensive list is given in DIN 55350 Part 13 [7] but a description which is more detailed and more intelligible for laboratory practice is given in Standard ISO 5725 (Precision of Measuring Processes) [6].

Hitherto, explanations have been dealing with the question as to which variations in results are to be expected if a measurement is repeated in always the same way, whilst nothing has been said about the repeat conditions. Since this is very important, it is advisable to define it in general. Two cases above all are differentiated: viz., repeat and comparative conditions. Repeat conditions are understood to be carrying out the test by the same process (method), on the same object (test sample) in the same location (laboratory), by the same technician, with the same equipment within a short period of time. Comparative conditions refer to the same process on the same object but at different locations, with different technicians and equipment.

It can be briefly formulated as follows:

repeat – in the same laboratory
comparative – between different laboratories (for example supplier and customer)

According to the standard, *precision* is the qualitative designation for the degree of agreement between results, and as a rule it is indicated by the standard deviation and/or variation coefficient. In keeping with the conditions, a distinction is made between

repeat and comparative precision
as well as repeat and comparative standard deviation etc.

Repeat precision is synonymous with the former term reproducibility.

The definition of repeat conditions is very narrow, so that in the case of comparisons, careful attention should be paid to the comparability of the conditions. Factors that adversely affect precision will quickly be found, by practical experience, and many restrictions can be dropped.

With routine work in the laboratory, it is important which technician carries out the test and which equipment is used for the example of tinctorial strength determination. The result (precision) will depend on the individual dyestuff and, of course, in the dyeing process employed. Experience shows that precision is largely identical in the case of specific groups of dyestuffs (for example, wool dyestuffs of the same type), but of course this experience must, in principle, first, be gained in each individual case.

It is plausible that, owing to the effect of different conditions, the value of comparative precision is clearly greater than that of repeat precision. Therefore, when test data are exchanged between suppliers and customers, poorer precision must be expected than that known from one's own laboratory. To improve conformity, a mutual agreement of the underlying conditions is required. But even when, for example, standardized test methods are used, a comparative standard deviation twice as large as the repeat standard difference is not uncommon [3].

In this section measuring processes have been discussed, whilst "measurements" were considered earlier. Each measurements is carried out by a specific process, however, so that the statements – bearing in mind the conditions – can be generalized for the process.

To sum up, it is found that the accuracy of a test result can be characterized by details of the precision, or, in the individual case, by the accuracy of measurement. In addition, the following should be pointed out: the statement relates only to a repeat of the test, but not to its correctness. It is therefore possible, that a result is, in fact, very accurate in the sense of precision, but not correct. In principle, therefore, the details of its *correctness* are also always part of the description of a result. For example, a tinctorial strength determination,

which has been carried out by means of an exhaust dyeing, is not necessarily valid for a continuous dyeing. Similarly, the validity for a tinctorial strength determination by an extinction measurement, which above all has to be carried out very accurately (in the sense of precision), cannot be expected of a dyeing from the outset, so that its correctness has to be determined only by appropriate comparisons.

Accuracy is a higher concept which comprises precision and correctness. It has not always been used in the present context, as mostly only precision has been meant. Accuracy has often been used in a general, colloquial sense in the interest of better comprehensibility.

4. Practical applications

The determination of tinctorial strength should serve as an example to explain the use of simple statistics, which can be applied to other assignments in the same way. For example, the individual stages in the determination of tinctorial strength can be studied separately, so that it is possible to evaluate how much each stage contributes to inaccuracy. Such a single stage is, amongst other things, photometric measurement and colorimetric assessment. Because of the unlevelness of dyeings, a larger number of individual measurements have to be carried out, according to the statistical evaluation of which, the uncertainty resulting from the measuring process and the surface unlevelness of the dyeing sample can be assessed. The proportion to be attributed only to the measurement can be determined quite simply. It can be determined by leaving the sample in the device and measuring it (at least) 20 times. It is so small (magnitude: variation coefficient 0.1%) with the measuring equipment available nowadays, that it can be disregarded compared with the variation caused during the dyeing process and the preparation of the test sample to be measured.

Similarly, the uncertainty arising from the surface unlevelness is determined by a relatively large number of measuring points (in the interest of a fairly great statistical certainty $>=20$) at the same dyeing. Empirically, it lies within the range 1–3% (variation coefficient) and is highly dependent on the preparation of the test sample. By means of this simple methodology, the effect of various substrates, for example, can also be examined – insofar as it has an effect on the surface unlevelness – by using the various substrates under otherwise identical conditions.

Apart from the determination of tinctorial strength, the determination of shade differences is an equally important task. According to colorimetric assessment, the total colour difference ΔE is determined – usually in accordance with the CIELAB formula – broken down into ΔC (chroma difference), ΔH (hue difference) and ΔL (lightness difference). The statistics can be applied to the overall colour difference ΔE in the same way. However, the same treatment is not correct for the components ΔC, ΔH, ΔL, as these are not independent of each other, and thus important pre-conditions for the simple application of the statistics are not met. Since the correct procedure in this case is somewhat complicated, and is hardly encountered in practice, it is only briefly mentioned here [8], and a warning is given against applying simple statistics without due consideration. If it is carried out, however, the result should be interpreted only as a rough estimate of the order of magnitude.

5. Statistical process control (SPC)

In the previous discussion, the time aspect of the progress of processes was not considered. However, in practice all processes have to be regarded as progressing chronologically. The repetition of a certain operation shall bring the ''same'' result as the previous execution. This also applies for the batch processes of dyeing and printing in the textile processing industry. As an aid for the *timewise* pursuance of the quality of a process, ''Statistical Process Control'' (SPC) can be used [9]. It has to be seen as an important element in quality assurance. SPC offers the possibility to separate random (natural) events which cause the deviation in a statistically controlled process from the systematical, unpredictable influences, i.e. the disturbances. A process can be regarded as being

under statistical control if only random events cause the deviations and the accompanying characteristics lie within the operation limits defined by the SPC process. Systematic errors can be recognised immediately by the simple and immediate evaluation and visualisation in diagrams of the data achieved in the process using the SPC method. Thus the necessary measures to eliminate failures can be initiated immediately. At the same time the factors that stabilize the production process can be identified, and thus a permanent improvement and reduction of costs can be reached. SPC methods are especially useful to investigate defined steps in the process.

Monitoring and control are the main functions of SPC. Timely intervention has to be emphasized, for it should not take place too early in the case of any change which is only a question of statistical fluctuations but it should happen as soon as possible if there are differences going beyond this. The rules of SPC are only exactly valid for normally distributed processes (Gaussian distribution), but they may also be applied to other kinds of distribution. However, in this case, they may possibly lead to greater errors, and detailed investigations are necessary.

SPC methods can be applied for the study of continuous characteristics (e.g., process temperature) as well as for discontinuous characteristics (e.g., counts of errors). Therefore, they are an ideal tool to reduce the range of deviations in measurement and testing processes in the laboratory or production. In general, the deviations in a test process are regarded to be in the magnitude of 1/10 of the tolerance range of the process step considered. In textile processing this is a problem insofar as the test methods applied (e.g. dyeing) are the same as in the production process. Their deviations often have a magnitude which is not sufficient for practical needs, as they exceed the visual acceptance limits.

A tried and tested working device of SPC is the so-called quality control chart which contains the continuously entered process control data in the form of a diagram represented over time. Process deviations immediately become clearly visible, and

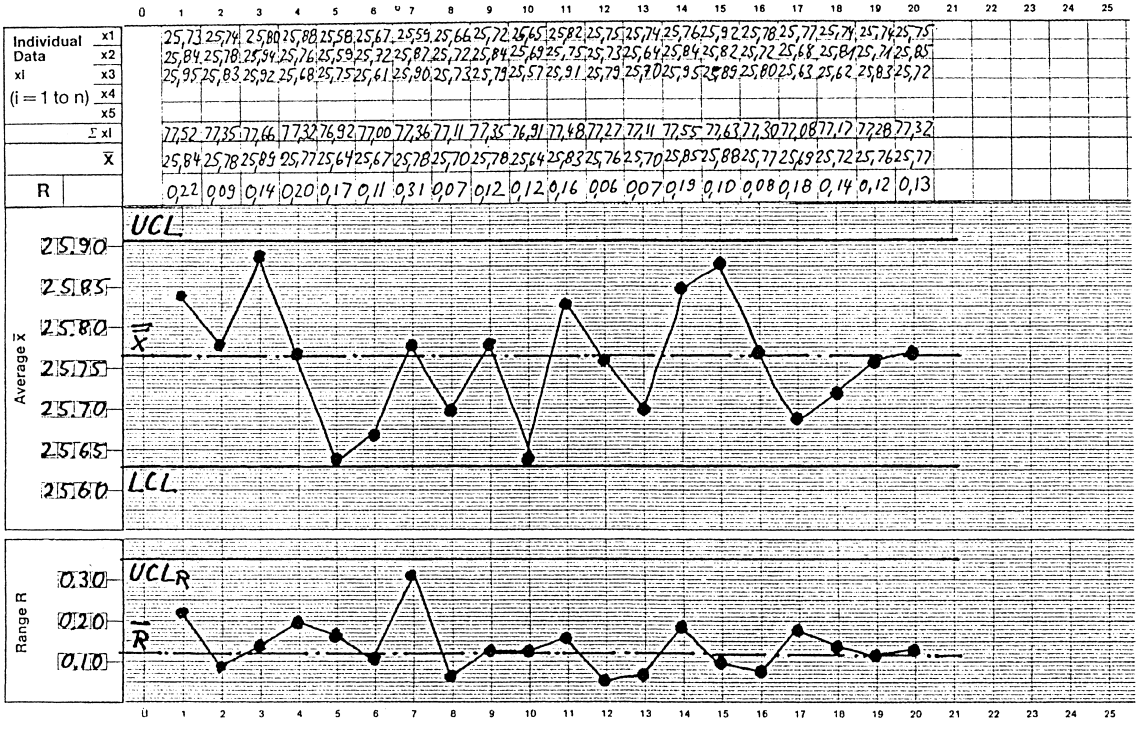

Fig. 3. $x/\ R$ – (individual value/range-) control chart.

can be dealt with by direct intervention possibly even during the running process.

As a basis to set up a control chart, the known characteristics (mean value $\bar{x}$, standard deviation s and range R) are used. The range is defined as the absolute value of the difference between two measurement values:-

$$R = |x_{i+1} - x_i|.$$

For processes in chemistry and textile processing measurement, values often are taken from a representative product sample. For these kind of processes, often a control chart with individual values x and the range R and the input intervention limits (UCL, LCL) is used. The line of the individual values then gives information on the position (actual value) of the process, while the range line gives an information on the process deviation (Fig. 3). Other control charts used are, for example, the x/s – (individual value/standard deviation-) control chart, or the $\bar{x}/R$ – (mean value/range-) control chart.

The simple procedure when working with SPC control charts is demonstrated by the following essential steps. A set of base data, e.g., from approximately 25 already available random tests is used as precursor for setting up a control chart. From these data, the preliminary upper and lower intervention limit (UCL, LCL) are calculated, e.g., with $\pm 3s$. The value of $\pm$3s is a general definition of the "natural" process limits, which indicates that only about one in 300 values (see Table 2) is to be expected outside of the intervention limits as a result of natural variation.

The control chart is completed in the described manner and each new test result is filled in *immediately* and connected with the previous one. The resultant lines for the position of the mean value $\bar{x}$, the range R and the standard deviation s can now be interpreted, and the preliminary intervention limits may possibly be corrected.

By means of given criteria it can be decided whether the process considered is under statistical control ("self-possessed") or not. If one of the following seven requirements comes true the process can no longer be regarded as being self-possessed. Then, special causes (defects) are present which have to be ended by appropriate measures (Fig. 4):-

Range chart R:
1. One measurement value (or more) out the UCL_R
2. Two thirds of the measurement values below $\bar{R}$

Measurement chart x
3. One measurement value (or more) outside of UCL or LCL ($3s$)
4. Seven successive measurement values rising or falling (trend)
5. Eight successive measurement values on one side of $\bar{x}$ (shift)
6. Two out of three successive measurement values outside of 2/3 of UCL or LCL ($2s$)
7. Four out of five successive measurement values outside of 1/3 of UCL or LCL ($1s$)

6. Processability analysis, ability indices

The result of a process with its characteristics should normally lie within given limits (specifications). The *Processability index* c_p serves to portray how well a process is able to meet these requirements. It is defined as the ratio of the limits of the specification to the "natural" process limits and is calculated as follows:

$$c_p = \frac{\mathrm{USL} - \mathrm{LSL}}{6s}.$$

Here USL is the Upper Specification Limit, LSL the Lower Specification Limit, s the standard deviation.

A c_p value of >1 therefore means that the $3s$ limits lie within the specification limits.

As the c_p value makes a statement independently of the position of the mean value of the process characteristic, compared with the mean value of the specification, some of the variations of the process characteristic can fall outside the specification limits if the mean value is displaced. A further c_{pk} index is therefore defined, which takes the shift(ing) of the mean value towards the

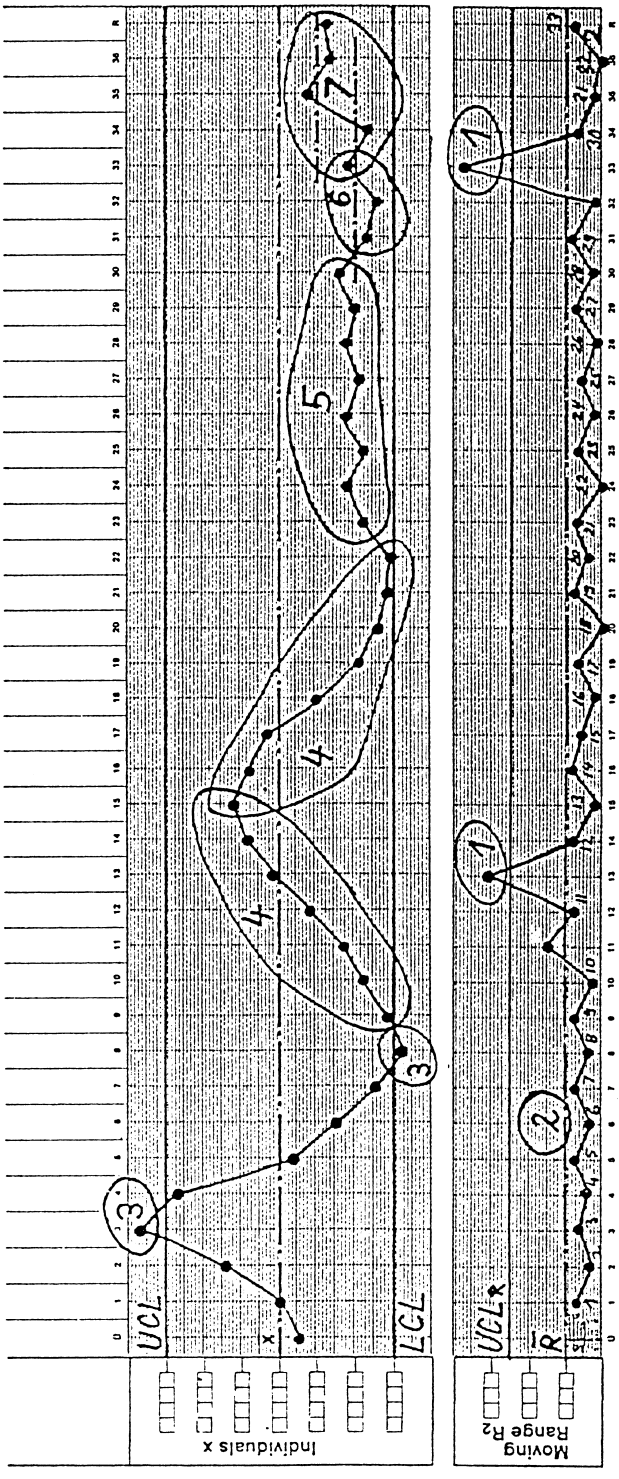

Fig. 4. Criteria for a self-possessed process.

target value into account, and includes the distance between the mean value and the nearest "critical" specification limit:

$$c_{po} = \frac{\mathrm{USL} - \bar{x}}{3s},$$

$$c_{pu} = \frac{\bar{x} - \mathrm{LSL}}{3s},$$

c_{pk} : the lower value of c_{po} or c_{pu}.

Here USL is the Upper Specification Limit, LSL the Lower Specification Limit, $\bar{x}$ the mean value, s the Standard deviation.

The demand for a c_{pk} index of $>= 1$ for a process is often encountered. This means that – because of natural variations – the specification limits should be exceeded only very rarely. The background to this demand is the wish to avoid any trouble in the further production process caused by the out-of-spec units with great reliability. The c_{pk} index has special significance in piece-goods production, where individual pieces are produced in large numbers, and where it is impossible to test each piece individually. A statistical description of the multitude of individual pieces is then virtually the only possibility of judging the quality of the whole batch, which is done by means of spot checks.

Analogous conditions exist in a continuous process, when characteristic data are not recorded continuously, but only discontinuously at intervals. Even then, a reliable statement on variations or out-of-spec units are only obtained by means of the statistical parameters.

The situation is different where a continuous process is also monitored continuously or when all units are tested (so-called 100% test), as is the rule in the case of chemical processes with batchwise production. In these cases, adherence to the specification can be ensured by the complete testing of all units and singling out of those out-of-spec. It is then no longer necessary to demand as large a c_{pk} index as possible for the process, in order to achieve a correspondingly low variation quota.

In general, it may be desirable to manage a process within narrow limits. It should in this case not be forgotten, however, that this involves additional expenditure, which must also be economically justifiable. If, in the case of a "100% test", it is ensured that the specification is satisfied at all times, an additional demand for a high c_{pk} index does not bring any further technical benefit and is therefore not justified.

7. Conclusion

In the introduction, attention was drawn to the importance of the exchange of information in the form of agreed specifications, as well as to the required critical evaluation of such stipulations. Since it is possible to judge the conformance to specifications only by tests, conformance can only be ensured by the additional consideration of the uncertainty of the test. This also plays a part in the testing of textile dyestuffs, as in may cases the testing uncertainty is of the same magnitude as that of the specification laid down. For instance, the measuring uncertainty in a tinctorial strength determination is equal to the tinctorial strength difference which leads to visible colour differences, especially in a sensitive combination of dyes and consequently it can really no longer be tolerated.

In the case of a test result at the specification limit, the true value lies outside the specification with a probability of 50%. In order to conform to the specification reliably, each test value must be at least a distance equal to the measuring uncertainty u from the specification limits. This means that the tolerance range laid down by the specification limits is reduced by the amount of twice the measuring uncertainty u and only the remainder is available for the variation of the process (characteristic).

In practice, the tolerance range laid down by the specification should not be used up by more than half of the measuring uncertainty u i.e., the tolerance range should be wider than four times the amount of the measuring uncertainty u.

As an example, it thus follows that in the case of the tinctorial strength with a measuring uncer-

tainty u of 2% in the tinctorial strength determination, a specification which is not narrower than 4%, based on the theoretical value, should be laid down. A narrow specification can realistically only be met if the measuring uncertainty can be reduced, as can often be achieved only by means of additional multiple tests.

Consideration of this point is a matter of course, but it appears that these interrelations are not known in the case of many a demand for narrow tolerances.

References

[1] Sachs L. Applied statistics. Berlin: Springer, 1992.
[2] Doerffel, Statistics in applied chemistry. Weinheim, 1987.
[3] Baumann W, et al. Melliand Textilberichte 1986;67:562.
[4] DIN 1319. Basic concepts of measuring technology. Concepts of measuring uncertainty and of the evaluation of measuring equipment and instruments, 1983.
[5] DIN ISO 10012. Quality guarantee requirements in the case of measuring equipment. Verification system for measuring equipment, 1992.
[6] DIN ISO 5725. Precision of measuring methods (processes)Determination of the repeat and comparison precision of stipulated methods of measuring by means of ring tests, 1988.
[7] DIN 55350. Concepts of quality guarantee and statistics. Part 13: Concepts regarding the accuracy of methods of determination and determination results, 1987.
[8] DIN 55600. Determination of the significance of colour differences of surface colours according to the CIEL ab-formula, 1988.
[9] Chambers DS, Weehler DJ. Understanding Statistical Process Control. SPC Press, 1986.