CHARACTERISTICS FOR JUDGING THE RELIABILITY
OF ANATYTICAL METHODS*

K.ECKSCHLAGER

Institute of Inorganic Chemistry,
Czechoslovak Academy of Sciences, 250 68 Prague - Řež

This work brings a critical evaluation of characteristics proposed earlier for evaluation of reliability of results obtained by analytical methods, and a new characteristic is proposed based on the information uncertainty of results. An example of the application of this characteristic to the results of the determination of Cu and Ni in KCl by atomic absorption photometry is described.

In evaluating the results of analyses it is necessary to consider their accuracy, *i.e.* the agreement among parallel determinations as well as their correctness, *i.e.* the agreement between the average result, $\bar{x}$, and the actual content, ξ, of the component in question. To evaluate the accuracy, we can use the standard deviation

$$s = \sqrt{[\sum(x_i - \bar{x})^2/(n-1)]}, \qquad (1)$$

where x_i means the result of i-th parallel determination. The relative standard deviation, $v = 100s/\xi$, or the quantity $\Gamma = \bar{x}/s$ according to Kaiser and Specker[1] is used more often. A measure for judging the correctness of results of analyses can be, *e.g.*, the quantity $t = (\bar{x} - \xi)\sqrt{n}/s$ or the finding on what significance level, α, the difference $|\bar{x} - \xi|$ is statistically significant[2], which can be obtained by comparing the quantity t with the critical value of the Student distribution, t_k. However, if we want to judge both the accuracy and correctness of results obtained by an analytical method, we meet a difficulty due to the fact that the value of t depends on the standard deviation s. This circumstance was not fully observed by McFarren, Lishka and Parker[3] who proposed the "total error" $T = 100(d + 2s)/\xi$, where $d = |\bar{x} - \xi|$, as a characteristic for judging the acceptability of analytical methods. Therefore, it was recommended[4] to replace the quantity T by $T_A = 100(d_A + 2s)/\xi$, where d_A is set equal to $|\bar{x} - \xi|$ only if this difference is statistically significant on a chosen significance level α; if, however, it is smaller than one half of the width of the reliability

interval for a chosen α (*i.e.* if it is statistically insignificant with respect to the value of the standard deviation, s, found for the results), then d_A is set equal to zero. Thus, the case where the difference $|\bar{x} - \xi|$ depends on a systematic error (*i.e.*, is statistically significant) is distinguished from the case where it is due to random scattering of x_i around ξ characterized already by the value of s. With the quantity T_A, however, the problem is to determine the significance level α for which the statistical significance of the difference $|\bar{x} - \xi|$ is to be found. In the original work[4], it is only noted that the significance level α must be strictly prescribed if the testing based on this characteristic is to be considered comparable. Another negative feature of the quantity T_A is the fact that its absolute value as a function of $|\bar{x} - \xi|$ is discontinuous in the point $|\bar{x} - \xi| = st_k/\sqrt{n}$. Therefore, in the present work a new quantity T_B is proposed which is analogous to T_A but based on a significance level α corresponding to the information uncertainty and is modified so that it exerts no discontinuity with changing $|\bar{x} - \xi|$.

RESULTS AND DISCUSSION

The information uncertainty U of normally distributed analytical results for which the probability density, $p(x)$, is equal to $\exp\left[-\frac{1}{2}(x - \mu/\sigma)^2\right]/\sigma\sqrt{2\pi}$, is given by the Shannon equation[5] of entropy $P = \log U$, *i.e.* by equation

$$P = \int p(x) \log p(x)\, dx = \log \sigma \sqrt{2\pi e} . \tag{2}$$

Since the information uncertainty of results of quantitative measurements has the meaning of length of the reliability interval[6] and since $\frac{1}{2}\sqrt{2\pi e} = 2.066$, we can introduce the following characteristic of both accuracy and correctness of results of an analytical method:

$$T_B = 100(2.066\sigma/\sqrt{n} + \vartheta)/\xi , \tag{3}$$

where $\vartheta = d - 2.066\sigma/\sqrt{n}$ represents that part of the difference $d = |\bar{x} - \xi|$ which exceeds the statistically significant value of uncertainty. The reason why it is preferable to relate the value of such a characteristic as T_B to the correct value of ξ and not to the found value of $\bar{x}$ is explained in the work of McFarren and coworkers[3].

If the quantity T_B is used, then for $d < 2.066\sigma/\sqrt{n}$ the value of ϑ is set equal to zero as in the case of T_A so that $T_B = 206.6\sigma/\xi \sqrt{n}$. For $d < 2.066\sigma/\sqrt{n}$ we set $\vartheta = d - 2.066\sigma/\sqrt{n}$, hence $T_B = 100d/\xi$. Both these relations are equivalent if $d = 2.066\sigma/\sqrt{n}$.

The characteristic T_B defined by the above equations could be used only if the value of σ were known. In practice, however, only its estimation, s, according to Eq. (*1*) is known. Then instead of $\frac{1}{2}\sqrt{2\pi e} = 2.066$ we use in the mentioned equations the

value of t_k for $\alpha = 0.039$ (corresponding to the information uncertainty of results) and for $v = n - 1$ degrees of freedom. These values were tabulated in ref.[6] and newly precised by computing on a computer. We use then the last two equations according to whether $d < st_k/\sqrt{n}$ or $d < st_k/\sqrt{n}$.

The meaning of the characteristic T_B is as follows: *a)* If $d \leqq st_k/\sqrt{n}$, the results obtained by means of a given analytical method are not subject to a systematic error which would exceed the information uncertainty of the results. Then the characteristic T_B is influenced only by the value of standard deviation as an expression of the fact that the value of $d = |\bar{x} - \xi|$ is due only to random scattering of the measured values of x_i around ξ.

b) If $d > st_k/\sqrt{n}$, the method is obviously subject to a systematic error, and it is the relative mean value of d which is decisive for judging the reliability of the analytical method under consideration.

As already mentioned, a disadvantage of the quantity T_A is its discontinuity with changing d for $d = |\bar{x} - \xi| = st_k/\sqrt{n}$. The characteristic T_B has no such disadvantage. The dependence of T_B on d for various values of s and n and for $\xi = 10$ is shown in Fig. 1; it is seen that the quantity T_B is for $d < st_k/\sqrt{n}$ the larger the less accurate are the results, but does not change up to $d = st_k/\sqrt{n}$. The number of parallel determinations, n, influences somewhat the value of T_B in this region, and for $n = \infty$ $T_B = 0$. The value of T_B increases linearly with d if $d > st_k/\sqrt{n}$ and depends neither on s nor on n.

When the characteristic T_B is used, then (similarly as with T and T_A) it is possible to find some "critical value" which would classify an analytical method as, *e.g.*, very reliable, reliable or unreliable. Such an attempt was made in ref.[3] with the use of the quantity T; when using this classification for methods suitable to determine
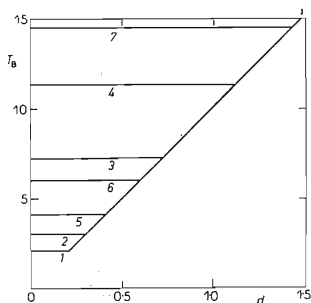

Fig. 1

Dependence of Characteristic T_B on $d = |\bar{x} - \xi|$

1 $s = 0.1$; $n = 100$; *2* $s = 0.1$, $n = 50$; *3* $s = 0.1$, $n = 10$; *4* $s = 0.1$, $n = 5$; *5* $\sigma = 0.2$, $n = 100$; *6* $s = 0.2$, $n = 50$; *7* $s = 0.2$, $n = 100$.

TABLE I

Values of T_B for Determination of Cu and Ni in KCl by Atomic Absorption Photometry

Metal	ξ	Amount of KCl, g	$\bar{x}$	s	T_B in case of calibration with	
					metal soln.	model sample
Cu	1·0	1	$0·86^a$	0·107	$14·50^c$	$14·50^c$
	10·0	5	$9·6^a$	0·351	4·74	4·74
Ni	10·0	1	$8·9^a$	0·443	11·00	5·99
	10·0	5	$7·3^a$	0·445	27·00	6·02
	10·0	1	$8·2^b$	0·439	18·00	5·94
	10·0	5	$7·8^b$	0·391	22·00	5·29

Extraction ^{a} with dithizone, ^{b} with dimethylglyoxime. ^{c} All values of T_B calculated with the aid of more accurate values of t.

vastly different contents of a component we conclude that every "critical" value of T applies only to a limited range of the values of ξ. Thus, a "critical" value which could be applied to the determination of concentrations close to 100% is too severe for methods of trace analysis, and *vice versa*. Hence, the applicability of this apparently universal characteristic is restricted to comparing methods used in determining roughly equal contents of one or more components or to find out whether with a dependent analytical method[7] the use of external (*i.e.* a calibration curve) or internal standards (*e.g.* method of standard addition) is more suitable, *etc*. The characteristic T_B can be also used to decide whether it is purposeful to use an operation which complicates considerably the method while improving the reliability of results only to a limited extent.

It was stated[8], for example, that in the determination of Cu and Ni by atomic absorption photometry in potassium chloride after extraction and reextraction, the results are somewhat lower if the calibration curve is not based on model samples containing the same amount of alkali metal salts as the sample, and if every model sample is not treated by the complete analytical procedure including extraction and reextraction. This means that the whole laborious and time-consuming calibration must be carried out for every determination anew. Whether this is necessary or not can be decided at best with the use of the characteristic T_B as shown in Table I. It is seen that in the determination of Cu the value of T_B does not change even in the case of calibration with model samples, *i.e.* that any small systematic negative error is statistically insignificant at a significance level $\alpha = 0·039$. On the other hand, in the determination of Ni, it is suitable to use model samples for calibration and subject them to the complete analytical procedure.

The fact that the characteristic T_B cannot be used in comparing with a generally valid critical value should not be considered as a disadvantage caused by its particular properties; it is rather a general feature of all characteristics based on the standard deviation, *i.e.* those that evaluate the accuracy of results[7]. It should be noted that the characteristic T_B serves to evaluate analytical methods whose results are not subject to systematic errors according to their accuracy, and methods subject to systematic errors according to the correctness of results, *i.e.* according to the difference between the actual and found values.

REFERENCES

1. Kaiser H., Specker H.: Z. Anal. Chem. *149*, 46 (1956).
2. Eckschlager K.: *Errors, Measurement and Results in Chemical Analysis*, p. 107. Published by SNTL, Prague - van Nostrand, London 1969.
3. McFarren E. A., Lishka R. J., Parker J. H.: Anal. Chem. *42*, 358 (1970).
4. Eckschlager K.: Anal. Chem. *44*, 878 (1972).
5. Shannon C. E.: *The Mathematical Theory of Communication*, Urbana Ill. 1949.
6. Eckschlager K.: This Journal *36*, 3016 (1971).
7. Eckschlager K.: Chem. listy *67*, 19 (1973).
8. Eckschlager K.: This Journal *34*, 1321 (1969).

Translated by K. Micka.