

CCA-1061

YU ISSN 0011-1643

545.3:541.13.21

Original Scientific Paper

The Possibility of Extending the Detection Function of an Analytical System to Lower Concentrations

Department of Analytical Chemistry, University of Cluj-Napoca,
Cluj-Napoca, Romania

C. Liteanu, I. Rîcă, and E. Hopîrtean

Received August 9, 1976

It is shown that the only way to increase the detection capacity of an analytical system is based on the use of the multitude values of the single analytical signal. This can be achieved in two modalities: 1) fixed multitude values (as mean or as sum), and 2) sequential multitude values (as signal sum and as frequencies sum). By using these procedures, the analytical detection can be applied under the classical detection limit at as low concentrations as desired, provided that the multitude of the individual values of the signal be sufficiently high.

From the distribution of probabilities $p(y_0)$ for the background noise and $p(y)$ for the analytical signal, as shown in Figure 1, we observe that the degree of overlap of the two probability fields increases as the concentration decreases.¹⁻³ Further, the only way to lower the detection limit is to reduce

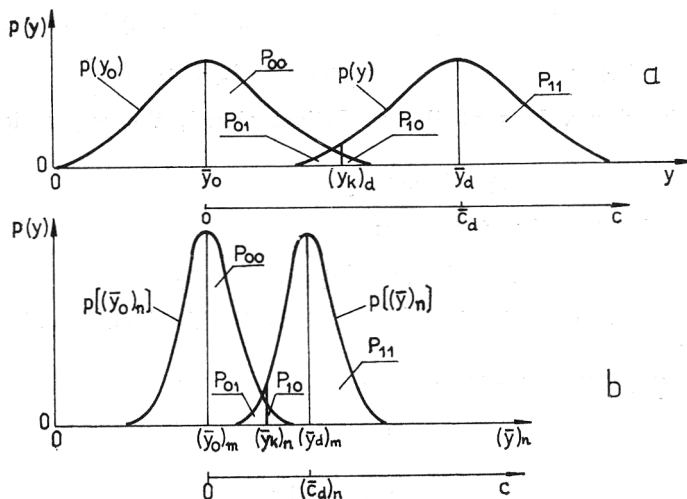


Figure 1. Two-step statistical model for defining the detection limit from single values (a), and from multitude values (b).

the scatter of the analytical signals about their mean. This can be done only by using the multitude values of the single analytical signal. This can be

achieved in two modalities: 1) using fixed multitude values (as mean or as sum), and 2) using sequential multitude values (as signal sum or as frequencies sum).

THE USE OF FIXED MULTITUDE VALUES

The Use of Mean Values

In the two-step model for defining the detection limit by operating with mean values $(\bar{y})_n$ (Figure 1b) to estimate the signal $(\bar{y}_d)_m$ corresponding to the detection limit $(\bar{c}_d)_n$, the decision level $(\bar{y}_k)_n$ will be based on the Neyman-Pearson criterion⁴ (as in the case of the use of single values y), i. e. for a certain value of the false detection probability P_{10} . For a normal distribution of the background signals y_0 and their mean values $(\bar{y}_0)_n$, based on the same considerations as in the case of the use of singular values of the signal,¹⁻³ for the detection level the following expression is obtained:

$$(\bar{y}_k)_n = (\bar{y}_0)_m + z_k \sigma_{(\bar{y}_0)_n} = (\bar{y}_0)_m + z_k \sigma_{y_0} / \sqrt{n} \quad (1)$$

where $(\bar{y}_0)_m$ is the generalized mean of the background fluctuations: $(\bar{y}_0)_m = \sum_{i=1}^N [(\bar{y}_0)_n]_i / N$, obtained from a great number of y_0 values, therefore $m = nN$; N is the number of mean values $(\bar{y}_0)_n$.

For the detection signal the following expression is obtained:

$$(\bar{y}_d)_m = (\bar{y}_k)_n + z_d \sigma_{(\bar{y})_n} = (\bar{y}_0)_m + z_k \sigma_{y_0} / \sqrt{n} + z_d \sigma_y / \sqrt{n} \quad (2)$$

where $(\bar{y}_d)_m = \sum_{i=1}^N [(\bar{y})_n]_i / N$ is the generalised mean of the $m = nN$ singular values of the analytical signal, namely of that N mean values $(\bar{y})_n$.

Taking the approximation $\sigma_{y_0} \approx \sigma_y$, the expression (2) becomes:

$$(\bar{y}_d)_m = (\bar{y}_0)_m + (z_k + z_d) \sigma_y / \sqrt{n} = (\bar{y}_0)_m + k_d \sigma_y / \sqrt{n} \quad (3)$$

Taking the singular values of the analytical signal, the expression (1) has the from:

$$y_k = \bar{y}_0 + z_k \sigma_{y_0} \quad (4)$$

and the expression (3) takes the form:

$$\bar{y}_d = \bar{y}_0 + (z_k + z_d) \sigma_y = \bar{y}_0 + k_d \sigma_y \quad (5)$$

The values z are obtained from table⁵ of the Laplace function $\Phi(z_k) = 0.5 - P_{10}$, and $\Phi(z_d) = P_{11} - 0.5$.

Based on the calibration functions:

$$y = \bar{y}_0 + bc \quad (6)$$

and

$$(\bar{y})_n = (\bar{y}_0)_m + bc \quad (7)$$

from eq. (3), (5), (6) and (7) it results then:

$$(\bar{c}_d)_n = \frac{k_d \sigma_y}{b \sqrt{n}} = \frac{\bar{c}_d}{\sqrt{n}} \quad (8)$$

namely

$$n = \left[\frac{\bar{c}_d}{(\bar{c}_d)_n} \right]^2 \tag{9}$$

However, in operating with mean values, the detection *characteristic*² will involve the parameter n too, alongside with the true detection probability P_{11} and the false detection probability P_{10} , and will have the form:

$$(\bar{c}_d)_n = f(P_{11}, P_{10}, n) \tag{10}$$

In conclusion, decreasing the detection limit by the use of the selection mean⁶ as the analytical signal, is not equivalent only to decreasing the background⁷ but to increasing the capacity of the analytical system used with regard to the differentiation of the analytical signal from the background noise.

In other words, to diminish the detection limit, so many measurements must be performed until the signal-to-noise ratio for a certain concentration level $(\bar{c}_d)_n$ assumes the value: $k_d = \frac{(\bar{y}_d)_m - (\bar{y}_o)_m}{\sigma_y / \sqrt{n}}$.

Figure 2 shows the correspondence between the decision parameter $k_d = f(P_{10}, P_{10})$ and the analytical signal $(\bar{y}_d)_m$ corresponding to the detection limit $(\bar{c}_d)_n$, for $P_{10} = 0.01$. Function $P_{11} = f(k_d)$ is obtained from eq 3 and tables of the Laplace function. For example, for $k_d = 2$, $P_{10} = 0.01$ and $n = 1$, $\Phi(z_k) = 0.5 - 0.01 = 0.49$, $z_k = 2.32$, $z_d = k_d - 2.32 = -0.32$, $\Phi(z_d) = 0.13$, $P_{11} = 0.5 - 0.13 = 0.37$. For $n = 2$, $z_k + z_d = k_k \sqrt{2} = 2.83$, $z_d = 2.83 - 2.32 = 0.51$, $\Phi(z_d) = 0.20$, $P_{11} = 0.5 + 0.20 = 0.70$.

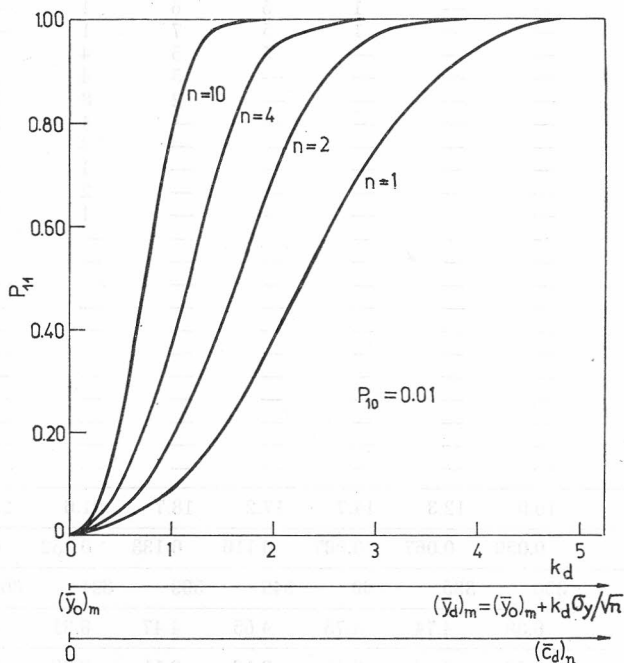


Figure 2. Displacement of the detection characteristic towards lower concentrations as a function of the number of measurements for a certain values of false detection probability $P_{10} = 0.01$.

It is observed that in passing from a single measurement to the mean of several measurements for a given value of k_d , the P_{11} true detection probability increases, hence the detection limit is lowered.

We shall present below some applications. Thus from the processing of a number of standardisation data (Table 1) obtained by repeated analyses of several standard steel samples ($c_{Mn} \leq 0.254\%$) on an automatic emission spectrometer, the following results are obtained:

TABLE I

Values of the Analytical Signal (Digits) of Eight Standard Steel Samples, on an Automatic Emission Spectrometer (ARL 31.000 type) Recorded 32 Times each for Manganese. Working Conditions: $\lambda_{Mn} = 263.98$ nm Argon Atmosphere, Intermittent Arc, Recording Time 5 s

y (digits)	Signal frequencies (number of samples)							
7	5	—	—	—	—	—	—	—
8	2	1	—	—	—	—	—	—
9	2	4	—	—	—	—	—	—
10	3	1	1	—	—	—	—	—
11	6	3	3	—	—	—	—	—
12	6	9	2	1	—	—	—	—
13	2	3	1	1	1	—	—	—
14	3	7	7	2	—	1	—	—
15	3	2	7	3	1	—	—	—
16	—	1	4	4	4	2	—	—
17	—	1	5	6	1	—	—	—
18	—	—	1	5	6	1	—	—
19	—	—	1	5	7	1	—	—
20	—	—	—	5	5	4	1	—
21	—	—	—	—	5	4	2	—
22	—	—	—	—	2	8	1	—
23	—	—	—	—	—	4	2	—
24	—	—	—	—	—	3	5	—
25	—	—	—	—	—	1	6	—
26	—	—	—	—	—	2	4	—
27	—	—	—	—	—	1	8	1
28	—	—	—	—	—	—	1	2
29	—	—	—	—	—	—	2	—
30	—	—	—	—	—	—	—	4
31	—	—	—	—	—	—	—	5
32	—	—	—	—	—	—	—	2
33	—	—	—	—	—	—	—	10
34	—	—	—	—	—	—	—	3
35	—	—	—	—	—	—	—	2
36	—	—	—	—	—	—	—	2
37	—	—	—	—	—	—	—	1
38	—	—	—	—	—	—	—	—
\bar{y} (digits)	10.9	12.3	14.7	17.2	18.7	21.6	25.2	32.3
% Mn	0.059	0.067	0.087	0.110	0.133	0.152	0.190	0.254
y (digits)	350	395	469	549	599	691	807	1033
s_y^2	6.38	4.74	4.75	4.65	4.47	8.31	4.99	4.68
s_y	2.53	2.18	2.18	2.16	2.11	2.88	2.24	2.16

$(\bar{y}_o)_m = 5.0$, $b = 107.1$, $\sigma_y = 2.1$. The calibration function will be:

$$y = 5.0 + 107.1 c \quad (11)$$

For $P_{10} = 0.025$, $z_k = 1.96$, and from (4) the detection level is: $y_k = 5.0 + 1.96 \times 2.1 = 9.12 \approx 9$ digits when the decision is reached on the basis of a single measurement, and $(\bar{y}_k)_n = 5.0 + 1.96 \times 2.1/\sqrt{n} = 5.0 + 4.12/\sqrt{n}$ digits, when the decision is formulated on the basis of the mean of n measurements.

In conclusion, the presence of mangan will be accepted (positive samples) based on a single measurement, when $y > 9$ digits, and when $y > 5.0 + 4.12/\sqrt{n}$ digits based on the mean result of n simple measurements, respectively.

If $P_{11} = 0.975$, $z_d = 1.96$, and from (5), $\bar{y}_d = 5.0 + (1.96 + 1.96) \times 2.1 = 13.23$ digits and $c_d = 0.0770\%$ Mn in the case of a single measurement, and $(\bar{y}_d)_m = 5.0 + 3.92 \times 2.1/\sqrt{n}$ digits (eq 3) and $(\bar{c}_d)_n = 0.077/\sqrt{n} \%$ Mn from the mean of n measurements. In the conditions for obtaining the equation (11), namely $n = 32$, $(\bar{y}_d)_m = 6.45$, so that $(\bar{c}_d)_n = 0.077/\sqrt{32} = 0.014\%$ Mn.

The confidence interval of the detection limit estimated from mean values is obtained as in the case of singular values³, by intersection of the confidence interval of the detection signal $(\bar{y}_d)_m$ with the dispersion band of the calibration function delimited by two hyperbolas.

The Use of the Sum Values

Considering that the sum values Σy have a normal distribution too, eq 4 (the detection level) becomes:

$$(\Sigma y)_k = \bar{\Sigma y}_o + z_k \sigma_{\Sigma y} \quad (12)$$

and eq 5 (the detection signal), becomes (the approximation $\sigma_{\Sigma y_o} \approx \sigma_{\Sigma y}$, namely $\sigma_{y_o} \approx \sigma_y$ is accepted):

$$(\bar{\Sigma y})_d = \bar{\Sigma y}_o + (z_k + z_d) \sigma_{\Sigma y} = \bar{\Sigma y}_o + k_d \sigma_{\Sigma y} \quad (13)$$

Further, the detection limit $(\bar{c}_d)_{\Sigma y}$ will be obtained from the calibration function (6), which will have the form:

$$\Sigma y = \bar{\Sigma y}_o + nbc \quad (14)$$

Since in the case of sum values $\sigma_{\Sigma y} = n \sigma_{(\bar{y})_n}$ and because $\bar{\Sigma y} = n(\bar{y})_n$, from eq 13 and 14 it results

$$(\bar{c}_d)_{\Sigma y} = \frac{k_d \sigma_y}{b \sqrt{n}} = \frac{\bar{c}_d}{\sqrt{n}} = (\bar{c}_d)_n \quad (15)$$

In this way the same value is obtained for the detection limit whether the mean values $(\bar{y})_n$ or the sum values $\bar{\Sigma y}$, are used. Thus, using data from Table 1 the following results are obtained: $\bar{\Sigma y}_o = 161$, $nb = 3428$, $s_{\Sigma y} = 11.9$, so that the estimate form of the calibration function (14) will be:

$$\Sigma y = 161 + 3428 c \quad (16)$$

Taking the same statistical conditions as in the case of singular values, namely: $P_{10} = 0.025$, $\Phi(z_k) = 0.475$, $z_k = 1.96$, $P_{11} = 0.975$, $\Phi(z_d) = 0.475$, $z_d =$

$\bar{c}_d = 1.96$, $k_d = 3.92$, from eq 13 there results $(\bar{\Sigma y})_d = 208$ and from eq 16 is obtained $(\bar{c}_d)_{32} = 0.014\%$ Mn $= (c_d)_{n=32}$.

To verify the approximation $\sigma_{y_o} \approx \sigma_y$, respectively $s_{y_o} \approx d_y$ (sample values), it is necessary to know the dependence:

$$s_y^2 = f(c) \tag{17}$$

The values $s_{y_o}^2$ for $c = 0$ and s_y^2 for $c = (\bar{c}_d)_n$ are calculated. Using F-test the hypothesis $H_o : s_{y_o}^2 = s_y^2$ will be verified.

With data from Table I and eliminating the value $s_y^2 = 8.31$, using the least squares method, there results:

$$s_y^2 = 5.40 + 3.51 c \tag{18}$$

so that for $c = 0$, $s_{y_o}^2 = 5.40$ and for $c = (c_d)_{32} = (c_d)_{32} = 0.014$, $s_y^2 = 5.45$.

Applying F-test, namely $s_y^2/s_{y_o}^2 = 1.01$, $F_{0.95;31,31} = 1.84$, the hypothesis $H_o : s_{y_o}^2 = s_y^2$, and $s_{y_o} = s_y$, is accepted.

THE USE OF SEQUENTIAL MULTITUDE VALUES (SEQUENTIAL STATISTICAL TESTS)

The basic feature of sequential tests is that the number of measurements necessary for decision-making, depends on the result of the observations. In other words, the number of measurements itself plays the role of a random variable.

We describe below the utilization of two sequential tests.

The Sequential Probability Ratio Test (Wald)⁸ (Sequential Sum Values)

From the two probabilities P_{11} (true detection) and P_{10} (first kind of false detection), and the relations $P_{11} + P_{01} = 1$ and $P_{10} + P_{00} = 1$, we calculate the ratios:

$$A = \frac{P_{11}}{P_{10}} = \frac{1 - P_{01}}{P_{10}} \text{ and } B = \frac{P_{01}}{P_{00}} = \frac{P_{01}}{1 - P_{10}} \tag{19}$$

on the basis of which the hypothesis H_o (component is absent) will be accepted if:

$$\sum_{i=1}^n y_i \leq \frac{\sigma_y^2 \ln B}{\bar{y}_d - \bar{y}_o} + n \frac{\bar{y}_d + \bar{y}_o}{2}, \tag{20}$$

or hypothesis H_1 (component is present) will be accepted if:

$$\sum_{i=1}^n y_i \geq \frac{\sigma_y^2 \ln A}{\bar{y}_d - \bar{y}_o} + n \frac{\bar{y}_d + \bar{y}_o}{2}, \tag{21}$$

or the experiment will continue if:

$$\frac{\sigma_y^2 \ln B}{\bar{y}_d - \bar{y}_o} + n \frac{\bar{y}_d + \bar{y}_o}{2} < \sum_{i=1}^n y_i < \frac{\sigma_y^2 \ln A}{\bar{y}_d - \bar{y}_o} + n \frac{\bar{y}_d + \bar{y}_o}{2} \tag{22}$$

For example, by processing a number of standardisation data obtained by recording 13 times the signals from 5 standard steel samples on an automatic

emission spectrometer, the following relation was found between the molybdenum content (for $c_{Mo} \leq 0.0102\%$) and the analytical signal (in digits).

$$y = 32.36 + 848 c \tag{23}$$

for which $\sigma_y^2 = 1.85$, and $\sigma_y = 1.36$. A steel sample was subjected to repeated measurements under similar analytical conditions to those providing the relation (23), to establish whether it contained Mo: $c \geq 0.001\%$ (hypothesis H_1) or not (hypothesis H_0).

Considering eqs 5 and 23, one obtains: $\bar{y}_d = 33.21$ digits for $c = 0.001\%$. Taking $P_{10} = 0.025$, $P_{11} = 0.975$, then $A = 39$ and $B = 0.0256$, so that eqs (20), (21) and (22) become:

$$\sum_{i=1}^n y_i \leq -8.0 + 32.8 n \text{ (hypothesis } H_0 \text{ is accepted)} \tag{24}$$

$$\sum_{i=1}^n y_i \geq 8.0 + 32.8 n \text{ (hypothesis } H_1 \text{ is accepted)} \tag{25}$$

$$-8.0 + 32.8 n < \sum_{i=1}^n y_i < 8.0 + 32.8 n \text{ (the experiment is continued)} \tag{26}$$

The results are listed in Table II and Figure 3.

TABLE II

Formulation of the Sequential Detection Decisions for a Steel Sample in Relation to Hypothesis H_0 (no Molybdenum Present) and Hypothesis H_1 ($c \geq 0.001\%$ Mo)

n	y	Sequential probabilities ratio test			Sequential frequentometric probabilities ratio test		
		$\sum_{i=1}^n y_i$	Decision level for hypothesis		N	Decision level for hypothesis	
			H_1	H_0		H_1	H_0
1	35	35	40.8	24.8	1	3.74	-3.14
2	31	66	73.6	57.6	1	4.04	-2.84
3	33	99	106.7	90.4	1	4.34	-2.54
4	32	131	139.2	123.2	1	4.64	-2.24
5	34	165	172.0	156.0	2	4.94	-1.94
6	32	197	204.8	188.8	2	5.24	-1.64
7	35	232	237.6	221.6	3	5.54	-1.34
8	34	265	270.4	254.4	4	5.84	-1.04
9	31	297	303.2	287.2	4	6.14	-0.74
10	35	332	336.0	320.0	5	6.44	-0.44
11	34	336	368.8	352.8	6	6.74	-0.14
12	32	398	416.6	385.6	6	7.04	0.16
13	33	431	434.4	418.4	6	7.34	0.46
14	34	465	467.2	451.2	7	7.64	0.76
15	32	497	500.0	484.0	7	7.94	1.06
16	35	532	532.8	516.8	8	8.24	1.36
17	35	567	563.6	549.6	9	8.54	1.66
18	34	601	596.4	582.4	10	8.84	1.96
19	35	636	629.6	615.1	11	9.14	2.26

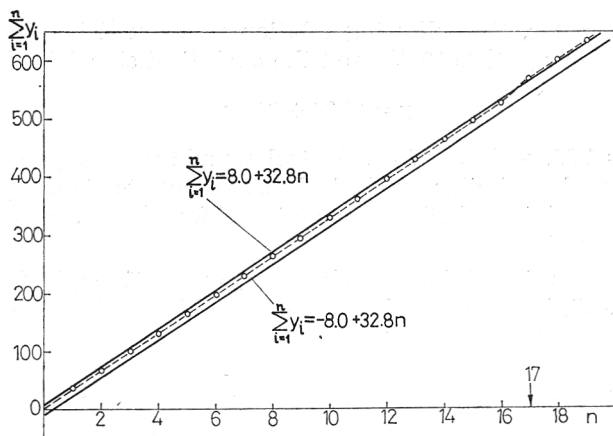


Figure 3. Sequential detection (molybdenum in steel) by means of the probability ratio test (sequential sum values).

It can be seen that the inequality (25) is satisfied starting from the 17th measurement and hence hypothesis H_1 is accepted, i. e. $c \geq 0.001\%$ Mo. For $P_{10} = 0.025$ and $P_{11} = 0.975$, $k_d = 3.92$ and therefore in agreement with equation (5), $\bar{y}_d = 32.36 + 3.92 \times 1.36 = 37.19$ and hence $\bar{c}_d = 0.0059\%$ Mn equation (23), so that from equation (9) it is seen that for the detection of a concentration of 0.001% Mo, $n = (0.0059/0.001)^2 = 34.8 \approx 35$ individual measurements would be necessary.

Sequential-Frequentometric Probability Ratio Test (Wald)⁸ Sequential Sum Frequencies

This test is applied to binomial distributions and therefore values 0 and 1 with probabilities P_0 and P_1 are considered as the random variable, depending on whether the hypothesis H_0 or H_1 is true. For this purpose, a certain value y_r of the signal y is considered as a reference, so that the discrete random variable will be 1 if $y > y_r$ (hypothesis H_1) or 0 if $y < y_r$ (hypothesis H_0).

If the total number of results is n and the number for which $y > y_r$ is N , the hypothesis H_0 will be accepted if:

$$N \leq B' = \frac{\log B - n \log \frac{1 - P_1}{1 - P_0}}{\log \frac{P_1}{P_0} - \log \frac{1 - P_1}{1 - P_0}} \tag{27}$$

the hypothesis H_1 will be accepted if:

$$N \geq A' = \frac{\log A - n \log \frac{1 - P_1}{1 - P_0}}{\log \frac{P_1}{P_0} - \log \frac{1 - P_1}{1 - P_0}} \tag{28}$$

and the experiment will continue if:

$$B' < N < A' \tag{29}$$

A and B are calculated from (19) and the probabilities $P_0(y < y_r)$ and $P_1(y > y_r)$ for the hypotheses H_0 and H_1 are obtained by means of the normal deviates $z_0 = (y_r - \bar{y}_0)/\sigma_{y_0}$ and $z_1 = (y_r - \bar{y}_d)/\sigma_y$, respectively, and the corresponding Laplace function values so that $P_0 = 0.5 - \Phi(z_0)$ and $P_{11} = 0.5 - \Phi(z_1)$.

Applying detection by this test for the model just described (eq 23), admitting the approximation $\sigma_{y_0} \approx \sigma_y$ and the same P_{10} and P_{11} values, than if the reference signal $y_r = 33.5$, $P_0 = 0.18$ and $P_1 = 0.42$. The following inequalities result for $c = 0.001\%$ Mo, and $\bar{y}_d = 33.21$:

$$N \leq -3.44 + 0.30 n \text{ (hypothesis } H_0 \text{ is accepted)} \quad (30)$$

$$N \geq 3.44 + 0.30 n \text{ (hypothesis } H_1 \text{ is accepted)} \quad (31)$$

$$-3.44 + 0.30 n < N < 3.44 + 0.30 n \text{ (the experiment continues)} \quad (32)$$

The results are presented in Table II and Figure 4.

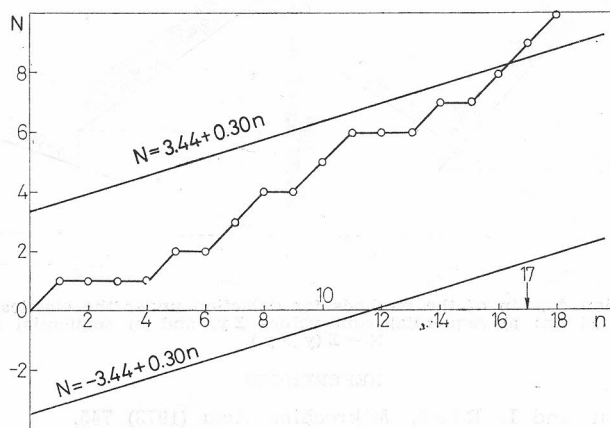


Figure 4. Sequential detection (molybdenum in steel) by means of the frequentometric probability ratio tests (sequential sum frequencies).

It is again found that the hypothesis H_1 is accepted starting from the 17th measurement. A fact known from the theory of sequential analysis is observed, namely that the sequential detection method is more efficient than the detection method using a fixed number of measurements of the analytical signal and requires fewer measurements.

CONCLUSIONS

The extent of the detection function of an analytical system under the classical detection limit (use of individual signals), by using the fixed multitude signals is given in Figure 5 and in the case of using the sequential sum values, in Figure 6.

Using these procedures, the analytical detection can be applied to as low concentrations as desired, provided that the multitude of individual values of the signal be sufficiently high.

The detection under the classical detection limit can be achieved frequentometrically too, by using the sum of the frequencies $y > y_0$, as the analytical signal.^{9,10}

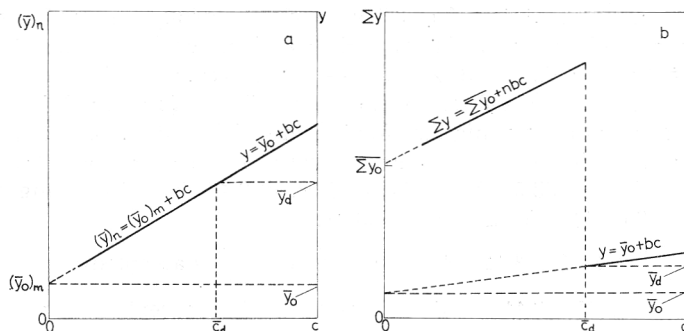


Figure 5. Application domain of the methods for detection under the classical detection limit (single values), based on: a) mean values $(\bar{y})_n$, and b) sum values Σy .

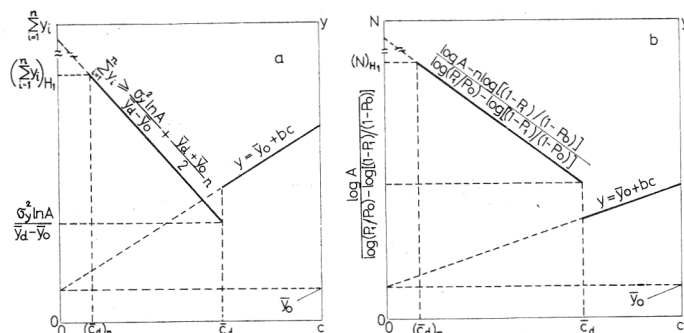


Figure 6. Application domain of the methods for detection under the classical detection limit (single values) based on: a) sequential sum values Σy , and b) sequential sum frequencies $N = \Sigma (y > y_c)$.

REFERENCES

1. C. Liteanu and I. Rică, *Mikrochim. Acta* (1973) 745.
2. C. Liteanu and I. Rică, *Mikrochim. Acta* (1975) 311.
3. C. Liteanu and I. Rică, *Pure Appl. Chem.* **44** (1975) 535.
4. J. Neyman and E. S. Pearson in C. W. Helstrom, *Statistical Theory of Signal Detection* Pergamon Press, New York 1960, p. 68.
5. P. D. Lark, B. R. Craven, and R. C. L. Bosworth, *The Handling of Chemical Data*, Pergamon Press, Oxford 1969, p. 335.
6. H. Kaiser, *Z. Anal. Chem.* **209** (1965) 1.
7. H. Kaiser, *Z. Anal. Chem.* **216** (1968) 80.
8. A. Wald, *Ann. Math. Stat.* **10** (1939) 299.
9. C. Liteanu and R. Alexandru, *Mikrochim. Acta* (1968) 639.
10. C. Liteanu, *Mikrochim. Acta* (1970) 715.

SAŽETAK

Mogućnosti proširenja detekcijskih funkcija analitičkog sistema na niske koncentracije

C. Liteanu, I. Rică i E. Hopârteanu

Povećanje detekcijskog kapaciteta u nekom analitičkom sustavu je moguće samo pomoću mnogostruke vrijednosti pojedinačnog signala. To se postiže bilo fiksnom bilo sekvencijskom mnogostrukom vrijednošću. Ako je mnogostrukost pojedinačnog signala dovoljno visoka tada se mogu određivati po volji niske koncentracije.

DEPARTMENT OF ANALYTICAL CHEMISTRY
UNIVERSITY OF CLUJ-NAPOCA
CLUJ-NAPOCA, ROMANIA

Prispjelo 9. kolovoza 1976.