Theory and practice of limits of detection, quantitation and decision in instrumental analyses

Akira Kotani, Ph.D.
Tokyo University of Pharmacy and Life Sciences
Yuzuru Hayashi, Ph.D.
Institute for FUMI Theory

Introduction

Widely and for several decades, detection limits (DLs) have been recognized as a figure of merit of vital importance and utilized in every discipline of analytical chemistry to ensure statistical reliability and practical suitability of analyses. A fundamental quantity underlying DLs is the standard deviation (SD) of response variables or measurements. First, this article introduces definitions and concepts of DLs and secondly, focuses on the externalization of DLs, i.e., how to obtain SD estimates in practice. Discussed are the advantages and disadvantages of methods for estimating SDs in instrumental analyses: the statistical approach with repeated experiments of real samples and an uncertainty theory, called FUMI theory (Function of Mutual Information), which can dispense with repeated measurements.

1. Detection limits and decision limits

DLs and decision limits are defined in ISO 11843-7[1] and JIS Z 8462-7[2]. A DL is a target quantity of detection, but not a criterion for judging whether or not a target material is detected in an analytical system. The presence of a material in a sample is judged by a decision limit.

1.1 Necessity of decision limits

In general, as a concentration of a target material decreases, its signal becomes smaller and finally makes it difficult to distinguish its shape from fluctuations of noise. A DL measurement, y_D, is often defined with an S/N (signal-to-noise ratio) as S/N = 3 for the sake of convenience. Let y be a

measurement or response, and x be a concentration (final quantity of analyses). A DL measurement, y_D, can be transformed into its corresponding DL concentration, x_D, through a calibration function ($y = f(x)$). If a measurement, y, is much more than y_D, an analyst can say that a material at a concentration of x_D is detected. However, the question now is not only the presence or absence of a material, but also a probability with which the material can be detected, when its measurement, y, is close to y_D.

If analytes of the DL concentration, x_D, called DL samples, are measured repeatedly, the probability of an observable y (or estimable x) rising above the DL is equal to that of falling below the DL. As long as a decision is made by comparing y with y_D or x with x_D, the probability of detecting a DL material in DL samples is 50 %. In order to achieve a higher probability of detection, e.g. 95 %, we need to introduce another standard, called a decision limit or critical value, y_C, which is less than y_D and x_D. A new rule of detection is that if $y<y_C$, a material of the concentration x_D is not detected and if $y \geqq y_C$, it is detected. The detection limit, y_D (i.e. x_D) is a target quantity for detection and the decision limit, y_C, is a criterion for judging detection.

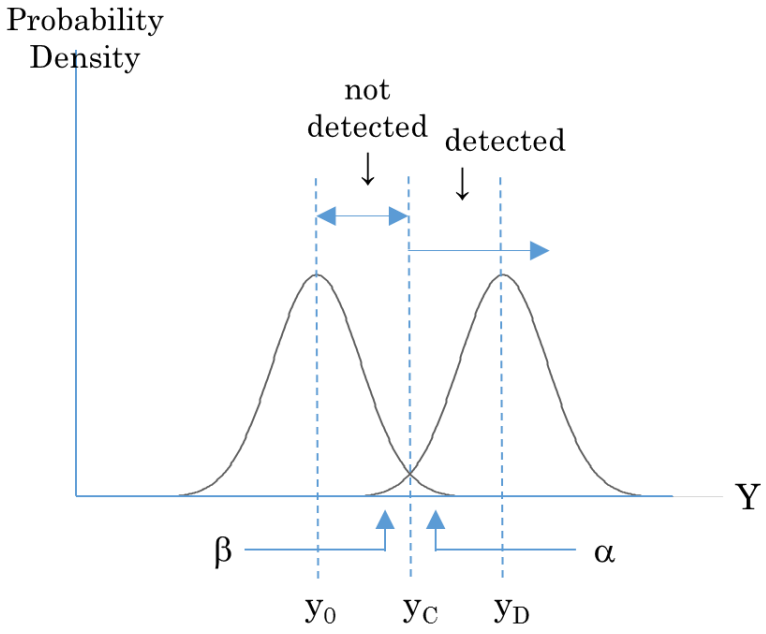

Figure 1　A concept of limits of detection and decision[1,2)]

The right curve of Fig. 1 illustrates a normal distribution of measurements when DL samples at a concentration of x_D are repeatedly measured. Its average is y_D and corresponds to x_D through a calibration function ($x_D = f^{-1}(y_D)$). The probability of a measurement, y, falling below y_C is β (= 5 %). We can see that as a criterion for judging detection decreases from y_D to y_C (see the last section), the probability of detection increases from 50 to 95 %. However, a problem occurs that blank samples can result in a measurement more than y_C. The left curve of Fig. 1 shows a distribution of blank measurements and the probability of a blank measurement exceeding y_C is α (= 5 %).

α is referred to as the probability of an error of the first kind (false positive) and β the probability of an error of the second kind (false negative). In a normal distribution with zero mean and SD of σ, the probability of a variable being more than $+1.65\sigma$ or less than -1.65σ is 5 %.

As shown in Fig. 1, the limits of decision and detection are defined, respectively, as

$$y_C = 1.65\sigma_y \quad (1)$$

$$y_D = 3.3\sigma_y \quad (2)$$

where σ_y denotes the SD of a normal distribution of measurements. It is assumed here that the distribution of measurements of blank samples is the same as that of DL samples (homoscedastic assumption). Since the concentration of blank samples (= 0) is close to x_D, the homoscedasticity is reasonable.

Let α and β be both 5 %. We can safely say that with a risk of at most 5 %,

- if $y < y_C$ or $x < x_C$, the analyte of the concentration x_D is not detected;
- if $y \geq y_C$ or $x \geq x_C$, the anaylte of the concentration x_D is detected.

The criterion, y_C, and rules, $<$ and $\geq$, are both essential for the concepts of detection limits. We should note that there is a possibility of detection, even if a measurement is less than the detection limit (see Fig. 1).

A target of DLs is usually a sample of an unknown concentration. When its measurement, y, is obtained, discussion of its corresponding concentration is possible through a calibration function, but makes no sense from the viewpoints of DLs. The question now is which state an analytical system characterized by the measurement, y, belongs to, blank or DL (see the distributions of Fig. 1). Other concentrations than a blank (basic state) or DL (reference state) are outside the purview of the definitions of DLs.

The concepts of limits of decision and detection (eqs 1 and 2) are very simple and easy to understand. However, analysts have to replace the theoretical SD, σ_y, (population SD) of eqs 1 and 2 with an estimate of σ_y (sample SD) to use these definitions in practice. The next section takes this subject.

2. Estimation of SD

Two methods for evaluating the SD, σ_y, of eqs 1 and 2 are described in this section.

2.1 Repeated measurements

The simplest equation for SD estimation is

$$\mathrm{SD} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1}} \qquad (3)$$

where y_i denotes the i-th measurement and $\bar{y}$ the average of n measurements. This SD estimate (eq 3) varies depending on a series of n experiments, though conducted under exactly the same experimental conditions. This scattering of SD estimates can be assessed by a well-known statistical theory, called the chi-square distribution.

Figure 2 shows a dependence of scattering of SDs estimated by eq 3 on the number of repeated measurements, n. Hereinafter, n consecutive measurements are called a series of measurements, which leads to an SD estimate. The 95 % confidence intervals imply that among 100 SD estimates which result from 100 series, 95 estimates fall between the upper and lower limits of the intervals and 5 estimates are outside the limits. For example, if n = 6, the 95 % intervals range

from 0.4 to 1.6 around the true value (= 1), indicating that 95 % of all the SD estimates scatter between ± 60 % of the true value. For ± 20 % SD scattering, more than 40 measurements are required (see Fig. 2).

In case of 6 measurements (n = 6), due to ± 60 % confidence intervals, an SD estimate can be half of a second estimate, though under the same experimental conditions. Then, an estimate of σ_y varies from series to series accordingly, and so do the limits of detection and decision defined by eqs 1 and 2.

The chi-square distribution shown in Fig. 2 points out a tradeoff between a small number of repeated experiments and high statistical reliability of SD estimates. This is a mathematical rule without exceptions. Furthermore, the repetition of experiments with real samples is unfavorable from the scientific and economical viewpoints. An alternative is an S/N as described in subsection 1.1. Unfortunately, however, the S/N cannot tell the probabilities of α and ß theoretically in most cases. That is, the numerical judgement of detection, as defined in eqs 1 and 2, is impossible with S/Ns.

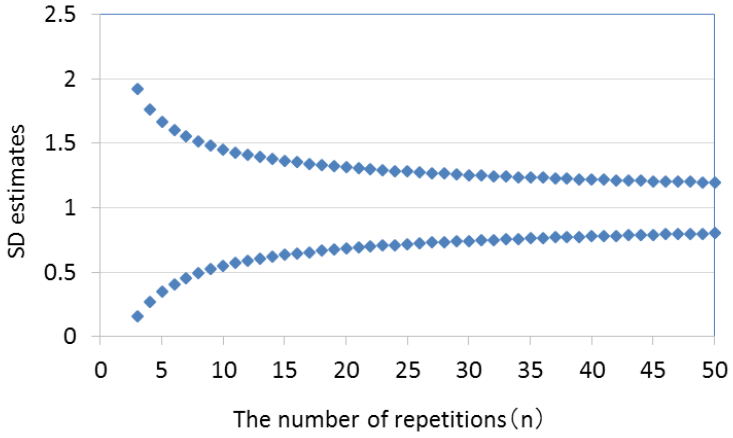

Figure 2 Dependence of 95 % confidence intervals of SD estimates on the number of repeated measurements (n)

$$\sqrt{\frac{\chi_\beta^2}{n-1}} \geq \frac{s}{\sigma} \geq \sqrt{\frac{\chi_\alpha^2}{n-1}}$$

denotes the 95 % confidence intervals. χ_α^2 and χ_β^2 denote the values of χ^2 with probabilities, 0.025 and 0.975, respectively.

2.2 FUMI theory

The FUMI theory can compensate for the theoretical shortcoming of S/Ns mentioned in the last subsection. The FUMI theory is based on a probabilistic model and can provide accurate SD estimates, σ(k), without repeated measurements of real samples. However, its applicability is restricted within some fields of analytical chemistry because of its strict foundations of mathematics. On the other hand, the statistical method (eq 3) has no practical limitation of use, but accurate SD estimates cannot be acquired, until a large number of repeated experiments are conducted. The FUMI theory and statistical method are complimentary in analytical chemistry.

Assumptions and principles on which the FUMI theory is constructed are:.

1. when a sample concentration is low at or near a DL, the most predominant source of measurement errors is background noise in instrumental output;
2. an error of an area measurement is the area created by the noise over the edge-to-edge domain of a signal;
3. the noise is modelled on mixed random processes of white noise and the first order autoregressive process, AR(1) (one of the Markov processes).

Therefore, the FUMI theory cannot apply to situations which are beyond the above assumptions and principles, e.g. other methods than instrumental analyses, quantitative analyses with large sampling errors, etc.

In authors' experience, the FUMI theory is effective for isocratic HPLC, gradient HPLC, internal methods of HPLC. GC and GC/MS are future issues.

The FUMI theory and statistical method pursue the common goal of a population SD, but their means are different. The latter assesses a feature (eq 3) of a population from an ensemble of experimental results, called a sample space. This ensemble can be constructed only by repeated experiments. In the FUMI theory, however, the population feature is described in terms of a fundamental equation (SD = σ(k)), the parameters of which are stochastically estimated (parametrized) from noise and signals of instrumental output. That is, a measurement SD is directly estimated from the cause of measurement errors (noise). Therefore, the FUMI theory can dispense with repeated measurements. Of course, it refers to a result from repeated measurements.

3. Limits of detection, quantitation and decision in practice

A complicated mathematical theory would not be practical without convenient software. The parametrization described in subsection 2.2 cannot actually be performed as easily as with spreadsheet software. This section takes commercial software as a technical tool. Figures 4 to 6 are all exports of the software.

3.1 Software for uncertainty of measurements

TOCO19 (Total Optimization of Chemical Operations) is a piece of software for automatically providing SD estimates for signals in a chromatogram. Among its output are limits of detection, quantitation and decision and precision profiles.

Figure 3 shows an overview of TOCO19 analyses. First, noise and signals are distinguished and secondly, necessary information for the analyses is extracted from them and analyzed by the FUMI theory to yield a result of σ(k).

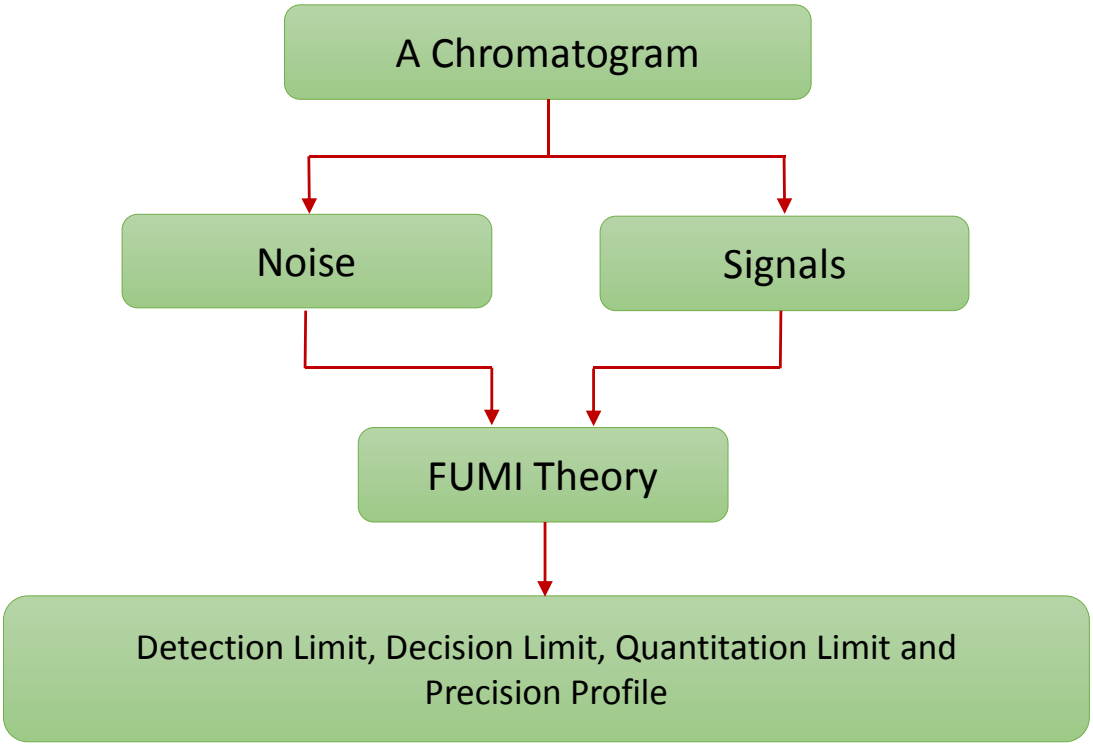

Figure 3 An overview of TOCO19

Figure 4 illustrates a chromatogram of organic acids in an HPLC system with a UV absorption detector. Here, an object for quantitative analysis is the third peak (acetic acid). The second peak is made up of overlapping peaks: one is malic acid and the other is unknown. The blue blocks denote signals automatically recognized by TOCO19. The line segments at Y = –0.15 are noise regions used by TOCO19 noise analyses and those at Y = –0.1 are signal regions eliminated from the noise analyses.

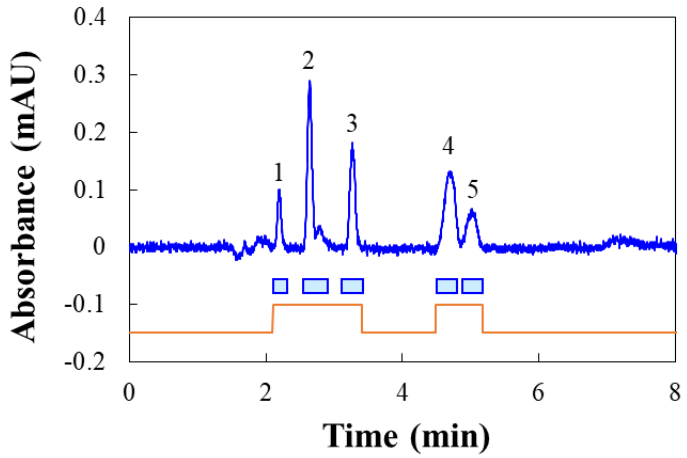

Figure 4　A chromatogram of organic acids in an HPLC system[3)]

The signals are (from left) system, malic acid, acetic acid, citric acid and succinic acid (10 mg/L each). Observed values are: σ_w = 0.00357, σ_m = 0.00176, ϕ = 0.246.

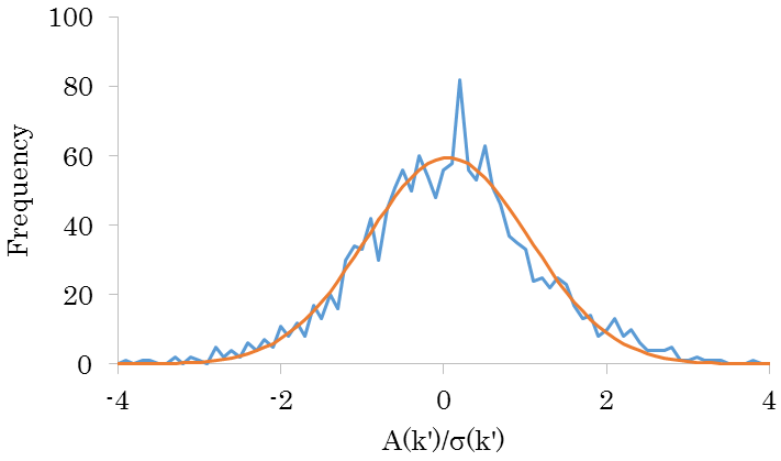

Figure 5　A histogram of measurement errors (areas created by noise)

Figure 5 shows a histogram of areas (A(k')) created by noise. The zigzag line denotes a distribution of noise areas observed over the noise regions of Fig. 4 and the smooth line a theoretical line of a normal distribution. The observed line is in excellent agreement with the theoretical one. This agreement is an evidence for successful uncertainty prediction of TOCO19, because the distribution of noise areas of Fig. 5 is equivalent to that of measurement errors as long as assumption 2 of subsection 2.2 holds true. The width, k', of the noise areas, A(k'), is fixed at 300 data points and the X-axis of Fig. 5 is standardized according to SD, σ(k'), of noise areas. The constant, k', is used for the noise analyses only and different from k of σ(k) which is a width of a target signal.

3.2 Signals for limits of detection, quantitation and decision

TOCO19 provides not only numerical values for limits of detection, quantitation, but also visual information about them. Figure 6 illustrates S/N aspects for the acetic acid peak in the quantitative analysis of Fig. 4. Its signal shape is approximated by a Gauss peak (normal distribution). The width, k, of a real peak corresponds to ± 3σ of the normal distribution. Noise in the noise regions of Fig. 4 is automatically analyzed by TOCO19 algorithms. The limits of decision and detection are defined by eqs 1 and 2, but the quantitation limit (QL) is regarded as $10\sigma_y$ according to the Japanese Pharmacopoeia. The population SD, σ_y, of eqs 1 and 2 is replaced with the sample SD, σ(k), evaluated by TOCO19.

As is well-known, an S/N is a convenient indicator for DLs and QLs. The S/Ns shown in Figs. 6A and C are close to the widely adopted values, 3 and 10, respectively. It is quite interesting for the distinct methods to yield comparable results. The reason is unknown, but will not be accidental.

Figure 6 shows an averaged signal on noise for the acetic acid peak of Fig. 4. When analytes of the x_D concentration are measured repeatedly, larger and smaller peaks than that of Fig. 6A will appear with the same probabilities (50 %), but smaller peaks than Fig. 6B will appear with only 5 % probability (β). On the other hand, blank samples will create peaks large than the peak of Fig. 6B with 5 % probability (α). In other word, the background noise will create a larger fluctuation than Fig. 6B with 5 % probability (α).

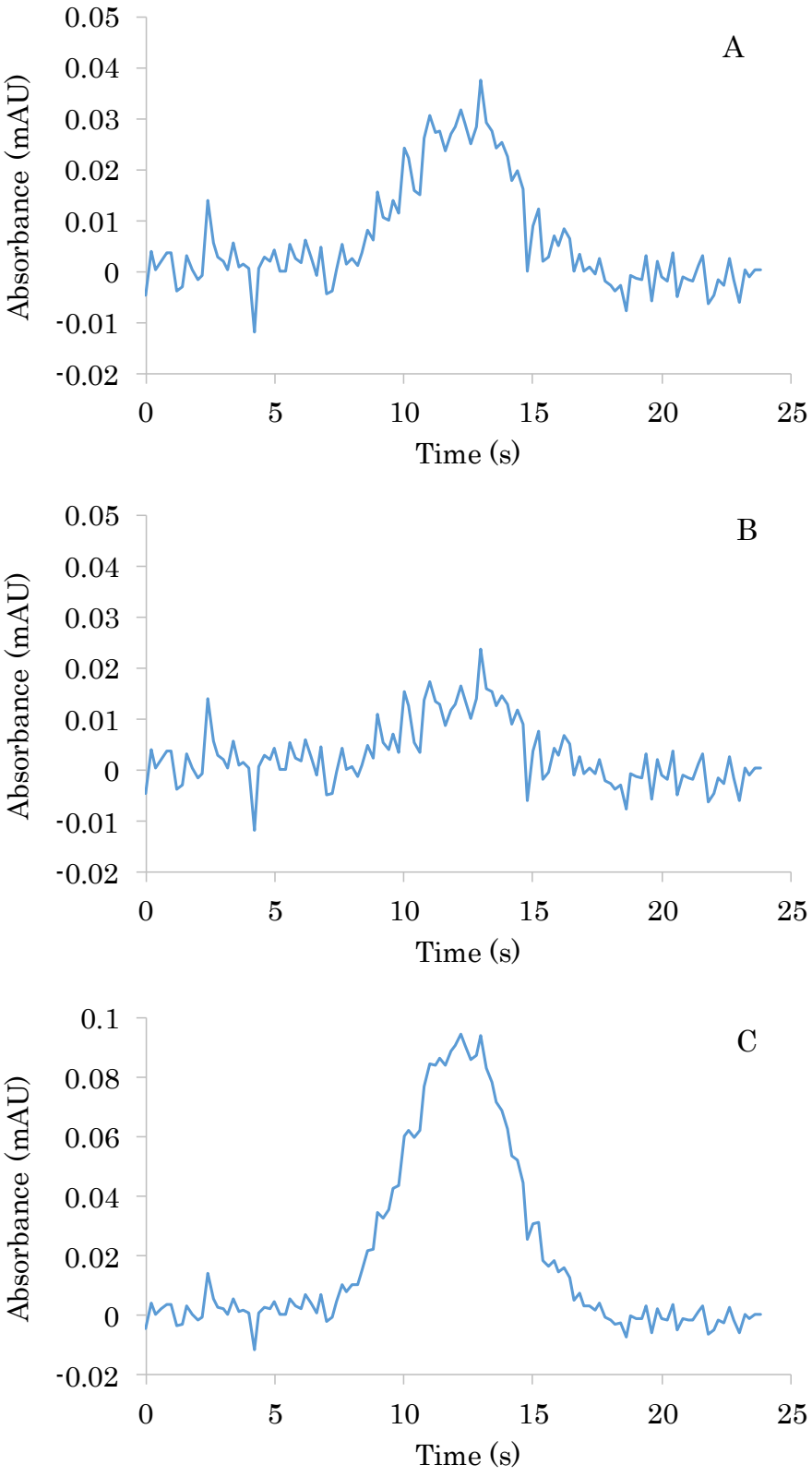

Figure 6 Signals for limits of (A) detection, (B) decision and (C) quantitation

Concentration of acetic acid: A 1.68 mg/L; B 0.84 mg/L; C 5.10 mg/L.

You might think that the signal of Fig. 6B is too small to be a threshold of detection. However, five percent (ß) is not a probability of success, but that of failure. This probability is as high as if you made a mistake once in twenty times at work, you would get depressed. Therefore, a decision limit signal is not far from noise (no work).

4. Precision profiles

A mathematical relationship between a sample concentration and relative standard deviation (RSD) of measurements (see Fig. 7), termed precision profiles by ISO[1)] and JIS[2)], helps survey entire structure of uncertainty in an analytical system. In general, the RSD of measurements decreases with increasing sample concentrations. At high concentrations, however, the RSD converges to a value attributed to other error sources than noise, e.g. volume errors of sample solutions injected into an instrument. Needless to say, near a DL, the major factor of errors is assumed to be background noise. In the precision profile of Fig. 7, the DL (= 1.68 mg/L) and QL (= 5.10 mg/L) can be spotted at the concentrations which correspond to 30 % and 10 % RSDs of measurements, respectively (see below).

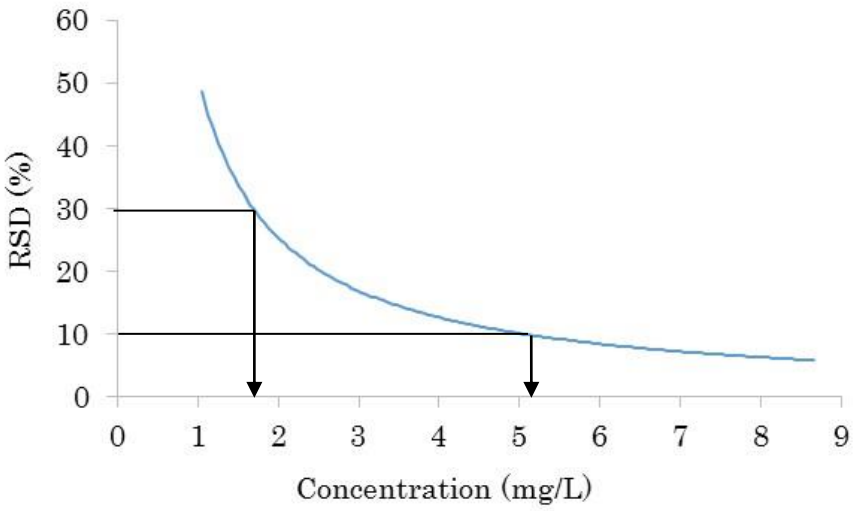

Figure 7　A precision profile for acetic acid in the analysis of Fig. 4

For the acetic acid peak of Fig. 4 (10 mg/L), RSDs of measurements estimated by TOCO19 and repeated experiments were observed to be 6.5 % (n = 1) and 6.4 % (n = 6), respectively (not shown). The 95 % confidence intervals of the latter range from 4.0 to 15.9 % RSD and contain the former RSD value. This statistical test indicates that TOCO19 can apply to the HPLC analyses with ultraviolet detection.

Equation 2 can be rewritten as[2)]

$$\sigma_y / y_D = 1 / 3.3 = 30\ \% \quad (4)$$

Taking into account that an SD divided by an average is an RSD, we can see that when samples of the DL concentration are measured repeatedly, the RSD of measurements is 30 %. In the similar manner, 10 % RSD can be derived for samples of the QL concentration ($QL = 10\sigma_y$).

Details of basic mathematical equations and concepts of TOCO19 can be found in refs. 3 and 4.

References

1) ISO 11843-7, Capability of detection - Part 7: Methodology based on stochastic properties of instrumental noise, ISO, (2012).

2) JIS Z 8462-7, Capability of detection - Part 7: Methodology based on stochastic properties of instrumental noise, JIS, (2018).

3) A. Kotani, S. Tsugu, H. Hakamata, Y. Hayashi, J. Chromatogr. A 1612 (2020)460644.

4) A. Kotani, H. Hakamata, Y. Hayashi, J. Chromatogr. A 1621 (2020) 461077.