

検出限界の定義と実例

FUMI 理論研究所

ある分析方法で検出できる最小の量が検出限界であり、分析化学における最も基本的かつ重要な概念のひとつである。検出限界は検出を誤認する確率（対象物質がサンプルに含まれていないのに検出されたとする誤判定、false positive）と検出を見落とす確率（false negative）に基づいて定義される。判定限界は検出限界量の物質が検出されたかどうかを判定する基準である。定量限界は十分な統計的信頼性をもって定量分析が行える最小量である。本稿は検出限界の考え方を紹介し、HPLC 分析における検出限界 (detection limit、DL)、判定限界 (decision limit)、定量限界 (quantitation limit、QL) の実例を挙げる。

第 1 章 検出限界の一般的定義

検出限界という概念が必要なのは、そもそも測定値がランダムな誤差を含み測定毎に変動するからである。そうでなければ、測定は無限小の量まで有効である。測定値のバラツキが最小の検出量を規定するのである。測定値のバラツキはじかに観測できるのであるから、その度合い（測定値の標準偏差 (SD)）から検出限界を定義することが早道である。しかし、測定値は電流、吸光度など用いた分析方法に特有な単位で記録されるため、この検出限界は一般的な利用に供しない。そこで、検量線 ($y = f(x)$) によって、分析の最終目的量 x に変換した検出限界を一般的な定義とする。本稿では最終目的量 x を濃度とする。

測定値がばらつくのであれば、検量線に変換した濃度推定値 x も同様にばらつく。すると、科学的合理性の観点から、濃度推定値の SD を使って検出限界をダイレクトに定義しようという主張が生まれる。

便宜的に、検出限界は検量線の Y 軸での定義と X 軸での定義が共存することになる。これらの定義では、何れにしても、検出限界は濃度単位 (X 軸の単位) で記述されている。

Y 軸での DL の定義 (y_D)

観測された測定値 y の統計量 (平均と SD) に基づいて Y 軸の DL を定義し、この DL 値を検量線によって変換し、X 軸の DL とする。

X 軸での DL の定義 (x_D)

濃度推定値 x の統計量に基づいて X 軸の DL をダイレクトに定義する。ただし、測定値 y の統計量を濃度推定値 x の統計量に変換する方法を前もって確立しておく必要がある。

確率論でよく知られている正規分布の再生性 (reproducibility) によれば、測定値 y が正規分布に従うならば、検量線 ($y = ax$) で変換した濃度推定値 ($x = y/a$) も正規分布に従う。理論的には、Y 軸での定義と X 軸での定義は同値であり、これらの記述の違いは本質的ではない。本稿では、測定値の分布は正規として話を進める。第 5 章では、HPLC の測定誤さが正規分布に従う実験的証拠を示す。

第2章 1つの確率を用いた定義

検出限界 (y_D と x_D) は、ブランクサンプルの平均を基点として、ブランク測定値または濃度推定値の SD の k 倍として定義される。

$$y_D = y_0 + k\sigma(y_0) \quad (1)$$

$$x_D = x_0 + k\sigma(x_0) \quad (2)$$

ここで、 y_0 と $\sigma(y_0)$ は、それぞれ、ブランクサンプルのくり返し測定値の平均と SD であり、 x_0 と $\sigma(x_0)$ は、それぞれ、 y_0 と $\sigma(y_0)$ を検量線 ($y = f(x)$) で変換した値である。

図1は検量線 ($y = ax$) とブランク測定値の正規分布を示している。測定値のバラツキは Y 軸方向であるから、正規分布は Y 軸方向に描かれている。バラツキの原因は測定、調製、サンプリングなどであるが、DL 濃度くらいの低濃度サンプルの機器測定では、それはバックグラウンドノイズであると考えてよい (後述)。ブランク測定値が図1のようにばらつくと仮定する。測定値のバラツキは正規でありランダムであるから、思いもよらないほど大きな測定値が出現する可能性がある。この出現確率を α 、その閾値を検出限界 y_D と定義する。 α は false positive あるいは第一種の過誤の確率とも呼ばれていて、実際には目的物質が含まれていないのに、含まれていると誤った判断をする確率である。

検出限界 y_D をブランク測定値の SD の 3 倍とすれば ($k = 3$)、検出限界以上のブランク測定値が現われる確率は 0.14% である ($\alpha = 0.14\%$)。つまり、ブランクサンプルの測定から、検出限界以上の値が現われる割合は千回に一回くらいである。一方、検出限界以下の測定値が得られる確率は 99.86% である。

図2はブランク測定値に対応する濃度推定値の正規分布を描いている。図1の検出限界 x_D は Y 軸の検出限界 y_D を検量線で X 軸の値に変換したものだが、図2の検出限界は濃度推定値の分布から、検量線を介さないでずばりと定義される。

式 (1) と (2) で $k = 3$ とすれば、確率 0.14% の過誤 (危険率) をもって、分析者は次のように言うことができる。

ある未知検体を測定した場合、

- 測定値が検出限界より小さければ、「何も検出されなかった」。
- 測定値が検出限界より少し大きければ、「何かが検出された」。

二番目の例では高い確率で、何かが検出されたと言えるが、その「何か」の濃度に言及することは危険な賭けである。たとえば、DL 濃度 x_D のサンプルを測ったとすれば、DL 以上の測定値もそれ以下の測定値もどちらも 50% の確率で出現する。捨ておけない物質の量をコイン投げで卜するのではお話にならない。次の章で示すように、検出限界の濃度にまで言及するには、もう一つの確率 β が要諦である。

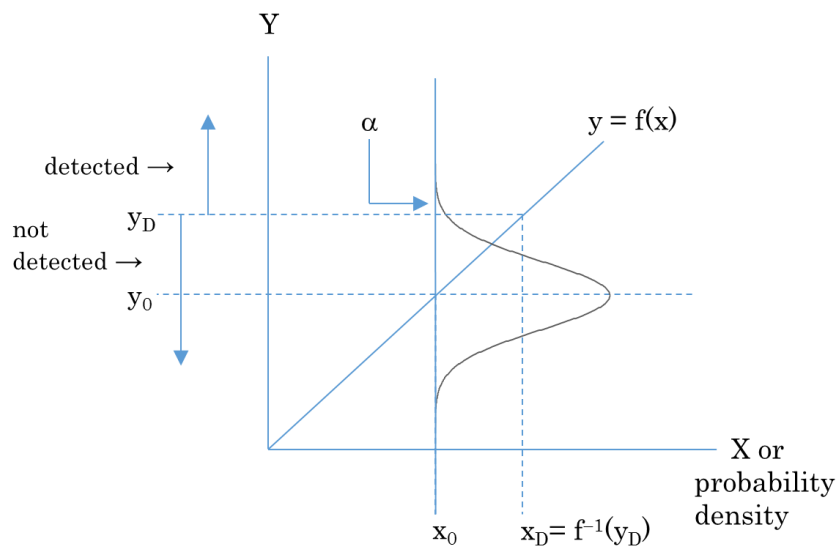


図1 DLのY軸における定義 (α)

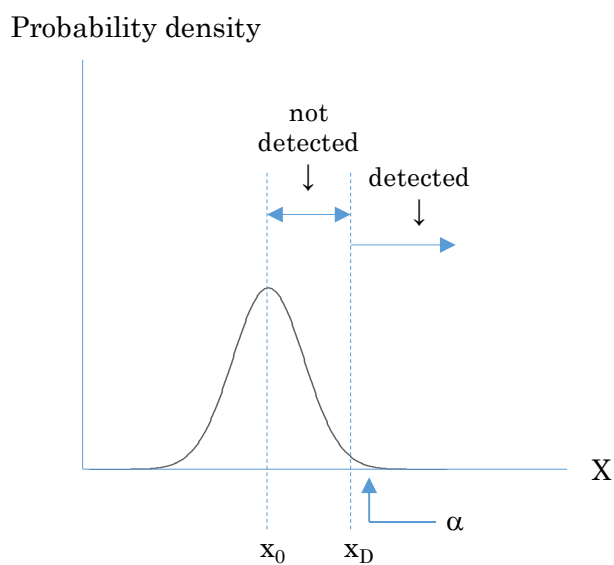


図2 DLのX軸における定義 (α)

第3章 2つの確率を用いた定義

検出限界濃度のサンプル (DL サンプル) の測定値に関する統計量を式 (1) と (2) に付け加えると、次の検出限界が定義される。

$$y_D = y_0 + k\sigma(y_0) + k\sigma(y_D) \quad (3)$$

$$x_D = x_0 + k\sigma(x_0) + k\sigma(x_D) \quad (4)$$

ここで、 $\sigma(y_D)$ と $\sigma(x_D)$ は、それぞれ、DL サンプルのくり返し測定値と濃度推定値の SD である。

図 3 と 4 では、二種類の統計量（正規分布）の導入によって、判定限界（critical value、 y_C と x_C ）が新たに定義されている。判定限界は JIS では限界値、分析化学では decision limit とも呼ばれている。

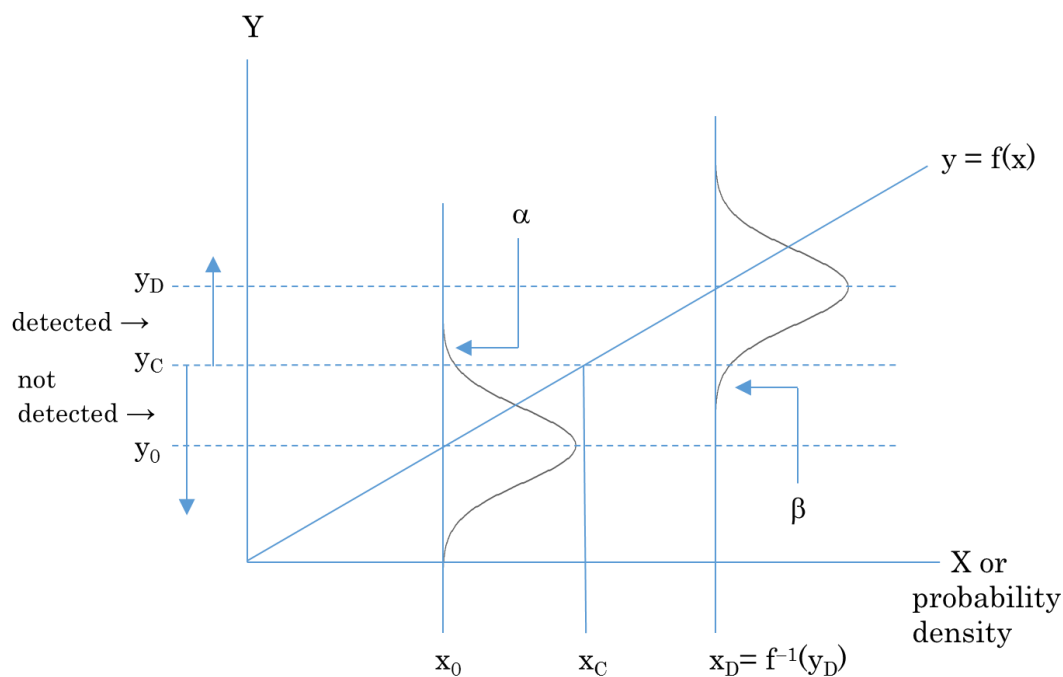


図 3 DL の Y 軸における定義 (α と β)

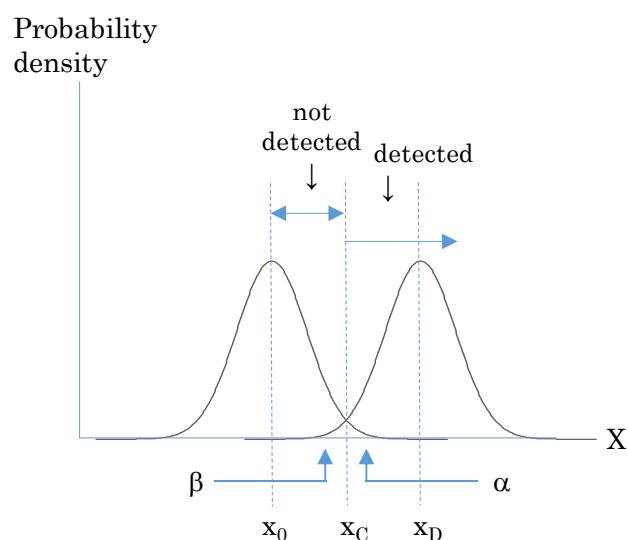


図 4 DL の X 軸における定義 (α と β)

ブランクサンプルを測定した場合、判定限界より大きい測定値が出現する確率が α であり、DL サンプルを測定した場合、判定限界より小さい測定値が得られる確率が β である。確率 β は false negative あるいは第二種の過誤の確率と呼ばれている。

$k = 1.65$ とすれば、 α と β はどちらも 5% である。確率 5% の過誤をもって、分析者は次のように言うことができる。

ある未知検体を測定した場合、

- 測定値が判定限界より小さければ、「DL 濃度のサンプルは検出されなかった」。
- 測定値が判定限界より少し大きければ、「DL 濃度のサンプルは検出された」。

二つの確率 (α と β) を用いることによって、濃度への言及が初めて可能になった。

測定結果が検出限界より小さくても、判定限界より大きければ、DL 濃度の物質の存在を指摘できるのは不思議な気がするかもしれない (第 5 章参照)。そうではあっても、しかつめらしい霊能者のコイン投げが信じられないならば、こちらを信じる方が得策である。

α と β は状況に応じて任意に設定することができる。本稿では、 $\alpha = \beta = 5\%$ とする。

第 4 章 便利な表記

測定値の SD はサンプル濃度に依存して変わるとする仮定 (非等分散の仮定) は一般性が高い。しかし、現実的には、DL 濃度はブランクサンプルの濃度に近いため、DL 測定値の SD はブランク測定値の SD に等しいと仮定できる (等分散の仮定)。等分散の仮定を採用すると、検出限界の式 (3) と (4) は、

$$y_D - y_0 = 2k\sigma(y_0) = 2k\sigma(y_D) \quad (5)$$

$$x_D - x_0 = 2k\sigma(x_0) = 2k\sigma(x_D) \quad (6)$$

とシンプルに書ける。ブランク測定値とブランク濃度をゼロ ($y_0 = 0$ 、 $x_0 = 0$)、測定値の SD を σ_y ($= \sigma(y_0) = \sigma(y_D)$)、濃度推定値の SD を σ_x ($= \sigma(x_0) = \sigma(x_D)$)、 $k = 1.65$ とすれば、

$$y_D = 3.3\sigma_y \quad (7)$$

$$x_D = 3.3\sigma_x \quad (8)$$

となる。これを少し変形すれば、次の式が得られる。

$$\sigma_y / y_D = 1 / 3.3 = 30 \% \quad (9)$$

$$\sigma_x / x_D = 1 / 3.3 = 30 \% \quad (10)$$

式 (9) と (10) の左辺は SD を平均で割ったものであるから、**相対標準偏差 (relative SD、RSD)** である。DL 濃度のサンプルのくり返し測定精度は **30%RSD** であると言える。これは検出限界のもう一つの定義である。すると、精度プロファイル (RSD-濃度プロット) から検出限界を探し出すことができる (第 5 章参照)。

第 5 章 実例

前章までは、検出限界と判定限界の理論的側面を考察した。本章は HPLC 分析における検出限界、判定限界、定量限界の実例を紹介する。検量線が線形ならば ($y = ax$)、式 (3) と (4) は同値であるので、以下では、コンテキストに応じてどちらかを参照する。

ブランク測定値をゼロ ($y_0 = 0$)、測定値の SD を $\sigma_y (= \sigma(y_0) = \sigma(y_D))$ 、 $k = 1.65$ とすれば、式 (3) は式 (7) に変形できる。ブランク測定値の SD の 3.3 倍が検出限界である。図 3 はこの状況を模している (分布に関しては図 4 の方が見やすい)。検出限界を実際に求めるときに、最も厄介な問題は標準偏差 σ_y の推定である。ISO 11843 シリーズは Part 1 で検出限界を定義しているが、Part 2 から Part 7 はすべて σ_y の推定法を論じている。検出限界の実際は測定値の SD を推定することが全てであると言っても決して大げさではない。

図 5 は 4 種類の酸を HPLC で測定したクロマトグラムであり、以下では、左から 4 番目のピーク (クエン酸) の精度に焦点を合わせる。測定値の SD はくり返し測定ではなく、FUMI 理論により推定する。理由は、後者はくり返し測定なしに、しかも正確に測定精度を求められるからである (参考文献：FUMI 理論)。FUMI 理論の解析は TOCO19 (開発中) を用いた。図 5 から図 10 は TOCO19 の出力である。

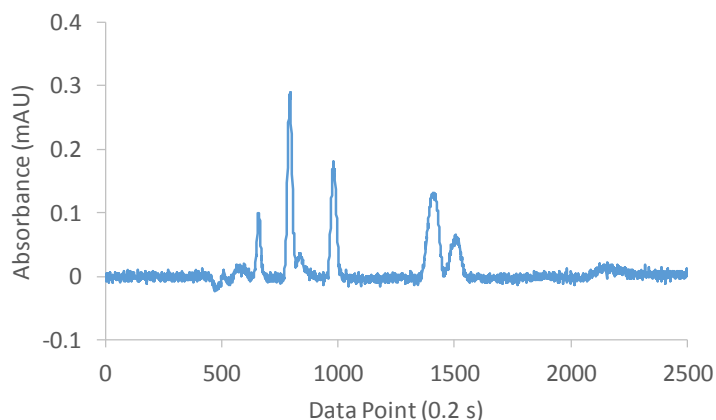


図 5 クロマトグラム

シグナル (左から) : システムピーク、リンゴ酸、酢酸、クエン酸、コハク酸 (それぞれ 10 mg/L)。東京薬科大学薬学部小谷明准教授との共同研究。

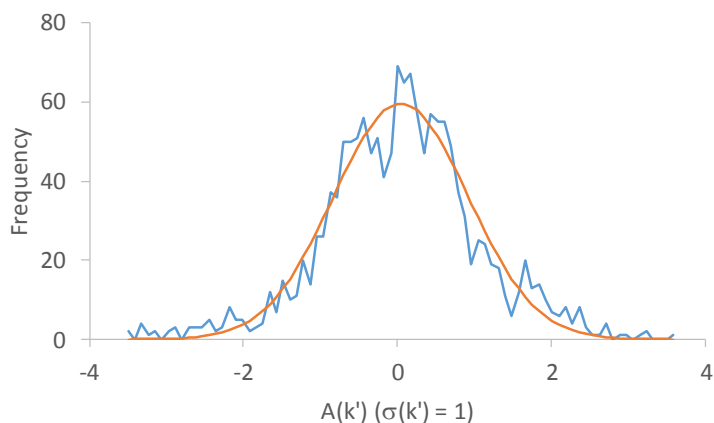


図 6 測定誤差の分布

X 軸：測定誤差の大きさ (SD = 1 として表示)。

図 6 は測定誤差のヒストグラムである。FUMI 理論では、低濃度のサンプルの測定誤差の主原因はバックグラウンドノイズである仮定している。測定値とはピーク両端の間にあるシグナル強度の積分値 (和) であるので、測定誤差はこの積分区間 (k') 内でノイズが作る面積 ($A(k')$) である。図 6 のギザギザの線が実際のベースライン (図 5) で観測された $A(k')$ の度数分布であり、滑らかな線が FUMI 理論から求めた理論線 (正規分布) である。図 6 のフィッティングの良さは、HPLC 測定の測定誤差 (ノイズが作る面積) が正規分布に従っている証拠である。図 6 のヒストグラムは図 1~4 の正規分布の実例である。面積測定値は各データポイントの測定値の和であるから、中心極限定理 (<http://www.fumi-theory.com/img/YH-normalD.pdf>) によって、正規分布に収束するのは当然であるとの見解は正しくない。なぜならば、一般に、分析機器のノイズはデータポイント間で相関があり各ポイントは確率的に独立ではないので、中心極限定理を適用できないからである。

図 7 は検出限界シグナルを実際のノイズ上に重ねたシミュレーションである。このシグナルの幅はクエン酸のピーク幅 (図 5) と同じであるが、その形はガウス型ピーク (正規分布) を採用している。推定された DL 濃度は 1.74 mg/L であるから、その DL シグナルは図 5 のクエン酸ピーク (4 番目) の 0.174 倍の大きさである。DL サンプルをくり返し測定すると、理論的には、平均として図 7 のピークが観測される。そうであれば、図 7 のピークより小さいピークが 50% の確率で観測されることになるので、DL 濃度のクエン酸の存在を 50% 以上の高い確率で主張するには、このような小さいピークも考慮する必要がある。この主張が容認される最小ピーク的面積 (濃度) が判定限界である (図 8)。

議論をさらに進める前に、定量限界について簡潔に述べておこう。図 9 は定量限界のシグナルを実際のノイズに重ねて描いている。本稿では、定量限界は測定値または濃度推定値の SD の 10 倍と定義している。すると、式 (7) から (10) と同様な議論によって、定量限界濃度のサンプルのくり返し測定値または濃度推定値の RSD は 10% となる。

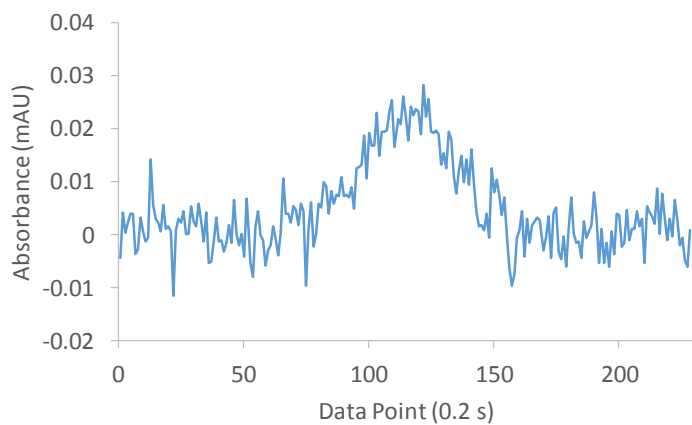


図7 検出限界シグナル (クエン酸)

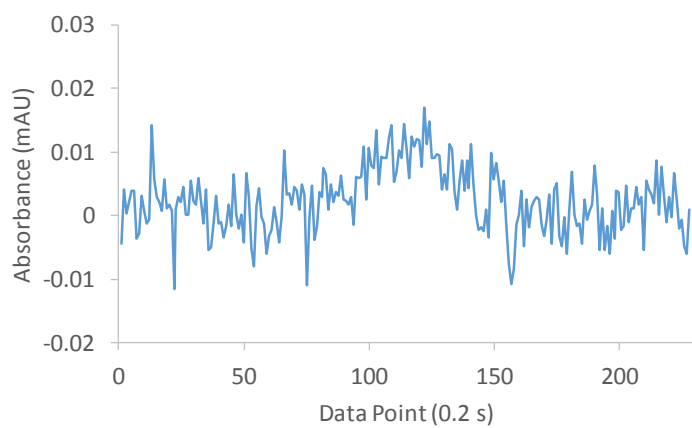


図8 判定限界シグナル (クエン酸)

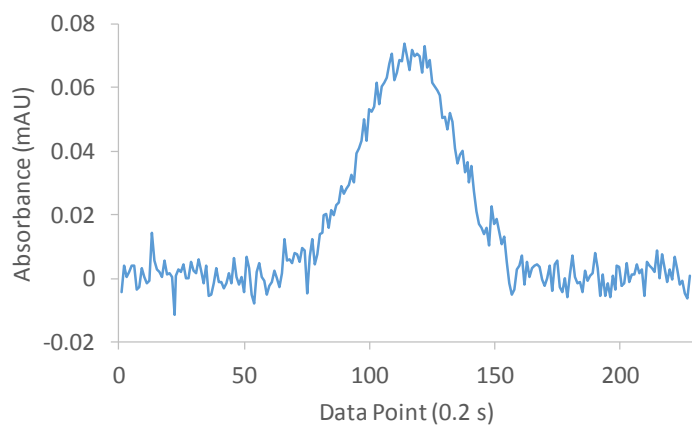


図9 定量限界シグナル (クエン酸)

検出限界と定量限界を S/N (signal-to-noise ratio) で定義することもある。検出限界は S/N=3、定量限界は S/N=10 が多く採用されている。図 7 と図 9 のシグナルの S/N はほぼこれらの値に近いと思える。S/N はいわば幾何学であり、確率 α と β を表すものではない。しかし、確率に基づく定義と S/N に基づく定義が、ほぼ一致していることは興味深い。

図 8 の判定限界ピークは DL 濃度のピーク (図 7) の半分の大きさである。この判定限界シグナルより少しでも大きなピークが観測されれば、我々は、確率 5% の過誤 (危険率) をもって、「DL 濃度のクエン酸が検出された」と断言できる。読者は、DL 濃度の物質の存在の判定に使うにはこのピーク (図 8) では小さすぎる、ノイズとあまり変わらないではないかと思われるかもしれない。以下で、すこし議論しよう。

$\alpha = 5\%$ とすれば、ブランクサンプルの測定で、判定限界 (図 8) より大きなピークが観測されるのは、理論的には 20 回中 1 回の割合である。問題は 20 回に 1 回を頻繁と思うかどうかである。千三つ (せんみつ) は 1000 回に 3 回しか本当のことを言わない嘘つきであるから、1000 回に 3 回だけ起こる事象は稀だと言えるだろう。千三つに比べれば、20 回に 1 回の割合は頻繁と思っても当然である。つまり、ブランクサンプルの測定からは、判定限界以上の値は頻繁に得られ、DL サンプルの測定からは、判定限界以下の値は頻繁に得られるのである。誤判定は頻繁であると思ってよい。

危険率 $\alpha = \beta = 5\%$ の場合、検出限界は $x_D = 3.3\sigma_x$ 、判定限界は $x_C = 1.65\sigma_x$ であり、誤判定は 20 回に 1 回起こる。そこで、危険率 $\alpha = \beta = 0.14\%$ 、検出限界を $x_D = 6\sigma_x$ 、判定限界を $x_C = 3\sigma_x$ とすれば、誤判定は 1000 回に 1 回しか起こらない。しかし、検出限界はほぼ 2 倍になる。この測定の目的をある製品中の不純物の検出と仮定すると、2 倍の違いは非常に大きい。新しい方法で前の方法における検出限界量 ($3.3\sigma_x$) の不純物を測定すると、約 40% の確率で不純物は見逃されてしまう ($x_C = 3\sigma_x$ であるから)。検出確率 60% はコイン投げの確率 (50%) とそれほど変わらないので、前の方法では高々 5% の見逃しであった検出限界量 ($3.3\sigma_x$) の物質は新しい方法ではほぼ検出されないとと言える。ある方法で検出されていた物質が他の方法で検出されなくなれば、消費者には不利である。疑わなければ、再検査はしないからである。要約すれば、少々の誤りをおかしても、検出限界を低く設定した方が目的に適っている。これは「虎穴に入らざれば虎子を得ず」の実践と言えるだろう。

図 10 はクエン酸測定の精度プロファイルであり、TOCO19 を用いて図 5 のクロマトグラムから推定したものである。上で議論したように、検出限界は 30%RSD を示す濃度、定量限界は 10%RSD を示す濃度である。図 10 では、矢印で示してある。精度プロファイルからは、検出限界と定量限界が視覚的に把握できる。なお、精度プロファイルは JIS の用語である (JIS Z 8462-5)。

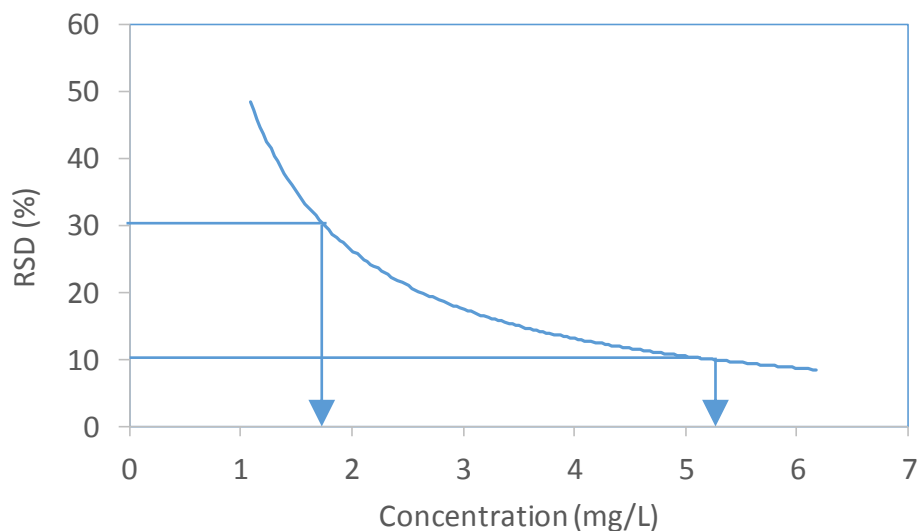


図 10 精度プロファイル

30%RSD を示す濃度 (1.74 mg/L) が検出限界、10%RSD を示す濃度 (5.26 mg/L) が定量限界である。図 7 と図 9 は、この定義に合致している。

第 2 章で議論したように、測定結果が検出限界以上ならば、ノイズでない何かを検出されたと判断する。「ノイズでない」と自信をもって言うためには、シグナルについての知識はいらないとしても、ノイズの性質は熟知していなければならない。例えて言えば、母親が赤ん坊の普段と違う泣き声を何かしらの病気のシグナルと察知できるのは、通常の泣き声に親しんでいるからである。常日頃観測できる現象に対して、稀にしか起こらない事象が異常として検知されるのである。異常をシグナルとすれば、通常はノイズである。ノイズの中に埋もれているシグナルを探し当てることは分析における検出であるが、新種生物の発見、新技術の発明、新理論の創造なども同じことであろう。知悉はその分水嶺を自ら超えるようだ。最後に、鎌倉時代初期の禅僧である道元の言葉を引用する。

「仏道をならふといふは、自己をならふ也。自己をならふといふは、自己をわするるなり」。

林 譲

平成 31 年 4 月 20 日

参考文献

検出限界

ISO 11843-1 Capability of detection-Part 1: Terms and definitions.

L.A. Currie, Pure & Appl. Chem., 1995, 67, 1699-1723.

G.L. Long and J.D. Winefordner, 1983, 55, 712A-724A.

H. Freiser and G.H. Nancollas, Compendium of Analytical Nomenclature, Definitive Rules 1987, IUPAC.

J.C. Miller and J.N. Miller、宗森信訳、データのとり方とまとめ方、共立出版株式会社。

FUMI 理論

ISO 11843-7 Capability of detection – Part 7: Methodology based on stochastic properties of instrumental noise.

JIS Z 8462-7 測定方法の検出能力—第 7 部：分析機器ノイズの確率論的性質に基づく方法

小谷明、FUMI 理論による HPLC の精度推定、<http://www.fumi-theory.com/kotani-f.html>。

林 譲、簡単な説明、<http://www.fumi-theory.com/simple-f.html>。