

分析機器における測定値の確からしさを知る方法について

林 譲[#]・松田りえ子

A Method for Knowing the Statistical Reliability of Instrumental Measurements

Yuzuru Hayashi[#] and Rieko Matsuda

The precision of instrumental measurements, usually expressed by SD (standard deviation) or RSD (relative standard deviation) is of importance in a society as well as in a world of analytical chemistry. For example, the detection limit and confidence intervals of calibration lines which are based on the precision are inevitable for regulatory science. As is well-known in statistics, however, a small number of repetitive experiments (e. g., 5) lead to a poor reliability of the precision. This presentation shows a method (called FUMI theory) which can provide a precision from a single measurement of noise and signal without any repetitive measurements. The statistical reliability of the precision from the FUMI theory corresponds to that from the repetitive method of 50 experiments.

Keywords: statistical reliability, instrumental analysis, FUMI theory
(Received May 30, 1997)

測定値のSD (標準偏差) またはRSD (対標準偏差) で表される定量精度は定量分析法を評価する規準として、特にレギュラトリーサイエンスの分野において、近年その重要性を急激に増している。分析値の信頼性保証の考え方や制度が分析化学の実践的な領域に導入され始めた現在の社会状況もこれに反映している。

通常、測定値のSDは繰り返し測定から求められる。しかし、この様にして求めたSD値の信頼性は高くないことが統計学的に知られている(カイ2乗分布より)。例えば、5回の繰り返し測定から求めたSD値は、3倍異なることがあっても不思議ではないので、ブランク試料の定数倍と定義される検出限界も同様に3倍違う値が得られることもありえる。レギュラトリーサイエンスとしては、検出限界の実験的決定の曖昧さは重要な問題である。

数多くの繰り返し測定(例えば50回)から求めたSD値の信頼性は高い。しかし、クロマトグラフィーなどで多くの繰り返し測定を行うことは、時間的な問題のために実現は殆ど不可能である。

そこで、我々は繰り返し測定によらない方法で測定値のSDを正しく「予測」する理論・方法を確立した。この理論はFUMI理論と呼ばれている(FUNCTION of Mutual Informationの略)。1つのクロマトグラムから予測された測定値

のSDの統計的信頼性は、50回の繰り返し測定から得られたSDの信頼性にほぼ等しい。ここでは、FUMI理論とそのレギュラトリーサイエンスへの応用を概説する。

1. FUMI理論

分析機器の定量精度を理論的に求める研究は欧米諸国では1940年代から始まり、現在でもよく行われている。日本ではこの様な研究例は殆どなかったが、FUMI理論が1994年に林・松田により提案された¹⁾。この理論では、分析機器のノイズとシグナルの確率論的性質から測定値のSDとRSDを予測する。サンプルが低濃度の時には、測定誤差は殆どバックグラウンドノイズに依存するので、検出限界などはこの理論から簡単に求めることが出来る。液体クロマトグラフィーの場合、高濃度のサンプルの測定誤差は注入誤差によるので、この誤差を考慮すれば、FUMI理論で広い濃度範囲の精度予測が可能である²⁾。

自然界の多くの現象が $1/f$ 揺らぎを示すことが知られている³⁾。 f は周波数であり、波のエネルギーが周波数に反比例するのがこの名称の由来である。 $1/f$ 揺らぎでは低周波成分のエネルギーが高周波成分のエネルギーより大きい。レッドウッドの木目、模様やピバルディの四季も $1/f$ 揺らぎを示すのは驚きであるが、殆ど全ての分析機器ノイズが $1/f$ 揺らぎを示しても不思議ではない。そして、実際そのとおりである²⁾。

$1/f$ 揺らぎは数学的に扱えないので(発散するため)、 $1/f$ 揺らぎを示す分析機器ノイズから定量精度を予測することは簡単ではなかった。FUMI理論では、 $1/f$ 揺らぎを数学的に解析可能な確率過程(ホワイトノイズとマ

[#] To whom correspondence should be addressed: Yuzuru Hayashi; Kamiyoga 1-18-1, Setagaya, Tokyo 158, Japan; Tel: 03-3700-1141 ext. 283; Fax: 03-3707-6950; E-mail: fumi@nihs.go.jp
本ステートメントは、日本薬学会第117回レギュラトリーサイエンス討論会(1997.4, 東京)にて発表した内容をまとめたものである。

ルコフ過程)の和で近似し、機器ノイズから定量精度を予測する。分析機器のノイズにホワイトノイズとマルコフ過程がそれぞれの程度含まれているかを、機器ノイズをフーリエ変換して決定する。一般に、ホワイトノイズとマルコフ過程の量は分析機器毎に異なるので、精度を求めたい機器のノイズを測定する必要がある。

定量精度(特にRSD)はシグナルの測定値が分からなければ計算することは出来ない。FUMI理論では、ノイズと同様にシグナルも測定し、高さ測定、全面積測定、部分積分測定などに伴う積分領域を決定しなければならない。

この様にして求めた、ノイズとシグナルのパラメータから定量精度を、繰り返し測定なしに、予測することが出来る。クロマトグラフィーでは、高さ測定と全面積測定ではどちらが精度が高いかが数十年に渡って議論されてきた。FUMI理論を用いればこの問題を解決することが出来る。どちらが良いかという一般的回答は得られないというのがFUMI理論の答えである。ノイズの性質は分析機器毎に異なるので、目的の分析機器のノイズとシグナルをFUMI理論で解析すれば上記の問題の答えが得られる。

2. 検出限界

第13改正日本薬局方の分析法バリテーションの定義では、ブランク試料の測定値の標準偏差の3.3倍を検出限界測定値としている。クロマトグラフィーではインテグレータを用いて定量するのが一般的であり、ブランク測定値の標準偏差を求めることは困難である。日本薬局方では、クロマトグラフ法では測定値の標準偏差の代わりにノイズ・レベルを用いることが出来ると記載してある。

実用的にはこれで十分であると思う。しかし、厳密に言えば、ノイズ・レベルからは検出限界の統計的性質が一意に導かれない。ノイズ・レベルが同じでも、測定値のSDが同じとは限らないからである。

日本薬局方にある統計的性質とは、検出限界における消費者及び生産者の危険率が5%以下となるように検出限界を定めることである。クロマトグラフ法に限ると、FUMI理論では、ベースラインノイズが作る「偽り」の測定値(面積値など)のSDをブランクの測定値のSDとする。このSD値はノイズをフーリエ変換することによって計算することが出来き、FUMI理論は検出限界に統計的性質を付随させる。例えば、ベースラインノイズの揺らぎをシグナルと見誤る確率が0.13%となるように検出限界を決定することが出来る。

3. 判別限界

濃度の異なるサンプルがある場合、このサンプルを測定で区別できる確率が95%であれば、この濃度差を95%判別限界という。測定で区別できるという意味は、高濃度のサンプルは低濃度のサンプルより大きい定量値を示すことである。判別限界(discrimination limit)は、検出限界を含む一般的な概念であるが、まだレギュラトリーサイエンスにおける応用例は少ない。ここでは、ヒトインスリンの純度試験の判別限界を述べる。

インスリンは分解するので、USPでは、分解物の濃度は全体の3%を越えてはならないと規定している。実際は、3%以上の分解物を含むインスリン製剤はまず存在しないが、仮想的な思考をしてみよう。検査機関は、3%以上の分解物を含むインスリン製剤は確実に検出し、却下したいであろう。しかし、測定は必ず誤差を含むので、例えば、3.2%の分解物を含む製剤の測定値が3%以上を示す確率は100%以下であろう。一方、5%の分解物を含む製剤を却下できる確率はほぼ100%であろう。

3.2%の分解物を含む製剤の測定値が3%以上を示す確率が95%であれば、この分解物濃度3.2%を95%判別限界と呼ぶ。一般に、判別限界は分析方法毎に異なり、その分析方法の信頼性を表す指標となる。例えば、95%判別限界が分解物濃度3.2%である方法Aは、分解物を3.5%で含む製剤を却下できる確率は95%以上である(100%に近いであろう)。しかし、95%判別限界が分解物濃度3.5%である方法Bは分解物濃度3.2%の製剤を却下できる確率は95%より小さい。この様に、判別限界が小さい分析法が良い分析法である。判別限界の上限を設定することは、分析法の信頼性を保証することにつながる。

キャピラリー電気泳動によるヒトインスリンの純度試験の判別限界の決定は、四方田らにより行われた(J. Chromatogr. B. 印刷中)。FUMI理論から得られた95%判別限界は、実験から得られた判別限界とよく一致した。この場合、繰り返し実験はFUMI理論を検証するために行われた。もし、繰り返し実験のみで正確な95%判別限界を求めようとすると、膨大な実験量が必要であろう。

文 献

- 1) Hayashi, Y., and Matsuda, R.: Analytical Chemistry, 66, 2874~2881 (1994)
- 2) 林 譲, 松田りえ子: ぶんせき, 195~200 (1995)
- 3) 武者利光: ゆらぎの世界, 講談社, 1993