

質量分析におけるシグナル、ノイズ、および検出限界

テクニカルノート

著者

Greg Wells、Harry Prest、
Charles William Russ IV、
Agilent Technologies, Inc.
2850 Centerville Road
Wilmington, DE 19809-1610
USA

概要

これまで、1回の測定から得られるクロマトグラム上のピークの S/N 比は、2つの異なる質量分析システムの性能を比較するための便利な性能指数として使われてきました。しかし、質量分析計の設計が進化してシステムノイズが非常に低くなったため、S/N 比に基づく性能比較は難しくなり、操作モードによっては不可能になりました。このことは、高分解能測定やタンデム MS 測定等の、バックグラウンドにイオンがないことが多く、ノイズが基本的にゼロである超低ノイズモードでの測定の場合には特に顕著です。複雑なマトリクス中の微量分析において装置のパフォーマンスを特徴づける手段としてよく使われる統計的な方法は、バックグラウンドノイズが高い場合でも低い場合でも装置のパフォーマンスを知るのに有効です。



Agilent Technologies

はじめに

一般に、分析化学における微量分析では、検出限界 (LOD) の確認が必要です。検出限界とは、物質がない状態のシステムノイズ (ブランク値) と物質のシグナルが区別できる物質の最低量のことです。質量分析計が微量分析に使用されることが増えており、これらの機器を使用する場合には分析対象成分の検出限界の評価に影響するファクターを理解することが重要になります。一般に使用される検出限界は数種類あります。これには、装置検出限界 (IDL)、方法検出限界 (MDL)、実用的定量下限 (PQL)、および定量下限 (LOQ) が含まれます。同じ用語を使用しているか、どんな定義を使用しているか、測定やキャリブレーションにどんなノイズを使用しているかなどの微妙な差異に応じて、LOD に違いがある場合があります。感度、ノイズ、S/N 比、検出限界など、機器性能を表す数についてはかなりの混乱があります。検出限界を評価、レポートする場合には、これらのファクターがどのように性能を示す数に寄与し、どのように決定されているかを理解することが重要です。バックグラウンドノイズをきわめて低く抑えながら個々のイオンを検出できるモードを備えた最新の質量分析計は、従来の検出限界の決定方法に対して新たな課題を与えています。

用語

- **装置バックグラウンドシグナル** – ブランク測定時の装置からのシグナル出力。一般には、アナログ/デジタルコンバータによってデジタル化される電圧出力です。
- **ノイズ (N)** – 装置バックグラウンドシグナルの変動。一般には、バックグラウンドシグナルの標準偏差として測定します。
- **分析対象成分シグナル (S)** – 物質の存在に対する機器のレスポンスの変化。
- **合計装置シグナル** – 分析対象成分シグナルと装置バックグラウンドシグナルの合計。
- **S/N 比 (S/N)** – ブランクで測定されたノイズに対する分析対象成分シグナルの比率。
- **感度** – ある特定の量のシグナルレスポンスから標準化された量を増加させたときに増加したシグナルレスポンスの量。一般には、検量線の傾きによって決定されます。(感度は、S/N 比や LOD などの用語と同じ意味で使用されることがありますが、この文書では、この分析定義の場合のみ感度という用語を使用します。)

装置検出限界 (IDL)

ほとんどの分析機器は、ブランク (分析対象成分を含まないマトリクス) の分析時にもシグナルを発生します。このシグナルを装置バックグラウンドレベルと呼びます。ノイズはバックグラウンドレベルの変動具合のことです。一般には、複数の連続ポイントのバックグラウンドシグナルの標準偏差を計算することにより測定されます。装置検出限界 (IDL) は、ある特定の統計的信頼区間の範囲内でノイズレベルと区別できるシグナルを生成するために必要な分析対象成分濃度です。おおよその LOD は S/N 比から得ることができます。

方法検出限界 (MDL)

ほとんどのアプリケーションでは、分析メソッドはクリーンな分析対象成分を分析するだけではありません。対象成分の検出のために、不要なマトリクス成分の除去、分析対象成分の抽出や濃縮、さらには誘導体化が必要かもしれません。装置で分析する前に分析対象成分をさらに希釈または濃縮することもあります。分析の手順が追加されると、その分誤差が発生する可能性は高くなります。検出限界は誤差を考慮して定義されるため、誤差が増加すると検出限界値が高くなります。このように分析の全手順を考慮した検出限界を MDL と呼びます。多くの場合、おおよその LOD はマトリクス中の分析対象成分の S/N 比から得ることができます。

定量下限 (LOQ) と実用的定量下限 (PQL)

ノイズから何かがわかって、必ずしも実際に存在する物質を特定の信頼度で知ることができるわけではありません。同じ分析対象成分を同じ機器と同じ条件で繰り返し測定した場合でも、機器におけるサンプル導入、分離、および検出プロセスの変動により結果は毎回わずかに異なります。LOQ は、分析対象成分の異なる 2 つの量の差をきっちりと判別できる限界値のことです。LOQ はラボによって大幅に異なる場合があるため、一般には実用的定量下限 (PQL) と呼ばれる別の検出限界が使用されます。統計学上、PQL と LOQ の間には特定の数学的な関係はありません。多くの場合、PQL は実質的に MDL の約 5~10 倍として定義されています。

S/N 比を使用した IDL および MDL の評価

一般に質量分析ではクロマトグラフをサンプル導入手段として使用します。クロマトグラフから得られる分析対象成分シグナルは、時間の関数としてほぼガウス分布に近い形状になります(図 1)。クロマトグラフを用いて質量分析計に分析対象成分を導入する場合、シグナルは一定ではなく、サンプルを示すすべてのポイントで同じ分析対象成分の量を表すわけではありません。S/N 比を使用して検出限界を評価する場合は、シグナルは一般にベースライン (\bar{x}_B) から上のシグナル最大値までの高さ(図 1 の S)とされ、ピークの下の部分でバックグラウンドノイズを評価します。ノイズを評価するための標準的な方法はベースラインノイズの peak-to-peak (最小値から最大値) の値を測定することです。この方法ではピークから離れた位置で、ピークからの 60 秒間、あるいはピークの前後で 30 秒間ずつベースラインノイズを測定します。最新のインテグレータとデータシステムの登場により、ノイズを評価するためのベースライン範囲は自動で選択され、ノイズは選択されたタイムウィンドウのベースラインの標準偏差 (STD) または二乗平均平方根 (RMS) として計算されます。ただし、1 回測定のみを使った IDL および MDL 評価のアプローチでは多くの場合に正確でないことが示されています。

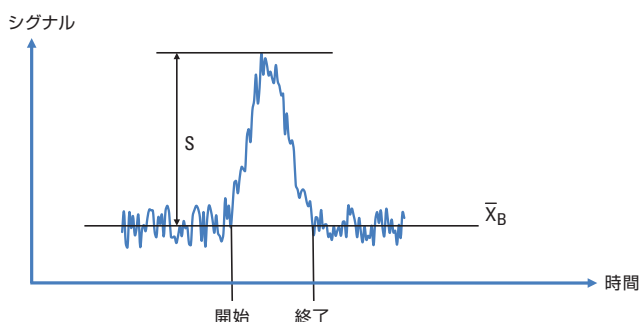


図 1. 時間の関数として分析対象成分のシグナルを示したクロマトグラム上のピーク。分析対象物の存在量は時間に依存している

分析対象成分の定量を行う場合、シグナルはバックグラウンドを差し引いたガウスピークの開始から終了までを積分したものとなります。この面積は単一サンプリングの測定値で、母集団(元のサンプル中の分析対象成分の量)の真の値を知るための推定量の一つです。分析対象成分の量は時間の関数として $C(t) = K C_0 F(t)$ として表現できます。 C_0 はテストサンプル中の分析対象成分の量、 $F(t)$ はクロマトグラム上のピークの形状(強度対時間)、 K はキャリブレーション係数です。分析対象成分の分析再現性がよく、分析対象成分量に対してレスポンスが直線である場合、比例定数である K は既知の分析対象成分量を使用したキャリブ

レーションによって決定できるため、ピーク面積を元の分析対象成分量の測定手段として使用することができます。同じサンプルの反復的な測定(面積)を行うと、母集団を表す実際の値について正規分布になる若干異なる一連のレスポンスが生成されます。サンプルとバックグラウンド両方における一連のシグナルのばらつきは、(1) 注入量のばらつき、(2) GC カラムに移動したサンプル量のばらつき、(3) バックグラウンドの量のばらつき、(4) イオン化効率のばらつき、(5) イオン源からのイオン抽出およびマスアナライザへの移動におけるばらつき、(6) 測定したイオン数を示す検出シグナルのばらつきのようなファクターにより起こります。最後のファクターは低イオン流量の場合には、たとえイオン検出器に導入されるイオン数が同じ場合でも起こり、出力されるシグナルには微妙に差が出てきます。なぜなら、検出器のレスポンスはイオンが検出器のどこにヒットするかによって変動するからです。クロマトグラム上のピーク的面積部分の決定も十分ばらつきに寄与します。ピークの開始、終了、およびバックグラウンドより下の面積部分の決定がすべて誤差に寄与します。これらのばらつきをまとめてサンプルノイズとみなします。つまり、出力シグナルのばらつきはサンプリングと検出プロセスのばらつきを合わせたものとなります。これらのばらつきに有限のイオン数の標準的なばらつき(イオン統計)が加わります。

最新の質量分析計システムは、バックグラウンドをほぼゼロにできるさまざまなモードで動作することができます。MS/MS、負イオン化学イオン化、および高分解能測定は、システムバックグラウンドシグナルがほぼゼロであることが多く、マトリクスからの化学的バックグラウンドがない場合には特にその傾向が顕著です。図 2 に、バックグラウンドが非常に低いクリーンなシステムにおけるオクタフルオロナフタレン (OFN) 標準試料の GC/MS によるマスキングクロマトグラムを示します。マスキングクロマトグラムの各ポイントは、セントロイド質量ピークの強度を表します。特定のセントロイド質量に十分なイオンがない場合は、結果として強度値がゼロとしてレポートされることがあります。非常にクリーンなシステムで MS/MS または負化学イオン化モード測定を行うと、観測可能なバックグラウンドイオンがなく、ノイズがゼロと計算される可能性があります。この状況は、イオン検出のしきい値を高くすると、より顕著になる場合があります。このような状況では、バックグラウンドノイズを増やさずに、イオン検出器のゲインを上げてシグナルレベルを上げることができます。分析対象成分のシグナルは増えますが、ノイズは増えません。これで S/N 比が上がったから検出限界がよくなったと考えるのは誤解であり、受け入れがたいものです。

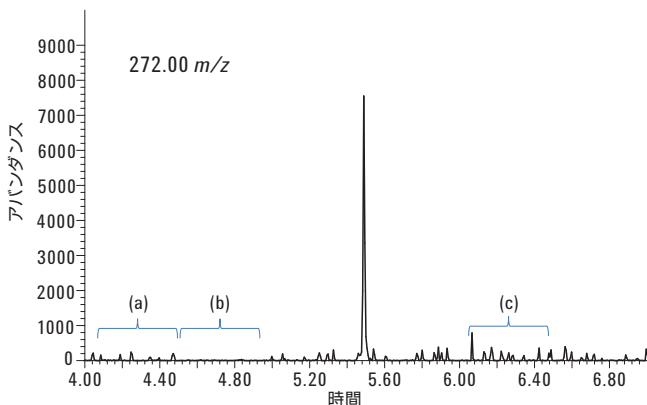


図 2. 1 pg OFN のフルスキャン分析における $m/z = 272$ のマスククロマトグラム。非常に小さい化学イオンノイズが示されている。

これには (1) バックグラウンドノイズを評価するにはベースラインのどの領域を選択すべきなのか (2) シグナルは増えるが、検出されるイオン数は増えないため、実際の検出限界は変わらないという 2 つの実際的な問題があります。バックグラウンドが高い場合とは異なり、バックグラウンドが非常に低く、ただしゼロではない場合、測定されるノイズはノイズの測定場所に大きく依存します。図 2 の (a)、(b)、および (c) というラベルの付いたバックグラウンドの領域では、RMS ノイズはそれぞれ 54、6、および 120 の RMS となります。この例では、ノイズを測定した場所が変わっただけで、S/N 比が 20 倍ほど変わってしまいます。このように、低レベルでばらつきのある大きなノイズがある場合には、S/N 比を検出限界の評価に使うのは明らかに有効ではなくなります。図 3 の MS/MS クロマトグラムのように、バックグラウンドノイズがゼロの場合は、状況がさらに不確実になります。この場合、ノイズはゼロで、S/N 比は無限になります。図 3 で観測されたノイズは、電気的なノイズだけで、これらはバックグラウンド中のイオンの存在によるノイズよりも数桁低くなります。

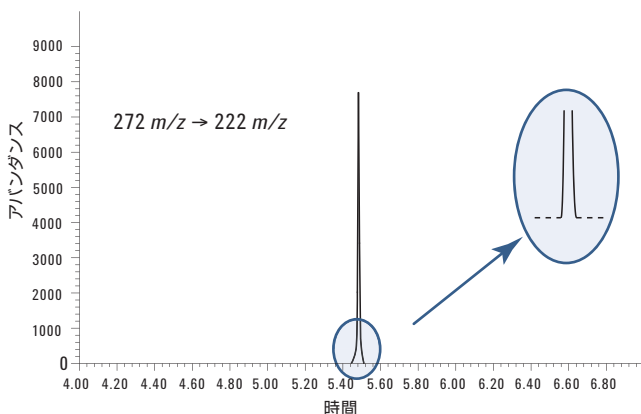


図 3. 化学イオンノイズのない 100 fg OFN の EI MS/MS のクロマトグラム ($272 m/z \rightarrow 222.00 m/z$)

新しい IDL および MDL の評価方法

クロマトグラフにおける分析については、より信頼性の高い IDL および MDL の評価方法は他にも数多くあります [2-7]。米国の場合、最も一般的な方法は、EPA Guidelines Establishing Test Procedures for the Analysis of Pollutants [2] です。欧州で一般に使用されている標準は、Official Journal of the European Communities, Commission Decision of 12 August 2002; Implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results [4] に記載されています。

この両方のメソッドは類似しており、複数の重複する標準試料を注入して測定システムでの不確かさを評価することが求められています。予想される検出限界 (ノイズレベルの 5~10 倍以内) に近い濃度を持つ同一のサンプルを数回、同等の回数のブランクとともに測定します。必要に応じて一定のバックグラウンドからの面積を取り除くために、分析対象成分の各測定値から平均ブランク値を減算します。しばしば、質量分析においてはその特徴より、いったんブランクからの寄与の大きさを確認された後、ブランクからの寄与は無視されたり、排除されたりします。このようにして、測定した一連の分析対象成分シグナルの標準偏差 (ベースライン減算したクロマトグラム上のピーク面積の標準偏差) が決定されます。ピーク面積のばらつきには、分析対象成分のシグナルノイズとバックグラウンドノイズ、注入間のばらつきが含まれるため、システムノイズとサンプリングノイズを区別しながら既知の信頼区間 (測定された面積がシステムノイズと統計的に区別できる確率) を用いて 1 回の測定の統計的意味を確立することができます。1 回のサンプル測定での S/N 比から IDL の評価をしようとしても、同じ分析対象成分の複数回測定からわかるサンプル測定間のばらつきやサンプリングノイズは得られません。

測定回数が少ない場合は ($n < 30$)、検定統計量 t_α の決定には Student t 分布の片側検定 [1] が使用されます。クロマトグラム上のピークの場合、最新のデータシステムでは、ベースラインより上のピーク面積がレポートされます (バックグラウンドは減算され、シグナルの変動に寄与しません)。したがって、IDL または MDL は、ゼロの母平均値 ($\mu = 0$) よりも統計的に大きいシグナル (ピーク面積) を与える分析対象成分量 \bar{X} として決定されます (詳細は付録 I を参照してください)。

$$IDL = \bar{X} - \mu = \bar{X} = t_\alpha \sigma_{\bar{X}} = t_\alpha S_{\bar{X}}$$

t_α の値は自由度を $n-1$ (測定回数 - 1) として $1-\alpha$ を測定値がゼロより大きくなる確率として Student t 検定の表から得た値で、 $S_{\bar{X}}$ は一連の測定の標準偏差で、サンプル平均値からの分布の真の標準偏差と同等のものとして扱われています。この議論を行うには、標準偏差とサンプリングノイズを算出する必要があるた

め 2 回以上の測定が必要になります。測定回数 n が多いほど、 t_α の値が小さくなり、IDL または MDL 算出における不確かさが小さくなります。

図 4 に、特定量 C_{STD} の標準試料の繰り返し測定により生成した平均値 \bar{X}_{STD} を中心とした測定値の分布を示します。標準偏差は、この分布の幅を測定したものです。 α は、測定値がゼロ以下になるパーセンテージまたは確率です。IDL は、測定が $1-\alpha$ の確率でゼロより大きくなるような平均測定値 \bar{X}_{IDL} に対応するサンプル量 C_{IDL} です。図 5 は、同じ α と同じ装置感度でもより標準偏差の小さい場合 (より高い精度の場合) を示しています。 α が同じままで測定の標準偏差がより小さくなると、同じパーセンテージだけゼロより大きくなる (同じ α) の分布を考えるとその平均値はより小さな値の方へ移動するため、IDL がより小さくなります。

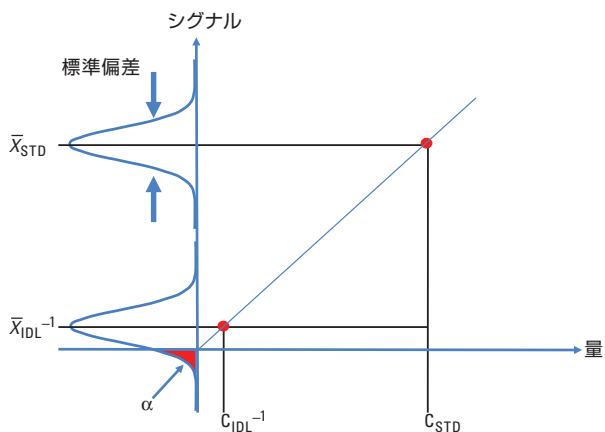


図 4. 装置検出限界 – 統計的に 0 より大きいシグナルを持つ分析対象成分の量。測定値のばらつきが大きい場合

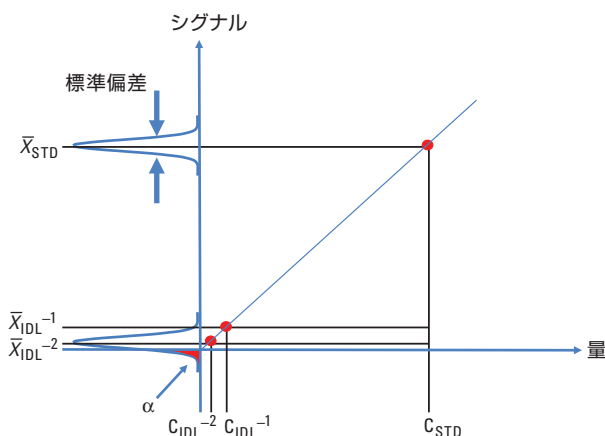


図 5. 装置検出限界 – 統計的に 0 より大きいシグナルを持つ分析対象成分の量。測定値のばらつきが小さい場合

例として、図 6 では 8 回の繰り返し注入 (自由度 7、 $n = 7$) を行いました。99 % ($1-\alpha = 0.99$) 信頼区間とすると、検定統計の値 (t_α) は 2.998 です。機器の検出限界を得るのが目的であるため、母平均をゼロと仮定します (この時 1 回の測定で得られる面積 \bar{X}_A が統計的に 99 % の確率でゼロより大きい値となります)。8 回測定における面積の平均値は 810 カウント、標準偏差は 41.31 カウント、相対標準偏差は 5.1 % で、IDL の値は次のように計算できます。

$$IDL = t_\alpha S_{\bar{X}}$$

$\bar{X}_A = (2.998)(41.31) = 123.85$ カウント。200 fg のキャリブレーション平均は 810 カウントなため、IDL は $(123.85 \text{ counts})(200 \text{ fg}) / (810 \text{ counts}) = 30.6 \text{ fg}$ になります。

相対標準偏差 (RSD) をレポートするデータシステムの場合、IDL は次のように計算できます。

$IDL = (t_\alpha)(RSD)(\text{amount standard})/100 \%$ 。前の例では、 $IDL = (2.998)(5.1 \%) (200 \text{ fg}) / 100\% = 30.6 \text{ fg}$ です。

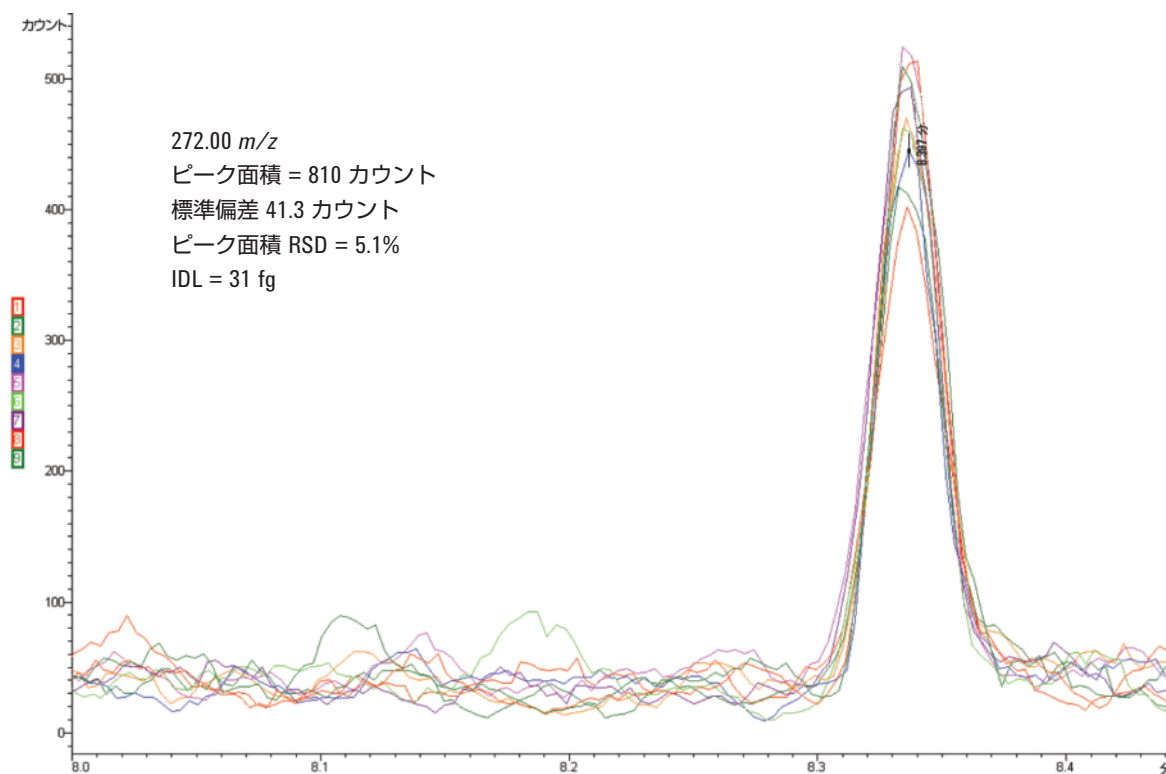


図 6. 200 fg OFN のフルスキャン分析における $m/z = 272$ のマスククロマトグラム。8回を繰り返し注入

まとめ

多くの場合、1回の測定から得られたクロマトグラム上のピークのS/N比を使用してIDLを計算する従来の方法は便利です。しかし、複雑なマトリクス中の微量分析によく使用される複数回注入の統計的方法を使用したほうがより実用的なIDLやMDLを計算できます。繰り返し注入の平均値と標準偏差を使用することにより、低濃度分析対象成分のレスポンス、分析対象物やバックグラウンド測定に含まれるばらつきの要因、試料導入やサンプリングプロセスによるばらつきの違いの統計的な意義を評価することができるようになります。このことは、バックグラウンドノイズがほぼゼロの最新の質量分析計で特に顕著になります。装置および方法検出限界を評価するための複数回注入メソッドは、バックグラウンドノイズが高くても低くても精密で統計的に妥当なものです。

参考文献

1. Statistics; D. R. Anderson, D. J. Sweeney, T. A. Williams; West Publishing, New York, 1996.
2. U.S. EPA - Title 40: Protection of Environment; Part 136 – Guidelines Establishing Test Procedures for the Analysis of Pollutants; Appendix B to Part 136 – Definition and Procedure for the Determination of the Method Detection Limit – Revision 1.11.
3. Uncertainty Estimation and Figures of Merit for Multivariate Calibration; IUPAC Technical Report, Pure Appl. Chem.; Vol. 78, No. 3, pp 633-661, 2006.
4. Official Journal of the European Communities; Commission Decision of 12 August 2002; Implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results.
5. Guidance Document on Residue Analytical Methods; European Commission – Directorate General Health and Consumer Protection; SANCO/825/00 rev. 7 17/03/2004.
6. Guidelines for Data Acquisition and Data Quality Evaluation in Environmental Chemistry - ACS Committee on Environmental Improvement; Anal. Chem. 1980, 52, 2242-2249.
7. Guidelines for the Validation of Analytical Methods used in Residue Depletion Studies; VICH GL 49 (MRK) – November 2009 Doc. Ref. EMEA/CVMP/VICH/463202/2009-CONSULTATION.

付録 I

シグナルをノイズと区別する方法

検出限界の評価は、分析対象成分のシグナルとバックグラウンドノイズを区別する方法に依存します。この評価方法は、仮説検定に基づいています。仮説検定は刑事裁判に似ています。刑事裁判では、被告人を無罪と仮定します。帰無仮説 H_0 は、無罪の仮定を表します。帰無仮説の反対である H_a (対立仮説) は、有罪の仮定を表します。刑事裁判の仮説は次のように記述されます。

H_0 : 被告人は無罪

H_a : 被告人は有罪

これらの競合する申し立て (仮説) を検査するために、裁判が行われます。裁判に提示された証拠がサンプル情報を提供します。サンプル情報が無実の仮定と矛盾しない場合、被告人が無実であるという帰無仮説は却下できません。一方、サンプル情報に無罪の仮定と矛盾がある場合は、帰無仮説を却下する必要があります。対立仮説を受け入れることができます。

クロマトグラム上のピークの場合、最新のデータシステムでは、ベースラインより上のピーク面積がレポートされます (つまり、一定のバックグラウンドは減算され、シグナルの変動に寄与しません)。分析対象成分の母集団の平均値 μ_A は、分析対象成分の捕集容器に格納されている分析対象成分の量の真値です。一回一回の測定は母平均を見積もるための測定値となります。一連の繰り返しサンプル測定の平均値 \bar{X}_A は、母平均 μ_A とほぼ近い値です。サンプルセットの平均値が、統計的に、特定の信頼区間内で母集団平均以上であるかどうかという質問に答える必要があります。IDL を得るのが目的であるため、母平均はゼロと仮定されます。この場合の拒否基準は次のとおりです。

- $H_0: \mu_A \leq 0$ シグナルの評価は、指定された信頼区間または有意水準内で、ゼロと違いがない。
- $H_a: \mu_A > 0$ シグナルの評価は、指定された信頼区間または有意水準内で、ゼロと違いがある。
- $\mu_A \leq 0$ の場合は H_0 を採択し、 $\mu_A > 0$ の場合は H_0 を棄却して対立仮説 H_a を採択する。

仮説の検定の統計的判断基準は次のとおりです: 検定統計 t_a が、確率表から得られた値よりも小さいかどうかです。表の値は、平均値をとるのに測定した回数 (n) と、平均値が母平均よりも大きいと判断される確率によって決められています。

$$t_a = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} < \text{表の値}$$

記号の意味: \bar{X}_A は、一連のサンプル測定の平均値です。

μ_A は、母集団の真値です。

$\sigma_{\bar{X}}$ は、一連のサンプル測定の標準偏差です。

t_a は検定統計量です。

$1-a$ は、サンプルセット平均が母平均と異なる確率です。

検定統計量 t の値は、測定回数 - 1 として定義される自由度の数値に依存します。測定回数が少ない場合 ($n < 30$)、 t_a の値は $n-1$ (測定回数 - 1) を使用した Student t 検定の表から得られます。たとえば、7 回の測定 (6 の自由度) で 99 % 信頼区間とすると (値がゼロではない確率が 99 % つまり $\alpha = 0.01$ で $\mu = 0$) では、表の値は 3.143 です。したがって、次の場合は、帰無仮説を採用します。

$$\frac{\bar{X} - \mu}{\sigma_{\bar{X}}} = \frac{\bar{X}}{S_{\bar{X}}} = t_a < 3.143$$

t_a がこれよりも大きい場合は、帰無仮説を棄却し、対立仮説 (つまり、バックグラウンド補正された分析対象成分シグナルが統計的にゼロではない) を、 $1-a$ の確からしさで採択する必要があります。母集団 $\sigma_{\bar{X}}$ の標準偏差の値は、通常はサンプル測定のセット $S_{\bar{X}}$ の標準偏差によって概算されます。4 回の測定 (自由度 3) において 97.5 % 信頼区間の検定統計の値は 3.18 です。8 回の測定 (自由度 7) において 99.0 % 信頼区間の検定統計の値は 2.90 です。これらの平均値は約 3 です。このため、一般的に IDL または MDL をおおまかに評価する場合には $\bar{X}_{IDL} = 3 S_{\bar{X}}$ が使われます。

詳細情報

アジレント製品とサービスの詳細については、アジレントのウェブサイト www.agilent.com/chem/jp をご覧ください。

www.agilent.com/chem/jp

アジレントは、本文書に誤りが発見された場合、また、本文書の使用により付随的または間接的に生じる損害について一切免責とさせていただきます。

本文書に記載の情報、説明、製品仕様等は予告なしに変更されることがあります。著作権法で許されている場合を除き、書面による事前の許可なく、本文書を複製、翻案、翻訳することは禁じられています。

アジレント・テクノロジー株式会社

© Agilent Technologies, Inc., 2011

Printed in Japan

May 3, 2011

5990-7651JAJP



Agilent Technologies