

連載 (講義)

数値データの取り扱い ～四捨五入から検出限界まで～ 第一回 データの丸め方と誤差の基礎

福島 整

独立行政法人物質・材料研究機構 共用基盤部門分析支援ステーション
〒305-0047 つくば市千現 1-2-1
fukushima.sei@nims.go.jp

(2010年5月21日受理)

分析業務にとって、数字の取り扱いは命である。有効数字やデータのばらつきは、ほとんどの現場担当者が無意識のうちに正当な方法で処理しているはずであるが、あらためて統計学の用語を目の前にすると自信を無くす人が多い。本講は、そのような人たちに自信を持たせ、あるいは体系的な知識を学習するための出発点と成ることを意図したものである。今回は、データの丸め方と、誤差の最も基礎的な部分について説明する。

The handling of numeric data

- from “round off” to “detection limit”-

I . How to round the numeric data, and the base of error

Sei FUKUSHIMA

Materials Analysis Station, Dept. Mat. Infrastructure, NIMS, 1-2-2 Sengen, Tukuba 305-0047, Japan
fukushima.sei@nims.go.jp

(Received: May 21, 2010)

One of the most important items for the analyst is the handling of the numerical data. The almost analysts handle and treat the numeric data, significant figure or dispersion for example, obtained from the measurement in the correct way unconsciously. However, there are many analysts who lose the self-confidence, when they wish to use the description of the statistical terminology. This lecture is intended giving a starting point for giving the self-confidence and learning a systematic knowledge to such people. In present part of this lecture, the way of rounding of the data and the most basic part of the error will be explained.

1. はじめに

断っておくが、数字の取り扱いの話は退屈である。新しい分析法や現場の測定に於ける目からウロコの工夫などのような、興奮させられる様な話はまず出てこない。それどころか、「あー、なんでこんな重箱の隅をつつくような話を」なんて思われる方

が多いであろう。しかし、現場の分析屋が提出する事実がともすると会社の経営を左右しかねず、その際にもっとも重要なのが「分析結果の数字をどう解釈するか」なのだという場合も確かに存在するのだ。そういうケースを、想像するのも悪くないであろう。現実の問題として、故障解析の結果を生産に反映さ

せる場合、分析結果の数字をどう解釈するかが歩留まりを左右するケースは結構多い。さらに、僅かな歩留まりの差が会社の収益に大きく影響することも、良くある話だからである。すなわち、いざとなったとき、分析屋は問題解決のキャスティングボードを握っている場合が確実にあるのだ。その時に、自分の出した結果から根拠を明確に判断を示せるのは、おそらく分析屋冥利に尽きよう。

言うまでもなく、分析業務に限らず数値データの取り扱いは、数字を扱う全ての分野において大変重要である。にもかかわらず、我が国の教育過程ではそれを比較的重要視してこなかった様に思える。この傾向は世界中でもあるようで、最近査読を頼まれる分析関係の投稿論文でも、ちょいちょい見かけるようになった。学術的分野では、実はその根本である客観性の基本は実験事実の再現性であり、したがって実験結果の数値に対する吟味がもっとも重要なのだが、安易な成果至上主義はそれを軽視する傾向を助長しているように思えてならない。まあ、愚痴はとりあえずこの程度に止めよう。

数値データの取り扱いの基礎的な部分は、すでにISO化されたものもあるが、基本的に統計学の基礎と異なるものではない。数値データの取り扱いの基本は、平均値とばらつきの取り扱いであることは言うまでもないが、これらについての基本的なISOには、ISO 5725 (1981年改訂)があり、それを基にしたJIS Z 8402「分析・試験の許容差通則」が制定されている。このISOは、ISO TC-69 "Application of Statistical Method"の中の"Measurement Method and Result" (SC6)により出されたものである。基本的には、統計学に基づいた数値データの扱い方を

ISOとしたと考えて良い。

一方、検出下限は感度等の議論でも有用であるが、通常は定量分析を行う作業の「しめ」の作業と位置づけられる場合が多い。これについては、ISO 11843が存在する。この制定は1997年 (ISO/TC69 /SC6/WG5)と比較的新しいことから、基礎概念に対する時代に即した検討が常に行われている事が想像できる。ICP/AES等での検量線による検出下限の議論は、ISO 11095 (1996年)として手順が示されている。これについては、日本では結構批判的な意見も多く、本稿の執筆時点では対応するJISは出ていない。以上について、図1に簡単にまとめておく。

本講の大部分は、このISO 5725 (JIS Z 8402)の内容に基づいたものと了解して差し支えない。すなわちここでは

- データの丸め方 (JISに基づく) と有効数字
- 誤差の扱い方 (誤差の伝搬則)
- 異常値の扱い方 (棄却検定)
- 測定の繰り返し回数
- 平均値と区間推定
- 差及び分散の検定
- 検量線の作成

の各項目について、なるべく公式や例題を用いて説明する。

また、検出下限については、1976年のIUPAC勧告とその流れを汲んだISO 11843 (1997年)の内容を紹介する。一方、表面化学分析 (ISO TC-201)においても、ピーク検出の重要な手順としての検出下限の議論が我が国の提案を出発点として進められている

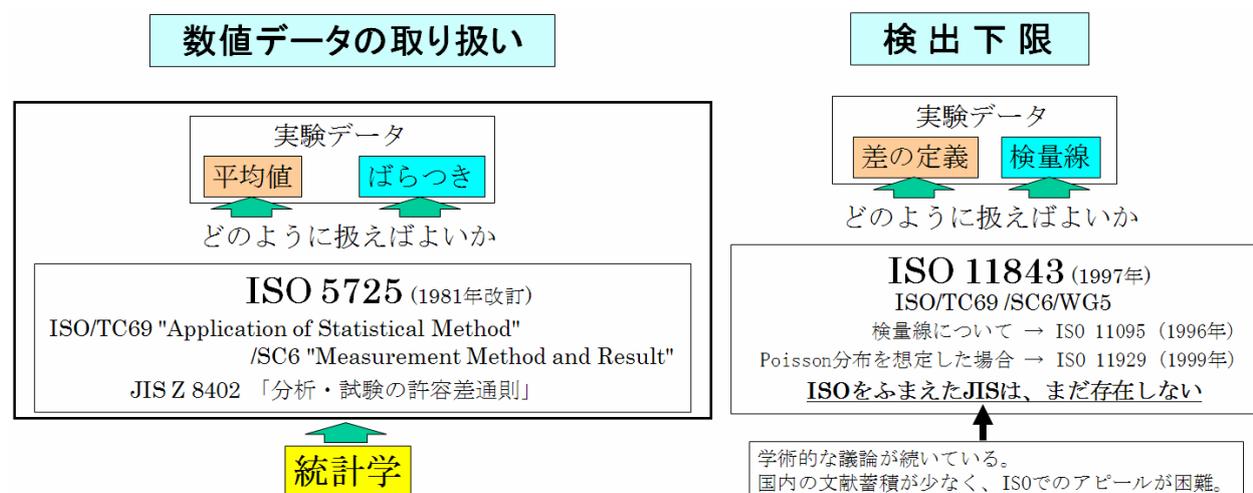


図1 数値処理及び検出下限と関連するISO規格

ことを付け加えておこう。

どの項目にしても、分析現場ではすでに当たり前としてなされていることをおさらいするようなものであるが、まとまった形の資料として提示された例は比較的少ないようであり、本稿が手順書の参考資料としてでも分析現場のお役に立てれば幸いである。

2. 生データの取り扱い

測定を1回行っただけで得た単独のデータでも、もちろんある程度の結論を得られることは否定しない。しかし、その結論の信憑性（信頼性ではない！）を与えているのは、測定条件・測定環境に関する情報、及び経験や理論を根拠とした様々な仮定が存在しているからこそである。

例えば、ある1回のXPSの測定で、二つの試料1及び2に対して、それぞれに対して同じ元素のピークをエネルギー位置でそれぞれ1回ずつ測定して、信号強度が I_1 、 I_2 のデータが得られたとする。この2つのデータから、含有量の大小を判断できるだろうか。数値処理の言葉で言い換えると、この2つの数字に意味のある（有為な）差があると見なせるだろうか。

例えば、「信号強度のばらつき（多くは標準偏差を考えているようである）は、強度の平方根だから．．．」等と考えて、二つの強度に差があるかどうか等を検討しようとするのも、一つの方針であろう。しかし、その陰には多くの議論が隠れている。

まず、測定値は本当に平均値として扱えるのだろうか？ また、その測定の与えるばらつきは、本当に平均値の平方根に従うと言えるのだろうか。特に後者は、大変悩ましい問題を含む。すなわち、装置の安定度をはじめとして、試料の均一性、試料表面の様子（なめらかさ、形状）など、測定に対して理想的な条件であったかどうかで、大きく変わる（通常、ばらつきは数倍程度大きくなってしまう）。

理論的に言うと、「ばらつき（分散）が平均値に等しい」のは、Poisson分布の性質から来ている。電子分光のような粒子計数による計測では、特に低い計数率の時には信号分布がPoisson分布に従うのがはっきり分かる。しかし、ある程度平均値の値が大きくと、Poisson分布は正規分布で近似できることが統計学の教えるところであり、「標準偏差が平均値の平方根で与えられる」という理論的な根拠となっているのである。

しかし、分析結果において客観性の高い信頼性（「信憑性」ではない）を示すことが分析者に課せ

られた使命である。そのためには、分析法や分析条件、あるいは試料に依存しない、客観的な信頼性を保証するための共通の手順が必須となってくる。ここに、データ取り扱いに対する標準化の根本的な意義が存在する。

したがって、本講での以下の議論は全て、1つの測定に関して常に「複数回の正しい計測」を行っていることを前提としている。

それでは、何回程度の繰り返しをすればよいのか？ それについては、いずれ本講の中で触れることにしよう。

2.1 有効数字

装置のデータ処理プログラムが出してくる数字に対して、どこまで信用できるか疑ったことはないだろうか？ 例えば%で示される結果であると、0.1%程度まで結果を示してくれる場合が多いであろう。この場合なら、一般に3桁の結果ということになる（10%を切ると、2桁の表示となるが）。この数字、そのまま信じて良いのだろうか。

3回程度全く同じ条件で同じ測定を繰り返して結果を出してみると、その目安を得ることが出来る。3回とも表示された数値が同じだったのであれば、その実験は少なくとも3桁の精度を確保している、（少なくとも有効桁数は3桁以上ある）と言って良い（しかし、正確かどうかは保証の限りではない）。

一方、3回とも結果の数字が違う場合はどう考えるのであろうか。小数点より上が常に同じ、というのであれば、「おそらく2桁の精度はありそうだな」と考える。そうでなければ．．．、「ひょっとしたらもっと良い測定条件があるのでは」と考えた方がよいかもしれない。

一般に、数値データは、有限の桁数で提示される。それでは、その桁（有効数字の桁数）はどの様に決めるか。ISO 5725では、3つの方法を示しているが、そのなかでも田口の方法あるいはHamakerの方法が推奨されている（この二つの方法は、ほぼ同じである）。

すなわち、分析精度に相当する測定値の標準偏差 σ_M （繰り返し測定を行って得られた測定値から得られた平均値を用いて計算された標準偏差でよい）に対して、 $\sigma_M/3$ の桁を有効数字の桁とする。

[例] 測定値から得られた $\sigma_M = 0.08 \rightarrow \sigma_M/3 \approx 0.03$ （小数点第二位までの数字）となるので、こ

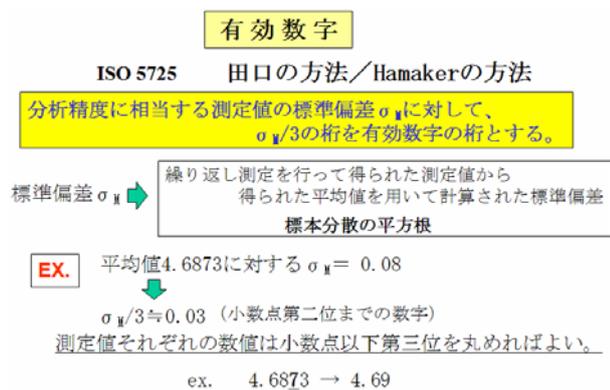


図2 有効数字の桁数の決定法

の場合の測定値それぞれの数値は小数点以下第三位を丸めればよい。

この内容を、図2にまとめておこう。

さらに、このようにして得られた測定値を用いた場合の最終的な平均値と標準偏差の求め方については、1951年にASTMより出されているManual on Quality of Materialsの勧告を具体的な手順として推奨している。これによれば(「推奨する」であって「すべきである」ではないことに注意)

- 平均値は、丸めた測定値より一桁多く示すことを推奨する。
- 標準偏差は多くても三桁以内で示すことを推奨する(三桁で十分)。平均値の桁数と比較し、桁数の少ない方にそろえることを推奨する。
- 乗算及び除算の結果は、計算に使用された数値のもっとも小さい桁数に1を足したけた数まで求めておくことを推奨する。

この内容を、図3に示す。「標準偏差は3桁以内」

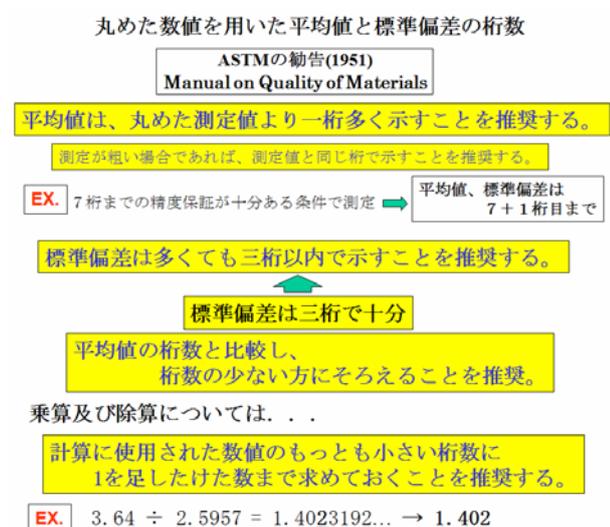


図3 有効桁数に数字を整えるためのASTMの勧告

という点に「え」と思われる方もおられると思うが、もともとデータの各値がかなり大きい場合であると、標準偏差の値も小数どころかかなり大きくなるという事を想像すれば、理解できるであろう。

表面分析に限らず一般に機器分析の世界では、ソフトウェアにより数値処理を行った場合、(もちろん測定条件によるものの)結果については大体有効数字2桁程度の場合が多い。したがって、ベテランの前でデータを示す場合は、3桁程度の数字で桁数を揃えておくと良い印象を与えられる(かも)。

これに対して、重さで分析する場合(湿式分析に於ける重量分析法などは)、有効数字が大変大きい。精密化学天秤の計量限界を思い起こして頂ければ、容易に理解できよう。したがって、精度が大変重要な場合、重量分析は大変有用な手段なのである。

2.2 数値の丸め方

さて、有効数字の桁数(有効桁数)が決まったら、データをこれに併せて整えねばならない。これについては、切り上げや切り捨てはもちろん意味があるが、最もなじみ深くかつ実用的な方法は「四捨五入」である。

普段の業務でも、なんとなく用いている「四捨五入」であるが、これについても規格があるのだ。ここでは、JIS Z 8401「数値の丸めかた」を示しておこう。

ポイントは、四捨五入の対象となっている桁の数字が「5」の場合である。

n桁の数字を作ることを考える。

- (n+1)桁目の数字が4以下ならば、切り捨てる。
- (n+1)桁目の数字が6以上であれば、切り上げる。
- (n+1)桁目の数字が5の場合であれば、以下のようになる。

→ (n+2)桁目の数字が無いか0のとき、n桁が偶数ならば切り捨てる。

→ (n+2)桁目の数字が無いか0のとき、n桁が奇数ならば切り上げる。

→ (n+2)桁目の数字が存在して0でないときは、切り上げる。

例えば、検討の結果、有効数字は3桁であった場合、上記方法でn=3であるから、例えば

- 4桁目の数字が4以下ならば、切り捨てる。

[例] $5.964 \rightarrow 5.96$ $5.960 \rightarrow 5.96$

・4桁目の数字が6以上であれば、切り上げる。

[例] 5.956 → 5.96 5.959 → 5.96

・4桁目の数字が5の場合は、以下のようにする
→ 5桁目の数字が無いか0のとき、3桁目が偶数ならば切り捨てる。

[例] 5.965_ → 5.96 5.965027 → 5.96

→ 5桁目の数字が無いか0のとき、3桁目が奇数ならば切り上げる。

[例] 5.955_ → 5.96 5.955081 → 5.96

→ 5桁目の数字が存在して0でないときは、切り上げる。

[例] 5.955138 → 5.96 5.955762 → 5.96

四捨五入したい桁の数字が5や6の時は、どうするか結構迷うものである。その時の判断として利用して頂ければと考える。また、そのためにも、測定データの桁数はある程度確保しておき、結果をまとめるときはきちんと丸めるという習慣をつけておきたいものである。

3. 誤差の扱い方

3.1 ばらつきと不確かさ

さて、「誤差の大きい測定」と「ばらつきの大きい測定」の意味は同じだろうか。この2つの表現は、どちらもよく使われるが、おおうにして混同されている。単純には、「ばらつきは誤差の一要因」と考えれば、違いがはっきりしよう。ばらつきが大きくても、それから推定された結果が真値に十分近いこともありうるからである(まあ、滅多にないが)。このばらつきや誤差を総称して、「不確かさ」(uncertainty)と呼ぶ。

実は不確かさの概念も、そう単純ではない。ISO国際文書「計測における不確かさの表現のガイド」[1]によれば、まず「不確かさ」の定義であるが、

「測定の結果に付随した、合理的に測定量に結びつけられ得る値のばらつきを特徴づけるパラメータ。」

として定義されている。現実世界では完全な測定はありえず、また真の値も知り得ない。得られるのは測定の結果としての値のみで、当然だがその値と「真の値」との距離(誤差)も知り得ない。そこで従来からの「知り得ない真の値から誤差だけ隔たった測定値」(いわゆる精確度)という仮想的な考え方を、用いずに測定の信頼性を表す(規定する)ために、まず「不確かさ」の定義が決められたのだ。

この定義では現実には得られる「値のばらつき」だけをとり、真の値については何も言及しない。したがって、明らかにこの定義だけでは「不確かさ」の要因(これが「パラメータ」だと思えば良い)をきちんと特定できない。つまり、どのようなものであれ、ばらつきに影響を与えている要因(すなわち、上記の定義を満たすパラメータ)なら、原理的には「不確かさ」として用いられる資格がある。では、これをどうやって扱おうとするのか。

このガイドにおいて不確かさは、「標準不確かさ」「合成標準不確かさ」および、「拡張不確かさ」の3つにわけて定義されている。

まず、不確かさとして標準偏差そのものを用いた場合、これを「標準不確かさ」と呼ぶ。

また、いくつかのばらつきを生ずる要因がわかっていて(要するに各パラメータの帰属ができていて)、測定値がそれぞれの要因によってどう影響を受けているかわかる(要するに、測定値がばらつきの要因となるものの関数と見なせる)場合、それらから計算される不確かさを「合成標準不確かさ」と呼ぶ。この場合でも、どの要因一つをとってみても、それから生ずるばらつきは標準偏差の式で計算できるという大前提がある。

ところが、実際の測定では、その他の評価しようのない要因が多い。したがって、要因の評価しようのない不確かさまで考慮した不確かさの定義が実用上必要となってくる。これが「拡張不確かさ」である。

つまり、評価しようのない要因を繰り込もうとすると、式では扱えなくなる。そのため、合成標準不確かさに対して係数 k をかけ、それを「拡張不確かさ」として定義するのである。ガイドによれば、 k は一般に2~3程度の値が取られる。それよりも、もっと重要な点は

「……すなわち、実際に得られる不確かさの数値は、理想的な系もしくは要因をすべて認識した系から得られることはきわめて希であることから、実際に数値として扱われている不確かさは、本来拡張不確かさと考えべきである。……」(筆者の要約)

という記述があることだ。

推奨もしくは指定された k の値はなく、大体このぐらいであると言う記述しかない。これがガイドたるゆえんでもあろうか。したがって現場では、通常の計算で得られる標準偏差は、この「拡張不確かさ」

に相当するものを見ているのだなと認識するのは、測定結果を解釈する上で必要なことであろう。

いずれにせよ、単に数式の取り扱いだけで十分であれば、統計学の教科書どおりにやればよく、国際規格などは必要ない。しかし、実際の現場では、統計学をそのまま適用するにはあまりに色々と乖離している点があるため、より現場に即した規格やガイドを国際的な合意の下に決めていっているのだということを感じ取って頂ければ、幸いである。

3. 2 誤差の用語とそのイメージ

とは言え、基本的な話を進めるには、どうしても統計学に戻ってしまう。また不確かさも、本来は拡張不確かさをを用いるべきであろうが、数式が（私にとっても）大変複雑であり、それこそ非実用的に見えるので、標準不確かさの世界でこの後も話を進めよう。

さて、「重箱の隅」の一つではあるが、用語について一言触れておきたい。それは、“accuracy”である。この用語は、分野と時代で定義が変わっている。詳しくは、例えば文献[2]に簡潔にまとめてあるが、訳語も「精確さ」と「正確さ」の2つがある。私のように古い時代の人間は、「正確さ」として勉強してきた。

しかし、本講の基本としている ISO 5725 (1992年)、あるいは JIS Z 8402 (1991年)では、これは“trueness”として定義され、「真度」という訳が当てられている。そして、この“trueness”と“precision”（「精度」）とを総合した概念を“accuracy”（「精確さ」）と定義しているのである。

しかし、例えば品質管理用語を定義している JIS Z 8101 (1981年)では昔のままであるし、計測用語を定義している JIS Z 8103 (1984年)では、欧文和文ともにわずかに違った用語が定義されている。

さて、この“trueness”（「真度」）と“precision”（「精度」）がどういう関係にあるかというイメージを図4に示す。

実は、統計学では「母集団」という概念が大変重要である。皆さんが「測定する」場合、統計学では測定データの無限に大きなデータの固まり（これを「母集団」と呼ぶ）を考え、ここからいくつか抜き出してくる（「サンプリングする」）事が、実際の測定に対応すると考えるのである。したがって、目的とする真の値は、この母集団の平均値が仮定されると言うて良い。ただし...

ここで気を付けなければならないのは、母集団の分布はどんなものか全くわからないので、図に示すような左右対称な一つ山の分布（ガウス分布、あるいは正規分布と呼ばれる）が仮定されるのだ、という事である。したがって、「母集団の平均値＝真の値」である保証は何処にもない。母集団の平均値＝真の値」が、真実正しいと誤解している方は結構多いが、その誤解を捨てないと正しい数値データの取り扱いは不可能だといって良い。今まで、このような仮定の下に色々な数値が処理されてきたが、実用上大きな障害となるような場合はほとんど無く、現実と良く即していることが経験的に裏打ちされてきているため、統計学は成り立っているのだと思って頂きたい。

ということで、実際の数値に対して統計学を応用する目的は、この母集団の平均値や分散を推定する事である。それによって、数値を定量的に様々な解釈に用いる「指標」を得ようとするのが数値処理の目的である。

3. 3 誤差の伝搬則

さて、概念論ばかりでもつまらないので、実際にすぐに使用できる話をしよう。

実測された数値は、精度（precision）と真度（trueness）の二つを考慮して扱わねばならないのは、ここまでお話ししてきたとおりである。図4に示すとおり、実測数値は、その測定を一定の条件である定回数だけ繰り返したときに得られるであろう分布を仮定した場合、その平均値と真の値の差を真度、分布の幅（有効数字影響を与えることは先に示したとおり）が精度であると考えるて良い。

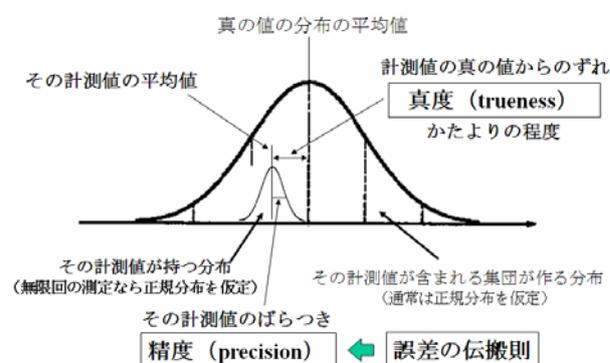


図4 真度(trueness)と精度(precision)

測定された数値をそのまま用いるだけでよいのであれば話は簡単なのだが、その平均値を用いて何らかの計算を行った結果を用いることも結構多いはずである。すなわち、平均値とばらつき（分散）を持つ数値データ同士について、平均値を演算した場合に、結果のばらつきをどう処理すべきだろうか。

これは、「誤差の伝搬則」という形で教科書に出てくる話である。公式もいくつか与えられるが、例えばかけ算やわり算の処理についてもきちんと公式を与えている教科書は、意外に少ない。

ここでも、完全な網羅は難しいので、具体的な演算に関しては主な公式を示しておくにとどめる。

また、あまり注意を喚起されることはないが、以下の式のいずれを使う上でも、その前提条件として、それぞれの数値のばらつきが互いに独立かつランダムであること（共分散が0と出来ること）がどの教科書にも書かれている。つまり、「2つの平均値それぞれをばらつかせている要因は、互いに無関係である」という仮定が必要だと言うことだ。さらに注意すべき事は、平均値のまわりにばらつきが比較的コンパクトにまとまっていること（分散自体が小さいこと）である。あまり分散が大きいと、共分散が0という近似の精度が悪くなるということにも注意しておこう。

式は、数学で言うところの変分の形で書かれているが、各式の δ (ナツカ)には、その(ナツカ)の標準偏差を代入すればよい。

・和と差

$$q = x + \dots + y - (u + \dots + w) \quad \dots(1)$$

の時

$$(\delta q)^2 = (\delta x)^2 + \dots + (\delta y)^2 + (\delta u)^2 + \dots + (\delta w)^2 \quad \dots(2)$$

$$\delta q \leq \delta x + \dots + \delta y + \delta u + \dots + \delta w \quad \dots(3)$$

・積と商

$$q = \frac{x \times \dots \times y}{u \times \dots \times w} \quad \dots(4)$$

の時

$$\left(\frac{\delta q}{q}\right)^2 = \left(\frac{\delta x}{x}\right)^2 + \dots + \left(\frac{\delta y}{y}\right)^2 + \left(\frac{\delta u}{u}\right)^2 + \dots + \left(\frac{\delta w}{w}\right)^2 \quad \dots(5)$$

$$\frac{\delta q}{|q|} \leq \frac{\delta x}{|x|} + \dots + \frac{\delta y}{|y|} + \frac{\delta u}{|u|} + \dots + \frac{\delta w}{|w|} \quad \dots(6)$$

・定数積

$$q = Bx \quad (B: \text{既知の定数}) \quad \dots(7)$$

の時

$$\delta q = |B| \delta x \quad \dots(8)$$

・べき乗

$$q = x^n \quad (n: \text{既知の定数}) \quad \dots(9)$$

の時

$$\frac{\delta q}{|q|} = |n| \frac{\delta x}{|x|} \quad \dots(10)$$

以上から、一変数関数の一般的な計算では

$$q = q(x) \quad \dots(11)$$

に対して

$$\delta q = \left| \frac{dq}{dx} \right| \delta x \quad \dots(12)$$

あるいは、例えば数値計算時に平均値 \bar{x} のまわりの誤差として

$$\delta q = q(\bar{x} - \delta x) - q(\bar{x}) \quad \dots(13)$$

と扱うことが出来る。また、さらに拡張して一般的な関数においては

$$q = q(x, \dots, z) \quad \dots(14)$$

に対して

$$(\delta q)^2 = \left(\frac{\partial q}{\partial x} \delta x \right)^2 + \dots + \left(\frac{\partial q}{\partial z} \delta z \right)^2 \quad \dots(15)$$

かつ

$$\delta q \leq \left| \frac{\partial q}{\partial x} \right| \delta x + \dots + \left| \frac{\partial q}{\partial z} \right| \delta z \quad \dots(16)$$

で、議論が出来る。

実は、(16)は実験を考える上で大変重要な点を示

唆している。すなわち、色々なデータを寄せ集めて計算して答えを求めようとする場合、データの中で飛び抜けてばらつきの大きなものが含まれていると、結果の誤差もそれに含まれてしまうということである。また、(16)のばらつきである δ (ナトカ)は、得られた値をそのまま代入する(平均値で割ったりしていない…規格化されていない)のであるから、利用する実験データの中で飛び抜けて大きな値を示す種類のものを用いるのであれば、結果の精度もそのデータが支配してしまうということである。

したがって、複数種類のデータによる計算を必要とする実験の場合、なるべくばらつきも数字の大きさも揃えておくのが、確からしい結論を得る上で重要であることがわかるであろう。

3. 4 平均値の精確さの意味

では、数値の精確さ(真の値からのずれ、及びばらつき)は、最終的にどのように考えたら良いであろうか。

統計学の基本にもどると、図4を基本にすると、図5に示すような考え方が基本になる。

すなわち、どんなデータであっても真の値1個が存在するのではなく、真の値に相当する平均値を与える無限個の要素からなる数値群(「母集団」)があると考える。この数値群は、様々な要因によるその数値群に固有のばらつきを形成している(とみなす)。真の値は、この母集団の平均値と対応させる。これを「母平均」と呼び、母集団のばらつきを「母分散」と呼ぶ。「母平均=真の値」ではない。多くは、正規分布が想定される(経験的にほぼ成立)。したがって、真の値をどの程度反映しているのか(どの程度ずれているか)は、真の値の代わりに仮定された母平均を用いて、何らかの論理的根拠に基づいた確率的な表現で表されることになる。「真の値に対する区間推定」や「信頼区間」などの概念はここから生まれる。

そうすると、実際に測定で得られた有限個の数値からなるデータは、正しくはかられた物であるなら「母集団から無作為に抜き取られたサンプル(標本)」ということになる。我々は、この標本をひねくり回して母集団を推定し、真の値を知ろうとするのである。したがって、データにより得られた平均値は「標本平均」、ばらつきは通常「標本分散」と呼ばれる。

あるデータ $\{x_1, x_2, \dots, x_N\}$ に対する平均値 \bar{x} を

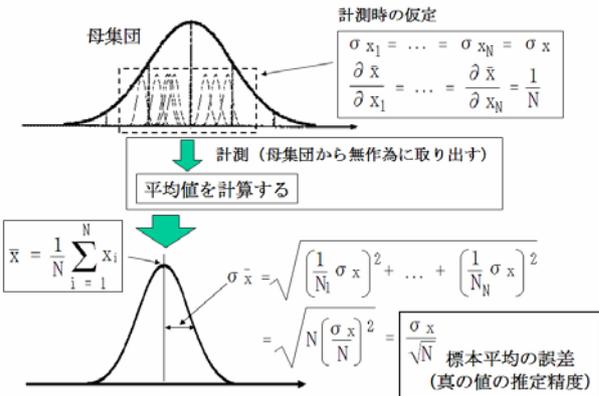
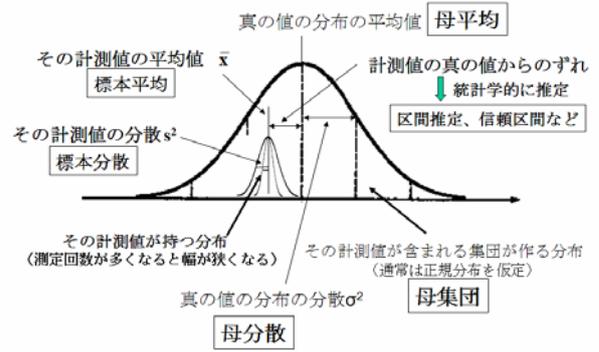


図5 標本平均、標本分散とそのふるまい

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad \dots (17)$$

で与えたとすると、そのばらつき $\sigma_{\bar{x}}$ は(15)より

$$\sigma_{\bar{x}} = \sqrt{\left(\frac{\partial \bar{x}}{\partial x_1} \sigma_{x_1}\right)^2 + \dots + \left(\frac{\partial \bar{x}}{\partial x_N} \sigma_{x_N}\right)^2} \quad \dots (18)$$

で与えられる。

図5で示すように、 $\sigma_{x_1} = \sigma_{x_2} = \dots = \sigma_{x_N} = \sigma_x$ が

つ(17)の微分から $\frac{\partial \bar{x}}{\partial x_1} = \frac{\partial \bar{x}}{\partial x_2} = \dots = \frac{\partial \bar{x}}{\partial x_N} = \frac{1}{N}$ であるとなれば

$$\begin{aligned} \sigma_{\bar{x}} &= \sqrt{\left(\frac{1}{N} \sigma_{x_1}\right)^2 + \dots + \left(\frac{1}{N} \sigma_{x_N}\right)^2} \\ &= \sqrt{N \left(\frac{\sigma_x}{N}\right)^2} \\ &= \frac{\sigma_x}{\sqrt{N}} \end{aligned} \quad \dots (19)$$

これが、標本平均の誤差（真の値の推定精度と考えればよい。データのばらつきではない）である。データ点数 N が大きければ大きいほど小さくなって、推定精度が良くなることがわかるであろう。

ちなみに、算術平均(17)による標本平均を不偏推定値と呼ぶことがあるが、これは各数値との差の二乗が最も小さくなる条件を満足するのが算術平均値であることから来る（最尤推定）。

一方、標本分散を不偏分散として扱うためには、標本平均と各数値の差の二乗和を、データ総数から 1 を引いた値で割る必要がある。これは、以下のようにして示される。母平均（いわば「真の値」）を μ としたとき、標本平均と各数値の差の二乗和について

$$\begin{aligned} \sum_{i=1}^N (x_i - \mu)^2 &= \sum_{i=1}^N \{(x_i - \bar{x}) + (\bar{x} - \mu)\}^2 \\ &= \sum_{i=1}^N (x_i - \bar{x})^2 + 2(\bar{x} - \mu) \sum_{i=1}^N (x_i - \bar{x}) + N(\bar{x} - \mu)^2 \\ &= \sum_{i=1}^N (x_i - \bar{x})^2 + N(\bar{x} - \mu)^2 \end{aligned} \quad \dots(20)$$

であることから、分散は二乗和の期待値である関係を用い（詳しくは、統計学の教科書を参照されたい）(20)式の両辺の期待値をとれば

$$\begin{aligned} E\left[\sum_{i=1}^N (x_i - \mu)^2\right] &= E\left[\sum_{i=1}^N (x_i - \bar{x})^2 + N(\bar{x} - \mu)^2\right] \\ \therefore \sum_{i=1}^N E\{(x_i - \mu)^2\} &= \sum_{i=1}^N E\{(x_i - \bar{x})^2\} + N \times E\{(\bar{x} - \mu)^2\} \end{aligned} \quad \dots(21)$$

したがって、 $\sigma^2 = E\{(x_i - \mu)^2\}$ 、 $s^2 = E\{(x_i - \bar{x})^2\}$ 及

び、次の関係 $\frac{\sigma^2}{N} = E\{(\bar{x} - \mu)^2\}$ から、(21)より

$$\begin{aligned} N\sigma^2 &= Ns^2 + \sigma^2 \\ \therefore \frac{\sigma^2}{N} &= \frac{s^2}{N-1} \end{aligned} \quad \dots(22)$$

となる。これで、標準偏差を計算するとき、なぜ分母から 1 を引くのかがおわかりになったかと思う。

また、以上の説明から、測定で得られる分散（標本分散）は、真の値（母平均）から求められるはずである値よりも常に若干小さめであることもわかっていく。感覚的に、「標準偏差の値よりも、なんとなく見かけよりもばらつきが大きいような気がする」といつも悩んでいる方は、データに関して正常な感覚をお持ちであると言えるのだ。

今回は、平均値に関する議論をした後に、検定の話も少々触れる予定である。

参考文献

- [1] 飯塚幸三監修, ISO 国際文書 計測における不確かさの表現のガイド 統一される信頼性表現の国際ルール, 日本規格協会(1996).
- [2] 藤森利美著, 分析技術者のための統計的方法 第2版, (社) 日本環境測定分析協会, 丸善(1995).