入門講座

測定の不確かさ評価について

—2.　不確かさの算出手順1—

城　野　克　広*

は じ め に

前回から4回に渡り，測定の不確かさについて解説しております．第1回は「不確かさとは何か」，今回と次回は「不確かさの算出手順」，第4回に「不確かさの活用」についてお話していきます．今回は，不確かさの一番大切なルールからお話しして，平均値と標準偏差の関係までを解説していきます．

2·1　2乗して足す

図2.1をご覧下さい．今，私の財布に1000円ほど入っているとします．ですが，1000円だったか，1300円だったか，900円だったか定かではありません．1000円がおおよその値としても，200円くらいのずれは驚くには値しません．

その財布をもって，本屋で本を購入します．本は600円のような気もするのですが，700円だったか，あるいは400円だったか，あいまいです．600円がおおよその値としても，150円くらいのずれはあってもおかしくなさそうです．

そして，本を購入したのち，焼きいも屋さんと出くわしました．焼きいもはひとつ200円．購入できるなら，ぜひ食べたいところです．財布の中身を確認すればよいのですが，たった200円のために財布の中身を確認するなんてプライドが許しません．

さて，この場合，焼きいもの購入に踏み切ってよいものでしょうか？

前回の解説で「あいまいな情報が複数あるとき，それらの両方が関係する事象についての確率を導き出すことは実は困難な問題」と述べました．つまり，上のクイズの厳密な答えは「検討もつかない」です．情報が不十分で決定的なことは言えません．

しかし，前回お話したように，そこをなんとかしてやろうというのが不確かさなのです．不確かさ評価の世界では，2つのあいまいな情報をまとめ上げるやり方は1つに決まっています．それは「**2乗して足す**」ということです．

もう少し具体的に言いますと，「普通のずれ」を2乗して足し，正の平方根を取ります．そうしますと，最終的に知りたい値の「普通のずれ」が求まるのです．（この「普通のずれ」は統計学で言う標準偏差のことです．それについては次節にて．）

上の例に適用してみましょう．財布にはざっと1000円．普通のずれが200円あります．そこから，600円くらいの本を購入．本の値段も150円ほどは普通にずれています．財布の残額はおおよそ1000円－600円＝400円．この400円の普通のずれを上の方法で求めましょう．

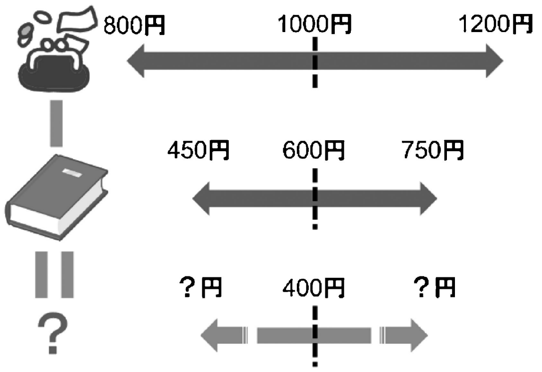


図2.1　「普通のずれ」が分かっている2つの数の差の「普通のずれ」はどのくらいか？上の場合，800－750＝50円から1200－450＝750円ではない．

* 産業技術総合研究所　物質計測標準研究部門主任研究員（〒305-8565 つくば市東1-1-1）
Things about Measurement Uncertainty—2. How to Evaluate It ?(1/2)—; Katsuhiro Shirono (National Institute of Advanced Science and Technology (AIST), Tsukuba)
Keywords; *uncertainty, guide to the expression of uncertainty in measurement* (*GUM*)
2014年6月4日受理[doi:10.2320/materia.54.360]

計算しますと，(200円の2乗)＋(150円の2乗)＝62500円^2＝(250円の2乗)です．つまり，財布に残っているお金はおおよそ400円です．しかし，250円多い650円でも飛び上がって喜んだりはしないし，逆に250円少ない150円でも悲嘆にくれるようなことではありません．

残額が150円というのが普通にあることなら，200円のやきいもを購入するわけにはいかないでしょう．ずいぶんおおざっぱな話ですが，一応結論を出すことができました．やきいもが食せないということは残念ではあるものの，意思決定という観点からは，「検討もつかない」よりだいぶんましと言えましょう．

この「2乗して足す」は，不確かさ評価を考える上で，最も重要な約束です．ぜひ，覚えておいてください．今回は2つしか不確かな要因は有りませんでしたが，これが3つでも4つでも同じことです．「2乗して足す」が基本なのです†．これは，誤差の伝ぱ(でんぱ)則にならって，**不確かさの伝ぱ則**と呼ばれています．

とは言え，2つ不確かな数があったら，それらを足したくなるのが人情ではないでしょうか．上の例で言えば，200円と150円を足して350円を普通のずれと考えたくなるかも知れません．どうして足してはいけないのでしょう．

200円と150円を足すということは，財布の中身はおおよその金額より200円多く，そして本の値段は150円安いことを想定している，あるいは，その逆(財布の中身は200円少なく，本の値段は150円高い)ということになります．前者は二つの不確かな値がともに財布の中身を大きくする方を考えており，後者は小さくする方を考えているわけです．

これはカジノでルーレットを2回続けて回したとき，(赤，赤)か(黒，黒)が出るのを想定しているようなものです．ところが，それ以外の場合，(赤，黒)や(黒，赤)の可能性もあるわけです．

赤2回や黒2回なら，簡単に目撃できる気もします．しかし，5つ不確かな要因があったらどうでしょう．それらのずれを単純に足すということになれば，これはルーレットで5回連続赤かもしくは5回連続黒が出るのを想定しているようなものです．この確率はわずかに16分の1．これは「普通の」確率とは言えないでしょう．

† きちんとした説明は難しいのですが，2乗が出てくることを計算結果から見てみたいと思います．財布の中身について，普通のずれのうちで最も多い場合(1200円)と少ない場合(800円)を考えます．本の値段についても，最も安い場合(450円)と高い場合(750円)を考えます．ここで起こり得る4つのパターンを考えてみましょう．
①財布の中身：1200円，本の値段：450円→750円
②財布の中身：1200円，本の値段：750円→450円
③財布の中身：800円，本の値段：450円→350円
④財布の中身：800円，本の値段：750円→50円
この4つの場合での平均的な値400円からのずれ(＋350円，＋50円，－50円，－350円)を2乗して，平均を取ってみましょう．これが，250円の2乗，つまり残額400円の普通のずれの2乗，になるのは決して偶然ではありません．2乗して足すことで，色々なパターンを考慮した結果としての普通のずれが求まるのです．

つまり，ひとつひとつの普通のずれをただ足してはいけないのは，それらが同じ方向にずれるということが普通に起きることではないからなのです．なぜ2乗するのかはさておき，なぜ足してはいけないかは何となく分かっていただけたでしょうか？ずれそのものではなく，その2乗を計算にうまく使うということは，不確かさ評価ではよく出てきます．ぜひ慣れていただきたいと思います．

2･2　標準偏差のミステリー

先の節で「普通のずれ」,「普通のずれ」と言ってきたものは，正しくは**標準偏差**と呼ばれるものです．数学嫌いの方は，こういう難しそうな単語が出てくるだけでぎょっとするかも知れませんが，これは普通(＝標準)のずれ(＝偏差)という意味でしかありません．具体的にはどういうものなのか確かめていきましょう．

例えば，ある長さの3つの実験データとして，(110 cm, 120 cm, 130 cm)という値が得られたとします．この平均は120 cmです．ずれはどのくらいかと考えると，110－120＝－10(cm), 120－120＝0(cm), 130－120＝＋10(cm)ということになります．

標準偏差は普通のずれですから，この3つの平均を計算したくなります．もちろん，平均というのは，データを足し算して，データの個数で割った値のことです．この場合は，(－10＋0＋10)/3＝0(cm)ということになります．すなわち，標準偏差はゼロということに……なりません．

それはそうでしょう．ばらつきが一定程度あるのに，普通のずれはゼロというのではいかにも不合理です．これは，110, 120, 130という組み合わせだからそうなるのではなく，上の手続きを踏めばいつでもゼロになります．標準偏差とは「ずれの平均」ではないのです．

それではどうしたものか…．「2乗したらいいんじゃないか？」と考え付いた方は，かなり統計学のセンスがあると言えます．実際，そのようにするのが良い結果を生むのです．

－10 cmの2乗は＋100 cm^2，0 cmの2乗は0 cm^2，＋10 cmの2乗は＋100 cm^{2}です．すべての値が正ですから，これらの和は幸運なことにゼロにはなりません．和を3で割って平均をとりますと66.6 cm^{2}ですね．

ただ，そもそもは長さの話をしていたはずなのに，最終的に長さの2乗，つまり面積の単位になっているのはいかがなものでしょう？やはり，ここは正の平方根を取っておいてあげましょう．これでようやく標準偏差8.2 cmが求まります．3つずれが－10, 0, ＋10 cmでしたから，それっぽい数値になったんではないでしょうか？

なんとか求まった…と思って安心した方には驚かすことになり，申し訳ありません．なんと，この話はウソです．実験データ(110 cm, 120 cm, 130 cm)を持っているとき，標準偏差を計算すると8.2 cmにはなりません．ではいくらになるか？これは10 cmになるのです．

上の話のどこかに間違いがあります．このミステリーの犯

人は，ミステリー・ドラマのセオリー通り，一番目立つ俳優ではありません．つまり，「2乗をする」というところはシロです．この手続きに問題はありません．

2時間の刑事ドラマですと，捜査が行き詰ったときに，家族の話題なんかが挿入されて，それが事件を紐解くヒントになりますよね．ここでも，ヒントになる挿話を挟んでみましょう．

主人公である刑事の娘さんの学校での成績がかんばしくありません．主人公は娘さんに，今度のテストで平均点よりもいい点を取ったら，ゲームを買ってあげるという約束をします．けれども，娘さんのテストの点は，またも学年平均の60点を下回る50点．ゲームで釣ってもだめか…と思っていたところ，娘さんは，「学年の平均の60点よりは悪かったけども，私が50点で，隣の家の太郎くんは30点，向いの家の次郎くんの点数は40点だから，うちのご近所の子の平均点は40点じゃない．私はこの近所じゃ頑張っている方だよ．ゲーム買ってくれてもいいじゃない！」などと言い出します．

まったく…ご近所の子の平均点ってなんだよ…．学年の平均と比べるに決まっているだろう．…待てよ．平均点…平均…そういえば，あの事件にも平均が関わっていたな．

さて，お察しになったでしょうか？先のミステリー，最初の方にしれっと登場した「この平均は 120 cm です」が犯人なのです．この 120 cm が本当の平均なのか？ということを突き詰めて考えて下さい．よくよく考えれば，120 cm は（110 cm, 120 cm, 130 cm）という限られた 3 つの値の平均値でしかありません．これは言うなれば，ご近所の子の平均点みたいなものなのです．

では，学年の平均点みたいなものはというと，測定においては実験を何十，何百，何千…理論的には無限回繰返したときの平均値ということになります．もちろん，これは知りようがありません．

統計学においては，本当の平均のことを**母平均**と言います．限られた数の得ることができた値の平均のことを**標本平均**と言います．母平均からのずれと，標本平均からのずれは異なります．**図2.2**に示したように，ひとつひとつの値が大きくても小さくても母平均は決まりきった値で，変化しません．一方で，標本平均の方はそうはいきません．ひとつの値が大きければ，それを使って計算する標本平均は，つられて大きくなります．このため，簡単に言うと，母平均からのずれよりも，標本平均からのずれの方が小さくなりやすいのです．

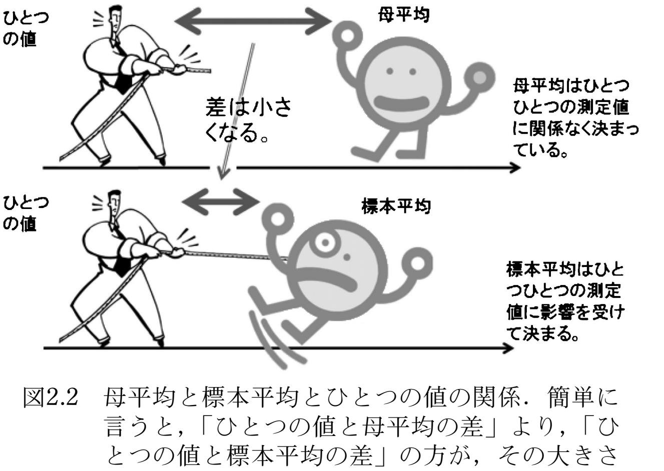


図2.2　母平均と標本平均とひとつの値の関係．簡単に言うと，「ひとつの値と母平均の差」より，「ひとつの値と標本平均の差」の方が，その大きさが小さくなりやすい．

どのくらい小さくなるかというと，標本平均からのずれの2乗を全てのデータについて足した和が，標準偏差の2乗を（データ数－1）倍した大きさになるくらい小さくなるのです[†]．データ数ではなく，（データ数－1）となっていることが，少し小さくなっていることに対応しています．このことを使ってあげれば，標準偏差をうまく計算することができます．

つまり上の計算の過程で，「和を 3 で割って平均をとりますと 66.6 cm^2 ですね.」というところを，「和を 2（＝3－1）で割りますと 100 cm^2 ですね.」とすれば，正しい手順になるということです．いわば，先のミステリーの結末は「この平均は 120 cm です」が首謀犯で，「和を 3 で割って」が実行犯というところでしょう．100 cm^2 の正の平方根は 10 cm です．これを標準偏差とするのが通常の手順なのです．

これはひとつの値が母平均からどのくらい離れているかの基準を与えてくれます．これを使って，どうやって繰返しの値のばらつきを表していくかについては，もう少し考えなくてはいけないことがあります．それについては次節にて．

2·3　平均値の標準偏差

前節では，標準偏差の計算の仕方について説明しました．標準偏差とは，「ある確率で現れる値のひとつひとつが，普通どのくらい母平均からずれるか」という指標です．要するに普通のずれということでしたね．

ところで，1 回の実験の値を報告することもあると思いますが，数回にわたっての平均値を報告することも多いと思います．その時の平均値というのは，どういうばらつきをもっているでしょうか？

何を言っているのかちょっと伝わりづらいかも知れませ

† 一応の理由を説明し（ようとしてみ）ましょう．

ずれの 2 乗和にどれほどの情報が含まれているかを考えてみましょう．本文中でみたように，ずれにはすべて足すと必ずゼロになるという制約が必ずあります．この制約から，1 つのずれが欠けても，それを復元することができると分かります．ということは，データ数を N とすると，標本平均からのずれは N 個あるけれども，それが持っている情報は（$N-1$）個分でしかないのです．

本文に記したように，ずれの 2 乗和は標準偏差の 2 乗の N 倍より小さくなります．どのくらい小さくなるか？実は，そのずれの持っている情報の量に比例して小さくなります．つまり，ずれの 2 乗和は標準偏差の 2 乗の（$N-1$）倍の大きさになっているのです．

これがよくされる一応の説明です．ただ，私自身がピンときていませんし，この説明で分からなくてもお気になさらずに．ちなみに，ずれにどれだけ制約がかかっていないかという意味で，（データ数－1）のことを**自由度**と呼びます．

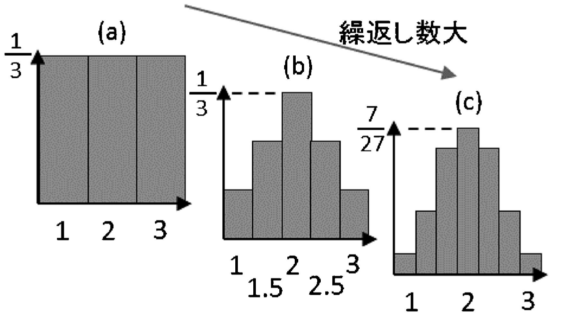


図2.3　1と2と3が等確率で現れる場合の（a）1回の測定の結果の確率分布，（b）2回繰返しの平均の確率分布，（c）3回の繰返しの平均の確率分布．縦軸はその値となる確率．範囲は1から3で変わらないが，繰返しが増えると分布が鋭くなる．

ん．具体的な数値で考えていきましょう．数値の1と2と3がそれぞれ3分の1の確率で現れる場合を考えてみましょう．どの数値がどれだけの確率で現れるかという様相を確率分布あるいは分布と言います．この確率分布の様子を**図2.3**(a)に示しました．

では，2回繰返しの平均にはどんな値が考えられるかというと，2つの数字の選び方には，(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3), (3, 1), (3, 2), (3, 3)という9つパターンがあり，それぞれに1, 1.5, 2, 1.5, 2, 2.5, 2, 2.5, 3という値が平均値として与えられることになります．

集計してみますと，1が1回，1.5が2回，2が3回，2.5が2回，3が1回表れています．この回数はそのまま，その値が平均として現れる確率に比例しますから，「平均値の確率分布」が分かります．これを図2.3(b)に示しました．もし，2回繰返した値の平均値を報告するならば，知りたいのはこの平均の確率分布の標準偏差です．ひとつひとつの値の標準偏差は，極端な話，分からなくてもよいのです．

さて，標本平均は1から3の範囲でばらついていますから，ばらつきの範囲は平均を取るまえの値，言うなれば1回繰返しの値の範囲と変わりません．一方で，2が現れる確率が1や3が現れる確率の3倍も大きいわけです．図を見る限りでも，標本平均の確率分布は1つの値の分布より尖っていると分かります．

さらに，繰返しを3回にすると，27通りの組み合わせができます．その標本平均の確率分布を図2.3(c)に示しました．より真ん中の値の確率が大きくなり，外側の確率が小さくなって，分布の形状が尖っていきます．2が現れる確率は1が現れる確率の実に7倍もあります．このように，繰返し回数を増やすと，分布はどんどん尖っていく…言い換えると，ばらつきは小さくなっていくのです．

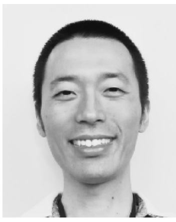
どのくらい小さくなるかと言うと，平均値の標準偏差はひとつひとつのデータの標準偏差を$\sqrt{\text{データ数}}$で割ったものになります†．この関係は，データ数が増えれば増えるほど標準偏差が小さくなることを意味しています．おそらくはみなさまの直感の通り，測定回数を増やすと，測定の精度が大きくなるのです．

もう少し付け加えるならば，$\sqrt{2}$と$\sqrt{3}$というのは，20％ほど違いますが，$\sqrt{100}$と$\sqrt{101}$は1％も変わりません．このことは2回実験するのと3回実験するのは，測定の精度に影響する大きな違いですが，100回実験するのと101回実験するのは，精度に大した違いを与えないということになります．これもわれわれの直感の通りでしょう．

ともあれ，繰返したいくつかの値の平均値の標準偏差は以下の手順で計算できます．ちょっと，ややこしいかも知れませんが，覚える価値のある手順と思います．

★繰返しの平均値の標準偏差の計算の手順
①　標本平均を計算する．
②　標本平均からの差の2乗を計算し，すべて足す．
③　それを(データ数−1)で割り，正の平方根を取る．
④　それを$\sqrt{\text{データ数}}$で割る．

今回は基本的な不確かさの伝ぱの方法と，繰返しによる不確かさ評価について説明してきました．次回は測定器のスペックなどの実験以外の情報による不確かさ評価などについて説明していきたいと思います．　（つづく）


城野克広

★★★★★★★★★★★★★★★★★★★★★★
2007年　東京大学大学院新領域創成科学研究科博士課程修了
2009年　産業技術総合研究所研究員
2013年3月・現職
専門分野：不確かさの評価の技術開発・比較試験の統計処理
◎活動内容などは下記Webサイトをご覧ください．
（https://staff.aist.go.jp/k.shirono/）
★★★★★★★★★★★★★★★★★★★★★★

† これは，不確かさの伝ぱ則から導くことができます．
まずは，平均をデータ数で割る前の値，つまりデータの和について考えてみましょう．データの和にはデータ数だけの不確かさの要因があります．不確かさの伝ぱ則によれば，標準偏差を2乗して足して，正の平方根を取れば，データの和の標準偏差となります．データ数がNならば，ひとつひとつのデータの標準偏差を$\sqrt{N}$倍したものが，データの和の標準偏差です．
次に，データの平均値の標準偏差を考えます．言うまでもなく，データの平均値とはデータの和をデータの数で割ったものです．あるばらつきのある値を定数で割った値の標準偏差がどうなるかは少し難しいですが，標準偏差もその定数で割ったものになります．
このことから，和の標準偏差をNで割ったもの，つまり，ひとつひとつのデータの標準偏差を$\sqrt{N}/N=1/\sqrt{N}$倍したものが平均値の標準偏差と分かります．