寄　稿

不確かさ評価で用いる統計的手法について

産業技術総合研究所　工学博士　田中秀幸

はじめに

不確かさを評価するためには，Aタイプ評価と呼ばれる実験データを統計的に扱い，ばらつきを標準偏差として表す手法と，Bタイプ評価と呼ばれる今までの知識などを用いて標準偏差を推定する手法を組み合わせ，個々のばらつきを表現する必要がある。

一般的にAタイプ評価は，誤差評価などでも行われていた繰り返し測定で得たデータの標準偏差で表すことが多いので最初の理解が進みやすい。また，Bタイプ評価については，今までの概念にはあまり無かった手法であるので，最初は理解しにくいものである。

しかし，特に試験における不確かさを取り扱うときには，個々の不確かさ要因について別々に実験を行うことが難しく，得られるデータがいくつかの不確かさ要因が組み合わさったものになっていることが多い。そのため，統計的手法でその合わさった不確かさ要因を個々の不確かさに分離する必要がある。

本解説では，この上に挙げた問題を解決する手法について説明することを目的にするが，この統計的手法を理解するだけではどのようにこの統計的手法を実際の測定結果に適応するかがわかりにくい。よってこの統計的手法をどのように適応させれば不確かさを算出できるのか，ということを目的に進めたいと思う。

1. 分散分析法

分散分析法とは，先ほど挙げた統計的手法である合わさった不確かさを個々の不確かさに分離する手法である。これを解説する前に分散，標準偏差について確認してみよう。

分散，標準偏差とはいわばばらつきの平均値を表したものであると言っても良いだろう。

分散$V(x)$，標準偏差$s(x)$は下式のように表すことができる。

$$V(x)=\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1} \tag{1.1}$$

$$s(x)=\sqrt{V(x)} \tag{1.2}$$

ここで，x_iはi番目の測定値でi=1～nである。$\bar{x}$は測定値の平均値を表している。

また，平均値$\bar{x}$の分散，標準偏差は，

$$V(\bar{x})=\frac{V(x)}{n} \tag{1.3}$$

$$s(\bar{x})=\frac{s(x)}{\sqrt{n}} \tag{1.4}$$

で表される。このが$s(\bar{x})$がAタイプの不確かさとして扱われる。

ここで，次のようなことを考えてみよう。

(例) ここに測定器A_1，A_2，A_3という3台の機械がある。この測定器が異なることによって現れるばらつきを評価したい。そのために実験を行い次の

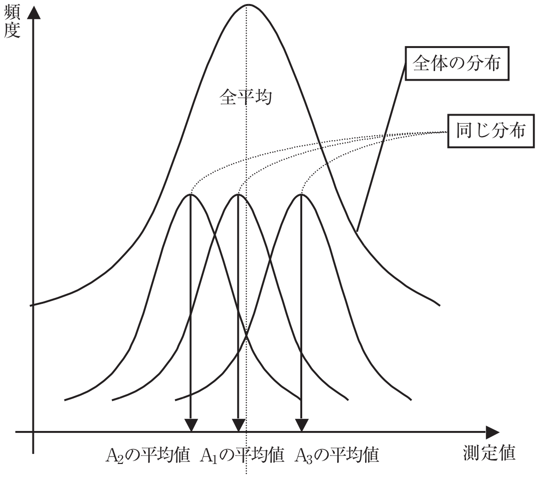

図1　データ構造その1

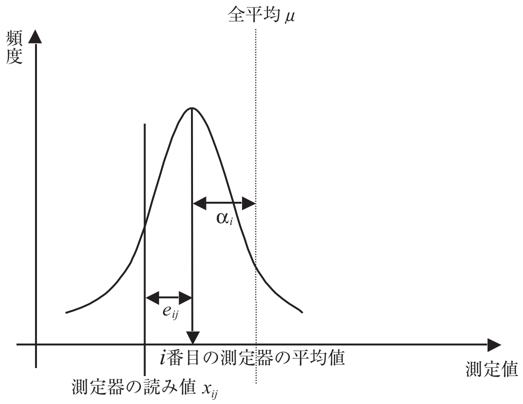

図2　データ構造その2

ような結果を得た。

測定器＼繰り返し	1	2	3	4	5
A_1	5.3	5.2	5.4	5.3	5.2
A_2	5.1	5.2	5.2	5.1	5.4
A_3	5.3	5.4	5.5	5.5	5.4

この結果をどのように処理すれば測定器A_1, A_2, A_3の違いを評価できるだろうか。

よく行われている方法を解説してみよう。まず，A_1，A_2，A_3の測定器それぞれで測定された5つのデータの平均値を算出して，その平均値を用いて分散，標準偏差を算出する方法である。すなわち，

	1	2	3	4	5	平均
A_1	5.3	5.2	5.4	5.3	5.2	5.28
A_2	5.1	5.2	5.2	5.1	5.4	5.20
A_3	5.3	5.4	5.5	5.5	5.4	5.42

全平均：$\bar{\bar{x}} = 5.30$　(1.5)

分散：$V(\bar{x}_i) = \dfrac{\sum_{i=1}^{3}(\bar{x}_i - \bar{\bar{x}})^2}{3-1} = 0.0124$　(1.6)

標準偏差：$s(\bar{x}_i) = \sqrt{0.0124} = 0.1114$　(1.7)

ここで，$\bar{x}_{i\cdot}$は各測定器で測った測定値の平均値である。また，i=1，2，3である。

この方法は理解しやすい。それぞれの測定器の平均値の分散，標準偏差を求めることによって，機械による違いを数値化したということである。しかし，この方法は多くのところで用いられているが，この方法では支障が出ることがある。この測定のデータ構造から考えてみよう。この測定のデータ構造は，

$$x_{ij} = \mu + \alpha_i + e_{ij} \tag{1.8}$$

と表すことができる。ここで，μは全平均，α_iは個々の測定器によるかたより，e_{ij}はどの装置にも現れる繰り返し誤差を表している。**図1**にこのデータ構造を模式的に表したものを示す。つまり，全平均μから測定器A_iが持つかたよりα_i離れたところに各測定器の平均値があり，更にそこから繰り返し誤差分e_{ij}離れたところに測定値x_{ij}が存在するということである（**図2**）。

では，先ほど算出した分散，標準偏差が何を表しているのかを知るために式を元にして考えてみよう。

$$\bar{\bar{x}} = \frac{\sum_{i=1}^{3}\sum_{j=1}^{5}(\mu+\alpha_i+e_{ij})}{3\times5} - \frac{15\mu+5\sum_{i=1}^{3}\alpha_i+\sum_{i=1}^{3}\sum_{j=1}^{5}e_{ij}}{15}$$

$$= \mu+\frac{1}{3}\sum_{i=1}^{3}\alpha_i+\frac{1}{15}\sum_{i=1}^{3}\sum_{j=1}^{5}e_{ij} = \mu+\bar{\alpha}+\bar{\bar{e}} \qquad (1.9)$$

$$\bar{x}_i = \frac{\sum_{j=1}^{5}x_{ij}}{5} = \frac{\sum_{j=1}^{5}(\mu+\alpha_i+e_{ij})}{5}$$

$$= \frac{5\mu+5\alpha_i+\sum_{j=1}^{5}e_{ij}}{5} = \mu+\alpha_i+\frac{1}{5}\sum_{j=1}^{5}e_{ij} = \mu+\alpha_i+\bar{e}_i \qquad (1.10)$$

式（1.6），（1.9），（1.10）より，

$$V(\bar{x}_i) = \frac{\sum_{i=1}^{3}(\bar{x}_i-\bar{\bar{x}})^2}{3-1} = \frac{\sum_{i=1}^{3}\{(\mu+\alpha_i+\bar{e}_i)-(\mu+\bar{\alpha}+\bar{\bar{e}})\}^2}{3-1}$$

$$= \frac{\sum_{i=1}^{3}\{(\alpha_i-\bar{\alpha})+(\bar{e}_i+\bar{\bar{e}})\}^2}{3-1}$$

$$= \frac{\sum_{i=1}^{3}(\alpha_i-\bar{\alpha})^2}{3-1}+\frac{\sum_{i=1}^{3}(\bar{e}_i+\bar{\bar{e}})^2}{3-1}+\sum_{i=1}^{3}(\alpha_i-\bar{\alpha})(\bar{e}_i+\bar{\bar{e}})$$

ここで，$\sum_{i=1}^{3}(\alpha_i-\bar{\alpha})(\bar{e}_i+\bar{\bar{e}})$ は装置と繰り返し誤差が独立であれば0になる。この独立という意味は，互いに影響しあわない，と考えればよい。ここでは深く触れない。興味がある方は，統計の教科書に載っている「相関」について調べていただきたい。

この項を0とすることができれば，

$$V(\bar{x}_i) = \frac{\sum_{i=1}^{3}(\alpha_i-\bar{\alpha})^2}{3-1}+\frac{\sum_{i=1}^{3}(\bar{e}_i+\bar{\bar{e}})^2}{3-1} \qquad (1.11)$$

となる。これを見て分かるように，$V(\bar{x}_i)$ として求めたものには，単に測定器間のばらつきだけではなく，繰り返し誤差の部分も含まれてしまっている。よって，技術的に繰り返し誤差を測定器間のばらつきと比べ非常に小さくできるのであればこの方法が良い。しかし，繰り返し誤差が無視できないほど大きな場合では，この方法だけでは，装置間のばらつきを過大評価してしまうことになる。分散分析法はこのように二つ以上のばらつきの要素が含まれているデータから各ばらつきに分解するものである。次にその計算方法を解説する。

今まで用いていた例を一般化してみよう。ばらつきを知りたいものがk個あり，繰り返しをn回行うこととする。そして，各要因の二乗和を以下のように求める。

$$S_A = \sum_{i=1}^{k}\sum_{j=1}^{n}(\bar{x}_i-\bar{\bar{x}})^2 \qquad (1.12)$$

$$S_e = \sum_{i=1}^{k}\sum_{j=1}^{n}(x_{ij}-\bar{x}_i)^2 \qquad (1.13)$$

$$S_T = \sum_{i=1}^{k}\sum_{j=1}^{n}(x_{ij}-\bar{\bar{x}})^2 \qquad (1.14)$$

ここで，$S_T = S_A + S_e$ が成り立つ。

また，各自由度は，

$$\phi_A = k-1 \qquad (1.15)$$

$$\phi_e = k(n-1) \qquad (1.16)$$

$$\phi_T = nk-1 \qquad (1.17)$$

となり，これも $\phi_T = \phi_A + \phi_e$ が成立する。

ここで表を作ってみると，

	二乗和	自由度	分散
因子A	S_A	ϕ_A	$V_A = S_A/\phi_A$
誤差e	S_e	ϕ_e	$V_e = S_e/\phi_e$
トータルT	S_T	ϕ_T	

となる。トータルの分散は重要ではないのでここでは省いてある。

更にこの分散が本当は何を表しているのかということを求めると，

$$E(V_A)=\sigma_e^2+n\sigma_A^2 \tag{1.18}$$

$$E(V_e)=\sigma_e^2 \tag{1.19}$$

となる。ここで，Eは「期待値」というものを表しており，簡単に言うと本来そうであるべきものを意味する。つまり，V_Aが本来表しているものは，$\sigma_e^2+n\sigma_A^2$ である，ということになる。また，σ_A と σ_e が何を表しているかということであるが，ここでは，Aの標準偏差，eの標準偏差であるととりあえずは理解しておいてほしい。ここの詳しい話は次節で行う。

式（1.18）を見ると，最初に計算したように，Aの分散には誤差が含まれている。よって，純粋なAの分散を求めようと思えば，

$$\sigma_A^2=\frac{V_A-V_e}{n} \tag{1.20}$$

で求めることができる。また，σ_e は繰り返し測定によるばらつきの標準偏差を表している。ここでは，ばらつきを見たい因子を1つしか取っていないが複数取ることもできる。去年行った壁の防音試験では因子を3つと繰り返しを行った分散分析を適用している。また，因子が二つまでならMicrosoft Excelにも分散分析の機能が入っている。この機能は通常のままではインストールされないが，「ツール」メニューの「アドイン」で，「分析ツール」にチェックをつけるとCDROMからインストールされる。そうすると，ツールに新しいメニューの「分析ツール」が追加され，それを選ぶと各種の統計ツールが使えるようになる。その中に分散分析も入っている。しかし，気をつけなければならないのは，分散分析の計算結果までは表示するが，分散の期待値は表示されないので自分で算出する必要がある。また，その他の統計解析ツールにも分散分析の機能を持つものが多くある。

2. 統計解析と不確かさ評価の結びつけ

不確かさ評価について一般の人たちと話してみるとよくある質問がある。それは，「統計もある程度理解した。また，不確かさについてもある程度理解した。しかし，統計と不確かさをうまく結びつけて考えられない。どうしたらよいか？」というものである。これに関して私は最近一つの答えを持つようになった。それは，母集団と標本についてきっちりと区別して考えられないことが原因である。

測定に特化して考えてみると，母集団とは測定を無限回行って得られるもので，標本とは，有限回の測定で得られるものである。言い換えると，測定とは，ある決まった母集団から測定を一回行うごとに一つの標本を取り出すという行為ということになる。ここで，今求めているものが，母集団についてのものなのか，それとも標本についてのものなのかの区別が曖昧であれば，不確かさ評価の時に何を求めているのかが分からなくなってしまう。この時に不確かさ評価と統計的手法の乖離が起きる。

Aタイプの不確かさの算出方法から見てみよう。

① 繰り返し測定を行ってデータをn個取る。

② そのデータから平均値,標準偏差を算出する。

③ 標準偏差を$\sqrt{n}$で割る。

このステップでAタイプの不確かさは見積もれるが，実はこの中に母集団と標本についての取り扱いが多く含まれている。その解説も一緒に書き込んでみると，

① 繰り返し測定を行って測定の母集団から標本をn個取り出す。

② そのデータから標本標準偏差，標本平均を求める。

③ 求められた標本平均が母平均であるとし，また標本標準偏差が母標準偏差であるとする。これによって，測定の母集団が推定される。

④ 得られた測定値は，この推定された母集団からn'個取り出した標本であると考える。

⑤ 推定された母集団からn'個の標本を取り出したときの平均値はn'個の標本を取り出すごとに変化する。その平均値の変化の度合いを考えると母標準偏差を$\sqrt{n'}$で割ったものになる。

⑥ 母標準偏差は標本標準偏差によって推定されているので，結局は標本標準偏差を$\sqrt{n'}$で割ったものが平均値の標準偏差となり，これがAタイプの不確かさになる。

このように，簡単に見えるAタイプの評価も実は母集団を推定して，その母集団から標本をn'個取り出したときの平均値の変動を表しているのである。よって，今求めたのは標本標準偏差であるのか，もしくは母標準偏差を推定したものであるのか，ということの区別がついていないと，複雑なAタイプの不確かさを見積もっている場合には混乱し，間違いを起こす可能性が多い。

また重要なのは，母集団を推定するまでの手順（①〜③）と，平均の標準偏差を求める手順（④〜⑥）は完全に一体のものではなく分離して考えることもできる。つまり，母集団の推定はある実験で行って，そこからn'個の標本を取り出したときに平均値がどのくらい変動するのかということを計算するのはまた違う実験によって行って良いということである。それを分かりやすくするために今回はAタイプの見積もり法を示したものの中でnとn'で区別している。

これにはどのようなメリットがあるかというと，もし実際の試験を行う条件が，繰り返し測定は3回だけで行う，ということになっていたとすると，3つの測定値から母集団を推定しなければならない。しかし，3つの測定値だけでは本当に母集団をうまく推定できているかが疑問である。そのような場合には，母集団をよく知るために不確かさ評価を行うときに数多くの繰り返し測定を行えばよい。そうすると母平均，母標準偏差がうまく推定できるであろう。そして測定値は，ここで推定された母集団からの標本であると考える。では今回の場合，3回の繰り返し測定から値を得るということになっているので，事前に求めておいた母標準偏差を$\sqrt{3}$で割ればその母集団から3つの標本を取り出したときの平均値の標準偏差を求めることができる。これがAタイプの不確かさとなる。

この方法は，母集団と標本の関係が頭の中で整理され区別をつけられていなければ使用することができないであろう。この方法をうまく使えば非常によい不確かさの推定ができる。しかし，この方法も用いるときに気をつけなくてならないことがある。事前に数多くの繰り返し測定を行うわけであるが，もしこの繰り返し測定を行うときの条件が，実際の試験等を行うときの条件と食い違いがあってはならない。違いがあると，事前の測定の母集団と実際の測定の母集団が異なる，ということになってしまう。そうすると，実際の測定が事前に求めた測定の母集団からの標本とは言えなくなってしまうのである。これでは，不確かさ評価がうまくできたとは言えない。

次に，この母集団，標本と分散分析の関係について解説する。1節の内容を思い出してほしい。分散の期待値を式（1.18），式（1.19）に示していた。これを式（2.1），式（2.2）として再掲する。

$$E(V_A)=\sigma_e^2+n\sigma_A^2 \tag{2.1}$$

$$E(V_e)=\sigma_e^2 \tag{2.2}$$

このσ_A，σ_eは母標準偏差のことを表している。すなわち，因子Aが原因で現れるばらつきの

母標準偏差と，繰り返し誤差が原因で現れるばらつきの母標準偏差を表しているということである。分散分析法を用いると，各因子が原因のばらつきの母標準偏差を知ることができる。では，この得られた母標準偏差（の推定値）をどのようにして用いればよいのかを考えてみよう。

最初に挙げた例をもう少し細かな部分を詰めて用いてみよう。

（条件）

・A_1，A_2，A_3という測定器を用いて測定を行っているが，一回の試験ではA_1，A_2，A_3のいずれか1つの測定器を用いて測定する。A_1，A_2，A_3は同じ頻度で使われている。

・普段の試験では5回繰り返し測定を行っているが，今回は不確かさを推定するので，良い母集団を推定するために10回の繰り返し測定を行った。

このような条件での不確かさはどのようになるか考えると，装置によって値が少し違うので，装置による不確かさを見積もる必要がある。今回の場合は，1回の実験でA_1，A_2，A_3のいずれか一つの測定器しか用いない。これは，装置の違いによるばらつきを表す母集団からある装置一つだけピックアップしている。つまり，母集団から1つだけ標本を抜き出しているということである。よって，装置の違いによる不確かさは，

$$u_A = \frac{\sigma_A}{\sqrt{1}} = \sigma_A \qquad (2.3)$$

となる。

一方，繰り返し測定による不確かさは，実際の試験では5回の繰り返しを行いその平均値を値としている。よって，繰り返し測定によるばらつきの母集団から5つのデータを取り出し，その平均値のばらつきが不確かさとなるので，

$$u_e = \frac{\sigma_e}{\sqrt{5}} \qquad (2.4)$$

となる。式（2.3）は3つの装置を用いているが$\sqrt{3}$で割らない。また，繰り返し測定の方は母集団を求めるのに10回の繰り返しを行っているが，$\sqrt{10}$で割らない。重要なのは，実際の試験でどのような実験を行っているかということなのである。

最後に

不確かさ評価を行うのに統計的な知識が必要であることは分かっていただけたと思う。しかし，不確かさ評価を行うために本当に重要なのは統計の知識ではない。一番重要な知識は，その試験，測定をどれくらいよく知っているか，ということである。あくまでも統計は道具でしかない。統計の専門家は不確かさの評価はできない。なぜなら測定に関しては何も知らないからである。これを読んだ方も統計は道具の一つにすぎない，という意識を持ち，その道具を使いこなしてほしい。良い道具を用いれば今までより良い結果を得ることができるはずである。