

不確かさ評価の基礎

名古屋大学 山本章夫

目次

1.はじめに	3
2.科学における予測の不確かさ	3
3. 誤差と不確かさ	4
3.1 誤差と不確かさの違い	4
3.2 誤差の種類	6
3.3 不確かさ評価の種類	8
4.確率密度関数	10
4.1 確率関数	11
4.2 確率密度関数	11
4.3 累積分布関数	13
5.期待値、分散、標準偏差	14
5.1 期待値	14
5.2 分散	15
5.3 標準偏差	17
5.4 モーメント	17
6.母集団の推定	17
6.1 母集団とは	17
6.2 平均値の不確かさ(標準偏差)	18
6.3 標準偏差の推定	19
6.4 信頼係数と信頼区間	21
6.5 推定臨界下限値増倍率	25
6.6 注意事項	25
7.共分散、相関係数、分散共分散行列	27
8.不確かさの伝搬	29

8.1 概要	29
8.2 多変数関数の誤差伝搬	30
8.3 二変数の場合の誤差伝搬.....	31
8.4 不確かさ伝搬の SANDWICH FORMULA	33
9.まとめ.....	35
謝辞.....	35
APPENDIX. 推定臨界下限増倍率の算出法	36
A.1 必要となる統計学知識	36
A.2 実効増倍率の母集団の平均値と標準偏差がわかっている場合	37
A.3 実効増倍率の母集団の平均値と標準偏差が不明であるが、これらが十分に精度良く推定できている場合	38
A.4 実効増倍率の母集団の平均値が不明であり、標準偏差のみがわかっている場合	40
A.5 実効増倍率の母集団の平均値も標準偏差も不明な場合	44

1.はじめに

次の質問に答えられるだろうか？

- ①不確かさと誤差の違いについて説明せよ
- ②複数回測定を行って得られた測定値の平均値の不確かさについて説明せよ
- ③不確かさには、性質の異なる二つの種類がある。これらについて説明せよ

①～③全てに正確に答えられる人はあまりいないのではないだろうか。原子力は、主として設計(計算)で製品の性能を保証している。そのため、計算結果がどれぐらい信頼できるか(疑わしいか)についての定量的な評価がきわめて重要になることは明らかである。不確かさ評価は、計算結果の信頼性を定量的に保証するために不可欠な考え方である。

本稿では、不確かさ評価を行うにあたって必要となる基礎知識について解説することを試みる。

2.科学における予測の不確かさ

科学の世界では、実世界における現象は、いくつかの過程をへて「予測」される。まず、測定結果に基づき、理論モデルが作成される。これらの理論モデルを(多くの場合は近似的な)解析モデルに置き直し、この解析モデルを用いて数値解析などにより予測結果を得る。このように、予測は複雑な手順を踏んで行われるものであり、その結果が「疑わしくなりえる」要素には以下のように様々なものがある。

- ①観察(測定)結果：現実世界の測定は有限回しか行うことができない。また、測定は、長さであれば定規、温度であれば温度計といった何らかの補助手段を用いて行われ数値化されるが、これらの補助手段により得られる値は、当然ながら厳密なものではない。また、観察したい物理量が直接測定できず、何らかの補正や換算係数を用いて測定値を得る場合もある。これらのことは、出発点となる測定結果にはある程度の「疑わしさ」が存在することを示している。
- ②理論モデル：測定結果に基づき理論モデルを構築する過程においても、我々の知識が完全であることはあり得ず、このために理論モデルには「疑わしさ」が必ず存在する。原子に関する理論モデルが変遷してきたことは、この一例である。
- ③解析モデル：理論モデルは、そのままでは解くことができない場合が多い。例えば、Boltzmann 輸送方程式はその一例である。そのため、理論モデルに基づき、解析モデルを作り、結果を得ることとなる。この過程では、様々な離散化、仮定、単純化などが行われるが、これらはいずれも予測の「疑わしさ」に結びつくものである。
- ④入力データ：解析を行うためには、解析モデルに与える入力データが必要になる。例え

ば原子炉の炉心解析では、燃料の組成、形状、温度、原子炉出力、と言ったパラメータが入力データとして必要となる。これらのパラメータは、全て測定により得られたものであることから、不確かさが必ず存在し、この不確かさは予測の「疑わしさ」に結びつく。

以上のことから、程度は様々であるものの、予測値は「疑わしい」ものであることがわかる。従って、予測値を有効に活用するためには、「ではどの程度疑わしいのか」に対して定量的に答える必要がある。これが不確かさ評価である。

3. 誤差と不確かさ

3.1 誤差と不確かさの違い

誤差と不確かさは、よく混同されるが、異なるものである。

誤差(error)：測定(計算)しようとする対象についての、測定(計算)された値と真の値との差(測定(計算)値－真の値)を言う。一般に、真の値は測定によっても得ることはできない。これは、測定にも不確かさが必ず含まれているからである。従って、定義通りの「誤差」は、永久に求めることはできない。しかし、十分に注意を払って得た測定値は、真の値に十分近いと考えて良いであろうから、炉物理の分野では、「計算値－測定値」を一般的に計算の誤差と呼んでいる。

不確かさ(uncertainty)：簡単に言うと、測定(計算)値のばらつきを特徴付けるパラメータである。別の言い方をすると、測定(計算)値の疑わしさを数値で表したものである。より正確には、測定(計算)の結果から得られる真の値の「候補」のばらつきを特徴付けるパラメータである。炉物理の分野では、不確かさは、誤差の標準偏差として評価することが一般的である。

あるパラメータに対する誤差 X_i に対し、誤差の平均(\bar{X})、標準偏差(不確かさ) $\sigma(X)$ 、「誤差の平均」の標準偏差 $\sigma(\bar{X})$ の一般的な計算式を以下にまとめる。ここで、 n は計算もしくは測定回数を示す。なお、標準偏差の計算方法の詳細については、5章で説明を行う。

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \quad (3.1)$$

$$\sigma(X) = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}} \quad (3.2)$$

$$\sigma(\bar{X}) = \frac{\sigma(X_i)}{\sqrt{n}} \quad (3.3)$$

■例1：ここに、ある決定論的な核特性解析コードがあるとする。5つの異なる臨界実験について解析を行った結果、実効増倍率は1.001, 1.002, 0.998, 1.003, 1.000となった。それぞれの臨界実験において測定された実効増倍率が1.000であるとき、この核特性解析コードの誤差と不確かさはどのように評価できるか？なお、実効増倍率の測定値の信頼性は十分に高く、不確かさは十分に小さいものとする。

測定値の信頼性が高いため、測定値を真の値と見なすと、誤差はそれぞれの炉心に対して、0.1% $\Delta k/k$, 0.2% $\Delta k/k$, -0.2% $\Delta k/k$, 0.3% $\Delta k/k$, 0.0% $\Delta k/k$ である。従って実効増倍率の誤差の平均値は(3.1)式より

$$\frac{0.001 + 0.002 + (-0.002) + 0.003 + 0.000}{5} = 0.0008$$

また標準偏差(不確かさ)は(3.2)式より

$$\sqrt{\frac{(1.001 - 1.0008)^2 + (1.002 - 1.0008)^2 + (0.998 - 1.0008)^2 + (1.003 - 1.0008)^2 + (1.000 - 1.0008)^2}{(5-1)}} \\ = 0.0019$$

この例における不確かさは、核データ、炉心解析コードのモデルや数値解析アルゴリズム、実験体系や使用されている材質の不確かさ s などに起因したものである。□

■例2：モンテカルロコードで、ある臨界実験(1体系)の解析を行った。初期乱数を変えて5回解析を行った結果、実効増倍率は1.001, 1.002, 0.998, 1.003, 1.000となった。この臨界実験における実効増倍率の測定値は1.000であり、十分に信頼性は高いものとする。このモンテカルロコードの誤差と不確かさはそれぞれいくらか。

誤差はそれぞれの炉心に対して、0.1% $\Delta k/k$, 0.2% $\Delta k/k$, -0.2% $\Delta k/k$, 0.3% $\Delta k/k$, 0.0% $\Delta k/k$ である。従って、五回の計算における誤差の平均値は以下のように計算できる。

$$\frac{0.001 + 0.002 + (-0.002) + 0.003 + 0.000}{5} = 0.0008$$

また、標準偏差は、

$$\sqrt{\frac{(1.001 - 1.0008)^2 + (1.002 - 1.0008)^2 + (0.998 - 1.0008)^2 + (1.003 - 1.0008)^2 + (1.000 - 1.0008)^2}{(5-1)}} \\ = 0.0019$$

となる。

この例では、一つの体系について、初期乱数を変えて計算を行っている。これら 5 回の計算から得られた誤差の標準偏差をモンテカルロコードの不確かさとするのは妥当であろうか？この例では、同じ体系を計算しているため、決定論的手法を用いれば、全く同じ実効増倍率が得られるはずである。しかし、モンテカルロ法では、統計誤差のため、毎回異なる実効増倍率が得られる。従って、5 回の計算結果から得られた標準偏差(0.0019)は、統計誤差に基づく不確かさを与えることとなる。

せっかく初期乱数を変えて 5 回も計算したのであるから、それを有効に活用したいと思うのは自然であろう。そのためには、5 回の計算の誤差の平均をモンテカルロ法の誤差とするのが良いと考えられる。そこで次に、誤差の平均値がどの程度確からしく見積もられているか、つまり、「誤差の平均値」の不確かさがどの程度あるかについて考えてみよう。今回は、5 回の計算により平均を求めているので、(誤差の標準偏差でなく)「誤差の平均値」

の標準偏差は(3.3)式より $\frac{0.0019}{\sqrt{5}} = 0.00086$ である。この結果から、この炉心に関して、こ

のモンテカルロコードの計算誤差が 1.0008 ± 0.00086 の範囲に入る確率が約 68%、 $1.0008 \pm 2 \times 0.00086$ の範囲に入る確率が約 95%、 $1.0008 \pm 3 \times 0.00086$ の範囲に入る確率が約 99.7%となる、と書きたいところであるが、これは正しくない。正しい評価については、6.3 節で説明する。□

ここで評価した誤差の平均値は、この炉心に対し、初期乱数を変えてモンテカルロコードの計算を多数繰り返し実行することにより低減することができる。つまり、「誤差の平均値」に不確かさが含まれているのは、今回の(5 回の)計算では、(初期乱数を変えて無限回モンテカルロコードの計算を繰り返した)母集団の一部をサンプリングしていることにより発生したものであると解釈できる。

3.2 誤差の種類

前節では、明示しなかったが、誤差には、偶然誤差と系統誤差の二種類がある。

偶然誤差：さいころを振ったとき出る目のように、本来、偶然に支配される要因によって発生する誤差のこと。計算に含まれる偶然誤差は、測定や計算を繰り返すことにより見積もることが可能である。なお、偶然誤差自体は、数多くの測定を行ったとしても減らすことはできない。ただし、誤差の平均値については、数多くの測定により、より正確な、言い換えるとより不確かさが小さい結果を得ることができる。放射性核種の崩壊、原子核と中性子の相互作用などは、偶然誤差を含んでいる。モンテカルロコードにおいて、中性子が原子核と反応するときの不確かさそのものは計算を繰り返し行っても低減することはできない。一方、実効増倍率や反応率などの統計誤差も偶然誤差に起因するものであるが、統計誤差は「平均値に対する不確かさ」であることから、計算回数(つまりヒストリー数)を増やすことによりこれを低減することができる。

系統的誤差：毎回の計算に同じ影響を与える誤差。この誤差の要因が明らかになっているかどうかは分からない。炉心解析においては、例えば拡散近似、メッシュ誤差、エネルギー群の縮約誤差などが系統誤差になり得る。計算に含まれる系統的誤差を評価するためには、(十分に信頼できる)測定値との比較、他の計算方法による計算結果との比較などを行う必要がある。

また、これらの誤差の定義に関連して、正確さに関する言葉の定義を以下に示す。

正確度(accuracy)：測定(計算)値が、真値にどの程度近いのかを示す尺度。系統的誤差の小さいことを示す。

精度(precision)：測定(計算)値を複数回行ったときの値のばらつきを示す尺度。偶然誤差が小さく再現性が高いことを示す。なお、真値とのずれは考慮していない。

図 3.1 にダーツを投げたときの結果を示す。(a)~(d)の分布は以下のように解釈できる。

- (a)偶然誤差小、系統誤差小。もしくは正確度高、精度高。
- (b)偶然誤差小、系統誤差大。もしくは正確度低、精度高。
- (c)偶然誤差大、系統誤差小。もしくは正確度高、精度低。
- (d)偶然誤差大、系統誤差大。もしくは正確度低、精度低。

一方、真の値が分からない場合、図 3.2 のようになり、偶然誤差が大きい小さいか、精度が高いか低いかは判断することができるが、系統誤差の大小、正確度の高低は判断できない。

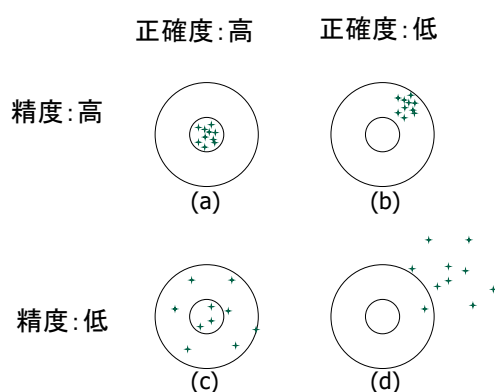


図 3.1 偶然(ランダム)誤差と系統的誤差

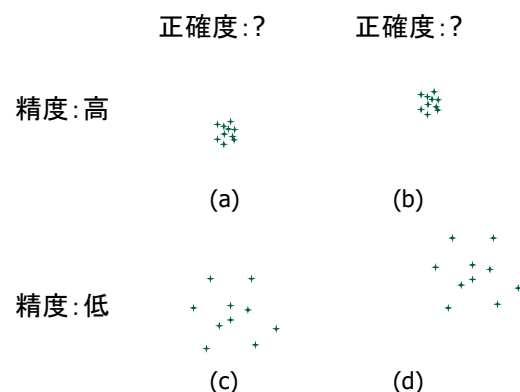


図 3.2 偶然(ランダム)誤差と系統的誤差 (真の値が分からない場合)

3.3 不確かさ評価の種類

不確かさ評価の方法は、誤差の種類によって異なる。以下に不確かさを評価する二つの方法を示す。

- ①タイプ A 評価法：計算を複数回実施し、統計的処理によって不確かさを推定する。
- ②タイプ B 評価法：タイプ A 以外の手法を用いて不確かさを推定する。

以下にいくつか例を示す。

- a)様々な炉心に対して決定論的手法による解析を実施し、各炉心に対する計算結果の誤差を統計処理して不確かさを評価する。(タイプ A)
- b)モンテカルロコードの計算を(初期乱数を変化させて)複数回実行し、その結果を統計処理して不確かさを評価する。(タイプ A)
- c)様々な炉心体系で拡散計算と輸送計算の結果を比較し、統計処理によって拡散計算の不確かさを推定する。(タイプ A)
- d)断面積の共分散データ(不確かさ)を用いて断面積をランダムにサンプリングし、複数の断面積を用いて炉心解析を実施し、その結果を統計処理して炉心特性の不確かさを評価する。(タイプ A)
- e)断面積の感度係数を用いて断面積データに含まれる共分散データを伝播させ、炉心特性の不確かさを評価する。(タイプ B)

タイプ A 評価法は、統計的手法を用いることから、偶然誤差に基づく不確かさを評価するための手法と見られることがある。しかし、a)や c)に示すように、タイプ A は必ずしも偶然誤差に対する不確かさ評価を行う手法ではなく、系統誤差に対する不確かさ評価に用いられることもある。

■例 3：巻き尺を使用して測定した 1 本のひもの長さの不確かさを評価する¹。

ステップ 1：どのような不確かさを考慮しなければならないかを検討する。

①巻き尺自体の不確かさ

- ・補正の必要性、校正がなされているか、校正で示された不確かさの定義
- ・巻き尺が伸びたり縮んだりしないか(校正してからの変化はどれぐらいか)
- ・温度や湿度などの環境が巻き尺の長さに影響を与えないか
- ・分解能はいくらか(ここでは 1mm と仮定)

②測定対象の不確かさ

- ・ひもはまっすぐ置かれているか

¹ 「不確かさの入門ガイド」、ASG104, 独立行政法人製品評価技術基盤機構、平成 26 年 5 月, p.21.

- ・ 温度や湿度などの環境がひもの長さに影響を与えないか
- ・ ひもの端は明確に分かるか

③測定方法の不確かさ

- ・ ひもの端と巻き尺の端をきちんと合わせられるか
- ・ 巻き尺とひものを並行に並べられるか
- ・ 同じ測定を同じ条件で繰り返すことができるか

ステップ 2：測定の実施。測定を 10 回行って、平均値として 5.017m、長さの標準偏差として、0.0021m が得られた。一般的には、これをひもの長さの不確かさとする場合も多いかも知れないが、これだけでは不確かさの一部でしかないことに注意。

ステップ 3：ステップ 1 で考慮した各条件の不確かさを評価する。

- ・ 巻き尺は校正済みで補正は不要。校正の不確かさ(測定誤差の標準偏差)はカタログ値から長さの 0.05% である。従って、校正の不確かさに基づく測定の不確かさは $5.017\text{m} \times 0.0005 = 0.0025\text{m}$ である。
- ・ 巻き尺の最小目盛りは 1mm であることから、目盛り読み取りの誤差は 0.5mm を超えることはないと推測できる。すなわち、目盛り読み取りの誤差は $\pm 0.5\text{mm}$ の範囲に入ると推測できる。目盛り読み取りの誤差を一様分布と見なすと、これに基づく不確かさ(標準偏差)は、一様分布の上限と下限の半値幅である 0.0005m を $\sqrt{3}$ で割ったもの、つまり $0.0005/\sqrt{3} \cong 0.0003\text{m}$ である²。
- ・ 巻き尺はまっすぐ置かれていると推定して、不確かさは 0m とする。
- ・ 巻き尺の温度や湿度に対する変化は無視できるとして、不確かさは 0m とする。
- ・ ひもをまっすぐに置くことは難しいため、実際の長さより短めに測定値が出やすい。この不確かさを長さの 0.1%、つまり 0.005m と想定する。この誤差がどのように分布しているかは分からないため、これを一様分布と仮定すると³、標準偏差は $0.005/\sqrt{3} \cong 0.0029$ となる。
- ・ 10 回の測定を行っているため、測定値の平均の不確かさは、 $0.0021/\sqrt{10} \cong 0.0007\text{m}$

ステップ 4：ステップ 3 で求めた不確かさが独立しているかどうかを確認する。

ここでは、これらの不確かさが独立しているとする。なお、不確かさに相関がある場合については、7 章で取り扱う。

ステップ 5：既知の補正を全て含んだ測定値を評価する。

² 一様分布の標準偏差の計算については、例 15 を参照のこと。

³ 短い側に測定値が出る可能性が高いと思われるが、不確かさとして、 $\pm 0.005\text{m}$ を見込み、上限値と下限値の半値幅としては 0.005m を仮定する。

今回は、ひもの曲がり起因する補正(系統誤差)0.005m を考慮して、 $5.017+0.005m=5.022m$ となる。

ステップ 6：全ての不確かさを合成する。

今回は、それぞれの不確かさに相関がないとしているので、合成された不確かさ⁴は

$$\sqrt{0.0025^2 + 0.0003^2 + 0.0029^2 + 0.0007^2} = 0.0039$$

ステップ 7：95%の信頼水準のもとで、ひもの長さの不確かさは $0.0039 \times 2 = 0.0078m$ となる⁵。

ステップ 8：以上の結果をまとめると、ひもの長さについて、次のような記載となる。

ひもの長さは $5.022 \pm 0.0078 m$ である。長さの不確かさは標準不確かさを 2 倍したものであり、信頼水準は 95%となる。この長さは、水平に置かれたひもを 10 回測定した値の平均値である。この結果は、測定時にひもが完全にはまっすぐに置かれていなかった影響を補正している。不確かさは「測定の不確かさに関する入門ガイド」による方法により測定した。

表 3.1 不確かさ計算のまとめ

不確かさの要因	値(±)	確率分布	除数	標準不確かさ
校正の不確かさ	0.005	正規	2	0.0025
分解能(巻き尺の最小目盛)	0.0005	一様	$\sqrt{3}$	0.0003
ひもの置き方	0.005	一様	$\sqrt{3}$	0.0029
10回繰り返し測定の平均値の標準不確かさ	0.0007	正規	1	0.0007
合成標準不確かさ		正規と仮定		0.0039
95%信頼水準での不確かさ		正規と仮定		0.0078

なお、上記の結果は、表 3.1 のような形でまとめることができる。□

4. 確率密度関数

不確かさを評価する際、対象とするパラメータがどのような確率で分布しているかはきわめて重要な情報である。以下では、対象とするパラメータが飛び飛びの値を取る(離散的な)場合と連続量である場合に分けて、確率の分布を表す関数について説明する。

⁴ 不確かさの伝搬については、第 8 章参照。

⁵ 95%については、例 8 を参照のこと。また、信頼水準については、6.3 節を参照。

4.1 確率関数

対象とするパラメータが離散的な場合、それぞれのパラメータを取る確率を表す関数を確率関数⁶と呼ぶ。確率関数 $p(x_i)$ は以下の性質を持つ。

$$\textcircled{1} \text{あるパラメータを取る確率は } 0 \sim 1 \text{ になるため、 } 1 \geq p(x_i) \geq 0 \quad (4.1)$$

$$\textcircled{2} \text{必ずどれかのパラメータを取るとることから、 } \sum_i p(x_i) = 1 \quad (4.2)$$

■例4：一つのさいころを振った。このさいころがいわゆる「グラサイ」⁷でなければ、1～6の目が出る確率は等しく1/6である。従って、サイコロの目に対する確率関数は、

$$p(x_i) = \frac{1}{6}, x_i = \{1, 2, 3, 4, 5, 6\}$$

として与えられる。この確率関数は①、②の性質を共に満たしている。□

■例5：さいころを振って偶数が出れば+1, 奇数が出れば-1だけx軸上を移動することを考える。原点から出発し、1回試行したときに位置 x_i をとる確率関数は

$$p(x_i) = \frac{1}{2}, x_i = \{-1, 1\}$$

$$p(x_i) = 0, x_i = \{0\}$$

2回試行したときには、

$$p(x_i) = \frac{1}{4}, x_i = \{-2, 2\}$$

$$p(x_i) = 0, x_i = \{-1, 1\}$$

$$p(x_i) = \frac{1}{2}, x_i = \{0\}$$

N(偶数)回試行したときには、

$$p(x_i = 2i - N) = {}_N C_{N-i} / 2^N, i = \{0, 1, 2, \dots, N\}, x_i = \{-N, -N + 2, \dots, N - 2, N\}$$

いずれの試行回数の場合についても、確率関数は(4.1)、(4.2)の性質を共に満たしている。□

4.2 確率密度関数

対象とするパラメータが連続量の場合、あるパラメータを取る確率を表す関数を確率密度関数とよぶ。確率密度関数 $p(x)$ は、以下の性質を持つ。

⁶ 離散的な変数に対する「確率密度関数」は確率関数、または確率質量関数と呼ばれる。

⁷ 特定の目が出やすくなっているいかさま用のさいころ。

①単位入力値あたりの確率として与えられる。これは、重さの密度が単位体積あたりの重量として定義されているのと同じである。

②正の値を取る。つまり、 $p(x) \geq 0$ (4.3)

③入力値の全範囲で積分すると、1になる。つまり、 $\int_{-\infty}^{\infty} p(x)dx = 1$ (4.4)

④ $a < x < b$ を取る確率は、 $P(a < x < b) = \int_a^b p(x)dx$ (4.5)

■例 6：10 分ごとに電車が来る駅に行ったときの待ち時間の確率密度関数は、待ち時間を分で表したときに

$$p(x) = \frac{1}{10}, \quad 0 \leq x < 10$$

$$p(x) = 0, \quad 10 \leq x$$

となる。この確率密度関数は、上記の(4.3), (4.4), (4.5)の性質を満たしている。□

さて、連続量の変数に対する確率密度関数を理解する上でやっかいなのが、確率密度関数の値が確率そのものを表しているのではないことである。これは④の性質に関連している。この点について理解するため、次の問いを考えてみよう。「密度 10g/cm^3 の物質がある。この物質の重さはいくらか？」

この問いに答えるためには、物質の体積が必ず必要である。確率密度関数→重さの密度、確率→重さのように関連づけて考えると、体積に対応する量を与えないと確率が計算できないことは明らかである。上記の問題で確率密度関数の値が $1/10$ であることは、例えば待ち時間が 3 分 30 秒である確率が $1/10$ であることを示しているのではない。 $1/10$ に待ち時間の範囲を乗じることにより、初めて確率を計算することができる。例えば、待ち時間が 3 分 30 秒から 4 分間の $1/2$ 分間である確率は、

$$\frac{1}{10} \times \frac{1}{2} = \frac{1}{20} = 5\%$$

であるし、待ち時間が 5 分から 10 分間の 5 分間である確率は

$$\frac{1}{10} \times 5 = \frac{1}{2} = 50\%$$

である。

■例 7：巨視的断面積 Σt の媒質に $x=0$ から中性子が入射した。 $x \sim x + dx$ でこの中性子が媒質と初めて衝突する確率は、以下のように求められる。

まず、位置 x まで衝突しない確率は、 $\exp(-\Sigma_t x)$ である。さらに、位置 x まで衝突せずに飛んできた中性子が $x \sim x + dx$ で媒質と衝突する確率は、 dx がきわめて小さいとして $\Sigma_t dx$

で与えられる。これから、 $x \sim x + dx$ で中性子が初めて媒質と衝突する確率は、 $\exp(-\Sigma_t x) \Sigma_t dx$ である。従って、中性子の衝突に関する(単位飛行距離あたりの)確率密度関数は、

$$p(x) = \Sigma_t \exp(-\Sigma_t x) \quad (4.6)$$

である。また、 x を $0 \sim \infty$ まで積分すると、

$$\int_0^{\infty} \Sigma_t \exp(-\Sigma_t x) dx = 1 \quad (4.7)$$

である。この確率密度関数は、①、②、③の性質を満たしている。なお、例 6 と同じように、位置 x で中性子が初めて媒質と衝突する確率が $p(x) = \Sigma_t \exp(-\Sigma_t x)$ なのではないことに注意。 $x \sim x + dx$ で初めて媒質と衝突する確率が $\exp(-\Sigma_t x) \Sigma_t dx$ で与えられるのである。また、④から $0 \sim X$ で衝突する確率は、

$$\int_0^X \Sigma_t \exp(-\Sigma_t x) dx = 1 - \exp(-\Sigma_t X) \quad (4.8)$$

である。中性子は距離 X までに衝突するか、しないかどちらかであることから、この結果は 1 から衝突しない確率である $\exp(-\Sigma_t X)$ を差し引いた結果になっている。□

■例 8 : 平均値 μ 、標準偏差 σ の正規分布に対する確率密度関数は、次式で与えられる。

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \quad (4.9)$$

全ての変数区間で積分すると

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx = 1 \quad (4.10)$$

例えば $\mu = 0, \sigma = 1$ の標準正規分布の場合、

$$P(-1 < x < 1) = \int_{-1}^1 \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} dx = 0.6826 \quad (4.11)$$

$$P(-2 < x < 2) = \int_{-2}^2 \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} dx = 0.9544 \quad (4.12)$$

$$P(-3 < x < 3) = \int_{-3}^3 \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} dx = 0.9974 \quad (4.13)$$

(4.11), (4.12), (4.13)は正規分布の場合、 $\mu \pm \sigma$ に全体の約 68%、 $\mu \pm 2\sigma$ に約 95%、 $\mu \pm 3\sigma$ に約 99.7%が含まれていることを示している。□

4.3 累積分布関数

ある確率分布関数 $p(x)$ が与えられたとき、 $x < \alpha$ である確率は、

$$F(\alpha) = P(-\infty < x < \alpha) = \int_{-\infty}^{\alpha} p(x) dx \quad (4.14)$$

で与えられる。これを累積分布関数という。例7で中性子が位置0~Xで衝突する確率を計算したが、これは累積分布関数になっている。

■例8: 偏差値は平均値が50, 標準偏差が10の正規分布を仮定した場合の成績分布である。偏差値の累積分布関数は、

$$F(\alpha) = \int_{-\infty}^{\alpha} \frac{1}{\sqrt{2\pi} \times 10} \exp\left\{-\frac{(x-50)^2}{2 \times 10^2}\right\} dx \quad (4.15)$$

で与えられる。従って、偏差値50以下である確率は

$$F(50) = \int_{-\infty}^{50} \frac{1}{\sqrt{2\pi} \times 10} \exp\left\{-\frac{(x-50)^2}{2 \times 10^2}\right\} dx = 0.5 \quad (4.16)$$

偏差値60以下である確率は

$$F(60) = \int_{-\infty}^{60} \frac{1}{\sqrt{2\pi} \times 10} \exp\left\{-\frac{(x-50)^2}{2 \times 10^2}\right\} dx = 0.8413 \quad (4.17)$$

偏差値90以下である確率は

$$F(90) = \int_{-\infty}^{90} \frac{1}{\sqrt{2\pi} \times 10} \exp\left\{-\frac{(x-50)^2}{2 \times 10^2}\right\} dx = 0.9999683 \quad (4.18)$$

である。(4.18)式から、偏差値90以上の人は(あくまでも成績が完全な正規分布になっている場合であるが)10万人中約3人である。□

5.期待値、分散、標準偏差

5.1 期待値

1~10の数字が書かれたカードが10枚、裏返しておかれている。このカードを一枚引き、1が出れば1000円もらい、2-10が出れば100円払う。さて、あなたはこの賭け事に乗るか？

この問題を考えるとき、重要なのは、「どれぐらい得すると期待できるか」である。これは期待値を用いることで評価できる。

ある離散的なパラメータ X_i に対して確率関数 $p(X_i)$ が与えられているとする。この場合、 X_i の期待値は

$$E[X] = \sum_i X_i p(X_i) \quad (5.1)$$

で与えられる。これは、平均値である。

■例9: 先ほどの賭の期待値は

$$1000 \times \frac{1}{10} - 100 \times \frac{9}{10} = 10 \text{円}$$

であり、平均的に得する賭になっている⁸。□

■例 10：さいころの目の期待値は

$$1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} = \frac{7}{2}$$

である。□

ある連続的なパラメータ X に対して確率関数 $p(X)$ が与えられているとする。この場合、 X の期待値は

$$E[X] = \int_{-\infty}^{\infty} Xp(X)dX \quad (5.2)$$

で与えられる。

■例 11：巨視的断面積 Σt の物質に中性子が入射したときに衝突せずに飛行する距離の期待値は確率密度関数が(4.6)式で与えられるため、

$$\int_0^{\infty} x \Sigma_t \exp(-\Sigma_t x) dx = \frac{1}{\Sigma_t} \quad (5.3)$$

である。これは中性子が衝突なしに飛行する距離の平均を示したものであり、平均自由行程と呼ばれる。□

5.2 分散

■例 12：1～10 の数字が一つずつ書かれた 10 枚のカードが裏返しておかれている。このカードを一枚引くが、二種類の賭ができる。

① 1～5 が出れば 100 円もらい、6～10 が出れば 50 円払う。

② 1～5 が出れば 1,000,001,000 円もらい、6～10 が出れば 1,000,000,500 円払う。

さて、あなたはどちらの賭に参加するか？①、②の期待値は

$$\text{①} : 100 \times \frac{5}{10} - 50 \times \frac{5}{10} = 25 \text{ 円}$$

$$\text{②} : 1,000,001,000 \times \frac{5}{10} - 1,000,000,500 \times \frac{5}{10} = 250 \text{ 円}$$

であり、②の方が明らかに有利である。では、②の賭をやるだろうか？一般的には、②に

⁸ 少年マガジンに連載中の「賭博霸王伝 零」においては、これでは失格である。なぜならば、「1～10 の数字が書かれたカードが 10 枚、裏返しておかれている」と書かれているだけであり、1～10 の数字が書かれているカードがそれぞれ 1 枚ずつあるとは限らないからである。普通の人にとってはこの手の話は、予測がつかないために確かに「不確かさ」である。一方、「賭博霸王伝 零」の主人公の宇海零にとっては、この手の話は「お見通し」であり、従って「不確かさ」ではなく、考慮すべき「系統誤差」なのかも知れない。

は手を出さないのではないかと思うが・・・□

例 12 は、ばらつきのあるデータを扱う場合、期待値のみでは情報として不十分であることを示している。すなわち、「どの程度データがばらついているか」を示す指標が必要である。分散は、データのばらつきを示す指標である。

ある離散的なパラメータ $X = (X_1, X_2, \dots)$ に対して確率関数 $p(X_i)$ が与えられているとする。この場合、 X_i の分散は

$$\text{var}[X] = E[(X - \mu)^2] = \sum_i (X_i - \mu)^2 p(X_i) \quad (5.4)$$

ここで、 $\mu = E[X]$ である。

■例 13：例 12 の①、②の賭の分散は

$$\textcircled{1} (100 - 25)^2 \times \frac{5}{10} + (-50 - 25)^2 \times \frac{5}{10} = 5625$$

$$\textcircled{2} (1,000,001,000 - 250)^2 \times \frac{5}{10} + (-1,000,000,500 - 250)^2 \times \frac{5}{10} \approx 10^{18}$$

であり、データのばらつきが全く異なる。つまり、②はばらつきが大きく「危険な賭」である。□

■例 14：あるデータ列 $\{1, 2, 3, 4, 5, 6\}$ の分散は、このデータ列の平均が 3.5 であることから、

$$\frac{(1 - 3.5)^2 + (2 - 3.5)^2 + (3 - 3.5)^2 + (4 - 3.5)^2 + (5 - 3.5)^2 + (6 - 3.5)^2}{6} \cong 2.917$$

である。□

ある連続的なパラメータ X に対して確率関数 $p(X)$ が与えられているとする。この場合、 X の分散は

$$\text{var}[X] = E[(X - \mu)^2] = \int_{-\infty}^{\infty} (X - \mu)^2 p(X) dX \quad (5.5)$$

で与えられる。

■例 15：一様な確率密度関数がある。

$$p(x) = \frac{1}{2}, \quad -1 < x < 1 \quad (5.6)$$

$$p(x) = 0, \quad \text{otherwise}$$

この分散は、

$$\text{var}[x] = \int_{-\infty}^{\infty} x^2 p(x) dx = \int_{-1}^1 \frac{1}{2} x^2 dx = \frac{1}{3} \quad (5.7)$$

である。□

5.3 標準偏差

標準偏差は、データのばらつきを表す指標の一つであり、分散の平方根として与えられる。

$$\sigma = \sqrt{\text{var}[X]} \quad (5.8)$$

例 8 で示したように、データが正規分布に従ってばらついている場合、全データの約 68% が $\mu \pm \sigma$ に、約 95% が $\mu \pm 2\sigma$ に、約 99.7% が $\mu \pm 3\sigma$ に含まれる。

5.4 モーメント

平均値と分散は、(5.2)式および(5.5)式からそれぞれ $E[X]$ 、 $E[(X - \mu)^2]$ で計算できる。これらの形は、いずれも

$$E[(X - c)^k] = \int_{-\infty}^{\infty} (X - c)^k p(X) dX \quad (5.9)$$

と一般化できる。 $E[(X - c)^k]$ を c まわりの k 次モーメントと呼ぶ。1 次から 4 次までのモーメントは、以下のように呼ばれる。

平均：原点周りの 1 次モーメント

分散：平均周りの 2 次モーメント

歪度：平均周りの 3 次モーメントであり、分布の偏りを示す指標である。負であれば左にゆがんだ分布、ゼロであれば左右対称、正であれば右にゆがんだ分布となる。

尖度：平均周りの 4 次モーメントであり、値が大きいほど正規分布に比べて裾が広がっている(裾が重い)。

6. 母集団の推定

6.1 母集団とは

我々が白鳥を「白い」と認識している。これは、すべての白鳥が白いと確認した結果ではなく、これまでに見た白鳥が白かったことから、「残りの白鳥も白いだらう」と推定した結果である。このように一部のサンプルから、残りの集団の性質を推定することを母集団の推定と呼ぶ⁹。

⁹ 「ブラック・スワン—不確実性とリスクの本質」、ナシーム・ニコラス・タレブ著は、母集団推定の失敗について書かれた好著である。

母集団とは、調べたい要素すべてからなる集団であるが、一般的に母集団の要素の数は非常に多く、全てを調べることはできない。従って、限られた調査結果から母集団を推定する必要が生じる。母集団推定の例としては、世論調査や工業製品の抜き取り調査があげられる。

6.2 平均値の不確かさ(標準偏差)

モンテカルロコードを用いてある炉心体系の計算を初期乱数を変えて n 回行った。この結果を用いて、実効増倍率の平均値と標準偏差を以下の式から計算した。

$$\mu_n = \sum_{i=1}^n \frac{X_i}{n} \quad (6.1)$$

$$\sigma_n = \sqrt{\text{var}[X_i]} = \sqrt{\sum_{i=1}^n \frac{(X_i - \mu)^2}{n}} \quad (6.2)$$

この平均値と標準偏差は、 n を非常に増やした場合に比べ、どの程度正しいだろうか？

まず、平均値について考えてみよう。実効増倍率が正規分布に従っていると仮定すると、各計算の無限増倍率は

$$X_i = \mu \pm \sigma \quad (6.3)$$

と書くことができる¹⁰。 n 回の計算の実効増倍率の平均値は

$$\mu_n = \sum_{i=1}^n \frac{X_i}{n} \quad (6.4)$$

である。各計算の実効増倍率のばらつきが独立であるとする、平均値のばらつき(平均値の標準偏差)を以下のように合成することができる¹¹。

$$\sigma_n = \sqrt{\sum_{i=1}^n \left(\frac{\sigma}{n}\right)^2} = \sqrt{n \frac{\sigma^2}{n^2}} = \frac{\sigma}{\sqrt{n}} \quad (6.5)$$

以上のことから、 n 回の計算を抜き出して平均値を求めた場合、この平均値の不確かさ(平均値の標準偏差)は、 $\frac{\sigma}{\sqrt{n}}$ となる。

■例 16：モンテカルロコードによる計算で得られる実効増倍率の真の標準偏差が 0.02 であったとする¹²。この場合、 n 回の計算で得られた実効増倍率の平均値に関する標準偏差は以

¹⁰ \pm を普通と異なった意味で使っていることに注意。 $x_i = \mu + \sigma, \mu - \sigma$ ではない。 x_i が μ を中心として標準偏差 σ でばらついているという意味である。

¹¹ 各項が $\frac{X_i}{n}$ であるため、各項のばらつきを $\frac{X_i}{n} = \frac{\mu \pm \sigma}{n}$ より $\frac{\sigma}{n}$ と見積もることができる。

¹² n 回の計算により得られた標準偏差でないことに注意。一般的には、母集団の標準偏差は

下の通りである。

$$n=1 \text{ の場合 : } \frac{0.02}{\sqrt{1}} = 0.02$$

$$n=4 \text{ の場合 : } \frac{0.02}{\sqrt{4}} = 0.01$$

$$n=9 \text{ の場合 : } \frac{0.02}{\sqrt{9}} \cong 0.006667$$

$$n=100 \text{ の場合 : } \frac{0.02}{\sqrt{100}} = 0.002$$

計算回数が増えると、平均値の予測精度が上がっていくが、これは直感的にもそうだとはいえるだろう。□

6.3 標準偏差の推定

では次に、6.2 節で扱った問題において、 n 回の計算結果から評価した標準偏差がどの程度正しいか検討してみよう。標準偏差の自乗である分散は、「それぞれの値と平均値との距離の二乗を平均したもの」を表している。従って、 $\frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2$ で計算すべきものである。

ここで、 μ は、モンテカルロ計算で得られた実効増倍率の母集団の平均値である。つまり、モンテカルロ計算を非常に数多く実行したときに得られる平均値となる。しかし、今回は n 回の計算のみを行っているので、 μ の値は分からない。そこで、次善の策として実施した n 回の計算で得られた実効増倍率の平均値 μ_n を使わざるを得ない。つまり、分散を

$\frac{1}{n} \sum_{i=1}^n (X_i - \mu_n)^2$ で評価せざるを得ない。

$f(c) = \frac{1}{n} \sum_{i=1}^n (X_i - c)^2$ という関数は、 $c = \mu_n = \frac{1}{n} \sum_{i=1}^n X_i$ の時に極小値を取る。しかし、 $\mu \neq \mu_n$ であることから(つまり、 n 回の実効増倍率の平均値と母集団の平均値が一致しないことから)、 $\frac{1}{n} \sum_{i=1}^n (X_i - \mu_n)^2$ は、 $\frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2$ に比べて必ず小さくなる。つまり、 n 回の計算で得られた実効増倍率から評価した分散は、必ず小さめに計算されるということである。

では、どの程度小さめに予測されるのだろうか。逆に、 n 回の計算の実効増倍率の標準偏差から、実効増倍率の母集団の標準偏差は推定できるだろうか。

n 回の実効増倍率の分散は、以下のように計算できる。これを標本分散という。

$$\text{標本分散 : } (\sigma_n)^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu_n)^2 \quad (6.6)$$

わからないが、ここでは簡単のため、母集団の標準偏差がわかっているものとしている。

一方、母集団の分散は、以下の式により推定可能である。これを不偏分散という。

$$\text{不偏分散} : (\sigma)^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \mu_n)^2 \quad (6.7)$$

■例 17: 例 16 で標本分散と不偏分散を比較する。実効増倍率を平均 1.01、標準偏差を 0.02、サンプル数(計算回数)を n としてランダムに 100 セット、 n 個の実効増倍率を「発生」させ、それぞれの標本分散と不偏分散を平均して求めた。結果を表 6.1 に示すが、標本分散は過小評価になっている一方、不偏分散にはこのような傾向は見られない。□

表 6.1 標本分散と不偏分散の比較

サンプル数	標本分散より計算した標準偏差	不偏分散より計算した標準偏差
5	0.0181	0.0202
10	0.0190	0.0201
40	0.0199	0.0202

以下では、不偏分散において、なぜ n でなく $n-1$ で割るのかを説明する。

(6.6)式の $\sum_{i=1}^n (X_i - \mu_n)^2$ に着目する。これは、以下のように変形できる。

$$\begin{aligned} (\sigma_n)^2 &= E\left[\sum_{i=1}^n (X_i - \mu_n)^2\right] \\ &= E\left[\sum_{i=1}^n \{(X_i - \mu) + (\mu - \mu_n)\}^2\right] \\ &= E\left[\sum_{i=1}^n (X_i - \mu)^2 + 2\sum_{i=1}^n (X_i - \mu)(\mu - \mu_n) + \sum_{i=1}^n (\mu_n - \mu)^2\right] \\ &= E\left[\sum_{i=1}^n (X_i - \mu)^2 - 2(\mu_n - \mu)\sum_{i=1}^n (X_i - \mu) + n(\mu_n - \mu)^2\right] \\ &= E\left[\sum_{i=1}^n (X_i - \mu)^2 - 2(\mu_n - \mu)\left(\sum_{i=1}^n X_i - n\mu\right) + n(\mu_n - \mu)^2\right] \\ &= E\left[\sum_{i=1}^n (X_i - \mu)^2 - 2(\mu_n - \mu)(n\mu_n - n\mu) + n(\mu_n - \mu)^2\right] \\ &= E\left[\sum_{i=1}^n (X_i - \mu)^2 - 2n(\mu_n - \mu)^2 + n(\mu_n - \mu)^2\right] \\ &= E\left[\sum_{i=1}^n (X_i - \mu)^2 - n(\mu_n - \mu)^2\right] \end{aligned} \quad (6.8)$$

(6.8)式の右辺第一項は、以下のように計算できる。

$$E\left[\sum_{i=1}^n (X_i - \mu)^2\right] = E[(X_1 - \mu)^2] + E[(X_2 - \mu)^2] + \dots + E[(X_n - \mu)^2] \quad (6.9)$$

$$= \sigma^2 + \sigma^2 + \dots + \sigma^2 = n\sigma^2$$

(6.9)式の右辺第二項は、以下のように計算できる。

$$E[n(\mu_n - \mu)^2] = n \times E[(\mu_n - \mu)^2] = n \times \frac{\sigma^2}{n} = \sigma^2 \quad (6.10)$$

従って、(6.8)式と(6.9)式を(6.7)式に代入し、

$$E\left[\sum_{i=1}^n (X_i - \mu_n)^2\right] = E\left[\sum_{i=1}^n (X_i - \mu)^2 - n(\mu_n - \mu)^2\right] = n\sigma^2 - \sigma^2 = (n-1)\sigma^2 \quad (6.11)$$

(6.11)式により

$$E\left[\frac{1}{n-1} \sum_{i=1}^n (X_i - \mu_n)^2\right] = \sigma^2 \quad (6.12)$$

であり、 $\frac{1}{n-1} \sum_{i=1}^n (X_i - \mu_n)^2$ の期待値が確かに母集団の分散 σ^2 になっていることが分かる。

すなわち、 n 個のサンプルを抜き出して分散を計算する場合、(6.7)式を用いることで、より確からしく母集団の分散を推定できることが分かる。(6.7)式は、サンプル数 n が限定されることによる分散の過小評価を補正しているという観点から「不偏」分散と呼ばれる。なお、注意する必要があるのは、(6.7)式で得られる分散はあくまでも推定値ということである。つまり、不偏分散にも不確かさがあるということである¹³。

6.4 信頼係数と信頼区間

解析コードを用いて核特性の解析を行った場合、真の値がどの範囲にあるのか、推定することは重要となる。例えば、臨界安全の分野において、使用している解析コードがどのような実効増倍率を算出するかを確認することは不可欠である。もちろん解析値には不確かさが含まれるため、ある確率である範囲に入る、という推定の仕方が望ましい。例えば、95%の確率で実効増倍率が 0.998~1.004 に入る、と言った推定である。この意味は、同じような推定を例えば 10000 回行った場合、約 9500 回は推定された範囲内に真の実効増倍率が存在するが、残り約 500 回については、この区間から外れている、ということである。このような推定を行う場合に用いる用語の定義を以下に示す。

信頼係数：どの程度の確率でその区間内に推定値が存在するかを示す係数。炉物理計算の分野では 95%がよく使用される。

信頼区間：信頼係数で与えられた確率で真の値が存在すると推定された範囲。

信頼限界：信頼区間の下限値と上限値。

¹³ 不偏分散の分散は $2\sigma^4 / (n-1)$ で与えられる。

(http://www.geocities.jp/ikuro_kotaro/koramu/335_mix.htm)

■例 18: モンテカルロコードを用いてある炉心体系の計算を初期乱数を変えて5回行った。得られた実効増倍率は 0.9986, 1.0282, 1.0211, 0.9804, 1.0022 であった。この炉心に対して、ヒストリー数を十分に増やした(=統計誤差を無視できる)計算を行った場合、得られる実効増倍率はどの程度になると推定できるだろうか。

実効増倍率の統計誤差は正規分布に従うため、実効増倍率の平均値も正規分布に従って変化する。実効増倍率の平均値は(3.1)式より、

$$\frac{0.9986+1.0282+1.0211+0.9804+1.0022}{5}=1.0061$$

また標準偏差(不確かさ)は(3.2)式もしくは(6.7)式より

$$\sqrt{\frac{(0.9986-1.0061)^2+(1.0282-1.0061)^2+(1.0211-1.0061)^2+(0.9804-1.0061)^2+(1.0022-1.0061)^2}{5-1}} \\ =0.0190$$

従って、平均値の標準偏差は(3.3)式より

$$\frac{0.0190}{\sqrt{5}}=0.0085$$

では、統計誤差を含まない実効増倍率の計算値はが $1.0061 \pm 1.96 \times 0.0085$ の範囲に入る確率は 95%になる¹⁴、と考えて良いだろうか。

ここで重要なのは、実効増倍率の平均値の標準偏差は 5 回の計算結果から推定された値であるということである。初期乱数を変化させて別に 5 回のモンテカルロ計算を行い、その結果から標準偏差を推定すると、0.0085 にはならないだろう、ということが想像できる。つまり、平均値の標準偏差自体に不確かさが含まれていることになる。この不確かさを考慮するためには、平均値の標準偏差に不確かさが無い場合に比べてより広い区間を考えなければならないことは明らかであろう。

実は、0.9986, 1.0282, 1.0211, 0.9804, 1.0022 という「実効増倍率」は、平均 1.01, 標準偏差 0.02 の正規分布から 5 回ランダムにサンプリングを行った結果である。従って、正確な(母集団の)標準偏差は 0.02 であり、これを「知っていた」とすると、5 回の平均値に対する標準偏差は、 $0.02/\sqrt{5}=0.0089$ となる。従って、統計誤差を含まない実効増倍率の計算値が $1.0061 \pm 1.96 \times 0.0089$ の範囲に入る確率が 95%となる。

実際に、「実効増倍率」を 5 回発生させ、平均値と平均値の標準偏差を求め、(平均値 $\pm 1.96 \times 0.0089$)の中に真の平均値(1.01)が入っているかどうかを 10000 回繰り返して確認してみた。この結果、(平均値 $\pm 1.96 \times 0.0089$)の中に真の平均値が入っている回数は 9494 回であり、入っていなかった回数は 506 回であった。つまり、真の平均値が(平均値 $- 1.96 \times$ 平均値の標準偏差 \sim 平均値 $+ 1.96 \times$ 平均値の標準偏差)に入っていると予測した場合、95%の確率で正し

¹⁴ $P(-1.960 < x < 1.960) = \int_{-1.960}^{1.960} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} dx = 0.95$ 。(4.9)~(4.13)も参照。

いといえるという結果になっている。

次に、5回の実効増倍率の「計算結果」から平均値の標準偏差を(不偏分散を通じて)推定し、(平均値 $\pm 1.96 \times$ 平均値の推定された標準偏差)の中に真の平均値が入っているかどうかを10000回繰り返し確認した。なお、10000回の推定において、標準偏差の最大値は0.0234、最小値は0.00029となり、推定された標準偏差自体にかなりのばらつきがあることが確認できた。真の標準偏差は先ほど述べたように0.0089である。また、(平均値 $\pm 1.96 \times$ 平均値の推定された標準偏差)の中に真の平均値が入っているかどうかを確認した場合、(平均値 $\pm 1.96 \times$ 平均値の推定された標準偏差)の中に真の平均値が入っていた回数は8824回であり、入っていなかった回数が1176回であった。これは、真の平均値が存在する区間の見積もりを明らかに過小評価していることを示している。

以上の結果から、平均値の標準偏差を不偏分散により推定する場合には、この平均値の標準偏差の推定値自体に不確かさがあるため、推定区間をより大きく取らなければならないことが分かる。なお、参考までに5回の実効増倍率を10000セット発生させ、それぞれのセットについて、実効増倍率の平均値および平均値の標準偏差の分布を計算したものを図6.1に示す。図6.1より、5回の実効増倍率の平均値(真値=1.01)は0.97~1.04という広い範囲に分布し、実効増倍率の平均値の標準偏差(真値=0.0089)は0.001~0.02以上と大きくばらついている。

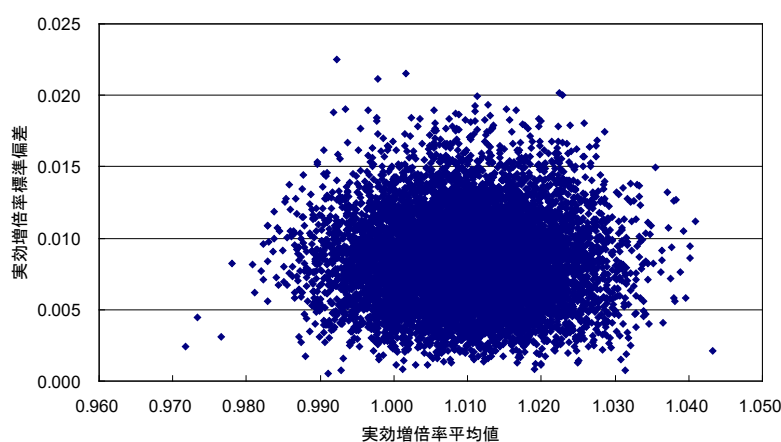


図 6.1 5回の計算で得られた実効増倍率の平均値(横軸)と平均値の標準偏差(縦軸)の分布

ここで扱っている問題は、Studentのt分布と呼ばれる確率分布により考慮することができる。t分布を用いてサンプル数 $n=5$ に対する95%の信頼係数に対する信頼区間を求めると、(平均値 $\pm 2.776 \times$ 平均値の推定された標準偏差)で与えられる。これは、先ほどの(平均値 $\pm 1.960 \times$ 平均値の推定された標準偏差)に比べて明らかに広い範囲である。この区間内に真の平均値が入っているかどうかを上記と同様に10000回確認すると、入っていた回数は9522回、入っていなかった回数は478回であった。すなわち、95%の信頼係数に見合った適正な信頼区間を与えているといえる。□

ここで、Student¹⁵の t 分布について簡単に説明しておく。パラメータ X_1, X_2, \dots, X_n が平均値 μ 、標準偏差 σ の正規分布に従うとする。まず、(6.13)式で定義されるパラメータ α を考える。

$$\alpha = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \quad (6.13)$$

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} : \text{平均値}$$

σ / \sqrt{n} : \bar{X} の真の標準偏差 (X_1, X_2, \dots, X_n から不偏分散により推定したものではなく、母集団の標準偏差から計算したもの)

α は、(6.5)式を考慮すると、平均値 0、標準偏差 1 の正規分布になる。これより、信頼係数 95% を満足する信頼区間は $-1.96 < \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} < 1.96$ 、すなわち、

$$\bar{X} - 1.96 \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.96 \frac{\sigma}{\sqrt{n}}$$

となる。つまり、母集団の標準偏差が分かっている場合には、

$\bar{X} - 1.96 \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.96 \frac{\sigma}{\sqrt{n}}$ により、信頼係数 95% における \bar{X} の信頼区間を推定することができる。

一方、母集団の標準偏差は一般には明らかでないため、不偏分散を通じて推定することとなる。この場合、

$$\alpha' = \frac{\bar{X} - \mu}{U / \sqrt{n}} \quad (6.14)$$

ここで、

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} : \text{平均値}$$

$$U = \sqrt{\frac{\sum_{i=1}^n (X_i - \mu)^2}{n-1}} : \text{不偏分散を通じて推定した母集団の標準偏差}$$

$$\frac{U}{\sqrt{n}} : \text{平均値の標準偏差の推定値}$$

¹⁵ Student さん、とは妙な名前と思われるが、これはペンネームとのこと。William Sealy Gosset さんがギネスビール社に勤めていたとき、ビールの品質管理の過程で t 分布についての知見を得たが、ギネスビール社は当時このような企業秘密を論文として公表することを許していなかった。そこで、Gosset さんは、Student というペンネームで t 分布について(会社にばれない形で)発表した。(Wikipedia より抜粋) ギネスを飲むときには、t 分布を思いだそう。

を対象とすることになるが、 α' は、正規分布でなく、 $t(n-1)$ 分布に従うことが知られている

¹⁶。 $n=5$ の場合、信頼係数 95% を満足する信頼区間は $-2.776 < \frac{\bar{X} - \mu}{U/\sqrt{5}} < 2.776$ 、すなわち、

$\bar{X} - 2.776 \frac{U}{\sqrt{5}} < \mu < \bar{X} + 2.776 \frac{U}{\sqrt{5}}$ となる。2.776 は、 t 分布表において、自由度 = $n-1 = 5-1$

=4、信頼係数 95% として与えられる値である。以下に代表的な α' の値を示す。

$n-1=4,$	信頼係数 95% : 2.776	信頼係数 98% : 3.747
$n-1=5,$	信頼係数 95% : 2.571	信頼係数 98% : 3.365
$n-1=10,$	信頼係数 95% : 2.228	信頼係数 98% : 2.764
$n-1=50,$	信頼係数 95% : 2.009	信頼係数 98% : 2.403
$n-1=240,$	信頼係数 95% : 1.970	信頼係数 98% : 2.342
$n-1=\infty,$	信頼係数 95% : 1.960	信頼係数 98% : 2.326

これより、 n が多くなるほど、係数が小さくなり、 $n=\infty$ の時には、 α' は平均 0、標準偏差 1 の正規分布の α と一致することが分かる。

6.5 推定臨界下限値増倍率

例 18 では、核計算コードの実効増倍率の平均値について、信頼区間の推定を行った。この例のように、核計算コードの平均誤差を知ることも重要である。一方、臨界安全の分野では、「核計算コードで計算した実効増倍率がいくつ以下であれば未臨界と判断して良いか」と言う問いに答えなければならない。「未臨界と判断できる実効増倍率の計算値」を推定臨界下限増倍率という。

仮に、母集団の平均値と分散が十分な精度でわかっているならば、推定臨界下限増倍率を求めることは容易である。例えば、母集団の平均が μ 、分散が σ^2 であれば、推定臨界下限増倍率を $\mu - 2\sigma$ と設定すれば、臨界になる確率は、 $(100\% - 95.44\%) / 2 = 2.28\%$ 以下になる。しかし、実際には、平均値と分散は標本(限られた数の計算結果)を通じて推定した値でしかなく、それぞれに不確かさが含まれる。このような状況の下で、推定臨界下限増倍率を求めるのは簡単なことではない¹⁷。推定臨界下限増倍率の評価方法については、別途 Appendix にまとめたので、そちらを参照して欲しい。

6.6 注意事項

限られた測定データから、母集団の確率分布を推定することは、もともと困難な問題である。炉心解析コードの不確かさの評価においては、「誤差が正規分布する」という仮定を

¹⁶ 詳細については、例えば東京大学教養学部統計学教室編、「統計学入門」、東大出版会、p.201 などを参照のこと。

¹⁷ 具体的な計算手順については、例えば 奥野浩、内藤淑孝、「臨界安全解析コードシステム JACS の計算誤差評価」、JAERI-M 87-057 参照。

置く場合が多いが、対象とする解析体系を注意深く選ばないとこのような前提は成り立たない。例えば、熱中性子体系の UO_2 燃料に対する実効増倍率の誤差分布と、高速中性子体系の MOX 燃料に対する実効増倍率の誤差分布では、誤差の平均値も標準偏差も異なることが予想できる¹⁸。従って、これらと同じグループとすると、正規分布を前提とした信頼区間推定において誤った結果を与えることになる。

図 6.2 は、連続エネルギーモンテカルロコード MVP と JENDL4.0 の組み合わせにより種々の臨界実験の実効増倍率を解析した結果である¹⁹。実効増倍率の実験値はいずれも 1.0 であり、MVP の統計誤差は $0.02\% \Delta k/k$ 程度と十分に小さい。

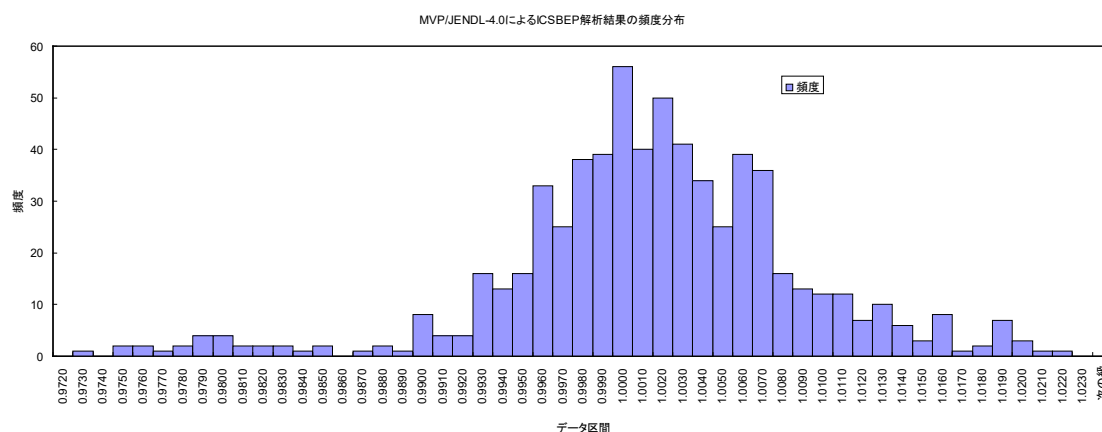


図 6.2 MVP+JENDL4.0 による臨界実験解析結果

図 6.2 を見ると、MVP+JENDL4.0 の解析結果はおおむね正規分布をしているように見えるが、燃料の種類別に分けると、図 6.3 のようになる。図 6.3 を見ると、例えば U233 体系は系統的に実効増倍率を過小評価しており、また、その分布も正規分布からは外れているように見える。従って、図 6.2 のデータによって MVP+JENDL4.0 の無限増倍率の信頼区間を評価した場合、U233 体系に対しては、適切な結果を与えないことが予想される。誤差が正規分布をなしていると仮定できるのは、あくまでも核特性が類似の体系のみであることに注意が必要である。

¹⁸ 例えば、「臨界安全評価関連コードの検証解析に関する報告書」、JNES/SAE09-031, 原子力安全基盤機構、平成 21 年に様々な体系における実効増倍率計算誤差の頻度分布が示されている。

¹⁹ 奥村啓介、長家康展、「JENDL-4.0 に基づく連続エネルギーモンテカルロコード MVP 用の中性子断面積ライブラリーの作成と ICSBEP ハンドブックの臨界性ベンチマーク解析への適用」、JAEA-Data/Code 2011/010 の結果を名古屋大学の遠藤知弘助教が整理したもの。

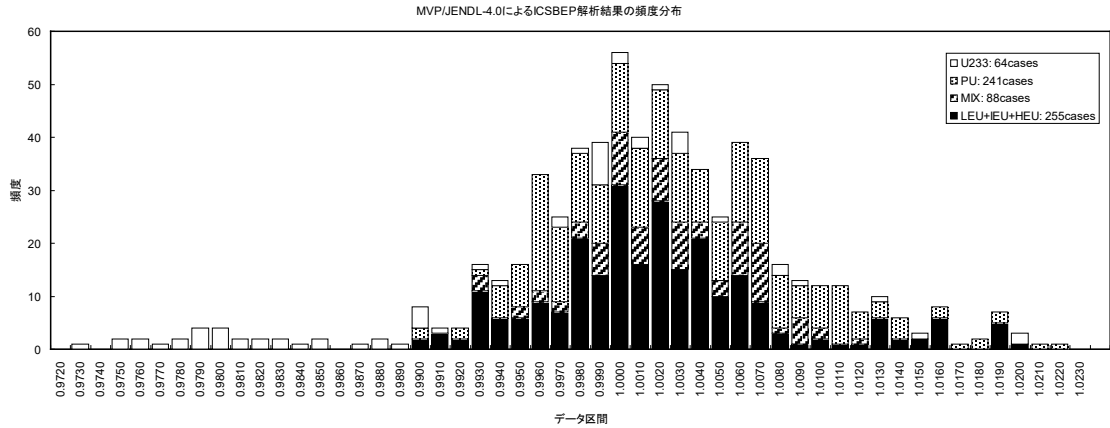


図 6.3 MVP+JENDL4.0 による臨界実験解析結果(燃料の種類別に分類)

7. 共分散、相関係数、分散共分散行列

ここまでは、それぞれの(一列の)データ列について平均や標準偏差の計算を行ってきたが、一般の不確かさ評価では、多変数を対象にして解析を行う。この場合、それぞれの変数の分布に加え、相互にどのような相関があるかということが不確かさの評価上重要になる。不確かさの合成については、次章で説明することとし、ここでは、変数間の相関を表す指標である共分散と相関係数を説明する。

二つのデータ列 $X = (X_1, X_2, \dots), Y = (Y_1, Y_2, \dots)$ がある。これらのデータの共分散は

$$\text{cov}[X, Y] = E[(X - \bar{X})(Y - \bar{Y})] = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) \quad (7.1)$$

として与えられる。また、

$$\alpha = \frac{E[(X - \bar{X})(Y - \bar{Y})]}{\sqrt{E[(X - \bar{X})^2]} \sqrt{E[(Y - \bar{Y})^2]}} = \frac{\text{cov}[X, Y]}{\sqrt{\sigma[X]^2} \sqrt{\sigma[Y]^2}} = \frac{\text{cov}[X, Y]}{\sigma[X] \sigma[Y]} \quad (7.2)$$

を相関係数とよぶ。 α が-1 に近くなるほど強い負の相関を表し、逆に 1 に近くなるほど強い正の相関を表す。また、0 の場合は無相関であることを示す。例えば、 $y_i = x_i$ であるとすると、

$$\alpha = \frac{E[(X - \bar{X})(X - \bar{X})]}{\sqrt{E[(X - \bar{X})^2]} \sqrt{E[(X - \bar{X})^2]}} = 1 \quad (7.3)$$

$Y_i = -X_i$ であるとする、

$$\alpha = \frac{E[(X - \bar{X})(-X + \bar{X})]}{\sqrt{E[(X - \bar{X})^2]} \sqrt{E[(-X + \bar{X})^2]}} = -\frac{E[(X - \bar{X})(X - \bar{X})]}{\sqrt{E[(X - \bar{X})^2]} \sqrt{E[(X - \bar{X})^2]}} = -1 \quad (7.4)$$

なお、定義から $\text{cov}[X, X] = \text{var}[X]$ であることに注意。

次に多くのデータ列 X_1, X_2, \dots , を考える²⁰。これらのデータ列のうち、 X_I と X_J の共分散は、

$$\text{cov}[X_I, X_J] = E[(X_I - \bar{X}_I)(X_J - \bar{X}_J)] = \frac{1}{n} \sum_{i=1}^n (X_{I,i} - \bar{X}_I)(X_{J,i} - \bar{X}_J) \quad (7.5)$$

ここで

$$\bar{X}_I = E[X_I], \quad \bar{X}_J = E[X_J]$$

である。これらの共分散を以下のように行列形式に並べたものが分散共分散行列である。

$$\mathbf{C} = \begin{bmatrix} \text{cov}[X_1, X_1] & \text{cov}[X_1, X_2] & \dots & \text{cov}[X_1, X_m] \\ \text{cov}[X_2, X_1] & \text{cov}[X_2, X_2] & \dots & \text{cov}[X_2, X_m] \\ \dots & \dots & \dots & \dots \\ \text{cov}[X_m, X_1] & \text{cov}[X_m, X_2] & \dots & \text{cov}[X_m, X_m] \end{bmatrix} \quad (7.6)$$

分散共分散行列の対角要素は、データ列 X_I の分散(標準偏差の自乗)となり、非対角要素は、データ列 X_I と X_J の共分散となる。定義上、分散共分散行列は、対称行列である。また、

$$\begin{bmatrix} \frac{\text{cov}[X_1, X_1]}{\sigma[x_1]\sigma[x_1]} & \frac{\text{cov}[X_1, X_2]}{\sigma[x_1]\sigma[x_2]} & \dots & \frac{\text{cov}[X_1, X_m]}{\sigma[x_1]\sigma[x_m]} \\ \frac{\text{cov}[X_2, X_1]}{\sigma[x_2]\sigma[x_1]} & \frac{\text{cov}[X_2, X_2]}{\sigma[x_2]\sigma[x_2]} & \dots & \frac{\text{cov}[X_2, X_m]}{\sigma[x_2]\sigma[x_m]} \\ \dots & \dots & \dots & \dots \\ \frac{\text{cov}[\ddot{X}_m, X_1]}{\sigma[x_m]\sigma[x_1]} & \frac{\text{cov}[\ddot{X}_m, X_2]}{\sigma[x_m]\sigma[x_2]} & \dots & \frac{\text{cov}[\ddot{X}_m, X_m]}{\sigma[x_m]\sigma[x_m]} \end{bmatrix} \quad (7.7)$$

は相関行列と呼ばれ、各要素の相関係数を示したものとなる。

■例 19 : 2012 年 1 月 4 日～2012 年 6 月 27 日までの間の TOPIX(東京株価指数)、金価格、為替レート(円/ドル)の変動の相関を評価する。期間中のこれらの値の変化を図 7.1 に、相関関係を図 7.2 に示す。

²⁰ $X_1 = \{X_{1,1}, X_{1,2}, \dots, X_{1,i}, \dots\}, X_2 = \{X_{2,1}, X_{2,2}, \dots, X_{2,i}, \dots\}, \dots$ である。

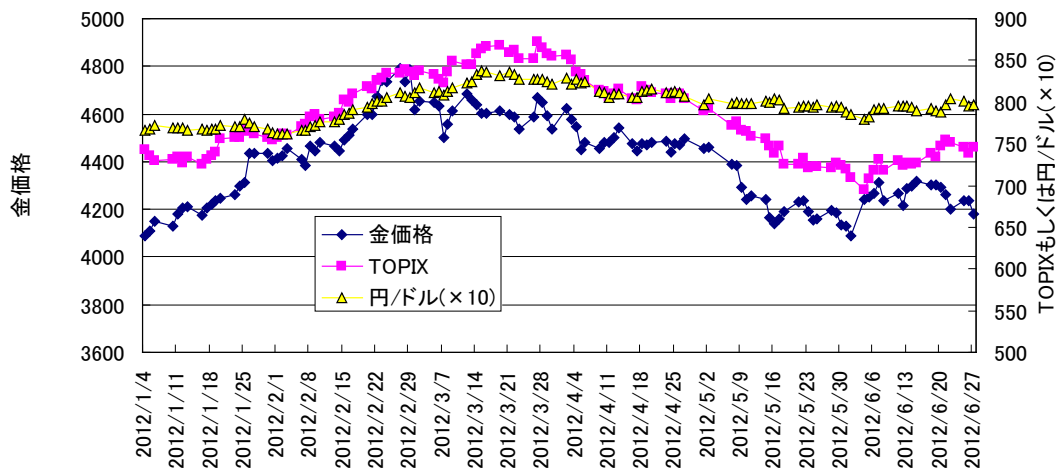


図 7.1 TOPIX、金価格、為替レートの変動

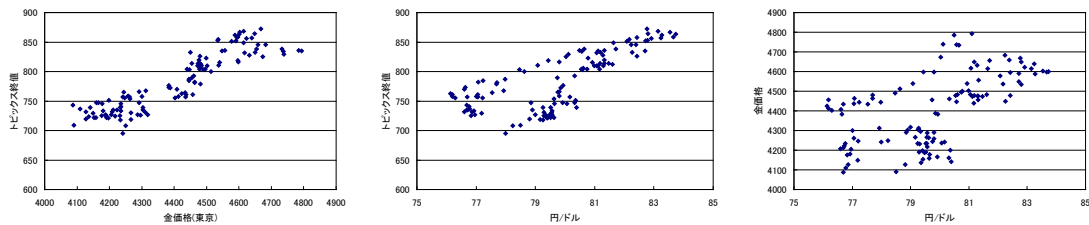


図 7.2 TOPIX(東京株価指数)、金価格、為替レート(円/ドル)の相関関係

また、これら 3 つのパラメータの分散共分散行列および相関行列を表 7.1 に示す。□

表 7.1 TOPIX(東京株価指数)、金価格、為替レートの分散共分散行列および相関行列

分散共分散行列				相関行列			
	金	トピックス終値	円/ドル		金	トピックス終値	円/ドル
金	32311			金	1.00		
トピックス終値	7854	2312		トピックス終値	0.91	1.00	
円/ドル	209	71.4	4.00	円/ドル	0.58	0.74	1.00

8.不確かさの伝搬

8.1 概要

ばらつきをそれぞれ有する二つのデータ列 X_i, Y_i がある。これらのデータ列の標準偏差は σ_x, σ_y であるとする。 $X_i + Y_i$ のばらつきはどのように評価できるだろうか。ここで、以下の三ケースについて考えてみよう。

- ① X_i, Y_i のばらつきが全く独立な場合
 - ② X_i, Y_i のばらつきに強い正の相関がある場合
 - ③ X_i, Y_i のばらつきに強い負の相関がある場合
- ①、②、③どのケースについて、 $X_i + Y_i$ のばらつきが大きくなるだろうか。②は、 X_i の

ばらつきが正側の時、 Y_i のばらつきも正側になるため、ばらつきが足し合わされ大きくなりそうである。最も極端なケースとして、 $Y_i = X_i$ であれば、 $X_i + Y_i$ の標準偏差は $2\sigma_x$ になる。一方、③は X_i のばらつきが正側の時、 Y_i のばらつきが負側になるため、ばらつきが打ち消しあい、小さくなりそうである。この極端なケースは、 $Y_i = -X_i$ であり、 $X_i + Y_i$ の標準偏差は零になる。①は、②と③の間に来ると推定される。

①、②、③のケースについて、 $X_i + Y_i$ の標準偏差は以下のように与えられる。これらの導出については後述する。

$$\textcircled{1} \sqrt{\sigma_x^2 + \sigma_y^2} \quad (8.1)$$

$$\textcircled{2} \sigma_x + \sigma_y \quad (8.2)$$

$$\textcircled{3} |\sigma_x - \sigma_y| \quad (8.3)$$

(8.1)式～(8.3)式を見ると、不確かさの大きさとしては、確かに②>①>③となっている。以下では、より一般的な形でこのような不確かさの伝搬について説明する。

8.2 多変数関数の誤差伝搬

ある多変数の関数 $f(X_1, X_2, \dots, X_m)$ がある。 X_1, X_2, \dots, X_m の不確かさ $\Delta X_1, \Delta X_2, \dots, \Delta X_m$ に起因する $f(X_1, X_2, \dots, X_m)$ の変化 Δf は、(8.4)式のように表すことができる。

$$\Delta f = \frac{\partial f}{\partial X_1} \Delta X_1 + \frac{\partial f}{\partial X_2} \Delta X_2 + \dots + \frac{\partial f}{\partial X_m} \Delta X_m = \sum_{k=1}^m \frac{\partial f}{\partial X_k} \Delta X_k \quad (8.4)$$

(8.4)式の両辺を二乗する。

$$(\Delta f)^2 = \sum_{l=1}^m \sum_{k=1}^m \frac{\partial f}{\partial X_l} \frac{\partial f}{\partial X_k} \Delta X_l \Delta X_k \quad (8.5)$$

(8.5)式の両辺の平均(期待値)を計算する²¹。

$$\begin{aligned} E[(\Delta f)^2] &= E[(f - \bar{f})^2] \\ &= E\left[\sum_{l=1}^m \sum_{k=1}^m \frac{\partial f}{\partial X_l} \frac{\partial f}{\partial X_k} \Delta X_l \Delta X_k\right] = \sum_{l=1}^m \sum_{k=1}^m \frac{\partial f}{\partial X_l} \frac{\partial f}{\partial X_k} E[\Delta X_l \Delta X_k] \\ &= \sum_{l=1}^m \sum_{k=1}^m \frac{\partial f}{\partial X_l} \frac{\partial f}{\partial X_k} E[(X_l - \bar{X}_l)(X_k - \bar{X}_k)] \end{aligned} \quad (8.6)$$

なお、 $\bar{f} = E[f(X_1, X_2, \dots, X_m)]$ 、 $\bar{X}_l = E[X_l]$ 、 $\bar{X}_k = E[X_k]$ である。

²¹ 変数が一つの際の平均値の計算は(5.2)式であった。多変数関数の場合には

$E[f(X_1, X_2, \dots, X_m)] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(X_1, X_2, \dots, X_m) p(X_1, X_2, \dots, X_m) dX_1 dX_2, \dots, dX_m$ となる。

(8.6)式より、(8.7)式が得られる。

$$E[(f - \bar{f})^2] = \sum_{l=1}^m \sum_{k=1}^m \frac{\partial f}{\partial X_l} \frac{\partial f}{\partial X_k} E[(X_l - \bar{X}_l)(X_k - \bar{X}_k)] \quad (8.7)$$

$E[(f - \bar{f})^2]$ は $f(X_1, X_2, \dots, X_m)$ の分散を、 $E[(X_l - \bar{X}_l)(X_k - \bar{X}_k)]$ は X_l と X_k の共分散を示している。従って、 $f(X_1, X_2, \dots, X_m)$ の分散は、変数の共分散 $E[(X_l - \bar{X}_l)(X_k - \bar{X}_k)]$ と偏微分値 $\frac{\partial f}{\partial X_l}, \frac{\partial f}{\partial X_k}$ から計算できる。なお、 $\frac{\partial f}{\partial X_l}$ は X_l が変化したときの f の変化量を示しており、 f に対する X_l の「効き方」を表しているので感度係数と呼ばれる²²。

8.3 二変数の場合の誤差伝搬

では、(8.7)式を二変数の場合について詳しく見てみよう。二変数の場合、(8.7)式は、

$$\begin{aligned} E[(\Delta f)^2] &= \frac{\partial f}{\partial X_1} \frac{\partial f}{\partial X_1} E[\Delta X_1 \Delta X_1] + \frac{\partial f}{\partial X_1} \frac{\partial f}{\partial X_2} E[\Delta X_1 \Delta X_2] \\ &\quad + \frac{\partial f}{\partial X_2} \frac{\partial f}{\partial X_1} E[\Delta X_2 \Delta X_1] + \frac{\partial f}{\partial X_2} \frac{\partial f}{\partial X_2} E[\Delta X_2 \Delta X_2] \end{aligned} \quad (8.8)$$

X_1, X_2 のばらつきに全く相関がない場合（共分散 = 0 の場合）、 $E[\Delta X_1 \Delta X_2] = E[\Delta X_2 \Delta X_1] = 0$ であることから、(8.8)式は(8.9)式の形に書くことができる。

$$E[(\Delta f)^2] = \frac{\partial f}{\partial X_1} \frac{\partial f}{\partial X_1} E[\Delta X_1 \Delta X_1] + \frac{\partial f}{\partial X_2} \frac{\partial f}{\partial X_2} E[\Delta X_2 \Delta X_2] \quad (8.9)$$

両辺の平方根を取って

$$\begin{aligned} \sigma(f(X_1, X_2)) &= \sqrt{\text{var}(f(X_1, X_2))} = \sqrt{E[(\Delta f)^2]} \\ &= \sqrt{\frac{\partial f}{\partial X_1} \frac{\partial f}{\partial X_1} E[\Delta X_1 \Delta X_1] + \frac{\partial f}{\partial X_2} \frac{\partial f}{\partial X_2} E[\Delta X_2 \Delta X_2]} \\ &= \sqrt{\left(\frac{\partial f}{\partial X_1}\right)^2 \text{cov}[X_1, X_1] + \left(\frac{\partial f}{\partial X_2}\right)^2 \text{cov}[X_2, X_2]} \\ &= \sqrt{\left(\frac{\partial f}{\partial X_1}\right)^2 \text{var}[X_1, X_1] + \left(\frac{\partial f}{\partial X_2}\right)^2 \text{var}[X_2, X_2]} \\ &= \sqrt{\left(\frac{\partial f}{\partial X_1} \sigma(X_1)\right)^2 + \left(\frac{\partial f}{\partial X_2} \sigma(X_2)\right)^2} \end{aligned} \quad (8.10)$$

²² 炉物理分野の感度解析で出てくる感度係数は、相対感度係数であり、 $\frac{X_l}{f} \frac{\partial f}{\partial X_l} = \frac{\partial f}{f} / \frac{\partial X_l}{X_l}$ で表される。これは、 X_l の相対的な変化量に対する f の相対的な変化量を示している。例えば、 X_l が 1% 変化したときの f の変化量を % 単位で示したものである。

であり、通常の誤差の合成になる。

■ 例 20 : X_1, X_2 の不確かさによる $f_1(X_1, X_2) = X_1 + X_2$ 、 $f_2(X_1, X_2) = X_1 X_2$ 、 $f_3(X_1, X_2) = X_1 / X_2$ の不確かさ(標準偏差)を計算する。なお、 X_1, X_2 の不確かさに相関はないものとする。(8.10)式を用いると、

$$\sigma(f_1(X_1, X_2)) = \sqrt{(\sigma(X_1))^2 + (\sigma(X_2))^2} \quad (8.11)$$

$$\begin{aligned} \sigma(f_2(X_1, X_2)) &= \sqrt{(X_2 \sigma(X_1))^2 + (X_1 \sigma(X_2))^2} \\ &= X_1 X_2 \sqrt{\left(\frac{\sigma(X_1)}{X_1}\right)^2 + \left(\frac{\sigma(X_2)}{X_2}\right)^2} \end{aligned} \quad (8.12)$$

$$\begin{aligned} \sigma(f_3(X_1, X_2)) &= \sqrt{\left(\frac{1}{X_2} \sigma(X_1)\right)^2 + \left(-\frac{X_1}{(X_2)^2} \sigma(X_2)\right)^2} \\ &= \frac{X_1}{X_2} \sqrt{\left(\frac{\sigma(X_1)}{X_1}\right)^2 + \left(\frac{\sigma(X_2)}{X_2}\right)^2} \end{aligned} \quad (8.13)$$

以上。□

次に、 X_1, X_2 の相関係数が α である場合を考える。(7.2)式より

$$\alpha = \frac{E[\Delta X_1 \Delta X_2]}{\sqrt{E[\Delta X_1 \Delta X_1]} \sqrt{E[\Delta X_2 \Delta X_2]}} \quad (8.14)$$

従って

$$E[\Delta X_1 \Delta X_2] = \alpha \sqrt{E[\Delta X_1 \Delta X_1]} \sqrt{E[\Delta X_2 \Delta X_2]} = \alpha \sigma(X_1) \sigma(X_2) \quad (8.15)$$

(8.8)式より

$$\begin{aligned} \sigma(f(X_1, X_2)) &= \sqrt{\frac{\frac{\partial f}{\partial X_1} \frac{\partial f}{\partial X_1} E[\Delta X_1 \Delta X_1] + \frac{\partial f}{\partial X_1} \frac{\partial f}{\partial X_2} E[\Delta X_1 \Delta X_2]}{\quad} + \frac{\frac{\partial f}{\partial X_2} \frac{\partial f}{\partial X_1} E[\Delta X_2 \Delta X_1] + \frac{\partial f}{\partial X_2} \frac{\partial f}{\partial X_2} E[\Delta X_2 \Delta X_2]}{\quad}} \\ &= \sqrt{\left(\frac{\partial f}{\partial X_1} \sigma(X_1)\right)^2 + \left(\frac{\partial f}{\partial X_2} \sigma(X_2)\right)^2 + 2\alpha \frac{\partial f}{\partial X_1} \frac{\partial f}{\partial X_2} \sigma(X_1) \sigma(X_2)} \end{aligned} \quad (8.16)$$

$\alpha = 1$ の場合

$$\begin{aligned}\sigma(f(X_1, X_2)) &= \sqrt{\left(\frac{\partial f}{\partial X_1} \sigma(X_1)\right)^2 + \left(\frac{\partial f}{\partial X_2} \sigma(X_2)\right)^2 + 2 \frac{\partial f}{\partial X_1} \frac{\partial f}{\partial X_2} \sigma(X_1) \sigma(X_2)} \\ &= \sqrt{\left(\frac{\partial f}{\partial X_1} \sigma(X_1) + \frac{\partial f}{\partial X_2} \sigma(X_2)\right)^2} = \left|\frac{\partial f}{\partial X_1} \sigma(X_1) + \frac{\partial f}{\partial X_2} \sigma(X_2)\right|\end{aligned}\quad (8.17)$$

$\alpha = -1$ の場合

$$\begin{aligned}\sigma(f(X_1, X_2)) &= \sqrt{\left(\frac{\partial f}{\partial X_1} \sigma(X_1)\right)^2 + \left(\frac{\partial f}{\partial X_2} \sigma(X_2)\right)^2 - 2 \frac{\partial f}{\partial X_1} \frac{\partial f}{\partial X_2} \sigma(X_1) \sigma(X_2)} \\ &= \sqrt{\left(\frac{\partial f}{\partial X_1} \sigma(X_1) - \frac{\partial f}{\partial X_2} \sigma(X_2)\right)^2} = \left|\frac{\partial f}{\partial X_1} \sigma(X_1) - \frac{\partial f}{\partial X_2} \sigma(X_2)\right|\end{aligned}\quad (8.18)$$

■例 21 : $f(X_1, X_2) = X_1 + X_2$ であれば、 $\frac{\partial f}{\partial X_1} = \frac{\partial f}{\partial X_2} = 1$ であることから、本章の最初に示した①、②、③のそれぞれのケースが正しいことが分かる²³。□

8.4 不確かさ伝搬の Sandwich formula

次に、核特性の感度解析で用いる不確かさ伝搬の式を導いておく。(8.7)式を f^2 で割ることにより、

$$\begin{aligned}E\left[\left(\frac{\Delta f}{f}\right)^2\right] &= \sum_{l=1}^m \sum_{k=1}^m \frac{X_l X_k}{f^2} \frac{\partial f}{\partial X_l} \frac{\partial f}{\partial X_k} E\left[\frac{\Delta X_l}{X_l} \frac{\Delta X_k}{X_k}\right] \\ &= \sum_{l=1}^m \sum_{k=1}^m \frac{\partial f / f}{\partial X_l / X_l} \frac{\partial f / f}{\partial X_k / X_k} E\left[\frac{\Delta X_l}{X_l} \frac{\Delta X_k}{X_k}\right] \\ &= \sum_{l=1}^m \sum_{k=1}^m S_{f, X_l} S_{f, X_k} E\left[\frac{\Delta X_l}{X_l} \frac{\Delta X_k}{X_k}\right]\end{aligned}\quad (8.19)$$

ここで $S_{f, X_l} = \frac{\partial f / f}{\partial X_l / X_l}$, $S_{f, X_k} = \frac{\partial f / f}{\partial X_k / X_k}$ である。これらは、脚注 15 で述べたように相対感度係数である。(8.19)式を行列形式で書くと

²³ 誤差をベクトルと考え、共分散をこれらのベクトルのなす角と考えるとわかりやすい。共分散(相関係数)=0 は二つの誤差ベクトルが直角であり、二つのベクトルの大きさは $\sqrt{|x|^2 + |y|^2}$ で合成される。共分散(相関係数)=1 は二つの誤差ベクトルが平行であり、二つのベクトルの大きさは $|x| + |y|$ で合成される。

$$E\left[\left(\frac{\Delta f}{f}\right)^2\right] = \begin{bmatrix} S_{f,X_1} & S_{f,X_2} & \dots & S_{f,X_m} \end{bmatrix} \begin{bmatrix} E\left[\frac{\Delta X_1}{X_1} \frac{\Delta X_1}{X_1}\right] & E\left[\frac{\Delta X_1}{X_1} \frac{\Delta X_2}{X_2}\right] & \dots & E\left[\frac{\Delta X_1}{X_1} \frac{\Delta X_m}{X_m}\right] \\ E\left[\frac{\Delta X_2}{X_2} \frac{\Delta X_1}{X_1}\right] & E\left[\frac{\Delta X_2}{X_2} \frac{\Delta X_2}{X_2}\right] & \dots & E\left[\frac{\Delta X_2}{X_2} \frac{\Delta X_m}{X_m}\right] \\ \dots & \dots & \dots & \dots \\ E\left[\frac{\Delta X_m}{X_m} \frac{\Delta X_1}{X_1}\right] & E\left[\frac{\Delta X_m}{X_m} \frac{\Delta X_2}{X_2}\right] & \dots & E\left[\frac{\Delta X_m}{X_m} \frac{\Delta X_m}{X_m}\right] \end{bmatrix} \begin{bmatrix} S_{f,X_1} \\ S_{f,X_2} \\ \dots \\ S_{f,X_m} \end{bmatrix} \quad (8.20)$$

である。これは

$$\mathbf{C}_{ff} = {}^t \mathbf{S}_f \mathbf{C}_{xx} \mathbf{S}_f \quad (8.21)$$

と表記することができる。 \mathbf{C}_{ff} は f の相対分散共分散行列(行列サイズ: 1×1)、 \mathbf{C}_{xx} は

X_1, X_2, \dots, X_m の相対共分散行列(行列サイズ: $m \times m$)、 \mathbf{S}_f は f の相対感度行列(行列サイズ $m \times 1$)である。(8.19)式は、共分散行列を感度行列で「挟み込む」形になっているため、炉物理の感度解析業界では"Sandwich Formula"などと呼ばれている。(8.20)式の平方根を取ることによって f の相対標準偏差を求めることができる。

次に考慮すべき関数が複数 M 個ある場合を考える。この場合、

$$\mathbf{S}_f = \begin{bmatrix} \frac{\partial f_1 / f_1}{\partial X_1 / X_1} & \frac{\partial f_2 / f_2}{\partial X_1 / X_1} & \dots & \frac{\partial f_M / f_M}{\partial X_1 / X_1} \\ \frac{\partial f_1 / f_1}{\partial X_2 / X_2} & \frac{\partial f_2 / f_2}{\partial X_2 / X_2} & \dots & \frac{\partial f_M / f_M}{\partial X_2 / X_2} \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_1 / f_1}{\partial X_m / X_m} & \frac{\partial f_2 / f_2}{\partial X_m / X_m} & \dots & \frac{\partial f_M / f_M}{\partial X_m / X_m} \end{bmatrix}$$

となり、 \mathbf{S}_f のサイズは $m \times M$ 、 \mathbf{C}_{xx} のサイズは $m \times m$ 、 \mathbf{S} のサイズは $M \times m$ であるので、

\mathbf{C}_{ff} は $M \times M$ の行列になる。 \mathbf{C}_{ff} の対角要素は、 X_1, X_2, \dots, X_m の不確かさに起因するそれぞれの関数の相対値の分散、非対角要素は断面積の不確かさに起因する二つの関数の相対値の共分散を示している。

■例 22: X_1, X_2, \dots, X_m が多群の微視的断面積、 f_i が少数群の集合体平均巨視的断面積であるとする。この場合、 \mathbf{S}_f は、多群の微視的断面積に対する少数群集合体平均巨視的断面積の感度係数を表す行列となる。この感度係数行列を何らかの方法で求めることができれば、

(8.19)式により少数群集合体平均巨視的断面積の分散と、共分散を求めることができる。この分散は、例えば1群の吸収断面積、2群の核分裂断面積の不確かさを示すものである。また共分散は、1群の吸収断面積、2群の核分裂断面積がどの程度の相関を持つかを表すことになる。□

■例 23 : x_1, x_2, \dots, x_m が少数群の集合体平均断面積、 f_i が核特性値(例：実効増倍率、出力分布、反応度係数、・・・)であるとすると、(8.21)を用いることで少数群集合体平均断面積の不確かさに起因して炉心核特性値にどの程度の不確かさを生じるか、また炉心特性値間にどの程度の相関が生じるのかを計算することができる。□

9.まとめ

不確かさの評価技術は、解析技術の進歩に伴ってより重要になる。しかし、不確かさ評価で必要となる計算は多くの場合複雑な統計処理の理論に基づいており、これらを理解することはそれなりにハードルが高い。そこで、本稿では、この不確かさ評価を実施する上で必要となる基礎的な事項をまとめ、不確かさ評価に取り組む上での手助けになる参考資料とすることを目指した。

謝辞

推定臨界下限増倍率の評価(Appendix)については、名古屋大学の遠藤知弘氏にまとめて頂いたものである。また、本稿の草稿について、大阪大学の北田孝典氏、北海道大学の千葉豪氏、名古屋大学の遠藤知弘氏に有益なコメントを頂いた。ここに記して感謝の意を表したい。

Appendix. 推定臨界下限増倍率の算出法

本付録では、推定臨界下限増倍率 k_{sub} の算出方法について補足説明を行う。ここでは n 個の臨界実験を解析して、実効増倍率 k_{eff} の標本として $k_i (i=1, \dots, n)$ を得たとして話を進める。

得られた n 個の k_{eff} の解析結果から、 k_{eff} の標本平均 \bar{k} および不偏分散 s^2 を以下のように推定することができる。

$$\bar{k} = \frac{1}{n} \sum_{i=1}^n k_i \quad (\text{A.1})$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (k_i - \bar{k})^2 \quad (\text{A.2})$$

ここで k_{eff} の母集団は平均を μ 、標準偏差 σ の正規分布に従うと仮定する。

■例題：あるモンテカルロコードで 10 個の臨界実験を解析したところ、

1.0412, 1.0236, 1.0198, 1.0297, 1.0087, 0.9741, 1.0195, 1.0200, 1.0243, 0.9818

という値が得られた。このモンテカルロコードの誤差が正規分布していると仮定できる場合、推定臨界下限増倍率を求めよ。(言い換えると、実効増倍率がいくら以下であれば未臨界と言えるかを明らかにせよ、ということである。)

なお、本計算の平均値 \bar{k} と標準偏差 s は以下のように求められる。

$$\bar{k} = \frac{1.0412 + 1.0236 + 1.0198 + 1.0297 + 1.0087 + 0.9741 + 1.0195 + 1.0200 + 1.0243 + 0.9818}{10} = 1.0143$$

$$s = \sqrt{\frac{(1.0412 - 1.0143)^2 + (1.0236 - 1.0143)^2 + (1.0198 - 1.0143)^2 + (1.0297 - 1.0143)^2 + (1.0087 - 1.0143)^2 + (0.9741 - 1.0143)^2 + (1.0195 - 1.0143)^2 + (1.0200 - 1.0143)^2 + (1.0243 - 1.0143)^2 + (0.9818 - 1.0143)^2}{(10 - 1)}} \\ = 0.0209$$

A.1 必要となる統計学知識

推定臨界下限増倍率を算出するに当たって、必要となる統計学の知識を以下で記述する。

(1) 標本平均 \bar{k} は、平均 μ 、標準偏差 σ/\sqrt{n} の正規分布に従う。

(2) 不偏分散 s^2 については、

$$z = (n-1) \frac{s^2}{\sigma^2} = \sum_{i=1}^n \frac{(k_i - \bar{k})^2}{\sigma^2} \quad (\text{A.3})$$

という量が自由度 $(n-1)$ のカイ二乗分布に従う。カイ二乗分布の性質より、 z の平均は $n-1$ 、分散は $2(n-1)$ となる。従って不偏分散 s^2 について考えた場合には、 s^2 の平均は σ^2 、分散 $2\sigma^4/(n-1)$ となる。 n が十分に大きくなるにつれて、 s^2 は平均 σ^2 、標準偏差 $\sqrt{2/(n-1)}\sigma^2$ の正規分布に近づくようになる。

- (3) 変数 X として平均 λ 、分散 1 の正規分布に従う量を考える。また、変数 Y として自由度 ν のカイ二乗分布に従う量を考える。ここで X と Y が統計的に独立である場合、以下で定義される量 T は、自由度 ν 、非心パラメータ λ の非心 t 分布に従う。

$$T = \frac{X}{\sqrt{Y/\nu}} \quad (\text{A.4})$$

A.2 実効増倍率の母集団の平均値と標準偏差がわかっている場合

まず、最も簡単なケースとして、実効増倍率の母集団の平均 μ と標準偏差 σ がわかっている場合を考える。例題として挙げた 10 個の実効増倍率の結果は、実は平均 1.01、標準偏差 0.02 の正規分布に従って、10 回ランダムにサンプリングを行った結果である。すなわち k_{eff} の母集団は $\mu = 1.01$ 、標準偏差 $\sigma = 0.02$ の正規分布に従っている。この母集団の平均と標準偏差を用いると、例えば $1.01 - 2 \times 0.02 < k_{eff} < 1.01 + 2 \times 0.02$ になる確率は 95.44% であることが言える。これは、 $k_{eff} < 1.01 - 2 \times 0.02$ になる確率(下側累積確率)が 2.28%($= (100\% - 95.44\%) / 2$)であることと等価である。従って、推定臨界下限増倍率 k_{sub} として、例えば 0.97(=平均値 - 2×標準偏差 = $1.01 - 2 \times 0.02$)にした場合、 k_{eff} が 0.97 未満であれば、臨界にならない確率は $100\% - 2.28\% = 97.72\%$ となる。逆に、臨界になる確率は 2.28% である(図 A.1)。これは、 k_{eff} が 0.97 未満となる解析値が 10000 ケースあった場合、そのうちの 228 ケースが臨界になることを示している。判断基準を 0.95(= $1.01 - 3 \times 0.02$)とした場合、つまり k_{eff} が 0.95 未満であれば、臨界にならない確率は先と同様の手順により 99.87% になる。以上で議論したように、臨界になる確率を零にはできないが、推定臨界下限増倍率 k_{sub} を小さく設定することにより、限りなく小さくできることを示している。

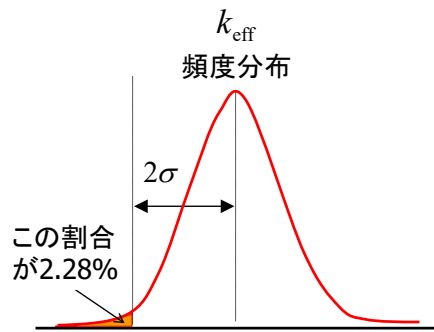


図 A.1 母集団の平均と標準偏差がわかっている場合の $k_{sub}(\mu, \sigma)$ 評価例

以上で述べた手順をまとめると、母集団の平均 μ と標準偏差 σ がわかっている場合については $k_{sub}(\mu, \sigma)$ を以下のような手順で設定することができる。

- ① 臨界になる確率 p を設定する。
- ② 実効増倍率 k_{eff} が平均 μ 、標準偏差 σ の正規分布に従うとして、正規分布の下側累積確率 p に対応するパーセント点 $-\alpha_p$ を求める。すなわち k_{eff} の分布を $k_{eff} \leq \mu - \alpha_p \sigma$ まで積分した累積確率が p に対応する。
- ③ 推定臨界下限増倍率を $k_{sub}(\mu, \sigma) = \mu - \alpha_p \sigma$ と設定する。このように設定することで、 k_{eff} の解析結果が $k_{sub}(\mu, \sigma)$ 未満であれば、臨界にならない確率が $(1 - p)$ となる。

以上の手順に基づいて例題の推定臨界下限増倍率 $k_{sub}(\mu, \sigma)$ を算出してみると、臨界になる確率を $p = 2.5\%$ と設定した場合には、それに対応するパーセント点は $\alpha_p = 1.96$ となるため、 $k_{sub}(\mu, \sigma) = 1.01 - 1.96 \times 0.02 = \underline{0.9708}$ となる。

A.3 実効増倍率の母集団の平均値と標準偏差が不明であるが、これらが十分に精度良く推定できている場合

しかし、実際には母集団の平均 μ と標準偏差 σ は未知であり、これらは計算結果(標本)の平均 \bar{k} と不偏分散 s^2 により推定される。ここで重要なことは、標本を用いた平均 \bar{k} も不偏分散 s^2 もいずれも不確かさを含んでいるということである。つまり、標本を取り直すと(別の体系を計算してみると)、平均と不偏分散としては、別の値が得られるということである。ただし、このような状況であったとしても、標本数 n が十分に多い場合には、平均と

不偏分散の不確かさを十分小さくすることができる。これは A.1 節で述べたように、標本数 $n \rightarrow \infty$ の極限で \bar{k} と s^2 の分散(標準偏差)が零に漸近することからも理解できる。従って、標本数が十分に大きく、平均と不偏分散の不確かさが十分に小さければ、図 A.2 に示すように、図 A.1 と(厳密には同等でないが)ほぼ同じようなやり方で推定臨界下限増倍率を求めることができそうである。

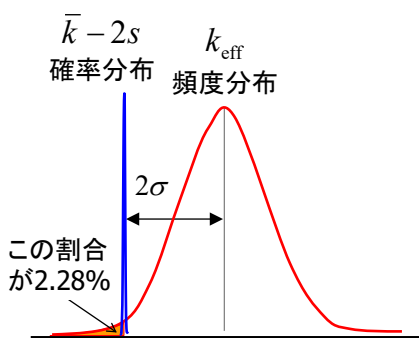


図 A.2 母集団の平均と標準偏差が未知、標本数が十分に大きい場合の k_{sub} 評価例

従って、標本数が十分に大きく \bar{k} と s^2 の推定結果が母集団の平均 μ と σ^2 に限りなく近い場合については、推定臨界下限増倍率 k_{sub} を以下のような手順で設定できそうである。

- ① n 個の実効増倍率 k_{eff} の解析結果から、平均 \bar{k} と不偏分散 s^2 を求める。
- ② 臨界になる確率 p を設定し、実効増倍率 k_{eff} が平均 μ 、標準偏差 σ の正規分布に従うとして、正規分布の下側累積確率 p に対応するパーセント点 $-\alpha_p$ を求める。
- ③ 推定臨界下限増倍率を $k_{sub} = \bar{k} - \alpha_p s$ と設定する。

以上の手順に基づいて例題の推定臨界下限増倍率 k_{sub} を算出してみると、 $p = 2.5\%$ と設定した場合には、 $k_{sub} = 1.0143 - 1.96 \times 0.0209 = 0.9733$ となる。この値は A.2 節で求められた、母集団の平均 μ と σ から推定された $k_{sub}(\mu, \sigma)$ の値 0.9708 に比較的近いものの、若干大きな値となっている。こうして求められた k_{sub} の値 0.9733 は、A.2 節で述べた μ と σ を用いた評価方法で考えた場合には、 $k_{sub}(\mu, \sigma) = 0.9733 = 1.01 - 1.835 \times 0.02$ に対応しており、実質的には α_p の値を小さめに設定してしまったことに対応する。すなわち、臨界になる確率が $p = 2.5\%$ となるように k_{sub} を設定したつもりであったが、 μ と σ のかわりに用いた \bar{k} と s の不確かさによって、臨界になる確率は 3.3% とわずかに増えてしまっている。

例題で示された 10 個の k_{eff} 解析結果の場合は、標本数 n が 10 個でも k_{sub} を(たまたま)それなりの値で設定することができた。参考までに、10 個で 1 セットの臨界実験の解析を別途 1000 ケース実施したことを想定してランダムサンプリングを行い、本節で述べた手法により各ケースで k_{sub} を算出した結果を図 A.3 に示す。図 A.3 に示したように、全 1000 ケースのうち約 56%が $k_{sub}(\mu, \sigma) = 0.9708$ よりも大きく設定される結果となり、これら約 56%のケースについては、「臨界になる確率が $p = 2.5\%$ 超えることがないように k_{sub} を設定できていない」こととなる。つまり、ここで示した手順が適用できるのは、あくまでも母集団の平均と標準偏差が「きわめてよく推定できている」場合のみであり、サンプル数が 10 程度のケースに使用することはできないことがわかる。

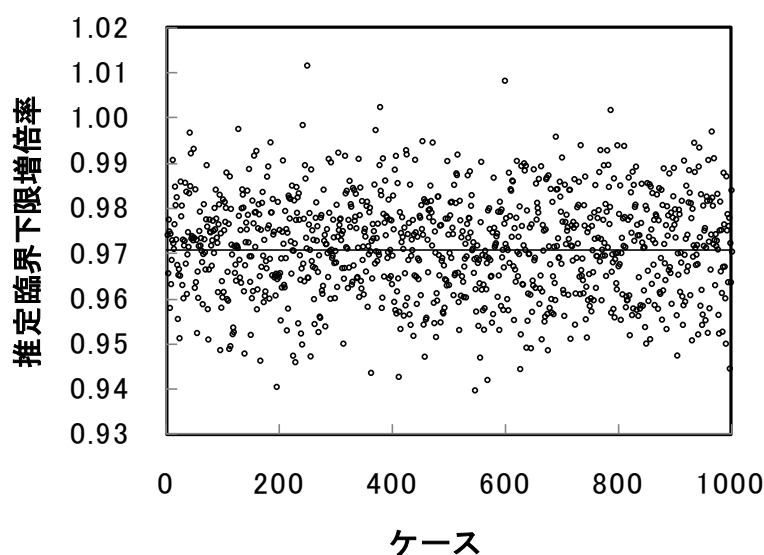


図 A.3 $\bar{k} - \alpha_p s$ により設定した k_{sub} の散布図(標本数 $n = 10$, $p = 2.5\%$)

A.4 実効増倍率の母集団の平均値が不明であり、標準偏差のみがわかっている場合

A.3 節では μ と σ のかわりとして \bar{k} と s を用いた k_{sub} の設定方法について述べたが、現実的には、 \bar{k} と s が μ と σ に等しいとみなせるほど十分な標本数を得ることはできない場合もある。この場合、後述の図 A.6 に示すように、(推定臨界下限増倍率の候補となり得る)「平均値 $- \alpha \times$ 標準偏差 ($\bar{k} - \alpha s$)」の値自体がある範囲でばらつくことになる。本節では、次のステップとして、実効増倍率の母集団の平均値が不明で (\bar{k} のみ不確かさがあり)、母集団の標準偏差がわかっているという仮想的な場合を考えることとする(図 A.4)

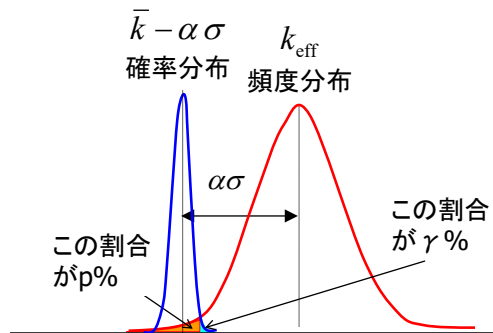


図 A.4 母集団の平均値が不明であるが、標準偏差が分かっている場合

この場合、標本平均 \bar{k} は平均が μ 、標準偏差が σ/\sqrt{n} の正規分布に従ってばらつくことになる。ここで仮に、実効増倍率の平均値 \bar{k} が真値 μ より高めに推定されたとする。例えば、 $1.01 + 2 \frac{0.02}{\sqrt{10}} = 1.0226$ という平均値が得られたとする(1.0226 より高い平均値 \bar{k} が得られる確率 γ は 2.28% である)。ここで、臨界になる確率 p を 2.28% となることを期待して、 $\alpha = 2$ として推定臨界下限増倍率 k_{sub} を $\bar{k} - \alpha\sigma = 1.0226 - 2 \times 0.02 = 0.9826$ と推定してみる。こうして得られた値は、A.2 節で述べた $\mu - 2\sigma$ で求められる臨界下限増倍率 $0.97 (= 1.01 - 2 \times 0.02)$ より明らかに大きい。すなわち、 k_{sub} として 0.9826 を設定してしまうと臨界になる確率は 2.28% より大きくなる(この場合、実質的には $\alpha_p = 1.37$ と設定したこととなり、臨界になる確率は 8.53% と大きくなる)。従って、 \bar{k} のばらつきまで考慮して、臨界になる確率が 2.28% 超えないようにするためには、 α の値を 2 よりも大きくとる必要がある。

A.2, A.3 のケースでは、目標になる実効増倍率 $\bar{k} - \alpha\sigma$ が幅を持たない「線」だった。従って、どれぐらいの確率で臨界になるか、ということだけを考えていれば良かった。しかし、今回は、目標になる実効増倍率 $\bar{k} - \alpha\sigma$ が「幅を持つ領域」になってしまっている。ということは、この「幅を持つ領域」の中で、どこに目標値(つまり、どれぐらいの確率で臨界になるかというしきい値の線)を置くか、ということも追加的に考えないといけないことになる。考えないとならないパラメータが二つに増えたことになる。図 A.4 の $\bar{k} - \alpha\sigma$ が図 A.4 に示す確率分布を持つ場合、どのあたりに目標値(どれぐらいの確率で臨界になるかというしきい値の線)を置けば良いだろうか。仮に $\bar{k} - \alpha\sigma$ の確率分布の中で、左側、つまり $\bar{k} - \alpha\sigma$ を小さい側に設定したとする。これは、 \bar{k} (実効増倍率の平均値)を低い側に見積もったこと、それだけ小さい実効増倍率でしか臨界にならないと評価したことと相当するため、非安全側の設定になると言える。従って、 $\bar{k} - \alpha\sigma$ の分布において、右側、つまり、 $\bar{k} - \alpha\sigma$ を大きい側に設定する必要がある。では、どこまで大きく設定すれば良いのだろうか

か。これは、ある確率 $\gamma\%$ を考えて、 $\bar{k} - \alpha\sigma$ が設定した値より大きくなる確率が $\gamma\%$ 以下になるように設定してやれば良い。このように設定したしきい値、あるいは目標値が臨界になる確率が $p\%$ の線に来るように α を調節してやれば良い。

以上の議論を踏まえると、 k_{sub} を決定する際には以下のような手順を踏むこととなる。

① n 個の標本値より平均値 \bar{k} を計算する

② 臨界になる確率 p を設定し、実効増倍率 k_{eff} が平均 μ 、標準偏差 σ の正規分布に従うと

して、正規分布の下側累積確率 p に対応するパーセント点 $-\alpha_p$ を求める。

③ $\bar{k} - \alpha\sigma \leq \mu - \alpha_p\sigma$ を満足するように、 α の値を設定することを考える。ここで、この条件式を以下のように式変形する。

$$\frac{\bar{k} - \mu}{\sigma/\sqrt{n}} \leq (\alpha - \alpha_p)\sqrt{n} \quad (\text{A.5})$$

④ (A.5) 式の左辺の量は、A.1 節で述べた \bar{k} の統計的性質より、平均が 0、標準偏差が 1 の標準正規分布に従うことがわかる。よって、 $\frac{\bar{k} - \mu}{\sigma/\sqrt{n}}$ の信頼区間として下側累積確率が $(1 - \gamma)$

となる区間(上側累積確率が γ となる区間) α_γ を、標準正規分布の上側累積確率 γ に対応するパーセント点から求める。

$$\frac{\bar{k} - \mu}{\sigma/\sqrt{n}} \leq \alpha_\gamma \quad (\text{A.6})$$

⑤ (A.5), (A.6) 式の右辺の値を比較することで、 $\bar{k} - \alpha\sigma \leq \mu - \alpha_p\sigma$ を満足するためには、パラメータ α の値を以下のように設定すればよいことが分かる。

$$\alpha = \alpha_p + \frac{\alpha_\gamma}{\sqrt{n}} \quad (\text{A.7})$$

⑥ こうして求められた α を使用して推定臨界下限増倍率を $k_{sub}(\bar{k}, \sigma) = \bar{k} - \alpha\sigma$ と設定する。

(A.7) 式からわかるように、実効増倍率の平均値として母集団の平均 μ のかわりに標本平

均 \bar{k} を用いた場合には、 \bar{k} の不確かさのぶんとして α_γ/\sqrt{n} だけ α の値を α_p より大きく設定すれば良いことが分かる。

上述した手順に基づいて例題の推定臨界下限増倍率 $k_{sub}(\bar{k}, \sigma)$ を算出してみる。
 $p = 2.5\%$ 、 $(1-\gamma) = 97.5\%$ の場合には、正規分布の下/上側累積確率のパーセント点より $\alpha_p = \alpha_\gamma = 1.96$ となるため、 $k_{sub}(\bar{k}, \sigma) = 1.0143 - \left(1.96 + \frac{1.96}{\sqrt{10}}\right) \times 0.02 = \underline{0.9627}$ と求められる。この値は、A.2節で求められた $k_{sub}(\mu, \sigma)$ の値0.9708より小さく、 \bar{k} の不確かさも含めて保守的に推定臨界下限増倍率を設定していると言える。手順④で(A.6)式の信頼区間を考えたことから理解できるように、もし仮に10個で1セットの臨界実験の解析を別途1000ケース実施したとすると、期待値として1000ケースのうち25のケース(=1000 \times γ =1000 \times 2.5%)について「設定した k_{sub} を用いた場合、臨界になる確率が $p = 2.5\%$ を超える可能性がある」といえる。参考までに、10個で1セットの臨界実験の解析を別途1000ケース実施したことを想定してランダムサンプリングを行い、各ケースで $k_{sub}(\bar{k}, \sigma)$ を算出した結果を図A.5に示す。図A.5に示したように、全1000ケースのうち26ケースがA.2で求めた「正しい」推定臨界下限増倍率である $k_{sub}(\mu, \sigma) = 0.9708$ よりも大きく設定される結果となり、その頻度は $\gamma = 2.5\%$ にほぼ等しいことが確認できる。つまり、推定臨界下限増倍率の「設定に失敗する」確率が予測した2.5%とほぼ一致していることが分かる。

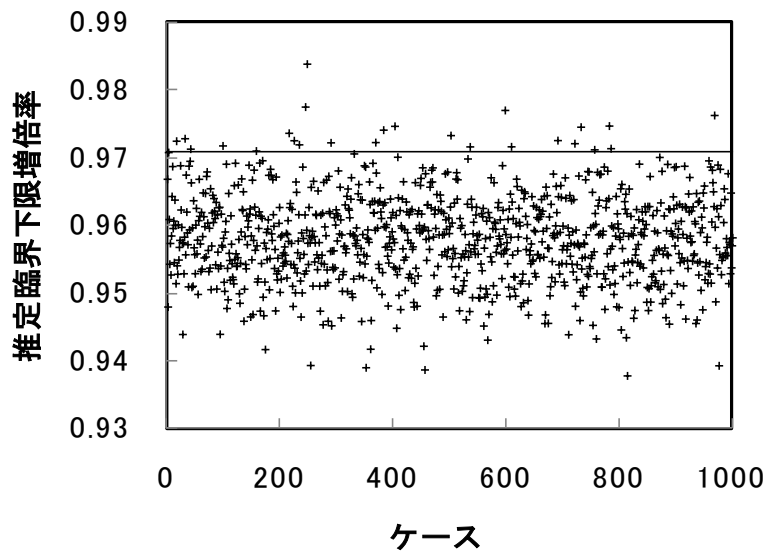


図 A.5 $k_{sub}(\bar{k}, \sigma)$ の散布図(標本数 $n = 10$, $p = \gamma = 2.5\%$)

A.5 実効増倍率の母集団の平均値も標準偏差も不明な場合

最後に、母集団の平均値も標準偏差も不明で、 \bar{k} と s の両方に不確かさがある場合について考える(図 A.6)。

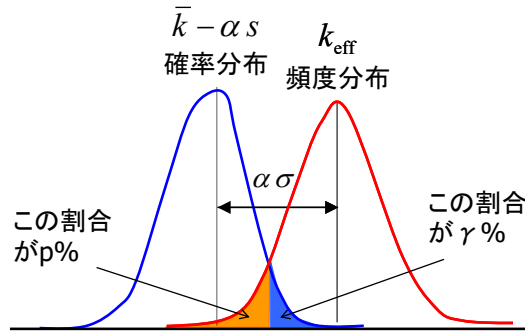


図 A.6 実効増倍率の母集団の平均値も標準偏差も不明な場合

この場合、A.4 節で述べた内容に加えて、 n 個の標本から推定された不偏分散 s^2 がカイ二乗分布に従うことも利用して、推定臨界下限増倍率 $k_{sub}(\bar{k}, s) = \bar{k} - \alpha s$ のパラメータ α を設定することとなる。文献(奥野, 内藤, "臨界安全安全解析コードシステム JACS の計算誤差評価", JAERI-M 87-057)を参考として、具体的な $k_{sub}(\bar{k}, s)$ の設定手順を以下にまとめる。

- ① n 個の標本値より \bar{k} と s^2 を求める。
- ② 臨界になる確率 p を設定し、実効増倍率 k_{eff} が平均 μ 、標準偏差 σ の正規分布に従うとして、正規分布の下側累積確率 p に対応するパーセント点 $-\alpha_p$ を求める。
- ③ $\bar{k} - \alpha s \leq \mu - \alpha_p \sigma$ を満足するように、 α の値を設定することを考える。ここで、この条件式を以下のように式変形する。

$$\frac{\frac{\bar{k} - \mu}{\sigma/\sqrt{n}} + \alpha_p \sqrt{n}}{\sqrt{\frac{(n-1)s^2}{\sigma^2} \frac{1}{n-1}}} \leq \alpha \sqrt{n} \quad (\text{A.8})$$

④(A.8)式の左辺の分子 $X = \frac{\bar{k} - \mu}{\sigma/\sqrt{n}} + \alpha_p \sqrt{n}$ は、A.1 節で述べた \bar{k} の統計的性質より、平均

が $\alpha_p \sqrt{n}$ 、標準偏差が 1 の正規分布に従う。また、(A.8)式の左辺の分母 $Y = \frac{(n-1)s^2}{\sigma^2}$ は

A.1 節で述べた s^2 の統計的性質より、自由度 $(n-1)$ のカイ二乗分布に従う。従って、A.1

節の(3)で述べた性質(A.4)式より、(A.8)式左辺の $X/\sqrt{Y/(n-1)}$ は自由度が $(n-1)$ 、非心パ

ラメータが $\alpha_p \sqrt{n}$ の非心 t 分布に従うことが分かる。よって、 $X/\sqrt{Y/(n-1)}$ の信頼区間

として下側累積確率が $(1-\gamma)$ となる区間 α_γ を、非心 t 分布の上側累積確率 γ に対応する

パーセント点から求める。

$$\frac{X}{\sqrt{Y/(n-1)}} \leq \alpha_\gamma \quad (\text{A.9})$$

⑤(A.8), (A.9)式の右辺の値を比較することで、 $\bar{k} - \alpha s \leq \mu - \alpha_p \sigma$ を満足するためにはパラ

メータ α の値を以下のように設定すれば良いことが分かる。

$$\alpha = \frac{\alpha_\gamma}{\sqrt{n}} \quad (\text{A.10})$$

⑥こうして求められた α を使用して、推定臨界下限増倍率を $k_{sub}(\bar{k}, s) = \bar{k} - \alpha s$ と設定する。

上述した手順に基づいて例題の推定臨界下限増倍率 $k_{sub}(\bar{k}, s)$ を以下で算出してみる。
 $p = 2.5\%$ の場合、 $\alpha_p = 1.96$ となる。また標本数 $n = 10$ の場合、 $X/\sqrt{Y/9}$ は自由度が 9、非心パラメータが 6.198 の非心 t 分布に従うので、下側累積確率 $(1-\gamma) = 97.5\%$ に対応するパーセント点が $\alpha_\gamma = 12.0195$ と求められる。よって、 $k_{sub}(\bar{k}, s)$ の推定で用いるパラメータ α は $\alpha = 12.0195/\sqrt{10} = 3.801$ と得られる。最終的に、推定臨界下限増倍率は $k_{sub}(\bar{k}, s) = 1.0143 - 3.801 \times 0.0209 = 0.9349$ と求めることができる。こうして評価された値は、A.4 節で求めた $k_{sub}(\bar{k}, \sigma)$ の値 0.9627 よりさらに小さな値であり、 \bar{k} だけでなく s の不確かさを見込んだぶんさらに保守的に k_{sub} を設定していると言える。参考までに、10 個で 1 セットの臨界実験の解析を別途 1000 ケース実施したことを想定してランダムサンプリングを行い、各ケースで $k_{sub}(\bar{k}, s)$ を算出した結果を図 A.7 に示す。図 A.7 に示したように、全 1000 ケースのうち 24 ケースが A.2 で求めた「正しい」推定臨界下限増倍率である $k_{sub}(\mu, \sigma) = 0.9708$ よりも大きく設定される結果となり、その頻度は $\gamma = 2.5\%$ にほぼ等し

いことが確認できる。つまり、推定臨界下限増倍率の「設定に失敗する」確率が予測した2.5%とほぼ一致していることが分かる。

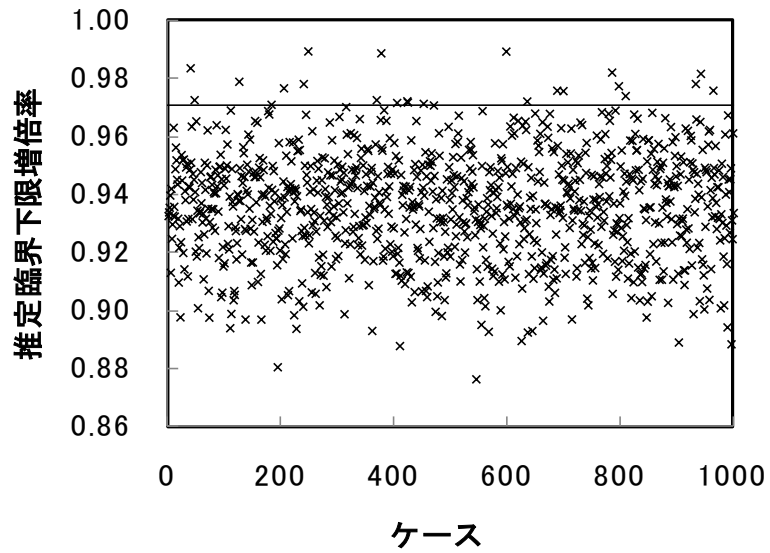


図 A.7 $k_{sub}(\bar{k}, s)$ の散布図(標本数 $n = 10$, $p = \gamma = 2.5\%$)