

1994.8

放射線測定と誤差

石松健二

第1章 統計量

1.測定値

どんな測定でも、同じ条件で測定すれば同じ値が得られそうに思われる。しかし、実際にそうなることば期待できない。その

一例として、同一条件で何度も測定した吸収線量の値を表1に示す。電離箱をファントム中の一定の位置に固定し、一定量の X 線を繰り返し照射して吸収線量を 50 回測定した結果である。測定値はもっとも大きい 127.0cGy ともっとも小さい 122.4cGy の間でばらつき、最大値と最小値には 4.6cGy の食い違いがある。このようなばらつきは、どんな測定にも本質的に存在する。したがって、測定結果から妥当な結論を引き出すためには、測定値のばらつきに対する理解と合理的な取り扱いを欠かすことができない。

表1 同一条件での測定値 単位: cGy

No.	1	2	3	4	5
0	126.0	124.9	123.8	124.7	126.2
5	124.7	125.9	125.0	125.4	123.7
10	124.2	125.7	124.9	125.0	124.9
15	123.0	125.8	123.9	125.3	125.1
20	124.3	123.6	125.8	126.4	125.2
25	125.3	125.3	123.9	125.7	125.8
30	125.5	125.4	125.4	125.3	126.4
35	125.3	125.2	123.0	124.9	125.3
40	122.4	126.3	125.0	125.5	125.1
45	124.6	124.3	127.0	124.4	123.2

2.平均値

いくつかの数値からその平均値を求めるのはありふれたことで、よく知られている。 n 個の測定値 x_1, x_2, \dots, x_n に対して、それらの平均値 \bar{x} は次の式で与えられる。

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} \quad (1)$$

平均値は測定値のばらつきの中心の位置を与える。このように、測定値のばらつきに関連していくつかの値が定義され、それらは統計値と呼ばれる。

3.分散(標準偏差)と不偏分散

ばらつきの中心の位置と共に、ばらつきの広がりも重要な性質である。残差平方和 S は測定値のばらつきの程度を示し、 n 個の測定値 x_1, x_2, \dots, x_n の平均値を \bar{x} とすると次の式で与えられる。

$$S = (x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2 \quad (2)$$

式の形から分かるように、測定値の数 n が大きくなれば S の値も大きくなる。そこで、 n の影響を除くために S を n で割った値 S/n を定義し、それを分散と呼ぶ。

$$\frac{S}{n} = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n} \quad (3)$$

よく知られている標準偏差 σ は、分散の平方根に等しい。

$$\sigma = \sqrt{\frac{S}{n}} = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}} \quad (4)$$

残差平方和 S を $n-1$ で割ったものを不偏分散 V という。

$$V = \frac{S}{n-1} = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n-1} \quad (5)$$

不偏分散も分散も測定値のばらつきの広がりを表すが、不偏分散のほうが合理的な面があり現在では不偏分散を使うことが多い。不偏分散の平方根 $\hat{\sigma}$ もよく使われ、場合によってはこれを標準偏差ということもある。

$$\hat{\sigma} = \sqrt{\frac{S}{n-1}} = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n-1}} \quad (6)$$

この資料では、少々面倒ではあるが σ を不変分散の平方根と呼び、標準偏差 σ と混同しないようにする。

標準偏差または不偏分散の平方根を平均値で割ったものを、変動係数 CV という。

$$CV = \frac{\sigma}{\bar{x}} \quad (\text{または, } CV = \frac{\hat{\sigma}}{\bar{x}}) \quad (7)$$

変動係数は百分率 (%) で表すことが多い。

4. 範囲

測定値のばらつきを示すのに、範囲を使うことがある。範囲 R は測定値 x_i 中の最大のもの $\max(x_i)$ と最小のもの $\min(x_i)$ の差である。

$$R = \max(x_i) - \min(x_i) \quad (8)$$

不偏分散の平方根などと違って、範囲はほとんど計算をしないで求めることができる。この簡便さが範囲の特徴で、測定値のばらつきを手早く見るのには便利な統計値である。しかし後に述べるように、範囲の値は測定値の数によって変化する性質がある。使うときには、この点に注意しなければならない。

5. 組分け

測定値の数が多くなるにつれて、統計値(平均値や不偏分散など)の計算は次第に面倒になってくる。計算の面倒さを少なくするために、従来から測定値の"組分け"が利用されてきた。コンピュータの普及によって計算の面倒さから解放された現在でも、組分けはもう一つの特徴のためによく利用されている。

表1の測定値の大きさを10種類程度に区別してみよう。区別した大きさの範囲を区間と呼び、区間の幅を級間隔と言う。区間の幅(級間隔)を0.5cGyに選べば、表1の全測定値は11の区間に分けられる。各区間に含まれる測定値の数を表2に示す。表中の j は区間の番号、 x_j は区間の測定値の代表値(区間の中心値を取る)、 f_j は区間に含まれる測定値の数(度数)を表す。このような表を度数表という。1番目の区間122.0-122.4cGyの代表値は122.2cGyに選び、その区間に含まれる測定値はすべて122.2cGyと考える。このように仮定すると、測定値の平均値 \bar{x} は次の式で与えられる。

$$\bar{x} = \frac{f_1}{n} x_1 + \frac{f_2}{n} x_2 + \cdots + \frac{f_m}{n} x_m \quad (9)$$

添字 m は区間の総数で、 n より十分に小さい。同様に残差平方和 S は次のようになる。

$$S = (x_1^2 f_1 + x_2^2 f_2 + \cdots + x_m^2 f_m) - n\bar{x}^2 \quad (10)$$

この方法では各測定値を代表値に置き換えて計算するので、正確計算結果(式1と式2)とはわずかに食い違う。しかし食い違いは小さく、実用上問題になることはない。

表2 度数表 (j : 区間番号、 x_j : 代表値、 f_j : 度数)

区間(cGy)	j	x_j (cGy)	f_j	区間(cGy)	j	x_j (cGy)	f_j
122.0 - 122.4	1	122.2	1	125.0 - 125.4	7	125.2	16
122.5 - 122.9	2	122.7	0	125.5 - 125.9	8	125.7	8
123.0 - 123.4	3	123.2	3	126.0 - 126.4	9	126.2	5
123.5 - 123.9	4	123.7	5	126.5 - 126.9	10	126.7	0
124.0 - 124.4	5	124.2	4	127.0 - 127.4	11	127.2	1
124.5 - 124.9	6	124.7	7			計	50

6. 度数分布 (ヒストグラム)

図1は、度数表(表2)をグラフにしたもので、ヒストグラムと呼ばれる。隣接して並んでいる長方形の柱の底辺の中心は区間の代表値(中心の値)と一致し、高さは度数を表している。ヒストグラムの底

辺の端の値は、この場合左から 122.0、122.5、123.0、
…、127.5cGy である。ヒストグラムを見れば、測
定値のばらつき具合を細かくしかも直感的に理解
することができる。

級間隔(区間の幅)は狭い方が良さそうに思われる
が、狭く取り過ぎると各区間の度数が減ってばらつ
きが大きくなり、かえって見づらくなる。データの
総数によるが、区間の数は 10~20 個程度に選ぶの
が常識的である。図の場合、11 区間に組分けして
あるが、データ総数が 50 個あるにもかかわらず、
やはり滑らかな分布をしているとはいえない。デー
タ数がさらに少なくなると、見やすいヒストグラム
を作ることは難しくなる。

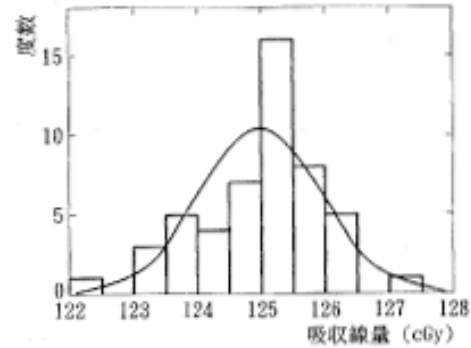


図1 測定値の度数分布

図中の曲線は後で述べる正規分布をデータに与え
はめたものである。正規分布曲線は平均値と標準偏差を与えれば描けるので、データが正規分布に
従うと分かっているならば、多くのデータ数が必要なヒストグラムよりも手軽に描くことができる。

7. m 組のデータ

今までに述べたのは、測定値全部をひとまとまりのデータとして平均値や不変分散の平方根を計
算する話であった。ここではデータが何組かのサブグループに分かれている場合について述べてお
く。

n 個の測定をもつサブデータが m 組あるとする。

$$\begin{aligned}
 &1 \text{ 組: } x_{11}, x_{12}, \dots, x_{1n} \\
 &2 \text{ 組: } x_{21}, x_{22}, \dots, x_{2n} \\
 &\dots \\
 &j \text{ 組: } x_{j1}, x_{j2}, \dots, x_{jn} \\
 &\dots \\
 &m \text{ 組: } x_{m1}, x_{m2}, \dots, x_{mn}
 \end{aligned} \tag{11}$$

任意の j 組の平均値 \bar{x} や残差平方和 S は式 1 と式 2 から次の式で与えられる。

$$\bar{x}_j = \frac{x_{j1} + x_{j2} + \dots + x_{jn}}{n} \tag{12}$$

$$S_j = (x_{j1}^2 + x_{j2}^2 + \dots + x_{jn}^2) - n\bar{x}_j^2 \tag{13}$$

測定全部の平均 \bar{x} と残差平方和 S は、次の式で計算することができる。

$$\bar{x} = \bar{x}_j = \frac{\bar{x}_1 + \bar{x}_2 + \dots + \bar{x}_m}{m} \tag{14}$$

$$S = (S_1 + S_2 + \dots + S_m) + n\{(\bar{x}_1^2 + \bar{x}_2^2 + \dots + \bar{x}_m^2) - m\bar{x}^2\} \tag{15}$$

全部の平均値 \bar{x} は、各組の平均値 \bar{x}_j を平均すればよい (\bar{x}_j は \bar{x}_j の平均値を意味する)。その一方で残
差平方和 S は、各組の残差平方和 S_j を加え合わせるだけでなく、さらに \bar{x}_j^2 を含んだ項を加えてやら
ねばならない。

8. 測定値と誤差

同じ条件で測定した吸収線量でも、測定する度に違った値が得られる。この章では、測定値のば
らつきの様子を表現する方法について述べてきた。では、何故測定値はいつもばらつくのだろうか。
測定値には測定誤差が含まれ、しかもその誤差が測定の度に変動するからである。

(1)偶然誤差:測定の方法をいくら改良しても、ある程度のばらつきが残るのは避けられない。しかし、
測定法が十分に洗練されると、ばらつきは偶然だけに支配されるようになる。このように偶然に支
配されて変動する誤差を偶然誤差ということがある。偶然誤差は確率の法則に従い、統計学を利用

して記述することができる。

(2)系統誤差:電子加速装置を起動して、前記のような吸収線量の測定を繰り返す場合について考えよう。放射線ヘッド中には熱を発生する部分があり、起動後 2~3 時間程度ヘッド中の温度は上昇を続け、その後ほぼ一定温度に達する。放射線ヘッド中にある線量モニタシステムの電離箱も同様な温度変化を受けるので、電離箱が密封形でなければ、線量モニタシステムの感度は温度の影響によって変化する。温度が変化する状態で吸収線量の測定を繰り返すと、測定値も当然温度の影響を受けて少しずつ変化する。これは偶然によるものではなく、線量モニタシステムの感度変化に起因する連続的な変化である。このように原因がはっきりした変化による誤差を系統誤差という。

系統誤差は統計学で記述することばできないものの、測定法や測定装置を洗練させると除去することができる。しかし、残っている系統誤差を偶然誤差と切り分けるためには、統計学に対する若干の知識と測定に対する経験とが要求される。

第2章 確率分布と期待値

9. 連続分布

測定値からいろいろな統計値(平均値や標準偏差など)を求めると、測定値のばらつきを定量化することができる。しかし、これらの統計値を十分に利用するためには、その意味や相互間の関連を知らねばならない。そのためには、測定値や統計値を理想化して、解析的に取り扱えるようにすると都合がよい。

ここでは表2の度数分布を例に挙げて、理想化と解析の手法を紹介する。まず、度数分布の意味を考えよう。任意の測定値が区間 j に含まれる確率 p_j は、

$$p_j = \frac{f_j}{n} \quad (16)$$

で与えられる。 f_j は区間 j に含まれる測定値の数(度数)、 n は測定値の総数である。一方、ヒストグラム(図1)の j 番目の柱の面積 A_j は、式16を考慮すると次のようになる。

$$A_j = (\text{柱の高さ}) \times (\text{区間の幅}) = f_j \times \Delta x = p_j n \Delta x \quad (17)$$

Δx は級間隔(区間の幅)である。つまり、ヒストグラムの柱の面積は測定値がその区間に含まれる確率 p_j に比例し(比例定数は $n\Delta x$)、測定値の確率分布を示すことが分かる。測定値 x_i は測定回数と同数の不連続な値しか存在し得ないが、解析的に取り扱うために連続な確率変数 x に置き換える。 x は連続した値を持つように理想化された測定値と考えることができるので、 x_i と同様に“測定値”と呼ぶことにする。また、 x は測定値 x_i のあらゆる値を取り得るので、測定値 x_i のすべてを含む集団(母集団)が x であるとも考えることもできる。それはまた、“測定”という概念そのものとも考えられる。この考え方によれば、測定値 x_i は測定 x を実行して得られた実現値ということになる。

測定値(または母集団) x の確率分布は、度数分布と違って連続分布になる。連続的な確率分布は、確率密度という関数で表される。測定値 x の確率密度を $f(x)$ とすれば、測定値 x (または x の実現値 x_i) が二つの値 x_1 と x_2 の間の値を取る確率は、

$$\int_{x_1}^{x_2} f(x) dx \quad (18)$$

で与えられる。図2は確率密度の分布の例を示し、式18は図の x_1 と x_2 の間のハッチを施した部分の面積に等しい。

x_1 と x_2 をヒストグラムの区間に一致させると、式18(図のハッチ部分の面積)と f_i/n (ヒストグラムの柱の面積に比例する)は共に測定値がその区間に現れる確率を示す。そこで、ヒストグラムと確率密度の曲線を重ねて描いてやれば、両方を直接比較することができる。図1中の滑らかな曲線は、確率密度 $f(x)$ に $n\Delta x$ を掛けた値の分布を示す。確率密度に $n\Delta x$ を掛けたのは、縦軸を調

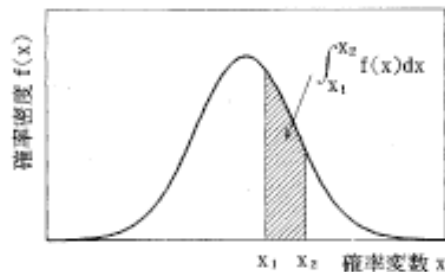


図2 確率と面積

整して度数分布と直接比較できるようにするためである(式 17 参照)。確率密度(確率分布)は度数分布を理想化したものといえることができる。

1 0. 正規分布 (確率分布)

確率分布の内でもっとも重要なのは正規分布で、大部分の測定値は正規分布に従うと考えてよい。正規分布は、平均 μ と標準偏差 σ が与えられると、確率密度の形(確率分布)を正確にいい表すことができる。正規分布に従う測定値 x の確率密度 $f(x)$ は、

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \quad (19)$$

で与えられる。 $f(x)$ は図 3 のような分布をもつ。図から分かるように、測定値 x の実現値は平均 μ と同じ値になる可能性(確率密度)がもっとも大きく、実現値が μ から離れるほど現れる可能性は小さくなる。

測定値 x の実現値が $\mu \pm \sigma$ の区間内の値を取る確率は式 18 で計算することができ、図の $\mu - \sigma$ と $\mu + \sigma$ を通る二つの縦線に囲まれた分布曲線の下での面積に等しい。実際に計算すると、その確率は μ や σ の値に関係なく 0.683 が得られる。たとえば $\mu = 125.0 \text{ cGy}$ 、 $\sigma = 1.0 \text{ cGy}$ とすれば、測定値が 124.0 cGy と 126.0 cGy の間の値を取る確率は 0.683 になる。もし $\sigma = 2.0 \text{ cGy}$ なら、 123.0 cGy と 127.0 cGy の間の値を取る確率が同じく 0.683 になる。つまり、標準偏差 σ が大きくなると、平均 μ から離れた値が相対的に現れやすくなることが分かる。

測定値 x の実現値が $\mu \pm k\sigma$ の区間内にある確率 p を同様に計算すると、表 3 に示す値が得られる。この表によると、測定値が平均から標準偏差の 2 倍以内($\mu \pm 2\sigma$ つまり $k=2$)にある確率は 0.955(約 95%)、2.5 倍($k=2.5$)以内にある確率は 0.988(約 99%)になる。

表 3 区間 $\mu \pm k\sigma$ を取る確率 p

k	0.5	1.0	1.5	2.0	2.5	3.0
p	0.333	0.683	0.867	0.955	0.988	0.997

1 1. 正規分布の描き方

正規分布をグラフに描くには、統計の教科書等に記載してある正規分布表を利用する。正規分布表に示してあるのは、

$$\phi(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \quad (20)$$

であって、 $t=0 \sim 4$ 程度の範囲で $\phi(t)$ の値が与えてある。 $\phi(t)$ は式 19 の μ をゼロ、 σ を 1 にした式で、 t と $\phi(t)$ から x と $f(x)$ の値を求めるには、次の換算をすればよい。

$$x = \mu + \sigma t, \quad f(x) = \frac{\phi(t)}{\sigma} \quad (21)$$

表(適当な教科書を見ていただきたい)から、 $t=0.00$ に対して $\phi(t)=0.3989$ が得られる。たとえば $\mu = 125.0 \text{ cGy}$ 、 $\sigma = 1.0 \text{ cGy}$ なら、式 21 によって、

$$x = \mu + \sigma t = 125.0 + 1.0 \times 0.00 = 125.0 [\text{cGy}]$$

$$f(x) = \frac{\phi(t)}{\sigma} = \frac{0.3989}{1.0} = 0.3989 [\text{cGy}^{-1}]$$

が得られる。同様に、 t のいろいろな値に対する $\phi(t)$ を読み取り、それを 0 と $f(x)$ 換算しグラフに描けばよい。

1 2. 期待値

測定値 x_i の平均値は式 1 または式 9 で求められる。確率変数 x に対応するように式 9 を修正すると、それは平均値の理想化といえる。得られる値を確率変数 x の期待値 $E(x)$ と呼び、 x の平均 μ に等しい。また、 $D(x) = \sqrt{E\{(x-\mu)^2\}}$ を x の標準偏差と呼び、 x の標準偏差 σ に等しい(これらの具体的な式は省略する)。

$$\text{平均: } \mu = E(x) \quad (22)$$

$$\text{標準偏差: } \sigma = D(x) = \sqrt{E\{(x - \mu)^2\}} \quad (23)$$

測定値と同様に、統計値 s (平均値 \bar{x} 、不変分散の平方根 $\hat{\sigma}$ 、範囲 R 等) も期待値 $E(s)$ や標準偏差 $D(s)$ を計算することができる。これらの統計値 s について計算した $E(s)$ と $D(s)$ を以下に示す。測定値 x は正規分布(平均 μ 、標準偏差 σ) をもつとする。

(1) 測定値の平均値 \bar{x} :

$$E(\bar{x}) = \mu \quad (24)$$

$$D(\bar{x}) = \frac{\sigma}{\sqrt{n}} \quad (25)$$

平均 μ や標準偏差 σ は直接測定できないので、式 24 を根拠にして、求めた平均値 \bar{x} を μ に等しいと推定するのである。平均値 \bar{x} を繰り返し求めると、その度に値はばらつく。式 25 は \bar{x} のばらつき(標準偏差)を与える。これは $\mu = \bar{x}$ と推定したときの誤差に当たる。

式 25 によると、平均値 \bar{x} の標準偏差 $D(\bar{x})$ は σ/\sqrt{n} である。 $D(\bar{x})$ の値は、測定の回数 n が大きくなると小さくなる。つまり、測定回数を増やすほど平均値 \bar{x} の誤差は減り、平均 μ からのずれが小さくなることが分かる。

(2) 不変分散の平方根 $\hat{\sigma}$:

$$E(\hat{\sigma}) = c\sigma, \quad E\left(\frac{\hat{\sigma}}{c}\right) = \sigma \quad (26)$$

$$D(\hat{\sigma}) = c_1\sigma, \quad c_1 = \sqrt{1 - c^2} \quad (27)$$

測定値 x_i の不変分散の平方根 $\hat{\sigma}$ を定数 c で割ったものは、測定値 x の標準偏差 σ に等しいと推定される。定数 c の値は測定値の数 n によって決まり、表 4 に示してある。 n が大きいと $c=1 - 1/4n$ と近似することができ、 $n=\infty$ では $c=1$ になる。 $D(\hat{\sigma})$ に関する定数 c_1 の値も表 4 に示してあり、 n が大きいと $c_1=1/\sqrt{2n}$ と近似され、 $n=\infty$ では $c_1=0$ になる。

このように、測定値 x_i の不変分散の平方根 $\hat{\sigma}$ から測定値 x の標準偏差 σ が推定される。しかし値は 1 に近いので、 $\hat{\sigma}$ をそのまま σ として使っても大きな誤差は生じない。

表 4 c の値

n	2	3	4	5	6	7	8	9	10	$n \rightarrow \infty$
c	0.798	0.886	0.921	0.940	0.952	0.959	0.965	0.969	0.97	$1 - 1/4n$
c_1	0.603	0.463	0.389	0.341	0.308	0.282	0.262	0.246	0.232	$1/\sqrt{2n}$

(3) 範囲 R :

$$E(R) = d\sigma, \quad E(R/d) = \sigma \quad (28)$$

$$D(R) = d_1\sigma \quad (29)$$

範囲を定数 d で割った値の期待値は、測定値 x の標準偏差 σ に等しい。範囲 R から x の標準偏差 σ を推定できることが分かる。定数 d や d_1 ($D(R)$ に関する定数) の値は測定値の数 n によって決まり、表 5 に示してある。

表 5 d の値

n	2	3	4	5	6	7	8	9	10
d	1.128	1.693	2.059	2.326	2.534	2.704	2.847	2.970	3.078
d_1	0.853	0.888	0.880	0.864	0.848	0.833	0.820	0.808	0.797

測定値数 n が大きいと、 R は $\hat{\sigma}$ に比べて余分にばらついて $\hat{\sigma}$ より不利になる。そのときには、全データを測定値 3~5 個を含むサブデータ (m 個) に分け、サブデータ毎に範囲 R_i を求めてその平均値 \bar{R} から σ を推定すればよい。 \bar{R} やその期待値等は次のようになる。

$$\bar{R} = \frac{R_1 + R_2 + \dots + R_m}{m}$$

$$E(\bar{R}) = d\sigma, \quad E(\bar{R}/d) = \sigma, \quad D(\bar{R}) = \frac{1}{\sqrt{m}} d_1 \sigma$$

測定値を $n=4$ 個ずつに分けると $d \approx 2$ になり、求めた R または \bar{R} を 2 で割れば標準偏差 σ が求められて都合がよい。

第3章 測定値のばらつき

1.3. 統計値 (μ と σ の推定)

測定値 x_i がどの程度ばらつくか、表 1 のデータで確かめてみよう。表中の測定値は小さい方の 2 桁にばらつきが認められる。50 個の測定値の内、最大値 $\max(x_i)$ は 48 番目の 127.0cGy、最小値 $\min(x_i)$ は 41 番目の 122.4cGy である。このデータから統計値を何個か計算すると、次のようになる。

(1) 平均値：式 1 から、

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{126.0 + 124.9 + \dots + 123.2}{50} = 125.0 \text{ [cGy]} \quad (30)$$

(2) 不変分散の平方根：式 6 から、

$$\begin{aligned} \hat{\sigma} &= \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n-1}} \\ &= \sqrt{\frac{(126.0 - 125.0)^2 + (124.9 - 125.0)^2 + \dots + (123.2 - 125.0)^2}{50-1}} \\ &= 1.0 \text{ [cGy]} \end{aligned} \quad (31)$$

(3) 範囲：式 8 から、

$$R = \max(x_i) - \min(x_i) = 127.0 - 122.4 = 4.6 \text{ [cGy]} \quad (32)$$

測定値 x_i から直接求めたこれらの値から、 x の平均 μ と標準偏差 σ を推定する。まず、 \bar{x} の期待値 $E(\bar{x})$ は x の平均 μ に等しい(式 24)ので、 μ の値を推定することができる。

$$\mu = E(\bar{x}) \approx \bar{x} = 125.0 \text{ [cGy]}$$

また、表 4 から $n=50$ に対して $C \approx 1 - 1/4n = 1 - 1/(4 \times 50) = 0.995$ を求め、式 26 から

$$\sigma = E\left(\frac{\hat{\sigma}}{c}\right) \approx \frac{\hat{\sigma}}{c} = \frac{1.0}{0.995} \approx 1.0 \text{ [cGy]} \quad (34)$$

が得られる。全部の測定値を使って求めたこの μ と σ の値は、もっとも確からしい値といえる。そこで、これらの値をまとめて次のように書いておこう。

$$\mu = 125.0 \text{ cGy}, \quad \sigma = 1.0 \text{ cGy} \quad (35)$$

また \bar{x} と $\hat{\sigma}/c$ の標準偏差 $D(\bar{x})$ 、 $D(\hat{\sigma}/c)$ は、式 25、式 26、式 27 と式 35 から

$$D(\bar{x}) = \frac{\sigma}{\sqrt{n}} \approx \frac{1.0}{\sqrt{50}} = 0.14 \text{ [cGy]} \quad (36)$$

$$c_1 = \sqrt{1 - c^2} = \sqrt{1 - 0.995^2} = 0.0999$$

$$D(\hat{\sigma}/c) \approx \frac{c_1 \sigma}{c} \approx \frac{0.0999 \times 1.0}{0.995} = 0.10 \text{ [cGy]} \quad (37)$$

になる。これらは、 \bar{x} と $\hat{\sigma}/c$ したがって式 35 の μ と σ の誤差を示すものと考えてよい。

式 24 によって、平均値 \bar{x} が μ に等しいと推定することができた。その点 \bar{x} は x と同様に正規分布をもち、 μ の値が $\bar{x} \pm D(\hat{\sigma}/c) = 125.0 \pm 0.14 \text{ cGy}$ の間にある確率は 0.683 である。一方、式 22 によって測定値 x の期待値も μ であり、一個の測定値 x_i を μ と推定することもできる。この場合には、 μ が $x_i \pm \sigma = x_i \pm 1.0 \text{ cGy}$ の間にある確率が 0.683 になる。同じ μ を推定するのにも、推定の誤差(標準偏差)は \bar{x} の方が 0.14cGy、 x_i の方が 1.0cGy と大きな違いがある。 μ の推定には、 \bar{x} の方がはるかに優れていることが分かる。

不変分散の平方根 $\hat{\sigma}$ の分布は正規分布とは違うことが分かっており、 σ が $\hat{\sigma}/c \pm D(\hat{\sigma}/c) = 1.0 \pm 0.10$ cGy の間にある確率は 0.683 にはならない。つまり、 $D(\bar{x})$ と $D(\hat{\sigma})$ は分布の違いによって性質がいくらか異なり、直接的に比較するのはあまり意味がない。

1.4. 統計値のばらつき

平均値 \bar{x} やその他の統計値のばらつき(誤差)は、測定値 x_i の数 n によって影響される。そこで、元になる測定値の数が分かるように、 n 個の測定値から求めた統計値 s の期待値、標準偏差を $E(s, n)$ 、 $D(s, n)$ 、また統計値自体も $\bar{x}(n)$ 、 $\hat{\sigma}(n)$ 、 $R(n)$ などと標記することにしよう。

表 1 のデータから 1 行分の測定値 5 個(たとえば No.1~5 の値)を選んで、一組とする。そうすると、全データを 10 個の組に分けることができる。各組ごとに平均値 $\bar{x}(5)$ 、不変分散の平方根 $\hat{\sigma}(5)$ 、範囲 $R(5)$ 、平均値 $\bar{x}(5)$ の標準偏差 $D(\bar{x}, 5)$ を求めると、表 6 に示す値が得られる。これらの統計値の値やその変動について調べてみよう。

(1) 平均値 $\bar{x}(5)$: 式 1 で計算した各組 5 個測定値の平均値 $\bar{x}(5)$ は、表 6 の $\bar{x}(5)$ の欄に示してある 0 それらの中の最大値は 125.6cGy、最小値は 124.6cGy である。 \bar{x} の期待値 μ のもっとも確からしい値は式 35 に示してある 125.0cGy で、この値に対して $\bar{x}(5)$ の値はすべて ± 0.6 cGy の範囲内にある。

このばらつきは妥当なものだろうか。 x のもっとも確からしい標準偏差 σ は式 35 から 1.0cGy、また $n=5$ である。これらの値と式 25 から、 $\bar{x}(5)$ の標準偏差

$$D(\bar{x}(5)) = \frac{\sigma}{\sqrt{n}} = \frac{1.0}{\sqrt{5}} = 0.45 \text{ [cGy]}$$

が得られる。表 3 によると、確率変数 x の値が標準偏差 σ の ± 2 倍範囲内にある確率は約 95% になる。したがって大部分の $\bar{x}(5)$ の値は、 $\mu = 125.0$ cGy を中心に $\pm 2 \times D(\bar{x}(5)) = \pm 0.9$ cGy の範囲内に収まる筈である。現実には全部が ± 0.6 cGy の中にあり、この程度のばらつきは十分に妥当といえる。

(2) 不変分散の平方根 $\hat{\sigma}(5)$: 式 6 で計算した各 5 個の測定値の $\hat{\sigma}(5)$ の値は、表 6 の $\hat{\sigma}(5)$ の欄に示すように 1.5cGy と 0.5cGy の間に分布している。

$\hat{\sigma}$ から x の標準偏差 σ を推定することができる。もっとも確からしい σ は、上記のように 1.0cGy である。これに対して個々の $\hat{\sigma}(5)$ から σ を推定すると、 $n=5$ の場合について σ の推定値のばらつきを実際に見ることができる。表 4 から $n=5$ に対して $c = 0.940$ を求め、その値を式 26 に入れると

$$\sigma = E\left(\frac{\hat{\sigma}(5)}{c}\right) \approx \frac{\hat{\sigma}(5)}{c} = \frac{\hat{\sigma}(5)}{0.940} \quad (38)$$

が得られる。表 6 の $\hat{\sigma}(5)$ の値からこの式で求めた σ の値が、表 7 の σ_s の欄に示してある。 σ_s の値は 1.6cGy から 0.5cGy の間でばらつき、 $\sigma = 1.0$ cGy を中心に ± 0.6 cGy の範囲内にある。この値を相対値に直すと、 $\pm 60\%$ になる。一方、上式によると $\sigma = \hat{\sigma}(5)/0.940 = 1.06 \hat{\sigma}(5)$ であり、 σ と $\hat{\sigma}(5)$ の違いは 0.06 倍(6%)に過ぎない。 $\hat{\sigma}$ をそのまま σ として扱って差し支えないことがよく分かる。

(3) 範囲 $R(5)$: 範囲 $E(5)$ の値は同じ組の $\hat{\sigma}(5)$ にほとんど比例しているが、それぞれの最大値と最小値の比は

$$\frac{\max(R(5))}{\min(R(5))} = \frac{3.9}{1.1} = 3.5, \quad \frac{\max(\hat{\sigma}(5))}{\min(\hat{\sigma}(5))} = \frac{1.5}{0.5} = 3.0$$

となり、ばらつきの程度(誤差)は範囲の方が少し多い。前に述べたように、測定値の数 n が大きいと範囲のばらつきは多くなる傾向があるので、範囲を求めるときには $n=4\sim 5$ に分割したサブデータを使った方がよい。また、 x の標準偏差は範囲からも推定することができる。表 5 から $n=5$ に対して $d=2.326$

表 6 5 個の測定値に対する統計量

No.	$\bar{x}(5)$	$D(\bar{x}, 5)$	$\hat{\sigma}(5)$	$R(5)$
1 - 5	125.1	0.44	1.0	2.4
6 - 10	124.9	0.37	0.8	2.2
11 - 15	124.9	0.24	0.5	1.5
16 - 20	124.6	0.51	1.1	2.8
21 - 25	125.1	0.50	1.1	2.8
26 - 30	125.2	0.34	0.8	1.9
31 - 35	125.6	0.20	0.5	1.1
36 - 40	124.7	0.43	1.0	2.3
41 - 45	124.9	0.66	1.5	3.9
46 - 50	124.7	0.63	1.4	3.8

を求め、式 28 に入れてやれば、

$$\sigma = E\left(\frac{R(5)}{d}\right) \approx \frac{R(5)}{d} = \frac{R(5)}{2.326} \quad (39)$$

が得られる。表 6 の $R(5)$ の値からこの式で σ を求めると、表 7 の σ_R の欄の値が得られる。不変分散の平方根 $\hat{\sigma}$ と範囲定から求めた x の標準偏差 σ_s と σ_R の値は、よく一致している。実用上どちらを使ってもよいが、原則的に精度のよい推定はどちらなのか検討しよう。

不変分散の平方根 $\hat{\sigma}(5)$ の期待値に関して式 26 と式 27 から次の式が得られる。

$$E\left(\frac{\hat{\sigma}(5)}{c}\right) = \sigma, \quad D\left(\frac{\hat{\sigma}(5)}{c}\right) = \frac{c_1 \sigma}{c}$$

もつとも確からしい σ は式 35 から 1.0cGy であり、また表 4 から $c=0.940$ 、 $c_1=0.341$ が得られ、 $\sigma_s \doteq \hat{\sigma}(5)/c$ の標準偏差は次のようになる。

$$D\left(\frac{\hat{\sigma}(5)}{c}\right) = \frac{0.341 \times 1.0}{0.940} = 0.36 \text{ [cGy]}$$

同様に、式 28 と式 29 から $\sigma_R \doteq R(5)/d$ の標準偏差を求めると、次のようになる。

$$D\left(\frac{R(5)}{d}\right) = 0.37 \text{ [cGy]}$$

不変分散の平方根 $\hat{\sigma}(5)$ から求めた標準偏差 σ_s の標準偏差(ばらつき)は 0.36cGy、範囲 $R(5)$ から求めた σ_R の標準偏差は 0.37cGy でほとんど一致している。さらに、 $n=10$ の場合について計算すると、 σ_s の標準偏差は 0.24cGy、 σ_R の標準偏差は 0.26cGy と範囲から求めた方がやや不利になる。

15. 1 回の測定と n 回の測定 測定の統計的な側面を考えよう。測定というのは、測定値すべてを含む母集団 x からでたらめに 1 個の測定値を取り出す(抽出する)作業と考えてよい。われわれが知りたいのは、測定値そのものよりも母集団の性質である。つまり、母集団(測定値) x の平均 μ と標準偏差 σ が重要になる。しかし、われわれには抽担された測定値しか見ることができない。これまでに述べた統計的な手法の目的は、測定値から母集団の平均と標準偏差を推定することに尽きる。

(1) 1 回測定と複数回測定 x の測定を 1 回実行して、1 個の測定値 x_0 を得たとする。この場合には、その測定値 x_0 を母集団の平均 μ に等しいと推定することになる。式 22 に示すように x の期待値が平均 μ だからである。しかし+測定を 1 回実行しただけでは、式 23 の x と μ の両方に x_0 を当てなければならぬので、同式の $x - \mu$ の項は意義がなくなり、標準偏差 σ を推定することはできない。ただし、測定によっては(たとえば放射線をカウント数で測定する場合)、1 個の測定値から標準偏差を推定できる場合もある。しかし、これは特殊な例で一般的ではない。

一方、 n 回の測定を実行し、その結果 n 個の測定値 x_1, x_2, \dots, x_n を得たとする。この場合にはもちろん、平均値 \bar{x} と不変分散の平方根 $\hat{\sigma}$ を計算することができる。そして、式 24 と式 26 によって μ と σ が得られる。母集団 x の性質を的確に推定するには複数回の測定が不可欠で、得られる情報は 1 回測定の場合の此ではない。

(2) 測定回数の効果 複数回の測定をするとき、測定回数 n のもつ意味について検討しよう。平均値 \bar{x} や不変分散の平方根 $\hat{\sigma}$ は、測定回数 n が小さくても大きくても式 1 や式 6 から確定的に求めることができる。したがって、それらの値から x の平均 μ と標準偏差 σ もまた確定的に求めることができる(式 24、式 26)。問題になるのは、得られた \bar{x} や $\hat{\sigma}$ の誤差(標準偏差)である。 \bar{x} の標準偏差は次の式で表される(式 25)。

$$D(\bar{x}) = \frac{\sigma}{\sqrt{n}}$$

つまり、 \bar{x} の標準偏差は測定回数 n の平方根に反比例すると分かる。 n が無限大あれば \bar{x} の標準偏

表 7 σ の推定

No.	σ_s	σ_R
1 - 5	1.1	1.0
6 - 10	0.9	0.9
11 - 15	0.6	0.6
16 - 20	1.2	1.2
21 - 25	1.2	1.2
26 - 30	0.8	0.8
31 - 35	0.5	0.5
36 - 40	1.1	1.0
41 - 45	1.6	1.7
46 - 50	1.5	1.6

σ_s : $\hat{\sigma}(5)$ から計算

σ_R : $R(5)$ から計算

差はゼロになる。 $n=\infty$ は母集団に含まれる測定値を全部抽出することを意味するので、 \bar{x} が母集団の平均 μ と完全に一致するのは当然である(ただし、実現はできない)。 n が小さくなって1に近づけば、 \bar{x} の標準偏差は x の標準偏差 σ (測定値 x_i の誤差)に近づく。測定回数 n を4回にすると、平均値 \bar{x} の誤差は1回の測定値 x_i の誤差の1/2ということになる。

一方、不変分散の平方根 $\hat{\sigma}$ の標準偏差は次の式で表される(式27)。

$$D(\hat{\sigma}) = c_1 \sigma$$

表4に示してあるように、定数 c_1 の値は測定回数 n が増加するにつれて小さくなる。したがって $\hat{\sigma}$ の標準偏差 $D(\hat{\sigma})$ も、 n が増えるにつれて減少する。以上のように、測定回数が多くなると平均値 \bar{x} も不変分散 $\hat{\sigma}$ も誤差が減少し、これらの値から推定される測定値(母集団) x の平均 μ と標準偏差 σ の確からしきは向上する。

第4章 測定データの解析

16. 回帰分析 これまでは、同じ条件で測定を繰り返して多数の測定値が得られたときの取り扱いについて考えてきた。そして、測定値にはばらつき(誤差)が避けられず、測定データを合理的に理解するには統計的な考え方が必要ことが分かった。

われわれが普通に行う実験では、同じ条件の測定を何度も繰り返すことはむしろ少ない。例えばTPRの測定について考えれば、ファントム表面からの探さを変えながら吸収線量を測定する。ファントム表面からの探さという変数は測定毎に変えられて、しかも測定結果には誤差が含まれる。このようなデータからできるだけ確からしい情報を引き出すために、回帰分析と呼ばれる手法が利用される。

(1) 最小二乗法による回帰分析 測定毎に値が変わる(われわれが値を決められる)変数 x_i と測定結果 y_i との間には、本来備わった関係がある。ここでは x_i と y_i の関係は次の式で完全に表現されると仮定する。

$$y_i = \alpha + \beta x_i + \varepsilon_i \quad (40)$$

α と β は未知の定数、 ε_i は y_i に含まれる誤差(ばらつき)で、正規分布(平均 $\mu=0$ 、標準偏差 σ)に従うとする。回帰分析とは、測定で得られた x_i と y_i を元にして定数 α や β の値を推定する手法をいう。

式40を ε_i について解くと、

$$\varepsilon_i = y_i - \alpha - \beta x_i \quad (41)$$

が得られる。 ε_i は誤差であるが、この式の α 、 β をその推定値 a 、 b で置き換えたもの $y_i - a - b x_i$ を残差と呼ぶ。そして残差の二乗の和

$$S_E = (y_1 - a - b x_1)^2 + (y_2 - a - b x_2)^2 + \cdots + (y_n - a - b x_n)^2 \quad (42)$$

が最小になるように推定値 a 、 b を決めるのが最小二乗法の原則である。

(2) 一次式への回帰分析 上の例では、誤差 ε_i がゼロならば x と y の関係は次のようになる。

$$y_i = \alpha + \beta x_i \quad (43)$$

このように x_i と y_i の関係が次式で表される場合には、 α と β の推定値 a 、 b を求めるのは比較的容易で次の手順で求めることができる。

$$x_S = x_1 + x_2 + \cdots + x_n, \quad y_S = y_1 + y_2 + \cdots + y_n \quad (44)$$

$$x_X = x_1^2 + x_2^2 + \cdots + x_n^2, \quad x_Y = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \quad (45)$$

$$\bar{x} = \frac{x_S}{n}, \quad \bar{y} = \frac{y_S}{n}, \quad S_X = x_X - \bar{x} x_S, \quad S_{XY} = x_Y - \bar{y} x_S \quad (46)$$

始めに測定データから式44を計算し、次にそれらから式45を計算し、最後に式46から a 、 b を求める。ほとんどの関数電卓は、この手順で一次式への回帰分析ができるようになっている。電卓に x_i と y_i を入れるだけで、 a と b が自動的に計算される。

(3) 相関係数 関数電卓で回帰分析を実行すると、相関係数も自動的に計算される。相関係数 r は次の式で与えられる。

$$r = \frac{S_{XY}}{\sqrt{S_X S_Y}} \quad (47)$$

S_{XY} と S_X は式 45 に与えてあり、 S_Y は次のとおりである。

$$y_Y = y_1^2 + y_2^2 + \cdots + y_n^2$$

$$S_Y = y_Y - \bar{y}y_s \quad (48)$$

相関係数は x と y がどの程度強く関連しているかを示す指標で、ゼロと 1 との間の値を取る。 x と y が全く無関係であれば相関係数はゼロになり、完全に対応していれば(式 40 の ε_i がゼロつまり誤差が全く無ければ)1 になる。したがって、 x と y の関連が強いほど(誤差が小さいほど)1 に近い値を取る。相関係数 r の二乗 r^2 を決定係数といい、次のような使い方をすることがある。たとえば、決定係数が 0.9 であれば、 y の値の 90% までが x によって決定されると考えるのである。

(4) 多項式の回帰分析 何か実験をするときと y 関係がはっきりしているとは限らない。むしろ、その関係を探するために実験を計画することが多い。回帰分析をするには x と y の関係式を与える必要があるので、 x と y の関係が判らない実験結果を回帰分析するには、 x と y の関係式を推定する工夫をしなければならない。

x と y の関係を理論的に導き出すのは、もっとも基本的なやり方である。たとえ粗い理論であっても、補正係数を導入して実験的にその値を決めてやれば、実験結果とよく一致する可能性がある。しかし一方では、実行が難しい一面もある。

上記とは別に、いろいろな関係をうまく表現できるような式を見つけて、その式で回帰分析をするやり方がある。この場合には、次のような多項式を使うことが多い。

$$y = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m \quad (49)$$

実験データから、この式の未知の定数 a_0 、 a_1 、 a_2 、 \cdots 、 a_m を求めるのである。実験データ数 (一对の x_i と y_i の数) n は、次数 m より 2 個以上多い必要がある。多項式への回帰分析を手で計算するのは難しく、普通はコンピュータを利用する。プログラムに少し手を加えれば、次の一連の式についても同じプログラムで回帰分析を実行できる。

$$y = b_0 + b_1 \log(x) + b_2 \{\log(x)\}^2 + \cdots + b_m \{\log(x)\}^m \quad (50)$$

$$\log(y) = c_0 + c_1 x + c_2 x^2 + \cdots + c_m x^m \quad (51)$$

$$\log(y) = d_0 + d_1 \log(x) + d_2 \{\log(x)\}^2 + \cdots + d_m \{\log(x)\}^m \quad (52)$$

これら 4 個の式のどれかを使うと、ほとんどの実験データをうまく処理することができる。この場合に注意しなければならないのは、どの形式の式が実験データとよく一致するかを見極めることである。回帰分析で式が確定すると(定数の値が決まると)、その式から x に対する y の値が求められ、それらは滑らかな曲線で表すことができる。実験データはその滑らかな曲線の回りにデタラメに分布する筈で、両方に何か系統的なずれが認められれば式の選択が不適當であったことになる。

また、多項式の次数 m は大きいほうがよいとは限らない。大き過ぎると x と y を表す曲線が細かく波打つ(測定値の誤差を忠実に表現する)。誤差まで表現するのは無意味で、実験データと系統的な食い違いが認められない範囲内で次数 m は小さいほうがよい。

これらの検討は、測定データ(点)と回帰式(曲線)とを同じグラフ上に描くことによって、容易に行うことができる。コンピュータプログラムでは測定データと回帰式との食い違いを示す表またはグラフが表示されるのが普通で、選んだ回帰式の形式の妥当さをそれらによって十分に検討しなければならない。

文献 1 と文献 2 は、多項式への回帰分析を TPR の計算に利用した例である。

17. 誤差の伝播 一つの変数 x の値を決めると他の変数 y の値も決まる、といった現象は身の回りにあふれている。たとえば、ファントム表面からの深さとその点での TPR の関係がそうである。もう少し複雑な関係もある。吸収線量を求めるには、電離箱で測定した電離量に吸収線量変換係数を掛けてやらねばならない。ところが吸収線量変換係数の値は、放射線のエネルギーを与えないと決まらない。つまり、放射線エネルギーに誤差があると、吸収線量変換係数に誤差が生じる。その吸収線量変換係数の誤差はまた、吸収線量に誤差を生じさせる。そのうえ、電離量測定の誤差もま

た、吸収線量に誤差を生じさせる。

以上のように測定値を計算処理すると、元の測定値や計算に使う定数の誤差に応じて計算結果にも誤差が生じる。この現象を誤差の伝播という。一般に、独立変数(互いに無関係に値を決められる変数) x_1, x_2, \dots, x_k の任意の関数 $y=f(x_1, x_2, \dots, x_k)$ を考えると、 x_1, x_2, \dots, x_k の誤差 $\sigma_1, \sigma_2, \dots, \sigma_k$ によって生じる y の誤差 σ_y は次の式で与えられる。

$$\sigma_y^2 = \left(\frac{\partial y}{\partial x_1}\right)^2 \sigma_1^2 + \left(\frac{\partial y}{\partial x_2}\right)^2 \sigma_2^2 + \dots + \left(\frac{\partial y}{\partial x_k}\right)^2 \sigma_k^2 \quad (53)$$

$\partial y / \partial x_i$ は、 x_i 以外の独立変数を定数と考えて y を x_i について微分することを意味する。

上の計算を二三の簡単な関係について実行すると、次のような結果が得られる。

$y = a_1 x_1 \pm a_2 x_2 \pm \dots \pm a_k x_k$ に対して

$$\sigma_y = \sqrt{(a_1 \sigma_1)^2 + (a_2 \sigma_2)^2 + \dots + (a_k \sigma_k)^2} \quad (54)$$

$y = a x_1^{\pm 1} x_2^{\pm 1} \dots \pm x_k^{\pm 1}$ に対して

$$\sigma_y = y \sqrt{\frac{\sigma_1^2}{x_1^2} + \frac{\sigma_2^2}{x_2^2} + \dots + \frac{\sigma_k^2}{x_k^2}} \quad (55)$$

誤差の伝播の計算を実際に適用した例には、文献3や文献4ある。

18. 感度解析 式53の右辺各項には、 x_i に関する y の偏微分係数(偏導関数)が含まれている。

それぞれの変微分係数を y に対する x_i の感度 s_i という。

$$s_i = \frac{\partial y}{\partial x_i} \quad (56)$$

x_i にそれぞれ誤差 σ_i があるとき、であった。関数 $y=f(x_1, x_2, \dots, x_k)$ に生じる誤差 σ_y の値を与えるが式53式53に式56を入れると、次の式が得られる。

$$\sigma_y = \sqrt{(s_1 \sigma_1)^2 + (s_2 \sigma_2)^2 + \dots + (s_k \sigma_k)^2} \quad (57)$$

この式が示すように、 $s_i \sigma_i$ は x_i の誤差 σ_i が y の誤差 σ_y に寄与する割合と考えることができる。

x_i の誤差 σ_i を小さくすれば、結果として y に生じる誤差 σ_y も小さくなる。しかし σ_i を小さくするには、優れた実験装置を使ったり手数を掛けたり、費用や人手が必要になる。それでは、費用や人手を最小限に押さえて σ_y を最小にするには、どうすればよいのだろうか。結果だけを示すと、次のようにしなければならない。

$$s_1 \sigma_1 = s_2 \sigma_2 = \dots = s_k \sigma_k \quad (58)$$

つまり、感度の高い x_i の誤差 σ_i を小さくすればよく。感度の低い x_j の誤差 σ_j 少々大きくても影響は少ない。一番効果的なのは、感度のもっとも高い測定値の誤差を押さえる努力で、感度の低い測定値の誤差を改善してもほとんど意味がない。

数値を挙げて説明しよう。二つの独立した測定値があり、一つの測定値 x_1 では $s_1=2$ 、 $\sigma_1=1$ 、他の x_2 は $s_2=1$ 、 $\sigma_2=1$ であるとする。 y の誤差 σ_y に対する寄与はそれぞれ $s_1 \sigma_1=2$ 、 $s_2 \sigma_2=1$ で、 x_1 の寄与の方が大きい。両方の誤差による y の誤差は、

$$\sigma_y = \sqrt{(s_1 \sigma_1)^2 + (s_2 \sigma_2)^2} = \sqrt{2^2 + 1^2} = 2.2$$

になる。その一方で、もし x_2 の誤差 σ_2 をゼロしたとすると、 y の誤差は

$$\sigma_y = s_1 \sigma_1 = 2$$

になる。前に述べたように、誤差 σ の標準偏差 $D(\sigma)$ (誤差の値のばらつき)は大きいのが普通である。そのために、 x_1 と x_2 の誤差の影響を含む $\sigma_y=2.2$ は、 x_2 の誤差を無くしたときの $\sigma_y=2$ に比べて違いくちがあるとはいえない。つまり、 y の誤差に影響するのはほとんど x_1 の誤差だけで、 x_2 の誤差をゼロにする努力は無駄である。逆に x_1 の誤差 σ_1 をゼロにすることができれば、 y の誤差は $\sigma_y=s_2 \sigma_2=1$ になる。感度の大きい x_1 の誤差を改善すると、このように大きな効果を期待することができる。

上に述べたように、いくつかの変数(測定値)が関与する量を求めるときには、それぞれ変数感度を知ることが効率的な実験をするのに重要である。このように各変数の感度を求める手続きを、感度

解析という。

電子線による電離量を測定して吸収線量を求めるときには、電離量測定の誤差の他に電子線エネルギーの誤差が吸収線量に誤差を生じさせる。文献 5 は吸収線量に対する電子線エネルギーの感度について検討したもので、電子線エネルギーの感度は著しく低いことが示されている。この事実を利用して、文献 4 は JARP が提唱する電子線エネルギーの測定法を簡略化する手段を提案している。簡略化しても、吸収線量の誤差をほとんど増加させないことが報告されている。

文献

1. 森川、石松：TPE 表の作成-最低限の測定データによる内挿-、日本放射線技師会雑誌、40(8)、929-938 (1993)
2. 森川賀根雄：TPR 表作成プログラム(TPR21)、(治療分科会セミナー資料)
3. 熊谷、石松、大塚、他：電子線の吸収線量測定についての一考察 -各種電離箱線量計の吸収線量変換係数-、日本放射線技術学会雑誌、47、8-14(1991)
4. 熊谷、石松、大塚、他：電子線の吸収線量測定についての一考察 -平均入射エネルギーの決定法-、日本放射線技術学会雑誌、47、15-21(1991)
5. 森川..電子線の吸収線量と平均エネルギー -吸収線量計算での誤差の伝播-、日本放射線技師会雑誌、40(6)、718-726(1993)