

第4部

実験データの統計的な解釈

Statistical Analysis of

Experimental Data

10. 測定値の解釈

Analysis of measured values

実験あるいは測定の結果何かの数値が得られた場合に、その数値は、正しい値とは違う値と考えるべきだと言われます ([補足 10.A](#))。

例えば、正確な温度計で温度を測定して、「40.9°C」と表示されたとします。正しい温度は40.85°Cより高く40.95°Cより低い温度らしいということでしょうけれど、“40.9”という「見かけの数値」と「正しい摂氏温度の（例えば“40.87...”という）数値」は違います。

有限の個数の数字や符号、ビット列などで数値を表現する限り、そのようなずれを避けることはできません。このようなずれのことが「丸め誤差」ラウンディング エラー rounding error と呼ばれることがありました。現実には、計測値を数値的に表現するための「ビット（メモリ素子）数」にかかるコストは急速に低減し続けているので、過去（コンピュータが使われる前）の「丸め誤差」の考え方と、現在の「丸め誤差」の考え方は、少し違います。

微弱な光やX線の強さを評価するときに使われる光子計数法 フォトン カウンティング メソッド photon counting method と呼ばれる計測技術があります。一定の時間内に検出器に入射した光子 フォトン photon のエネルギーをパルス性の電気信号に変換し、高い確率で検出できる信号強度（例えばCMOSレベル）にまで信号レベルを増幅し、増幅された電気パルスを実用的なデジタル計数回路を使って数える方法がとられます。光子計数法による光強度測定の結果として得られる数値は「光子の数」という意味を持つ離散的な値です。知りたかったことは「検出器に入射される光の強さ（単位時間あたりに検出素子に到達する光のエネルギー）」であり、本来は連続量のはずなのだとしても、物理現象として光のエネルギーは光子（光の量子）フォトン photon として

実体化されるので、光子計数法では、有限の測定時間で「光の強さ」の正しい値を直接知ることはできません。

一方で、計数法による強度測定に関わる「誤差の意味」のかなりの部分は、既に良く知られています。「測定誤差」と呼ばれることの内容には、なにかと扱いにくいことも多く含まれていますが、計数法に由来する誤差は「測定誤差」の中でも、数学的には比較的扱いやすいものです（[補足 10.B](#)）。

また、計数法の誤差は「丸め誤差」と似た性格を持ちます。「丸め誤差」も含めて、このタイプの誤差がクオンタイゼーション エラー量子化誤差 quantization error と呼ばれる場合もありました。

「誤差 error」と「偏差 deviation」の違いは、あまりはっきりとはしません。一般的には「真の値」あるいは「正しい値」からのずれのことが「誤差」と呼ばれる傾向があります。実験値や観測値が「予想値」からずれていた場合に、予想値の方が正しいことがわかっていれば実験値が「誤差」を持つことになるかもしれませんが、実験値の方が正しい値に近く、予想値が間違っている場合には、その「ずれ」を実験値の「誤差」と呼ぶ人はいないでしょう。

本来は「正しい値」がわかっていない時、あるいは予想値の不確かさが強い場合ほど、その「実験」には価値があります。予想値の正しいことがわかっているのであれば、そもそも「実験」をする必要はないはずです。「予想値」と「実験で観測された値」の違いを「誤差」と決めつけるのは好ましい態度ではありません。

また、観測値が統計的なばらつきを示す時に、それが物理現象として合理的に説明される場合や、むしろその統計的な変動の大きさが本質的に重要な情報を含む場合もあります。そのような場合に、それがなんらかの「『真の値』からのずれ」であったとしても「誤差」という言葉を使うのはふさわしくありません。

「偏差」・「ずれ」という言葉は「誤差」と言う言葉より少し弱い意味合いで使われます。状況に応じて「誤差」「偏差」「ずれ」「違い difference」と言う言葉を使い分けるように意識すると良いでしょう。

現実的な測定装置・計測技術を使って測定をした場合の誤差としては、「系統誤差」「統計誤差」と呼ばれる2つのタイプの誤差があるとされます。系統誤差の小さいことを「**確度 accuracy**が高い (accurate)」, 統計誤差の小さいことを「**精度 precision**が高い (precise)」と呼ばれる傾向がありました。

10-1 系統誤差と統計誤差 Systematic & random errors

何回も同じ測定を繰り返したときに、同じようにずれるようなずれ方のことはシステムティック デヴィエーション系統的な偏差 systematic deviation あるいは**系統誤差 systematic error**と呼ばれます。系統誤差はこうせい キャリブレーション較正 calibration により取り除くべきと言われます。

同じ実験を繰り返した時に、ずれの方向も大きさもバラバラなずれ方をすることはスタティスティカル統計的な偏差 statistical deviation あるいは**ランダムな偏差 random deviation**と呼ばれます。統

計誤差を抑えるためには、繰り返し実験をして複数の数値（データ）を得てから、平均などの代表値を得るべきと言われます。

「系統誤差（偏差）」 「統計誤差（偏差）」と呼ばれることについて、多くの人が持つイメージを [Fig. 10.1.1](#) と [Fig. 10.1.2](#) に示します。

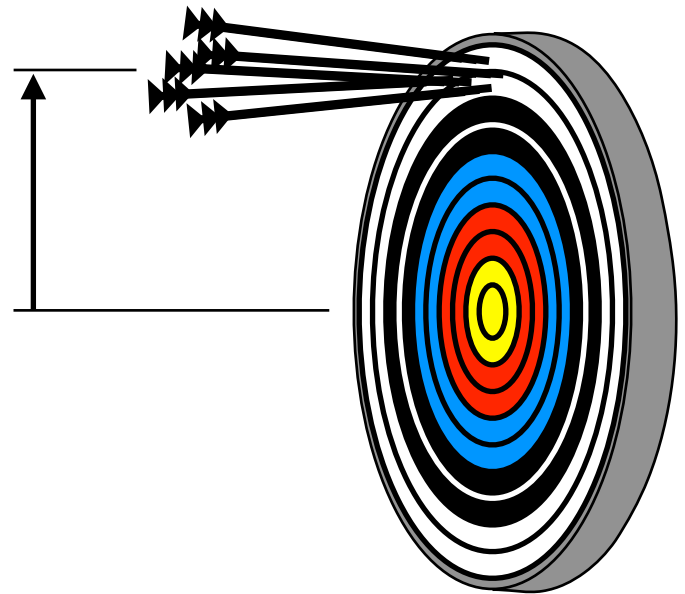


Fig. 10.1.1 「系統的な偏差」のイメージ

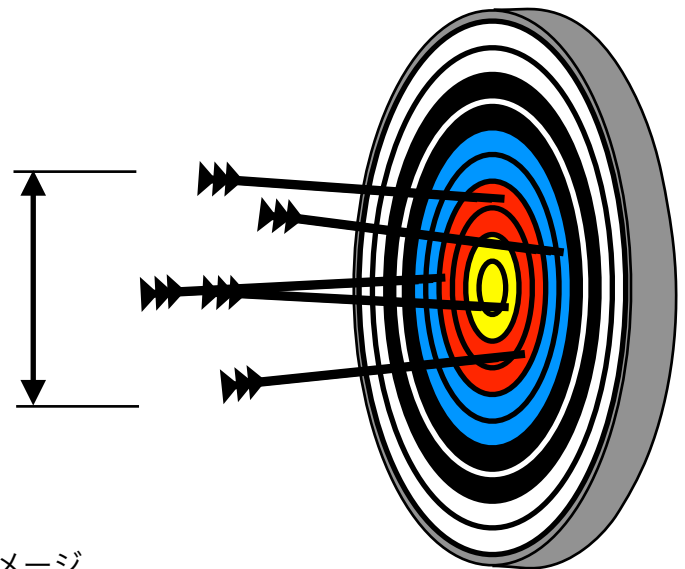


Fig. 10.1.2 「統計的（ランダム）な偏差」のイメージ

10-2 平均と分散・標準偏差 Mean, variance & standard deviation

同じ条件で複数回の実験をして、複数の数値からなるデータ $\{X_1, X_2, \dots, X_n\}$ が得られたとします。

統計的な誤差の影響を減らすために、このデータからどのように代表値 リプレゼンタティブ representative value ヴァリュー を求めるべきでしょうか？それは単純な問題ではなく、データの性格や、そのデータから何かの方法で求めた代表値をどのような目的・意図で利用するかによって変わります。そのうちで平均（mean または average） \bar{X} ：

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{1}{n} \sum_{j=1}^n X_j \quad (10.2.1)$$

として計算した結果を代表値とする方法は、実験データを解析する場合に、最も高い頻度で使われる方法です。操作者 (operator) が意図していなくても、装置を使った測定の場合

には、測定装置製造会社が暗黙のうちにそのような処理を施したデータを出力するようにしています。

式 (10.2.1) を使った計算で求められる \bar{X} の値は**標本平均** sample meanとも呼ばれます。

個々の数値 X_j の出現するイベント (事象) が**平均** m , **標準偏差** standard deviation σ (ギリシャ文字のシグマ) デヴィエーション であるような何かの確率分布に従い、その確率密度函数が $f(x)$ と表されるとします。「確率密度函数が $f(x)$ である」という意味は、「小さい数値 dx に対して、観測値 X_j が $x \leq X_j < x + dx$ の範囲の値をとる確率が、『 $f(x) dx$ 』で表される」ということです。この仮定は、数式を使えば、

$$\int_{-\infty}^{\infty} f(x) dx = 1 \quad (10.2.2)$$

$$\int_{-\infty}^{\infty} x f(x) dx = m \quad (10.2.3)$$

$$\int_{-\infty}^{\infty} (x - m)^2 f(x) dx = \sigma^2 \quad (10.2.4)$$

のように表現されます。

式 (10.2.3) で表されるような m の値を**真の平均**と呼ぶことができます。式 (10.2.4) で表される「標準偏差の平方 (自乗)」 σ^2 は**分散** variance ヴェアリアンス と呼ばれます。

データの個数が有限ならば、式 (10.2.1) を使って計算される標本平均 \bar{X} は式 (10.2.3) で表される真の平均 m から「必ずずれた値」になります。

それでは、標本平均 \bar{X} は真の平均 m からどれくらいずれるのでしょうか？

それぞれの観測値 X_j の出現するイベント (事象) が互いに独立だとします。本当は、繰り返し測定の結果が互いに独立とみなせない場合があることにも注意するべきですが、ここではそのことは棚上げにしておくことにします。イベント出現の独立性を仮定できる場合には、全てのデータの出現のしかた $\{x_1, x_2, \dots, x_n\}$ について、標本平均の値を計算して、確率をかけて足し合わせたものが、標本平均の期待値になるはず。このことから標本平均 \bar{X} の期待値 $\langle \bar{X} \rangle$ は、

$$\begin{aligned} \langle \bar{X} \rangle &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{1}{n} \sum_{j=1}^n x_j \right) f(x_1) f(x_2) \dots f(x_n) dx_1 dx_2 \dots dx_n \\ &= \frac{1}{n} \sum_{j=1}^n \int_{-\infty}^{\infty} x_j f(x_j) dx_j = m \end{aligned} \quad (10.2.5)$$

と計算され、「真の平均」 m に等しくなります。このような性質から、式 (10.2.1) によって算出される標本平均 \bar{X} を**不偏推定値** unbiased expectation アンバイアスト エクスペクテイション と呼ぶことができます。

標本平均 \bar{X} が真の平均 m からどの程度ずれるかは、標本平均の分散の期待値を計算すればわかります。この期待値は

$$\langle (\bar{X} - m)^2 \rangle = \frac{\sigma^2}{n} \quad (10.2.6)$$

となります ([補足 10.2.A](#))。

標本 X_j の本当の分散は σ^2 で表されます。式 (10.2.6) から「標本平均の分散」 $\langle (\bar{X} - m)^2 \rangle$ は標本分散の $1/n$ 程度の値になることがわかります。このことが「統計誤差を抑えるためには、繰り返し実験をして**平均**などの**代表値**をとれば良い」と言われることの根拠になります。一回の測定の統計誤差の大きさが σ の場合に、 n 回測定を繰り返して平均をとれば、その平均値の統計誤差の大きさは σ/\sqrt{n} になることが期待されるからです。なお、標本平均の標準偏差 (分散の平方根) のことが標本平均の**標準誤差 standard error** と呼ばれることも一般的です。

以下の式で表される値を**標本分散 sample variance** と呼び、ここでは $(\Delta X)^2$ と表すことにします (この表し方は必ずしも一般的ではありません)。

$$(\Delta X)^2 = \frac{1}{n-1} \sum_{j=1}^n (X_j - \bar{X})^2 \quad (10.2.7)$$

式 (10.2.7) では、右辺で和を n で割るのでなく $n-1$ で割っていることに注意してください。

このことは、真の平均 m を標本平均 \bar{X} で代用して「偏差の自乗の平均値」の期待値を求めると、

$$\langle (X - \bar{X})^2 \rangle = \left\langle \frac{1}{n} \sum_{j=1}^n (X_j - \bar{X})^2 \right\rangle = \frac{n-1}{n} \sigma^2 \quad (10.2.8)$$

となって、必ず真の分散 σ^2 より小さい値 (偏^{かたよ}った値) になってしまうからです。それに対して式 (10.2.7) で計算される標本分散 $(\Delta X)^2$ の期待値には

$$\langle (\Delta X)^2 \rangle = \sigma^2 \quad (10.2.9)$$

という関係が成立します。標本値から「真の分散」を推定したいのであれば、式 (10.2.7) を使うことが、不偏推定 (偏^{かたよ}っていない推定) ということになります。式 (10.2.7) を使って計算される標本分散 $(\Delta X)^2$ は**不偏分散 unbiased variance** あるいは**不偏標本分散 unbiased sample variance** とも呼ばれます。ただし、標本数がかなり多く (n が大きく) なければ不偏標本分散 $(\Delta X)^2$ の値は真の分散 σ^2 から大きくずれた場合をとることが普通です。その一方で、標本数 n が多くなれば「不偏標本分散」 $(\Delta X)^2 = \langle (X - \bar{X})^2 \rangle / (n-1)$ と「偏差の自乗の平均」 $\langle (X - \bar{X})^2 \rangle / n$ とは近い値になります。結局のところ「不偏標本分散」と「偏差の自乗平均」との違いは、実際には深刻な問題というほどではありません。

少ない (n の小さい) 試行しかしていないのに、誤差まで含めて正しい値を知れていると思いたがる心理の方が、現実的には深刻な問題です。

標本数 n は有限の個数でしかありえないので、標本分散 $(\Delta X)^2$ は、真の分散 σ^2 から必ずずれた値になります。標本分散 $(\Delta X)^2$ が真の分散 σ^2 からどの程度ずれるかも、「『標本分散』の分散」の期待値を計算すれば推定することが (特定の条件では) 可能です。導出の過程は省略しますが、独立性を仮定できる場合には、標本分散の標準誤差は

$$\Delta [(\Delta X)^2] \approx \left\{ \frac{1}{n^2} \sum_{j=1}^n (X_j - \bar{X})^4 - \frac{1}{n} \left[\frac{1}{n} \sum_{j=1}^n (X_j - \bar{X})^2 \right]^2 \right\}^{1/2} \quad (10.2.10)$$

という式で見積もられると考えられています (Ida, 2008)。

標本分散の平方根 $\Delta X \equiv \sqrt{(\Delta X)^2}$ は**標本標準偏差** sample standard deviation と呼ばれます。標本標準偏差は真の標準偏差 σ からずれた値になりますが、真の標準偏差 σ の推定値としては不偏推定になっており、この値 ΔX のことが単に**標準偏差** standard deviation と呼ばれることも少なくありません。

10-3 代表値の選び方 Choice of representative values

複数の数値からなるデータ $\{X_1, X_2, \dots, X_n\}$ からどのような**推定値** エスティメイトド esitimated value あるいは**代表値** レプリゼンタティブ representative value を求めるべきかは、データの性格や求めた代表値をどのような目的で利用するかによって変わります。このことを一般化するため、**ノルム** norm という考え方を使うことのできる場合があります。推定値あるいは代表値、**予想値** エクスベクテド expected value を r とした時に、次数 ν (ν はギリシャ小文字のニュー nu) の偏差 (ずれ) のノルムは以下のように定義されます。

$$L_\nu \equiv \left(\sum_{j=1}^n |X_j - r|^\nu \right)^{1/\nu} \quad (10.3.1)$$

ノルムとは、何かの意味で「大きさ」「長さ」「距離」のようなものを表し、データから「真の値」を推定するためには、偏差のノルム (隔^{へだ}たり) を最小化する解を求めれば良いと考えられています。

実験データの解析のためには、式 (10.3.1) で表される L_ν のうち、 $\nu = 2$ の場合の偏差のノルム、 L_2 (エル・ツー) ノルムを最小化する解の用いられることが多いのですが、常に偏差の L_2 ノルムを最小化する値を代表値として選ぶことが最適であるとは限りません。画像認証のように $\nu < 2$ のノルムの用いられる場合も、ゲーム戦略・危機戦略 (リスク・マネジメント) のように $\nu > 2$ のノルムが用いられる場合もあります。

また**最頻値** モード mode のように、ノルムを使わない代表値の選び方もあります。

10-3-1 L_2 ノルムの最小化・平均・最小自乗法

次数 $\nu = 2$ のノルムのことは L_2 (エル・ツー) ノルムと呼ばれ、^{ユークリッド} Euclid 空間での距離に対応する値を意味します。実験データを解析して推定値を得るためには、偏差の L_2 ノルムを最小化する方法が、最も良く使われる方法です。これは「 L_2 ノルムの自乗」 L_2^2 を最小化することと同じことで、

$$L_2^2 = \sum_{j=1}^n (X_j - r)^2 \quad (10.3.1.1)$$

と書き直せば、これは「偏差の自乗の和」を最小にすることと同じです。このことから、この推定法は^{リスト スクエアズ メソッド}最小自乗法 least-squares method あるいは^{エスティメーション}最小自乗推定 least-squares estimation と呼ばれる場合もあります。以下の関係：

$$\frac{dL_2^2}{dr} = 2 \sum_{j=1}^n (X_j - r) = 2n \left(\frac{1}{n} \sum_{j=1}^n X_j - r \right) \quad (10.3.1.2)$$

から、 L_2 ノルムの自乗あるいは偏差の平方和 L_2^2 を最小にする代表値の選び方は

$$r = \frac{1}{n} \sum_{j=1}^n X_j \quad (10.3.1.3)$$

となります。代表値の選び方としては、**平均** (標本平均) が**最小自乗解**であり、**不偏推定解**になります。

10-3-2 L_1 ノルムの最小化・中央値・ロバスト推定

現実の測定では、測定値として「本来の値」から極端にずれた値が突発的に出現することがあります。例えば、光の強度と光子のエネルギー (光の波長) を同時に測定しようとする時に、^{センサー}「検出器 sensor が、貫通力の高い宇宙線の被曝を受けて、異常なレベルの高エネルギー光子による信号パルスを出力したかもしれない」ようなことがありうることは、否定できません。現実に筆者が主に日本の「高エネルギー加速器研究機構」で実験のようなことをした経験では、計測装置の示す突発的な異常値は、^{エレクトロマグネティック}電磁誘導 (electromagnetic induction; EMI) ノイズ由来であったり、^{インダクション イーエムアイ}ライン (電源系統) ノイズ由来であったり、^{せいきかんかいろう ポジティブ フィードバック サーキット}検出器系の回路構成のうちの正帰還回路 positive feedback circuit の暴走に由来していたり、^{オペレータ}制御ソフトウェアやオペレーション・システムの不具合に由来したり、操作者 operator の誤操作によるが多かったと思います。そうだとすると、そのように突発的なノイズを排除することが困難である場合もあるでしょうし、それを経験することが、ないわけではないかもしれません。

現在の日本の高校数学レベルの教育では、^{ひん}頻度は低くても「実験データの一部が極端に予想値から違っている」ことがあり、それは「^{はずち}外れ値」とされます。この場合に L_2 ノルムを代表値として使ってしまうと、出現頻度の低い外れ値であっても、その影響をかなり受けることになり、測定結果が予想値から大きく変わることになってしまいます。

それに対して、次数が2より小さい ($\nu < 2$) ノルムを使えば外れ値の影響を受けにくくなり、そのようなノルムを利用する方法が、**ロバスト推定** (robust = 「頑健な」「頑固な」という意味) と呼ばれます。ロバスト推定は、一般的には不偏推定にはならず、ある意味で「^{かたよ}偏った」結果しか得られません。しかし「データが、低い頻度で突発的な強い誤差要因の影響を受ける」ような場合に有効であることには間違いがなく、データ解析が「不偏推定でなければいけない」と言う明確な理由もありません。

ロバスト推定の中で極端な例は、次数 $\nu = 1$ のノルム、 L_1 (エル・ワン) ノルムを用いる方法で、**画像認識** (指紋認証・虹彩認証・顔認証) などで用いられると言われます。式 (10.3.1) から、偏差の L_1 ノルムは

$$L_1 = \sum_{j=1}^n |X_j - r| \tag{10.3.2.1}$$

と表されます。かりに $X_1 \leq X_2 \leq \dots \leq X_n$ だとして、 $X_m \leq r \leq X_{m+1}$ とすれば、

$$\begin{aligned} L_1 &= \sum_{j=1}^m (r - X_j) + \sum_{j=m+1}^n (X_j - r) \\ &= mr - \sum_{j=1}^m X_j + \sum_{j=m+1}^n X_j - (n - m)r \\ &= (2m - n)r - \sum_{j=1}^m X_j + \sum_{j=m+1}^n X_j \end{aligned} \tag{10.3.2.2}$$

と表されることとなります。



Fig. 10.3.2.1 マンハッタン島 (ニューヨーク) の道路網



Fig. 10.3.2.2 ニューヨークのタクシー (イエローキャブ)

式 (10.3.2.2) から、 $m \leq n/2$ の場合には、代表値 r が大きくなるほど L_1 ノルムは小さくなり、 $m \geq n/2$ の場合には、代表値 r が大きくなるほど L_1 ノルムは大きくなることがわかります。

L_1 ノルムを最小にする代表値の選び方は、**中央値**あるいは**中位数**、**メジアン** メディアン median と呼ばれる値を求めることと同じことです。

データ数 n が奇数の場合には、小さい順から大きい順へ（あるいは、大きい順から小さい順へ）数値を並べ直した時の、 $(n+1)/2$ 番目の値 $X_{(n+1)/2}$ が中央値になり、 L_1 ノルムを最小にする解は一意に確定します。データ数 n が偶数の場合には、中央値は普通 $(X_{n/2} + X_{n/2+1})/2$ として定義されます。しかし、本来は L_1 ノルムを最小にする解 r は $X_{n/2} \leq r \leq X_{n/2+1}$ の関係を満たす任意の値をとることができて、一意には確定しません。

L_1 ノルムは「マンハッタン・ノルム」と呼ばれることがあります。ニューヨークのマンハッタン島（あるいは日本の札幌や京都）のような格子状の道路網を使って移動する場合に、タクシーを使って目的地に到達するルートには、最短のものが複数存在し、このことも L_1 ノルムを最小化する解が一意に確定しないことに対応しています (Fig. 10.3.2.1, Fig. 10.3.2.2)。

有限個数の標本から中央値を求めようとしても、その値は**真の中央値**とは異なるものになることは、標本の「平均」を求める場合と同じことです。個々の数値 X_j の出現するイベント（事象）が、確率密度関数 $f(x)$ の確率分布にしたがうときに、真の中央値 x_{median} は、

$$x_{\text{median}} = F^{-1}\left(\frac{1}{2}\right) \quad (10.3.2.3)$$

$$F(x) \equiv \int_{-\infty}^x f(t) dt \quad (10.3.2.4)$$

と表されます。式 (10.3.2.4) で定義される関数 $F(x)$ は**累積分布関数** キュミュラティブ cumulative distribution function と呼ばれます。累積分布関数の逆関数 $F^{-1}(x)$ が簡単な数式で表されるのであれば「真の中央値」 $x_{\text{median}} = F^{-1}(1/2)$ を計算することは容易ですし、コンピュータを使って計算することが前提であれば、例えば二分法などのアルゴリズムを使って方程式 $F(x_{\text{median}}) = 1/2$ を解くだけのことです。

「箱ひげ図」 box plot を描くときに用いられる**第1四分位点** ロウワー クォーター lower quartile、**第3四分位点** upper quartile はそれぞれ $F^{-1}(1/4)$ と $F^{-1}(3/4)$ と表されます。第1四分位点の値も第3四分位点の値も有限個数の標本点から求めた値は「真の値」からずれた値になるのも同様です。

中央値には、標本が序列化されていれば決まるので、数値データであっても、数値そのものには意味がなく、数値の大小（順位）のみが意味を持つという特徴もあります。

10-3-3 L_∞ ノルムの最小化・中点値・ミニマックス推定

ロバスト推定とは逆に、 $\nu > 2$ のノルムを用いる代表値の選び方・推定法もあります。特に次数が無限大 $\nu \rightarrow \infty$ の場合、

$$L_\infty = \lim_{n \rightarrow \infty} \left[\sum_{j=1}^n |X_j - r|^\nu \right]^{1/\nu} = \max \left\{ |X_j - r| \right\} \quad (10.3.3.1)$$

と表すことができ、 L_∞ (エル・むげんだい) ノルムは、「偏差 (の絶対値) の最大値」と同じこととなります。最大偏差を最小にする推定法を**ミニマックス mini-max 推定**と呼びます。 $\{X_j\}$ の最小値が X_{\min} 、最大値が X_{\max} であれば、ミニマックス解は

$$r = \frac{X_{\min} + X_{\max}}{2} \quad (10.3.3.2)$$

となります。この値は**中点値**あるいは**ミッドレンジ mid-range** と呼ばれることがあります。が、いずれもあまり一般的ではありません。

ミニマックス法は、数値計算ライブラリにおける**関数近似**や、**危機管理 risk management**、**ゲーム戦略**などで使われます。

ミニマックス法は偏差の大きさが有限 (有界) の場合にしか使えません。確率密度関数 $f(x)$ が

$$\begin{cases} f(x) > 0 & [x_{\min} \leq x \leq x_{\max}] \\ f(x) = 0 & [x < x_{\min}, x_{\max} < x] \end{cases} \quad (10.3.3.3)$$

のように有限な区間 $[x_{\min}, x_{\max}]$ でのみゼロでない値をとることが前提となります。

10-3-4 最頻値・モード

最頻値 mode はノルムを用いない代表値の決め方で、出現頻度が最大となる値を代表値とし、「議事の採決」や「選挙」などで用いられる「多数決」と同じことです。選択の対象が「尺度化された数値」に限定されないという意味では、「汎用性の高い手法」と言う性格があります。また、頻度が低く代表値から大きくずれた測定値 (外れ値) の影響を受けず、「外れ値に対してロバスト (頑固) な性質を持つ」とも言えます。ただし、実際には、測定値が統計的な変動を示す場合にその影響を敏感に受け、むしろ「脆弱性 vulnerability」の高い代表値の選択法」という性格の強い面があります。

個々の数値 X_j の出現するイベント (事象) が、「確率密度関数が $f(x)$ と表されるような確率分布」にしたがうとき、最頻値 x_{mode} は「確率密度関数 $f(x)$ が最大値をとるときの x の値」を意味します。

測定の対象が連続的な物理量（長さ・質量・体積など）に相当するもので、測定値 X_j の有効ビット数が多く、実質的に連続量とみなせるような場合には、実測のデータ $\{X_1, X_2, \dots, X_n\}$ 中のすべての数値は 1 回出現するだけになるはずであり、物理量についての観測データから最頻値を推定する方法は単純ではありません。

最頻値を推定するための一つの方法に、**ヒストグラム histogram**（頻度の棒グラフ）を作成し、ビン **区間 bin** の中点にあたる値をとる方法があります。ただし、この場合、どのような区間を設定するべきかについて何らかの選択をしなければならず、その選択の仕方によって推定される最頻値も変わります。

ヒストグラムのビン bin の位置と幅をどのように設定すれば良いかには、色々な提案があります。そのうちの一例として、標本数 n 、標本の最小値 X_{\min} 、最大値 X_{\max} に対して、ビンの数 k と幅 h を

$$k = \left\lceil \sqrt{n} \right\rceil \quad (10.3.4.1)$$

$$h = \frac{X_{\max} - X_{\min}}{k} \quad (10.3.4.2)$$

と選ぶ「平方根選択」と呼ばれる方法があります。式 (10.3.4.1) 中の $[x]$ の記号は 天井関 **数 ceiling function** と呼ばれ、「 x を したまわ 下回らない最小の整数」（切り上げ）を意味します。

平方根選択を用いるとしても、ビンの幅 h として、式 (10.3.4.1) と式 (10.3.4.2) で求められる値に近く 10^n あるいは 2×10^n 、 5×10^n (n は整数) などと表される値に決めることも普通の選択です。そのようにすれば結果が見やすくなるだけでなく、じっ 十進法で数値を表記した時に必要となる文字数を減らせるので、記憶・記録媒体に必要なコストを節約できる面もあります。

10-4 モーメントとキュムラント moment & cumulant

測定値の平均や分散・標準偏差から測定値の統計分布に関する推定を行う考え方をさらに拡張して、モーメントやキュムラント、わいど せんど 歪度、尖度などを使う場合があります。

任意の関数 $f(x)$ の k 階の そ **粗モーメント raw moment**（原点まわりのモーメント） μ'_k （ギリシャ文字のミューにプライム記号、 k の添え字；ミュー・プライム・ケー）と キュムラント **キュムラント cumulant** κ_k （ギリシャ文字のカッパに k の添え字；カッパ・ケー）は、以下の式で定義されます。

$$\mu'_k \equiv \left. \frac{\partial^k M(\theta)}{\partial \theta^k} \right|_{\theta=0} \quad (10.4.1)$$

$$\kappa_k \equiv \left. \frac{\partial^k K(\theta)}{\partial \theta^k} \right|_{\theta=0} \quad (10.4.2)$$

$$K(\theta) \equiv \ln M(\theta) \quad (10.4.3)$$

$$M(\theta) \equiv \int_{-\infty}^{\infty} e^{\theta x} f(x) dx \quad (10.4.4)$$

式 (10.4.3) で定義される函数 $K(\theta)$ は **キュムラント母函数** ジェネレーティング cumulant generating function と呼ばれ、式 (10.4.4) で定義される函数 $M(\theta)$ は **モーメント母函数** moment generating function と呼ばれます。

函数 $f(x)$ のゼロ階粗モーメント μ'_0 (ミュー・プライム・ゼロ) とキュムラント κ_0 (カッパ・ゼロ) は

$$\mu'_0 = M(0) = \int_{-\infty}^{\infty} f(x) dx \quad (10.4.5)$$

$$\kappa_0 = K(0) = \ln M(0) \quad (10.4.6)$$

と表され、 $f(x)$ が普通の意味での確率密度函数であれば $\int_{-\infty}^{\infty} f(x) dx = 1$ の関係から、

$$\mu'_0 = M(0) = 1 \quad (10.4.7)$$

$$\kappa_0 = K(0) = 0 \quad (10.4.8)$$

となります ([補足 10.4.A](#))。以下 $f(x)$ が確率密度函数であることを前提とします。

一般的に k 階の粗モーメント μ'_k は

$$\mu'_k = \int_{-\infty}^{\infty} x^k f(x) dx = \langle x^k \rangle \quad (10.4.9)$$

と書いて、確率変数 x の k 乗の平均に等しいと言う関係があります ([補足 10.4.B](#))。

1 階のキュムラント κ_1 は、

$$\kappa_1 = \langle x \rangle \quad (10.4.10)$$

となつて **平均** $\langle x \rangle$ に等しく、2 階のキュムラント κ_2 は、

$$\kappa_2 = \langle x^2 \rangle - \langle x \rangle^2 = \langle (x - \langle x \rangle)^2 \rangle \quad (10.4.11)$$

となつて **分散** に等しいという関係があります。3 階のキュムラント κ_3 は、

$$\kappa_3 = \langle x^3 \rangle - 3\langle x^2 \rangle \langle x \rangle + 2\langle x \rangle^3 = \langle (x - \langle x \rangle)^3 \rangle \quad (10.4.12)$$

と表され、4 階のキュムラント κ_4 は、

$$\kappa_4 = \langle x^4 \rangle - 4\langle x^3 \rangle \langle x \rangle - 3\langle x^2 \rangle^2 + 12\langle x^2 \rangle \langle x \rangle^2 - 6\langle x \rangle^4 = \langle (x - \langle x \rangle)^4 \rangle - 3\langle (x - \langle x \rangle)^2 \rangle^2 \quad (10.4.13)$$

と表されます ([補足 10.4.C](#))。

確率分布の k 階の **中心モーメント** central moment μ_k は、平均を $m = \langle x \rangle$ としたときに

$$\mu_k \equiv \int_{-\infty}^{\infty} (x - m)^k f(x) dx = \langle (x - m)^k \rangle \quad (10.4.14)$$

として定義される値であり、単に「モーメント」と呼ばれるときには、式 (10.4.14) で表される中心モーメント μ_k を意味することがむしろ普通です。

2階のキュムラント κ_2 は2階の中心モーメント $\mu_2 = \langle (x - \langle x \rangle)^2 \rangle$ に等しく、3階のキュムラント κ_3 は3階の中心モーメント $\mu_3 = \langle (x - \langle x \rangle)^3 \rangle$ に等しくなります。

3階キュムラントあるいは3階中心モーメントは、函数の形状の非対称性を特徴付ける値です。ただしこの値は標準偏差 σ の3乗にも比例するので、函数の形状の非対称性のみを特徴付けるために、

$$\gamma_1 \equiv \frac{\mu_3}{\sigma^3} = \frac{\kappa_3}{\kappa_2^{3/2}} \quad (10.4.15)$$

として定義される^{わいど スキューネス}**歪度 skewness** という値も用いられます。

実測のデータから統計分布の3階キュムラントや3次中心モーメント、歪度を推定することが可能な場合があります。ただし、これらについて「意味のある数値」を得るためには、分散を推定する場合より、さらに多くの標本数が必要となるので、3階以上のキュムラントを推定した値の「不偏性 unbiasedness」^{アンバイアストネス} について議論することは、実質的には無意味な場合が多いでしょう。

4階キュムラント κ_4 と4階の中心モーメント μ_4 、2階の中心モーメント、 μ_2 の間には、

$$\kappa_4 = \mu_4 - 3\mu_2^2 \quad (10.4.21)$$

の関係があります。

4階キュムラントあるいは4階中心モーメントは、函数の形状の^{とが}尖り方を特徴付ける値です。これらの値は標準偏差 σ の4乗あるいは分散の2乗に比例するので、函数の形状の尖り具合のみを特徴付けるために、

$$\gamma_2 \equiv \frac{\kappa_4}{\sigma^4} = \frac{\mu_4}{\sigma^4} - 3 \quad (10.4.22)$$

として定義される^{せんど カートウシス}**尖度** (kurtosis あるいは ^{エクセス エクセス カートウシス}excess, excess kurtosis) と呼ばれる値の用いられる場合があります。日本工業規格 (Japanese Industrial Standards; JIS, Z8101-1: 2015) では、国際標準機構規格 (ISO, 3534-1: 2006) と同様に、4次の中心モーメント μ_4 と標準偏差 σ の4乗の比 μ_4/σ^4 で表される値が「**とがり**」(kurtosis) と定義されています。

10-5 測定の意味 Meaning of measurement

実験あるいは**測定**の目的は、多くの場合に何かの**物理量**について**正しい値**を知ることと思われるのは普通でしょう。**量子論** quantum theory では、観測可能な物理量は**オブザーバブル** observable と呼ばれ、その正しい値の現れ方は、何かの**統計分布** statistical distribution に従うとされます ([補足 10.5.A](#))。

測定によって現れる観測値の出現のしかたが、正しい統計分布 (観測可能な物理量の本質的な統計分布) とは別の統計分布に従うのは普通のことです。真値と観測値の統計分布の間の**平均値**の違いが**系統誤差**と呼ばれ、**分散**あるいは**標準偏差**の違いが**統計誤差**と呼ばれると見ることがもできます。

実験的な測定の研究の目的は「物理量の現れ方の正しい統計分布を知ること」とも言えますが、「平均」＝「1階キュムラント」と「分散」＝「2階キュムラント」とが一致するだけでは、その統計分布のすべてを知ったことにはなりません。3階以上のすべてのキュムラントも確定しなければすべてを知れたことにはならないはずですが、ただし、現実に測定値を何かの目的で利用するためには、平均値が最も重要であり、分散がその次に重要で、3階以上のキュムラントは、階数が高くなるほど重要性が低くなると考えて良いでしょう（[補足 10.5.B](#)）。

10-6 畳込 Convolution

分光測定 spectroscopic measurement や回折・散乱測定 diffraction/scattering

measurement の結果が、試料固有の分光特性や試料固有の構造秩序による本質的な分光・回折強度図形をそのまま表すものでなく、測定装置に由来する^{ぼや}暈け smearing やピーク位置ずれ peak shift, 強度図形の変形 deformation などの影響を受けたものになることは、古くから当然のこととして受け入れられていました。

試料固有の性質による強度図形が函数 $f(x)$ として表され、装置の影響による^{ぼや}暈けが函数 $g(x)$ で表されるとします。比較的単純な場合には、観測される強度図形を表す函数 $h(x)$ は函数 $f(x)$ と函数 $g(x)$ の^{たたみこみ}畳込 convolution として表されます。その関係は、

$$h(x) = f(x) * g(x) = \int_{-\infty}^{\infty} f(x-y) g(y) dy \quad (10.6.1)$$

と表されます。このとき、函数 $g(x)$ は装置函数 instrumental function と呼ばれます。式 (10.6.1) の関係は、ディラックのデルタ函数 Dirac delta function $\delta(x)$ ([補足 10.6.A](#)) を使えば、

$$h(x) = f(x) * g(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(x-y-z) f(y) g(z) dy dz \quad (10.6.2)$$

と表現することもできます。

畳込 $h(x) = f(x) * g(x)$ のフーリエ変換 Fourier transform $\mathfrak{H}(\xi)$ (“ \mathfrak{H} ” はアルファベット大文字 “H” のフラクトゥール字形, ξ はギリシャ文字小文字のグザイ xi)

$$\mathfrak{H}(\xi) \equiv \int_{-\infty}^{\infty} h(x) e^{2\pi i \xi x} dx \quad (10.6.3)$$

は、成分函数 $f(x), g(x)$ のフーリエ変換 $\mathfrak{F}(\xi), \mathfrak{G}(\xi)$ (“ $\mathfrak{F}, \mathfrak{G}$ ” はアルファベット大文字 “F, G” のフラクトゥール字形)

$$\mathfrak{F}(\xi) \equiv \int_{-\infty}^{\infty} f(x) e^{2\pi i \xi x} dx \quad (10.6.4)$$

$$\mathfrak{G}(\xi) \equiv \int_{-\infty}^{\infty} g(x) e^{2\pi i \xi x} dx \quad (10.6.5)$$

の積に等しいという関係

$$\mathfrak{H}(\xi) = \mathfrak{F}(\xi) \mathfrak{G}(\xi) \quad (10.6.6)$$

が成立します (補足 10.6.B)。この関係は^{たたみこみ}畳込定理 convolution theorem と呼ばれます。

一般的にフーリエ変換の逆変換 (逆フーリエ変換 inverse Fourier transform) が

$$f(x) = \int_{-\infty}^{\infty} \mathfrak{F}(\xi) e^{-2\pi i \xi x} d\xi \quad (10.6.7)$$

と表されること (補足 10.6.C) と、式 (10.6.6) に示す畳込定理の関係から、

$$f(x) = \int_{-\infty}^{\infty} \frac{\mathfrak{H}(\xi)}{\mathfrak{G}(\xi)} e^{-2\pi i \xi x} d\xi \quad (10.6.8)$$

の関係が導かれます。式 (10.6.3) と式 (10.6.5)、式 (10.6.8) を組み合わせれば、畳込としての実測強度データ $h(x)$ に対して、装置関数 $g(x)$ が既知であれば、装置の影響を受ける前の本質的な強度図形 $f(x)$ が得られることとなります。一般的には、このような処理のことが^{ぎやくたたみこみ}逆畳込 deconvolution と呼ばれます。

しかし、現実には、実測の強度データに対して逆畳込処理を施しても、うまく行く例は多くありません。それは実測の強度図形 $h(x)$ が誤差や統計的な変動を伴うからです。

Stokes (1948) は実測の粉末 X 線回折強度図形に対して逆畳込処理を施して、逆畳込が一見うまく働くように見える例を示していますが、うまく行っているように見える理由は、実測強度のフーリエ変換 (フーリエ係数) $\mathfrak{H}(\xi)$ のうち、 $|\xi|$ の比較的大きい部分の値を強制的に 0 とみなす処理を施しているからです。このような処理は「逆畳込処理」をする人たちの使う^{じょうとうしゅだん}常套手段であり、時系列強度データの場合には、多くの場合に「フーリエ変換の高周波数成分は『雑音 noise』によるものであり『意味がない』から切り捨てる」などと「もっともらしく聞こえる理由づけ」のされることが普通です。

補足

(補足 10.A) 実験データと正しい値との違い (↔)

「科学的な手法で得られた実験データが示すことは客観的な事実である。すべての判断・主張は、科学的な手法で検証可能な証拠 evidence のみに基づいた明確な論理によって導かれるべきであり、それ以外の方法で『正しい値』を知ることはできない」という考え方はあります。そのように現代的な意味と同じ意味での実証主義 positivism は、19 世紀初頭のフランスの思想家・社会学者オーギュスト・コンテ Auguste Comte によって明確に主張されたと言われます。

一方で社会学 sociology の分野では、コンテの社会学的な実証主義を拒否する考え方も提唱されるようになり、19 世紀ドイツ出身の哲学者・革命家カール・マルクス Karl Marx (唯物史観 historical materialism, 資本

論, 社会主義, 共産主義などで有名) や 20 世紀オーストリアの哲学者・数学者エドムント・フッサール Edmund Husserl の現象学のように, 反実証主義 antipositivism (否定主義 negativism) に分類される考え方もあり, その考え方は現代の思想や教育にも, 強い影響を与えています。

歴史的には, 17 世紀フランスの哲学者・数学者であったルネ・デカルト René Descartes の思想, フランスの哲学者 ニコラ・ド・マルブランシュ Nicolas de Malebranche, オランダの哲学者バールーフ・スピノザ Baruch Spinoza (汎神論), ドイツの哲学者・数学者・科学者ゴットフリート・ライプニッツ Gottfried Wilhelm Leibniz (モノド, 予定調和説で有名) らの思想は, 理性主義 (合理主義) rationalism として知られます。

一方で 17 世紀英国の哲学者・法学者・政治家フランシス・ベーコン Francis Bacon (「知識は力なり」とい言葉・イドラ概念で有名) の提唱した思想・認識論 epistemology, 17 世紀英国の哲学者トマス・ホブズ Thomas Hobbes (社会契約説で有名), 17 世紀英国の哲学者ジョン・ロック John Locke (自由主義の父と呼ばれる) らの思想は経験主義 empiricism として知られます。

18-19 世紀のドイツの哲学者イマヌエル・カント Immanuel Kant (認識論においてコペルニクスの転回 Copernican revolution をもたらしたとされる) が理性主義と経験主義の対立を解決し統合したとも言われますが, それほどでもないとも言われます。 (↔)

(補足 10.B) 光子計数法の統計誤差 (↔)

光子計数法による測定結果の統計的な性質は, 概ねポアソン分布 Poisson distribution で説明できます。ポアソン分布については 11 章で取り上げます。平均 λ のポアソン分布は, 1 階以上のキュムラントがすべて λ に等しいと言う性質を持ちます。 (↔)

(補足 10.2.A) 標本平均の分散の期待値 (↔)

式 (10.2.1) :

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{1}{n} \sum_{j=1}^n X_j \quad (10.2.1)$$

として計算される標本平均 \bar{X} が真の平均 m からどの程度ずれるかは, 標本平均の分散の期待値を計算すればわかります。この期待値は

$$\begin{aligned} \langle (\bar{X} - m)^2 \rangle &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{1}{n} \sum_{j=1}^n x_j - m \right)^2 f(x_1) f(x_2) \dots f(x_n) dx_1 dx_2 \dots dx_n \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{1}{n} \sum_{i=1}^n x_i - m \right) \left(\frac{1}{n} \sum_{j=1}^n x_j - m \right) f(x_1) f(x_2) \dots f(x_n) dx_1 dx_2 \dots dx_n \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{i=1}^n \frac{x_i - m}{n} \sum_{j=1}^n \frac{x_j - m}{n} f(x_1) f(x_2) \dots f(x_n) dx_1 dx_2 \dots dx_n \\ &= \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_i - m) (x_j - m) f(x_i) f(x_j) dx_i dx_j \\ &= \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \langle (x_i - m) (x_j - m) \rangle \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \begin{cases} \sigma^2 & [i=j] \\ 0 & [i \neq j] \end{cases} \\
&= \frac{1}{n^2} \sum_{i=1}^n \sigma^2 \\
&= \frac{\sigma^2}{n}
\end{aligned} \tag{10.2.A.1}$$

となります。 (↔)

(補足 10.4.A) 粗モーメントとキュムラント, モーメント母関数とキュムラント母関数 (↔)

統計分布の確率密度関数が $f(x)$ と表されるとき, その統計分布の k 階粗モーメント μ'_k と k 階キュムラントは, 式 (10.4.1)–(10.4.4) :

$$\mu'_k \equiv \left. \frac{\partial^k M(\theta)}{\partial \theta^k} \right|_{\theta=0} \tag{10.4.1}$$

$$\kappa_\nu \equiv \left. \frac{\partial^k K(\theta)}{\partial \theta^k} \right|_{\theta=0} \tag{10.4.2}$$

$$K(\theta) \equiv \ln M(\theta) \tag{10.4.3}$$

$$M(\theta) \equiv \int_{-\infty}^{\infty} e^{\theta x} f(x) dx \tag{10.4.4}$$

によって定義されます。式 (10.4.1) の表現は, モーメント母関数 $M(\theta)$ が

$$M(\theta) = 1 + \mu'_1 \theta + \frac{\mu'_2 \theta^2}{2} + \frac{\mu'_3 \theta^3}{3!} + \dots = 1 + \sum_{k=1}^{\infty} \frac{\mu'_k \theta^k}{k!} \tag{10.4.A.1}$$

のように θ の冪級数として表現される同じことであり, (10.4.2) の表現は, キュムラント母関数 $K(\theta)$ が

$$K(\theta) = \kappa_1 \theta + \frac{\kappa_2 \theta^2}{2} + \frac{\kappa_3 \theta^3}{3!} + \dots = \sum_{k=1}^{\infty} \frac{\kappa_k \theta^k}{k!} \tag{10.4.A.2}$$

と表されることと同じことです。

式 (10.4.1) と式 (10.4.4) とから,

$$\mu'_k \equiv \left. \frac{\partial^k}{\partial \theta^k} \int_{-\infty}^{\infty} e^{\theta x} f(x) dx \right|_{\theta=0} \tag{10.4.A.3}$$

とし, 式 (10.4.2) と式 (10.4.3) とから,

$$\kappa_\nu \equiv \left. \frac{\partial^k}{\partial \theta^k} \ln \int_{-\infty}^{\infty} e^{\theta x} f(x) dx \right|_{\theta=0} \tag{10.4.A.4}$$

としても構いません。ただし, 実際には, 式 (10.4.A.3) や式 (10.4.A.4) のようにモーメント母関数やキュムラント母関数を介さない直接的な表現をするより, 母関数を介した表現をする方が便利な場合があります。

例えば最も基本的な確率分布モデルとも言える正規分布 normal distribution の確率密度関数は, 平均 m , 標準偏差 σ の場合に

$$f_{\text{normal}}(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right] \quad (10.4.A.4)$$

と表されるので、モーメント母関数 $M_{\text{normal}}(\theta)$ は式 (10.4.4) から

$$\begin{aligned} M_{\text{normal}}(\theta) &= \int_{-\infty}^{\infty} e^{\theta x} \frac{1}{\sqrt{2\pi\sigma}} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right] dx = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{\infty} \exp\left[-\frac{(x-m)^2}{2\sigma^2} + \theta x\right] dx \\ &= \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{\infty} \exp\left[-\frac{(x-m)^2 - 2\sigma^2\theta x}{2\sigma^2}\right] dx \\ &= \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{\infty} \exp\left[-\frac{(x-m-\sigma^2\theta)^2 - (m+\sigma^2\theta)^2 + m^2}{2\sigma^2}\right] dx \\ &= \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{\infty} \exp\left[-\frac{(x-m-\sigma^2\theta)^2 - 2m\sigma^2\theta - \sigma^4\theta^2}{2\sigma^2}\right] dx \\ &= \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{\infty} \exp\left[-\frac{(x-m-\sigma^2\theta)^2}{2\sigma^2} + m\theta + \frac{\sigma^2\theta^2}{2}\right] dx \\ &= \exp\left(m\theta + \frac{\sigma^2\theta^2}{2}\right) \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{\infty} \exp\left[-\frac{(x-m-\sigma^2\theta)^2}{2\sigma^2}\right] dx = \exp\left(m\theta + \frac{\sigma^2\theta^2}{2}\right) \quad (10.4.A.5) \end{aligned}$$

となり、式 (10.4.A.5) と式 (10.4.3) から正規分布のキュムラント母関数 $K_{\text{normal}}(\theta)$ は、

$$K_{\text{normal}}(\theta) = m\theta + \frac{\sigma^2\theta^2}{2} \quad (10.4.A.6)$$

となります。式 (10.4.A.6) と式 (10.4.A.2) とを比較すれば、正規分布の 1 階・2 階キュムラント $(\kappa_1)_{\text{normal}}$ 、 $(\kappa_2)_{\text{normal}}$ について、

$$(\kappa_1)_{\text{normal}} = m \quad (10.4.A.7)$$

$$(\kappa_2)_{\text{normal}} = \sigma^2 \quad (10.4.A.8)$$

となり、正規分布の 3 階以上のキュムラント $(\kappa_3)_{\text{normal}}$ 、 $(\kappa_4)_{\text{normal}}$ 、… はすべて 0 になることがわかります。

逆に「正規分布とは、3 階以上のキュムラントの値がすべて 0 になるような統計分布である」として、正規分布を特徴付けられることができます。すべての階数の粗モーメント $\mu'_k(\mu'_1, \mu'_2, \dots)$ を決めることができるのであっても（それでその統計分布を特定できることになるのですが）、粗モーメント $\mu'_k(\mu'_1, \mu'_2, \dots)$ を使うより、キュムラント $\kappa_k(\kappa_1, \kappa_2, \dots)$ を使う方が便利な場合は、他にも多くあります。 (↔)

(補足 10.4.B) 粗モーメントと確率変数の^{べき}冪平均 (↔)

統計分布の確率密度関数が $f(x)$ と表されるとき、その統計分布の k 階粗モーメント μ'_k は、式 (10.4.1) と (10.4.4) :

$$\mu'_k \equiv \left. \frac{\partial^k M(\theta)}{\partial \theta^k} \right|_{\theta=0} \quad (10.4.1)$$

$$M(\theta) \equiv \int_{-\infty}^{\infty} e^{\theta x} f(x) dx \quad (10.4.4)$$

によって定義されます。式 (10.4.1) と (10.4.4) とから、

$$\mu'_k \equiv \left. \frac{\partial^k}{\partial \theta^k} \int_{-\infty}^{\infty} e^{\theta x} f(x) dx \right|_{\theta=0} \quad (10.4.B.1)$$

$$\mu'_1 \equiv \left. \frac{\partial}{\partial \theta} \int_{-\infty}^{\infty} e^{\theta x} f(x) dx \right|_{\theta=0} = \left. \int_{-\infty}^{\infty} x e^{\theta x} f(x) dx \right|_{\theta=0} = \int_{-\infty}^{\infty} x f(x) dx = \langle x \rangle \quad (10.4.B.2)$$

$$\mu'_2 \equiv \left. \frac{\partial^2}{\partial \theta^2} \int_{-\infty}^{\infty} e^{\theta x} f(x) dx \right|_{\theta=0} = \left. \int_{-\infty}^{\infty} x^2 e^{\theta x} f(x) dx \right|_{\theta=0} = \int_{-\infty}^{\infty} x^2 f(x) dx = \langle x^2 \rangle \quad (10.4.B.3)$$

$$\mu'_3 \equiv \left. \frac{\partial^3}{\partial \theta^3} \int_{-\infty}^{\infty} e^{\theta x} f(x) dx \right|_{\theta=0} = \left. \int_{-\infty}^{\infty} x^3 e^{\theta x} f(x) dx \right|_{\theta=0} = \int_{-\infty}^{\infty} x^3 f(x) dx = \langle x^3 \rangle \quad (10.4.B.4)$$

$$\mu'_4 \equiv \left. \frac{\partial^4}{\partial \theta^4} \int_{-\infty}^{\infty} e^{\theta x} f(x) dx \right|_{\theta=0} = \left. \int_{-\infty}^{\infty} x^4 e^{\theta x} f(x) dx \right|_{\theta=0} = \int_{-\infty}^{\infty} x^4 f(x) dx = \langle x^4 \rangle \quad (10.4.B.5)$$

のような関係は自然に導かれます。 (↔)

(補足 10.4.C) 1階から4階のキュムラントと確率変数の^{べき}平均 (↔)

統計分布の確率密度関数が $f(x)$ と表されるとき、その統計分布の k 階粗モーメント μ'_k と k 階キュムラントは、式 (10.4.1)–(10.4.4) :

$$\mu'_k \equiv \left. \frac{\partial^k M(\theta)}{\partial \theta^k} \right|_{\theta=0} \quad (10.4.1)$$

$$\kappa_k \equiv \left. \frac{\partial^k K(\theta)}{\partial \theta^k} \right|_{\theta=0} \quad (10.4.2)$$

$$K(\theta) \equiv \ln M(\theta) \quad (10.4.3)$$

$$M(\theta) \equiv \int_{-\infty}^{\infty} e^{\theta x} f(x) dx \quad (10.4.4)$$

によって定義されます。

式 (10.4.1)–(10.4.4) の定義に従って、1階の粗モーメント μ'_1 とキュムラント κ_1 は

$$\begin{aligned} \mu'_1 &= M'(0) = \left. \frac{\partial}{\partial \theta} \int_{-\infty}^{\infty} e^{\theta x} f(x) dx \right|_{\theta=0} = \left. \int_{-\infty}^{\infty} e^{\theta x} x f(x) dx \right|_{\theta=0} \\ &= \int_{-\infty}^{\infty} x f(x) dx = \langle x \rangle \end{aligned} \quad (10.4.C.1)$$

$$\begin{aligned} \kappa_1 &= K'(0) = \left. \frac{\partial K(\theta)}{\partial \theta} \right|_{\theta=0} = \left. \frac{\partial \ln M(\theta)}{\partial \theta} \right|_{\theta=0} = \left. \frac{\partial M(\theta) / \partial \theta}{M(\theta)} \right|_{\theta=0} = \frac{M'(0)}{M(0)} \\ &= \int_{-\infty}^{\infty} x f(x) dx = \langle x \rangle \end{aligned} \quad (10.4.C.2)$$

と表されます。

2階の粗モーメント μ'_2 とキュムラント κ_2 は、

$$\mu'_2 = M''(0) = \left. \frac{\partial^2}{\partial \theta^2} \int_{-\infty}^{\infty} e^{\theta x} f(x) dx \right|_{\theta=0} = \left. \frac{\partial}{\partial \theta} \int_{-\infty}^{\infty} x e^{\theta x} f(x) dx \right|_{\theta=0} = \left. \int_{-\infty}^{\infty} x^2 e^{\theta x} f(x) dx \right|_{\theta=0}$$

$$= \langle x^2 \rangle \quad (10.4.C.3)$$

$$\begin{aligned} \kappa_2 = K''(0) &= \left. \frac{\partial^2 K(\theta)}{\partial \theta^2} \right|_{\theta=0} = \left. \frac{\partial^2 \ln M(\theta)}{\partial \theta^2} \right|_{\theta=0} = \left. \frac{\partial}{\partial \theta} \left[\frac{M'(\theta)}{M(\theta)} \right] \right|_{\theta=0} = \left. \frac{M''(\theta)M(\theta) - [M'(\theta)]^2}{[M(\theta)]^2} \right|_{\theta=0} \\ &= \frac{M''(0)M(0) - [M'(0)]^2}{[M(0)]^2} = \langle x^2 \rangle - \langle x \rangle^2 = \langle (x - \langle x \rangle)^2 \rangle \end{aligned} \quad (10.4.C.4)$$

となります。

3階の粗モーメント μ'_3 とキュムラント κ_3 は

$$\mu'_3 = M'''(0) = \langle x^3 \rangle \quad (10.4.C.5)$$

$$\begin{aligned} \kappa_3 = K'''(0) &= \left. \frac{\partial^3 K(\theta)}{\partial \theta^3} \right|_{\theta=0} = \left. \frac{\partial^3 \ln M(\theta)}{\partial \theta^3} \right|_{\theta=0} = \left. \frac{\partial^2}{\partial \theta^2} \left[\frac{M'(\theta)}{M(\theta)} \right] \right|_{\theta=0} \\ &= \left. \frac{\partial}{\partial \theta} \left\{ \frac{M''(\theta)M(\theta) - [M'(\theta)]^2}{[M(\theta)]^2} \right\} \right|_{\theta=0} = \left. \frac{M'''M + M''M' - 2M''M'}{M^2} - \frac{(M''M - M'^2) \times 2M'}{M^3} \right|_{\theta=0} \\ &= \left. \frac{M'''M - M''M'}{M^2} - \frac{2M''M'M - 2M'^3}{M^3} \right|_{\theta=0} = \left. \frac{M'''M^2 - M''M'M - 2M''M'M + 2M'^3}{M^3} \right|_{\theta=0} \\ &= \left. \frac{M'''M^2 - 3M''M'M + 2M'^3}{M^3} \right|_{\theta=0} = \langle x^3 \rangle - 3\langle x^2 \rangle \langle x \rangle + 2\langle x \rangle^3 = \langle (x - \langle x \rangle)^3 \rangle \end{aligned} \quad (10.4.C.6)$$

となります。

4階の粗モーメント μ'_4 とキュムラント κ_4 は

$$\mu'_4 = M''''(0) = \langle x^4 \rangle \quad (10.4.C.7)$$

$$\begin{aligned} \kappa_4 = K''''(0) &= \left. \frac{\partial^4 K(\theta)}{\partial \theta^4} \right|_{\theta=0} = \left. \frac{\partial^4 \ln M(\theta)}{\partial \theta^4} \right|_{\theta=0} = \left. \frac{\partial}{\partial \theta} \left(\frac{M'''M^2 - 3M''M'M + 2M'^3}{M^3} \right) \right|_{\theta=0} \\ &= \left. \frac{M''''M^2 + 2M''''M'M - 3M''''M'M - 3M''^2M - 3M''M'^2 + 6M''M'^2}{M^3} \right. \\ &\quad \left. - \frac{3M'(M'''M^2 - 3M''M'M + 2M'^3)}{M^4} \right|_{\theta=0} \\ &= \left. \frac{M''''M^2 - M''''M'M - 3M''^2M + 3M''M'^2}{M^3} - \frac{3M'(M'''M^2 - 3M''M'M + 2M'^3)}{M^4} \right|_{\theta=0} \\ &= \left. \frac{1}{M^4} \left(M''''M^3 - M''''M'M^2 - 3M''^2M^2 + 3M''M'^2M - 3M''M'M^2 + 9M''M'^2M - 6M'^4 \right) \right|_{\theta=0} \\ &= \left. \frac{1}{M^4} \left(M''''M^3 - 4M''''M'M^2 - 3M''^2M^2 + 12M''M'^2M - 6M'^4 \right) \right|_{\theta=0} \\ &= \langle x^4 \rangle - 4\langle x^3 \rangle \langle x \rangle - 3\langle x^2 \rangle^2 + 12\langle x^2 \rangle \langle x \rangle^2 - 6\langle x \rangle^4 \end{aligned} \quad (10.4.C.8)$$

となります。 (↔)

(補足 10.5.A) 観測可能な物理量の統計分布 (↔)

例えば位置 position や運動量 momentum は観測可能な物理量 observable です。ただし、固体の中での特定の原子 (原子核) の位置は (実験・観測によって決められると思われるかもしれませんが), 実験・観測で知ることができるのは, 原子の平均的な位置や, 平均位置からのずれの統計分布の特徴だけです。 (↔)

(補足 10.5.B) 高階キュムラントの重要度の低下 (↔)

例えば, 平均 $m = 10\,000$ のポアソン分布に従う場合, 標準偏差は $\sigma = 100$ と見積もられます。一万個の数値データ $\{X_1, X_2, \dots, X_n\}$ ($n = 10\,000$) から標本平均 \bar{X} を

$$\bar{X} = \frac{1}{n} \sum_{j=1}^N X_j \quad (10.5.B.1)$$

として求める場合, 標本平均の標準偏差は

$$\frac{\sigma}{\sqrt{n}} = \frac{100}{\sqrt{10\,000}} = 1$$

と予想されます。また, ポアソン分布では, すべての階数のキュムラントが平均に等しくなるはずですが。

モンテカルロ法を使って, 平均 10 000 であるような一万個のポアソン分布に従うランダムな数値を百組発生させ, 各組ごとの標本平均の平均と標準偏差を求めると, 確かにその結果は「平均 10 000.0, 標準偏差 1.0」となりました。同じように 2 階キュムラントの平均と標準偏差を求めると「平均 10 018, 標準偏差 138」となり, 三階キュムラントは「平均 12 800, 標準偏差 23 000」, 四階キュムラントは「平均 135 000, 標準偏差 5 200 000」となりました。

このように, 光子計数法による強度測定など, ポアソン分布に従う統計的な物理量の観測値を用いる場合, 高階のキュムラントになるほど有効数字が急激に低減します。常識的な測定で「意味のある数字がほとんど得られない」ことについて細かい議論をしても意味がないので, 「高階のキュムラントほど重要性が低くなる」こととなります。 (↔)

(補足 10.6.A) ディラックのデルタ関数 (↔)

ディラックのデルタ関数 Dirac delta function $\delta(x)$ は, 高校までの数学では「関数」に分類されるものではありません。日本では超関数と呼ばれることがあります。ディラックのデルタ関数の定義の仕方には, 色々な方法がありますが, 以下の式 (10.6.A.1), (10.6.A.2) で表される性格を持つものとするのが理解しやすいかもしれません。

$$\delta(x) = \begin{cases} +\infty & [x = 0] \\ 0 & [x \neq 0] \end{cases} \quad (10.6.A.1)$$

$$\int_{-\infty}^{\infty} \delta(x) dx = 1 \quad (10.6.A.2)$$

ディラックのデルタ関数は, 原点以外での値は 0 なのに, $[-\infty, \infty]$ の範囲で積分すれば積分値が 1 となるような関数です。

ディラックのデルタ関数を用いることによって, 量子力学が古典力学と無理なく対応づけられます。現代的な科学 (物理学・化学など) について理解するためには必須の知識であるとも言えます。制御工学や信号処

理の分野でもインパルス関数 impulse function として用いられるなど、理学・工学の分野で広く使われます。 (↔)

(補足 10.6.B) 畳込定理 (↔)

ディラックのデルタ関数 Dirac delta function $\delta(x)$ を使って畳込^{たたみこみ}を式 (10.6.2)

$$h(x) = f(x) * g(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(x - y - z) f(y) g(z) dy dz \tag{10.6.2}$$

のように表現します。畳込 $h(x)$ のフーリエ変換 Fourier transform $\mathfrak{F}(\xi)$ が

$$\begin{aligned} \mathfrak{F}(\xi) &= \int_{-\infty}^{\infty} h(x) e^{2\pi i \xi x} dx = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(x - y - z) f(y) g(z) dy dz e^{2\pi i \xi x} dx \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} \delta(x - y - z) e^{2\pi i \xi x} dx \right] f(y) g(z) dy dz \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{2\pi i \xi (y+z)} f(y) g(z) dy dz = \int_{-\infty}^{\infty} e^{2\pi i \xi y} f(y) dy \int_{-\infty}^{\infty} e^{2\pi i \xi z} g(z) dz \\ &= \mathfrak{F}(\xi) \mathfrak{G}(\xi) \end{aligned} \tag{10.6.B.1}$$

ただし,

$$\mathfrak{F}(\xi) \equiv \int_{-\infty}^{\infty} f(x) e^{2\pi i \xi x} dx \tag{10.6.B.2}$$

$$\mathfrak{G}(\xi) \equiv \int_{-\infty}^{\infty} g(x) e^{2\pi i \xi x} dx \tag{10.6.B.3}$$

となるように、「畳込定理」の関係は簡単に導けます。 (↔)

(補足 10.6.C) フーリエ変換と逆フーリエ変換 (↔)

函数 $f(x)$ のフーリエ変換 Fourier transform が

$$\mathfrak{F}(\xi) \equiv \int_{-\infty}^{\infty} f(x) e^{2\pi i \xi x} dx \tag{10.6.C.1}$$

と表されるとします。

$$\begin{aligned} \int_{-\infty}^{\infty} \mathfrak{F}(\xi) e^{-2\pi i \xi x} d\xi &\stackrel{\text{Eq.(10.6.C.1)}}{=} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y) e^{2\pi i \xi y} dy e^{-2\pi i \xi x} d\xi \\ &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} e^{2\pi i \xi (y-x)} d\xi \right] f(y) dy \stackrel{\int_{-\infty}^{\infty} \exp(2\pi i \xi x) d\xi = \delta(x)}{=} \int_{-\infty}^{\infty} \delta(y - x) f(y) dy \\ &= f(x) \end{aligned} \tag{10.6.C.2}$$

の関係から、フーリエ変換の逆変換（逆フーリエ変換 inverse Fourier transform）が

$$f(x) \equiv \int_{-\infty}^{\infty} \mathfrak{F}(\xi) e^{-2\pi i \xi x} d\xi \quad (10.6.C.3)$$

と表されることを確認できます。 ([↔](#))

参考文献 10

Ida, T. (2008). “Statistical properties of measured X-ray intensities affected by counting loss,” *J. Appl. Crystallogr.* **41**, 1019–1023. [doi: [10.1107/S0021889808030732](https://doi.org/10.1107/S0021889808030732)]

Stokes, A. R. (1948). “A numerical Fourier-analysis method for the correction of width and shapes of lines on X-ray powder photographs,” *Proc. Phys. Soc.* **61**, 382–391. [doi: [10.1088/0959-5309/61/4/311](https://doi.org/10.1088/0959-5309/61/4/311)]