

## 4 実験データの処理と統計学

すべての実験データには不確定要素が含まれるが、統計学を使うことで結論を導き出すことが出来る。統計学は実験計画を立てる時にも使われ、実験に必要な精度に応じて、実験装置による公差（許容誤差）や必要な測定時間などを考え、その分析結果から実現可能性を考えることが出来る。

ここでは、原子核・素粒子物理の実験に関するものに限定して統計学を学ぶ。初歩的な確率と組み合わせの理論の知識はあるものとする。

### 4.1 確率分布の特徴

統計学は確率過程を扱う。例えば、時間  $T$  の間の放射線源の崩壊数やサイコロなど、試行ごとに結果が変動するものが対象である。確率過程の結果は確率変数  $x$  によって表され、 $x$  の範囲は過程におけるすべての取りうる値である。

例) サイコロ

確率変数  $x = 1, 2 \dots 6$ ; サイコロの出目

確率分布関数  $P(x) = 1/6$ ; ある  $x$  の目が出る確率

過程によって確率変数は離散的な場合と連続的な場合がある。 $x$  が離散的な場合、 $P(x_i)$  は各  $x_i$  の値を取る確率を表す。連続的な場合、 $P(x')$  は、 $x' \in (x, x + dx)$  となる確率が  $P(x)dx$  となるような確率密度である。

#### 4.1.1 累積分布

$x_1 \leq x \leq x_2$  の範囲での確率分布は、累積分布関数

$$P(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} P(x)dx \quad (4.1)$$

で与えられる。ここで  $P(x)$  は連続型であるとし、離散型の場合は、

$$P(x_1 \leq x \leq x_2) = \sum_{i=1}^2 P(x_i) \quad (4.2)$$

で与えられる。慣例により、連続型、離散型ともに以下のように規格化される。

$$\int P(x)dx = 1 \quad (4.3)$$

$$\sum_i P(x_i) = 1 \quad (4.4)$$

これは、取りうるすべての  $x$  の確率を足すと 1 になることを意味している。そのため、 $P(x_i)$  及び  $\int P(x)dx$  は 0 以上 1 以下である。

#### 4.1.2 期待値

確率変数や確率変数関数の期待値、特に後者は重要な意味を持つ。 $P(x)$  で分布する確率変数を  $x$  とすると、 $x$  の期待値は

$$E[x] = \int xP(x)dx \quad (4.5)$$

で与えられ、これは単に平均の標準形である。積分は取りうるすべての  $x$  について行う。離散型では、

$$E[x] = \sum_i x_i P(x_i) dx \quad (4.6)$$

となる。 $x$  の関数  $f(x)$  の期待値も同様に、

$$E[f(x)] = \int f(x) P(x) dx \quad (4.7)$$

で表される。

簡単のため、以下、この章では連続型と仮定して議論していく。特に断りがない限り、離散型は  $\int$  を  $\sum$  に置き換えたものである。

#### 4.1.3 分布のモーメント:平均と分散

確率分布はそのモーメントにより特徴づけられる。 $r$  を整数として、いくつかの固定点  $x_0$  の  $r$  次のモーメントは、 $(x - x_0)^r$  の期待値で定義される。

実際に重要なモーメントは1次と2次のものである。最も重要な、1次の  $x_0 = 0$  でのモーメントは

$$\mu = E[x] = \int x P(x) dx \quad (4.8)$$

であり、これは(4.5)式と同じ  $x$  の平均である。

(4.8)式の平均と、繰り返し測定から計算される平均とは区別しなければならない。前者は理論上の分布から計算した理論上の平均値を意味し、後者は標本(サンプル)から得た実験的な平均値である。4.4.2章で見るように、標本平均は理論平均の概算である。以下、この章では  $\mu$  を理論平均とする。

次に特徴的な量は、以下の2次のモーメントである。

$$\sigma^2 = e[(x - \mu)^2] = \int (x - \mu)^2 P(x) dx \quad (4.9)$$

この  $\sigma^2$  は分散、 $\sigma$  は標準偏差である。分散は平均からのズレの二乗平均であり、標準偏差により分布の横幅を評価し確率変数  $x$  の平均からの変動がわかる。これもまた平均と同様、理論値と実験値を区別しなければならない。(詳しくは4.4章)

3次のモーメントは歪度と呼ばれ、対称、非対称を評価するときに使うこともあるが、3次以上は小さな情報しかないので滅多に使われない。

#### 4.1.4 共分散

これまで確率変数が1つの場合を考えてきたが、より一般的に多変量(多変数)の場合について考える。多変量分布での確率密度関数は  $P(x, y, z, \dots)$  である。

独立した確率変数  $x, y, \dots$  の平均、分散についてはこれまでと同様の定義を用いる。(ただし積分はあらゆる変数に関して行う)そして、3つ目の重要な量として  $x$  と  $y$  の共分散

$$\text{cov}(x, y) = E[(x - \mu_x)(y - \mu_y)] \quad (4.10)$$

を定義する。ここに、 $\mu_x, \mu_y$  は  $x, y$  の平均である。共分散はあらゆる変数のペアについて定義され、3変量分布(確率変数  $x, y, z$ )の場合、3つの共分散  $\text{cov}(x, y), \text{cov}(x, z), \text{cov}(y, z)$  が存在する。

共分散は 2 つの変数の間の線形相関を評価するもので、相関係数  $\rho$  で表されることが多い。

$$\rho = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y} \quad (4.11)$$

$\sigma_x, \sigma_y$  は  $x, y$  の標準偏差。相関係数は  $-1 \leq \rho \leq 1$  の値をとり、完全な線形相関のとき  $|\rho| = 1$  となる。変数が独立なら  $\rho = 0$  である。ただし逆は成り立たず、 $\rho = 0$  のときに言えるのは  $x, y$  に線形相関が無いということだけである。実際、線形でなければ  $y = x^2$  のように相関があっても  $\rho = 0$  となる。

ここで、変数が独立とは  $P(x, y) = P_1(x)P_2(y)$  が成り立つという意味である。

## 4.2 いくつかの一般的な確率分布

確率分布にはたくさんのものがあるが、理論的分布のグループに分類した場合、多くの物理問題は驚くほど少ないグループで記述、もしくは近似出来る。特に、二項分布、ポアソン分布、ガウス分布の 3 つの適用範囲は非常に広い。この節ではこれらの分布について学ぶ。

### 4.2.1 二項分布

コイントスのような、各独立な試行の結果が成功か失敗の 2 択であるような場合を考える。この時、 $N$  回試行して  $r$  回成功する確率は、次の二項分布で与えられる。

$$P(r) = \frac{N!}{r!(N-r)!} p^r (1-p)^{N-r} \quad (4.12)$$

ここで  $p$  は一回の試行で成功する確率である。

(4.12) 式は図 4.1 のように離散的な値を持つ。(図は様々な  $N, p$  での二項分布の例) (4.8), (4.9) 式より平均と分散を計算すると、

$$\mu = \sum_r r P(r) = Np \quad (4.13)$$

$$\sigma^2 = \sum_r (r - \mu)^2 P(r) = Np(1-p) \quad (4.14)$$

$P(r)$  は二項展開の  $r$  番目の項に他ならず、

$$\sum_r \frac{N!}{r!(N-r)!} p^r (1-p)^{N-r} = [(1-p) + p]^N = 1 \quad (4.15)$$

となることより、(4.12) 式は規格化されていることが分かる。この累積分布はやや複雑な形をしており、項の和という形では記述出来ない。項が多すぎなければ、各項を別々に計算してから和を取ればいいが、そうでなければ、累積二項分布の表計算によって求める。

$p$  が小さすぎず  $N$  が十分大きければ二項分布はガウス分布で近似できる。このガウス分布の平均、分散は (4.13), (4.14) 式のものである。実際、 $p \geq 0.05$  で  $N$  が 30 以上ではガウス分布はかなり良い近似になる。ただし、ガウス分布は連続型なので、近似によって離散型という二項分布の特徴が消えてしまうことに注意する。 $p \leq 0.05$  で  $Np$  が有限な場合では、二項分布はポアソン分布に近似することが出来る。

#### 4.2.2 ポアソン分布

二項分布の平均  $\mu = Np$  が有限となるよう注意して、 $p \rightarrow 0, N \rightarrow \infty$  の極限を取ったものがポアソン分布である。(4.12) 式に  $p = \mu/N$  を代入すると、

$$\begin{aligned}
 P(r) &= \frac{N!}{r!(N-r)!} \left(\frac{\mu}{N}\right)^r \left(1 - \frac{\mu}{N}\right)^{N-r} \\
 &= \frac{1}{r!} N(N-1)\cdots(N-r+1) \left(\frac{\mu}{N}\right)^r \left(1 - \frac{\mu}{N}\right)^{N-r} \\
 &= \frac{\mu^r}{r!} \left(1 - \frac{1}{N}\right) \left(1 - \frac{2}{N}\right) \cdots \left(1 - \frac{r-1}{N}\right) \left(1 - \frac{\mu}{N}\right)^{-r} \left(1 - \frac{\mu}{N}\right)^N \\
 N \rightarrow \infty \text{ で} \\
 &\quad \left(1 - \frac{1}{N}\right) \left(1 - \frac{2}{N}\right) \cdots \left(1 - \frac{r-1}{N}\right) \left(1 - \frac{\mu}{N}\right)^{-r} \rightarrow 1 \\
 &\quad \left(1 - \frac{\mu}{N}\right)^N \rightarrow e^{-\mu}
 \end{aligned}$$

以上より、ポアソン分布は、

$$P(r) = \frac{\mu^r e^{-\mu}}{r!} \quad (4.16)$$

となり、これは離散型である。1 回の試行でこの通りの結果が出る可能性はとても小さいが、放射性崩壊や粒子の反応といった試行回数の多いものはかなりよく記述される。

例)  $^{137}\text{Cs}$  の崩壊

半減期: 27 年

単位時間あたりの崩壊確率  $\lambda = \ln 2/27 = 0.026/\text{year} = 8.2 \times 10^{-10} \text{s}^{-1}$

実験に使う  $^{137}\text{Cs}$  の量は  $1\mu\text{g}$ : 含まれる原子核数は  $10^{15}$  個

各原子核がそれぞれ試行を行うので、平均崩壊数  $\mu$  は上の 2 つ積、 $8.2 \times 10^5/\text{s}$  となる。

ここで、(4.16) 式は  $N, p$  を含まず、その積である平均  $\mu$  だけの関数である。即ち、 $N, p$  以外の変数で  $\mu$  を表現しても問題ない。単位時間あたりの発生確率  $\lambda$  と単位計測時間  $t$  を使えば、 $\mu = \lambda t$  と書け、この時 (4.16) 式は次式で表される。

$$P(r) = \frac{(\lambda t)^r e^{-\lambda t}}{r!} \quad (4.17)$$

次に、(4.8) 式より  $\mu$  は確かに平均であるが、(4.9) 式より分散を求めてやると、

$$\sigma^2 = \mu \quad (4.18)$$

が成り立つことが分かる。よって標準偏差は  $\sigma = \sqrt{\mu}$  となる。

図 4.2 は様々な  $\mu$  でのポアソン分布である。非対称であるため、ピークや最大値は平均値に対応していないことに注意する。ただし、 $\mu$  の増加に伴い対照的になっていき、 $\mu \geq 20$  では  $\sigma^2 = \mu$  であるようなガウス分布に近似できる。今回の近似でも、ポアソン分布は離散型、ガウス分布は連続型であることに注意する。

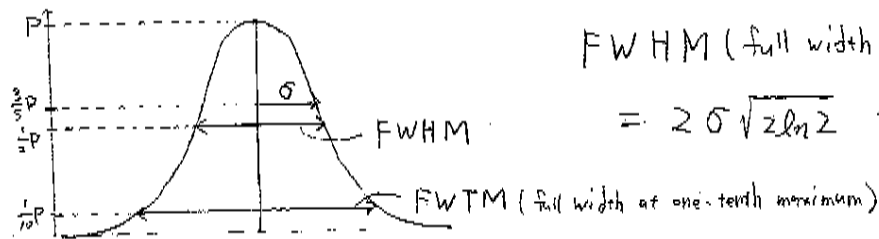
### 4.2.3 ガウス分布 (正規分布)

ガウス分布は統計学において最も重要な役割を果たし、最も使われている分布である。測定誤差、特に機器誤差はガウス分布によって記述される。したがってガウス分布は真の分布の良い近似になる。

ガウス分布は連続的で、確率密度は以下の式で与えられる。

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (4.19)$$

$\mu$ : 平均,  $\sigma$ : 標準偏差,  $\sigma^2$ : 分散



FWHM (full width at half maximum)

$$= 2\sigma\sqrt{2\ln 2} = 2.35\sigma \quad (4.20)$$

また  $\mu=0, \sigma^2=1$  のとき標準正規分布と言う。このときの変数を  $z$  で表して

$$z = \frac{x-\mu}{\sigma} \quad (4.21)$$

ガウス分布の実用的特徴は  $\sigma$  による積分間隔でのガウス分布下の面積であり、これにより、確率  $x$  が  $\mu$  からのずれが  $\pm\sigma$  以下の範囲に含まれる確率が 68.3%,  $\pm 2\sigma$  以下で 95.5%,  $\pm 3\sigma$  以下で 99.7% となっている。

4.2.4 カイ2乗分布 検定

カイ2乗分布は理論値と実験値を一致させる際に便利である。カイ2乗は次のように定義される。平均 $\mu_i$ 、標準偏差 $\sigma_i$ のガウス分布に従う独立のランダム変数 $x_i$ が $n$ 個あるとして、

$$u = \sum_{i=1}^n \left( \frac{x_i - \mu_i}{\sigma_i} \right)^2 \quad (4.22)$$

の和で表わされるものがカイ2乗である。ただしここでは $u = \chi^2$ としていく。また $u$ は以下の分布で表わされる

$$P(u) du = \frac{\left(\frac{u}{2}\right)^{\frac{\nu}{2}-1} \exp\left(-\frac{u}{2}\right)}{2 \Gamma\left(\frac{\nu}{2}\right)} du \quad (4.23)$$

$\Gamma\left(\frac{\nu}{2}\right)$ : ガンマ関数,  $\nu$ : 自由度 (整数)

(4.23)では

$$\mu = \nu, \quad \sigma^2 = 2\nu \quad (4.24)$$

である。

カイ2乗は理論的平均からの $x_i$ のばらつきを特徴づける。 $x_i$ がガウス分布のおりに分布するならば、平均して、それぞれが約1で $u = \nu$ になる。カイ2乗分布の有用性は仮定を試すために使えることである。測定値と仮定理論平均とのカイ2乗を計算することで、測定値において仮定した平均とのばらつきが妥当であるかどうか分かる。もしあつてもないカイ2乗の値が得られたら、使用された理論パラメータを疑わなければならぬ。

## 4.3 測定誤差と測定過程

実験における様々な測定は不確かさや誤差に左右される。測定誤差はサンプリング誤差として表れ得るが、その前にまず系統誤差と偶然誤差とを区別する必要がある。

### 4.3.1 系統誤差

系統誤差はデータの欠陥よりの不確かさである。例として、目盛りがゼロに合っていない電圧計の使用が考えられる。ゼロに合っていないので当然値がずれるので、それを考慮しなければならぬ。しかし目盛りをゼロに合わせようとしてもピッタリゼロに合うことはなく限界がある。

系統誤差で重要なことは、同じ機器によって得られた測定データは同じ方向に同じ量ずれることである。このことは測定値にはらつく偶然誤差とは対照的である。したがって系統誤差は同じ状況下でデータの集まりがとられたと見なせるとき最も重要である。あいにく、一貫した方法で系統誤差を扱い、分析することはできない。それぞれの実験はたいてい独立のもののみなされ、誤差の大きさを単独で評価できそうな原因を特定するのはとても難しい。

### 4.3.2 偶然誤差

偶然誤差は統計学の理論で取り扱うことができる。偶然誤差の不確かさは機器の不正確と観測される現象の固有の統計学的性質、またはその方から生じ、偶然誤差には統計誤差と機器誤差がある。

統計誤差 ... ランダムプロセスの測定から生じる誤差

実験者は全てを測定するのはとても無理なので、限られた大きさのランダムサンプルや真の値の評価としてのサンプルパラメータを使うことで、母集団や分布のパラメータを決定しようとする。この例は、放射性崩壊、陽子-陽子散乱などにおいてである。これらの過程ではすべて量子力学の確率的な法則によって振る舞われるので、与えられた期間での崩壊や散乱の数はランダム変数である。これらの過程での間心値は理論的確率分布の平均であり、1単位時間あたりの崩壊

や散乱の数を測定するとき、この分布からのサンプルが取られる。いいかえれば、変数  $X$  が  $X_1$  になる。繰り返し測定で  $X_2, X_3, \dots$  と得られていく。これらのデータから実験者はその平均値を評価できるようになるが、サンプルが限られているので不確かさがあり、これが測定誤差を表す。

## 機器誤差

テーブルの長さを定規で測定するような実験を考える。できるだけ正確に定規を読み取って一連の繰り返し測定をしてみる。すると値がずれることに気づき、もし柱状図で結果をプロットすると、形を成し始める明確な分布の概観が見られる。値が異なっているのは多くの小さな要素の結果であり、それは実験者にコントロールされずに、測定ごとに変わる。例えば、接合部のゆがみ、温度変化による収縮や膨張、実験者が目盛りとゼロに合わせ損ねるなど。これは全て機器誤差の元であり、機器という言葉には観測者も含まれる。支配下における要素が多くなれば、ずれの大きさはより小さくなり、機器がより正確であると言える。現実にはありえないが、完全な機器という理想状態では分布は測定された量の真の値に集中する  $\delta$  関数になる。

したがって、固定量の測定は抽象的な、つまり機器の不正確によって決定される理論分布からサンプルをとることである。機器誤差のほかに全ての場合作、分布はガウス分布である。系統誤差がない場合を仮定すると、ガウス分布の平均は機器の精度に比例した測定量の真の値と標準偏差に等しくなければならない。



#### 4.4 サンプルとパラメータの見積もり 最尤法

サンプルとは実験的手法であり、未知の分布のパラメータについて情報が得られるような方法である。大事なものは、

- サンプルが代表的なものでかつ(偏りのないこと)
- 一見正しくないデータがあっても除外しないこと

次に、得られたデータから最もふさわしい値を見積もる。ここでの最もふさわしい値とは見積もりと実際の値のずれ(分散)が最も小さいものことである。

見積もりにおける問題

- ① 最もふさわしい見積もりを決定すること
- ② 見積もりの不確かさを決定すること

以下では最尤法について記述する。

##### 4.4.1 サンプルモーメント

分布に基づきサンプル  $x_1, x_2, x_3, \dots, x_n$  について考える。その理論的な平均と分散はそれぞれ  $\mu, \sigma^2$  があるとす。これはサンプル母集団として知られている。  
(標本)

標本平均  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$  (相加平均) (4.25)

$n \rightarrow \infty$  とすると  
 $\mu = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n x_i$  (理論平均) (4.26)

標本分散  $s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$  (4.27)

$n \rightarrow \infty$  とすると  
 $\sigma^2 = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$

標本が多変数  $x, y$  の2変数の場合

標本共分散  $COV(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$  (4.28)

これからの極限をとると、理論的な共分散に収束する。

### 4.4.2 最大法

$$\text{条件付き確率 } P(A|B) = \frac{P(A \cap B)}{P(B)}$$

最大法はサンプルの理論分布が既知の場合にのみ適用できる。

(多くの物理実験の場合は、ガウス分布かポアソン分布である。)

ここでは、一般的に考え、標本が  $f(x|\theta)$  という分布に従うものとする。

(標本  $n$  個、観測値  $x_1, x_2, \dots, x_n$ )

分布関数のパラメータである  $\theta$  を定めたい。

まず、尤度関数  $L(\theta(x))$  は

$$L(\theta(x)) = f(x_1|\theta) f(x_2|\theta) \dots f(x_n|\theta) \quad (4.29)$$

$L(\theta(x))$  は  $x_1, x_2, \dots, x_n$  の一連の値が得られる確率であることがわかる。

$L$  が最大となるような  $\theta$  が、求めたい  $\theta$  の値である。すなわち、

$$\frac{dL}{d\theta} = 0$$

となる  $\theta$  を求めればよい。場合には、 $L$  が  $L$  の対数が取れやすくて、

$$\frac{d(\ln L)}{d\theta} = 0$$

とすることもできる。

この方程式を解いて得られた  $\hat{\theta}$  をパラメータ  $\theta$  の最大推定量という。

$\hat{\theta}$  は  $x_i$  を含む関数なので、 $x_i$  が変われば当然  $\hat{\theta}$  も変わる。

$\hat{\theta}$  の分散を考えよう。

$$\sigma^2(\hat{\theta}) = \int (\hat{\theta} - \theta)^2 L(\theta(x)) dx_1 dx_2 \dots dx_n \quad (4.32)$$

これは一般的な公式ではあるが、簡単な場合を除いて解析的に解くのは不可能。

より簡単に解くには、 $n$  がとても大きいときにはのみ使える近似法を用いると、

$$\sigma^2(\hat{\theta}) \approx - \left( \frac{d^2 \ln L}{d\theta^2} \right)^{-1} \quad (4.33)$$

(パラメータ  $\theta$  が 2つ以上あるとき、行列)

$$U_{ij} = - \frac{\partial^2 \ln L}{\partial \theta_i \partial \theta_j} \quad (4.34)$$

を用いて、近似分散は

$$\sigma^2(\hat{\theta}_i) \approx (U^{-1})_{ii} \quad (4.35)$$

と表せる。

$\sigma^2(\hat{\theta})$  の式'では  $\hat{\theta}$  の平均値を理論値  $\theta$  としたが、これは  $n$  が無限の場合にしか適用されない。~~よもやかしが、実際には  $\hat{\theta}$  が~~

この性質をもっている推定量は不偏である。

実際には  $\hat{\theta}$  が不偏性を備えていなくても  $\sigma^2(\hat{\theta})$  の式'は適用できる。

最大法にはもう一つ便利な性質があり、もし  $u = f(\theta)$  という量があるとき、 $u$  の最良の見積りは  $\hat{u} = f(\hat{\theta})$  と表せることである。

次に、実際に最大法をポアソン分布とガウス分布に当てはめてみる。

#### 4.4.3 ポアソン分布に対する推定量

$n$ 個のサンプル  $x_1, x_2, x_3, \dots, x_n$  が平均  $\mu$  のポアソン分布から得られたものとする。

この場合の尤度関数は

$$L(\mu|x) = \prod_{i=1}^n \frac{\mu^{x_i}}{x_i!} \exp(-\mu) = \exp(-n\mu) \prod_{i=1}^n \frac{\mu^{x_i}}{x_i!} \quad (4.36)$$

対数をとって、

$$L^* = \ln L = -n\mu + \sum x_i \ln \mu - \sum \ln x_i! \quad (4.37)$$

微分して、0に結果を設定すると

$$\frac{dL^*}{d\mu} = -n + \frac{1}{\mu} \sum x_i = 0 \quad (4.38)$$

よって、

$$\hat{\mu} = \frac{1}{n} \sum x_i = \bar{x} \quad (4.39)$$

すなわち、 $\mu$  の最大推定量は標本平均と一致する。

$\bar{x}$  の分散は (4.33) から求めることができる。ここで (4.9) 式'の定義すると

$$\sigma^2(\bar{x}) = E[(\bar{x} - \mu)^2] \quad (4.40)$$

これを標本平均に適用すると

$$E[(\bar{x} - \mu)^2] = E\left[\left(\frac{1}{n} \sum x_i - \mu\right)^2\right] = \frac{1}{n^2} E\left[\left(\sum x_i - n\mu\right)^2\right] = \frac{1}{n^2} E\left[\sum (x_i - \mu)^2\right] \quad (4.41)$$

合計の平方に拡張すると

$$\left[\sum (x_i - \mu)\right]^2 = \sum_i (x_i - \mu)^2 + \sum_{i \neq j} (x_i - \mu)(x_j - \mu) \quad (4.42)$$

(4.42) 式の期待値をとると、第2項が消えて、

$$\sigma^2(\bar{x}) = \frac{1}{n^2} E[\sum (x_i - \mu)^2] = \frac{1}{n^2} \sum E[(x_i - \mu)^2] = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n} \quad (4.43)$$

これは分布に依らない一般的な結果である。

ポアソン分布では  $\sigma^2 = \mu$  があり、推定ポアソン平均の誤差は、

$$O(\hat{\mu}) = \sqrt{\frac{\mu}{n}} \approx \sqrt{\frac{\hat{\mu}}{n}} = \sqrt{\frac{\bar{x}}{n}} \quad (4.44)$$

ここで推定値  $\hat{\mu}$  は理論値  $\mu$  が置き換えられる。

#### 4.4.4 ガウス分布に対する推定量

$n$  個の標本が、ガウス分布と同じ分布に従うとき、その尤度関数は

$$L = \prod_{i=1}^n \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x_i - \mu)^2}{2\sigma^2}\right] \quad (4.45)$$

対数をとると、

$$L^* = \ln L = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2} \sum \frac{(x_i - \mu)^2}{\sigma^2} \quad (4.46)$$

$\mu, \sigma^2$  それぞれについて推定量を求めるため、それぞれが偏微分  $L^*$  に等しく 0 とすると、

$$\frac{\partial L^*}{\partial \mu} = \sum \frac{x_i - \mu}{\sigma^2} = 0 \quad (4.47)$$

$$\frac{\partial L^*}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2} \sum \left(\frac{x_i - \mu}{\sigma}\right)^2 \frac{1}{\sigma^2} = 0 \quad (4.48)$$

(4.47) を解くと、

$$\hat{\mu} = \frac{\sum x_i}{n} = \bar{x} \quad (4.49)$$

つまり、ガウシアン分布の理論平均の ~~最良の推定量~~ は、標本平均である。  
 最良推定量

(4.43)の一般的な結果から、推定の不確かさは

$$\sigma(\bar{x}) = \frac{\sigma}{\sqrt{n}} \quad (4.50)$$

これは平均の標準偏差である。すなわち、 $n$ が大きい(測定回数が多い)ほど、 $\bar{x}$ (標準偏差)の誤差が小さくなる。

$\sigma$ は未知である。(4.48)を解いて  $\sigma^2$ を推定すると

$$\hat{\sigma}^2 = \frac{1}{n} \sum (x_i - \mu)^2 \approx \frac{1}{n} \sum (x_i - \bar{x})^2 = S^2 \quad (4.51)$$

これは偏った推定量であり、 $E[S^2] = \sigma^2 - \sigma^2/n = (n-1)\sigma^2/n$  ともなる。

$n$ が大きければ、真の分散  $\sigma^2$  に近づくので問題ないが、 $n$ が小さいとき、 $\sigma^2$  は  $S^2$  よりも小さい値が推定されることになる。

理由は簡単で、小さなサンプルでは平均から大きく外れたサンプルが得られることがまれであるから、サンプルの分散が小さめに見積もられることにある。

よって、実用的には

$$\hat{\sigma}^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} \quad (4.52)$$

という推定量が用いられる。

$\hat{\sigma}^2$ の分散と  $\hat{\sigma}$ の標準偏差は計算すると、

$$\sigma^2(\hat{\sigma}^2) = \frac{2\sigma^4}{n-1} \approx \frac{2\hat{\sigma}^4}{n-1} \quad (4.53)$$

$$\sigma(\hat{\sigma}) = \frac{\sigma}{\sqrt{2(n-1)}} \approx \frac{\hat{\sigma}}{\sqrt{2(n-1)}} \quad (4.54)$$

となる。

## 4.4.5 加重平均

ある同位量を精度の異なる計器で測定する場合、平均や標準偏差は今までと同いように決まらぬので、それぞれの  $x_i$  に対し、誤差  $\sigma_i$  に応じて重みをかけて計算する。

同位平均  $\mu$  で、異なる標準偏差  $\sigma_i$  を持つガウス分布に従うサンプル  $x_1, x_2, \dots, x_n$  (標準) において、尤度関数は (4.45) 式から、

$$L = \prod_{i=1}^n \frac{1}{\sigma_i \sqrt{2\pi}} \exp \left[ -\frac{(x_i - \mu)^2}{2\sigma_i^2} \right]$$

最大法を用いて、これを計算すると、

$$\frac{\partial (\ln L)}{\partial \mu} = \sum_i \left( \frac{x_i - \mu}{\sigma_i^2} \right) = 0$$

$$\therefore \mu = \frac{\sum x_i / \sigma_i^2}{\sum 1 / \sigma_i^2} \quad (4.55)$$

(4.55) から加重平均の誤差は

$$\sigma^2(\hat{\mu}) = - \left( \frac{\partial^2 (\ln L)}{\partial \mu^2} \right)^{-1} = \frac{1}{\sum 1 / \sigma_i^2} \quad (4.56)$$

(4.55)、(4.56) は  $\sigma_i$  が全て等しい場合は、(4.49)、(4.50) に一致する。

## 4.5 適用例

### 4.5.1 一連の測定による平均と誤差

#### 例 4.1

あるものの長さの測定結果が表のようになったとする。ここで、器具による測定はガウス分布に従う。よって、平均は (4.49) から

$$\bar{x} = 17.61533$$

一方標準偏差が (4.52) から  $5.855 \times 10^{-3}$  となるので平均からの誤差は (4.50) から

$$\sigma(\bar{x}) = 5.855 \times 10^{-3} / \sqrt{15}$$

よって

$$x = 17.616 \pm 0.002$$

### 4.5.2 異なる誤差の統合データ

#### 例 4.2

ミューオンの寿命の測定について表のように異なる誤差のデータが得られたとする。これについては加重平均をとり、(4.55)(4.56) から

$$\tau = 2.1970 \pm 0.0006 \mu\text{s}$$

となる。

### 4.5.3 カウント率とその誤差

#### 例 4.3

$^{22}\text{Na}$  線源を測定する検出器からの 1 分間でのカウント数が表のようになったとする。これは Poisson 分布に従うので (4.39) と (4.44) より

$$\bar{\mu} = 2205.6$$

$$\sigma(\bar{\mu}) = \sqrt{\frac{2205.6}{5}} = 21$$

ここで 1 分ずつ区切るのではなく 5 分一気に測ったとみるとどうなるか? 5 分間当たりのカウント数は合計の 11028 で誤差は測定が一回なので  $\sigma = \sqrt{11028} = 105$  これを 5 で割ってやると先ほどと同じ  $2206 \pm 21$  となる。

(Null)

#### 4.5.4 帰無実験：カウントや観測のないときの信頼限界の設定

何かの反応や崩壊を時間  $T$  の間測定し、その反応や崩壊を一度も確認できなかったとする。このとき、その寿命はどうなるか？まず、平均反応率を  $\lambda$  とすると時間  $T$  の間に反応が観測れない確率は、

$$P(0 | \lambda) = \exp(-\lambda T) \quad (4.57)$$

これを (全ての  $\lambda$  を一般的に等価としたうえで) 時間  $T$  の間に反応が観測されないときの  $\lambda$  に対する確率分布と捉えると、 $\lambda \leq \lambda_0$  である確率は (4.1) から規格化を考慮に入れて

$$P(\lambda \leq \lambda_0) = \int_0^{\lambda_0} T \exp(-\lambda T) d\lambda = 1 - \exp(-\lambda_0 T) \quad (4.58)$$

これは 0 から  $\lambda_0$  に対する信頼レベル: *confidence level* (CL) として知られ、(4.58) より

$$\lambda_0 = \frac{1}{T} \ln(1 - \text{CL}) \quad (4.59)$$

##### 例 4.4

50g の  $^{82}\text{Se}$  を 100 日間観測して、ニュートリノレスなダブル  $\beta$  崩壊が起こるかどうかを調べるとする。(この反応はレプトン数の保存により通常は禁止されているが、最近の理論によると許されるらしい。) 器具の検出効率は 20%。その結果、反応は見られなかったとする。このとき CL の値を 90% とすると (4.59) から

$$\lambda \leq \lambda_0 = -\frac{1}{100 \times 0.2} \ln(1 - 0.9) = 0.115 \text{ day}^{-1}$$

ここで 50g の  $^{82}\text{Se}$  中の原子核の数は

$$N = \frac{N_a}{82} \times 50 = 3.67 \times 10^{23}$$

よって一個の原子核当たりでは観測時間が  $N$  倍されたとみて (4.59) より

$$\lambda \leq \frac{0.115}{3.67 \times 10^{23}} = 3.13 \times 10^{-25} \text{ day}^{-1}$$

よって寿命は

$$\tau \geq 8.75 \times 10^{21} \text{ years} \quad 90\% \text{ CL}$$

##### 例 4.5

発信機から  $N$  回信号を送ってあるシステムによって  $N$  回全てが観測できたとする。このときこのシステムの効率の下限を考える。



まず、効率を  $\epsilon$  とすると  $r$  回観測する確率は

$$P(N, r) = {}_N C_r \epsilon^r (1 - \epsilon)^{N-r}$$

よって

$$P(N, N) = \epsilon^N$$

よって  $N$  回信号が観測される時  $\epsilon = \epsilon'$  である確率は、 $\epsilon$  が定数である限りにおいては

$$P(\epsilon = \epsilon') \propto \epsilon'^N$$

規格化して

$$P(\epsilon) = (N + 1)\epsilon^N$$

よって下限  $\epsilon_0$  にたいして

$$CL = \int_{\epsilon_0}^1 P(\epsilon') d\epsilon' = 1 - \epsilon_0^{N+1}$$

$$\epsilon_0 = (1 - CL)^{1/(N+1)}$$

#### 4.5.5 カウント間の時間分布

平均崩壊率  $\lambda$  の放射線源を観測して時間  $T$  の間崩壊が観測されない確率は (4.57) で与えられる。これを崩壊が観測されなかったときの観測時間  $T$  の確率密度としてとらえ規格化してやると

$$P(T) = \lambda \exp(-\lambda T) \quad (4.60)$$

となる。