

1998年9月22日

<<第13週 ポアソン分布とガウス分布，ガウス積分と統計>>

<記法上の注意>

以下の文中で， a^2 は a の2乗を意味する． b_n や b_{n+1} はサブスクリプト(下付文字)が n や $n+1$ であることを意味する．

<目次>

試行と確率
 集合と確率
 平均と期待値
 確率と濃度
 時間平均と空間平均とエルゴードの定理
 二項分布
 ガウス分布
 ポアソン分布
 確率変数
 平均値と高次モーメント
 標準偏差と分散
 中心極限定理
 多変数ガウス分布関数
 最小2乗法と χ^2 法
 乱数とモンテカルロ法

<試行と確率>

いよいよ本講義の最終回になった．ここで，自然科学一般で最も重要な数学概念である，確率(probability)と統計(statistics)を学ぼう．

まず確率の概念の復習をしよう．確率という概念は，パスカルの時代に確立された．特に，確率の理論に貢献した人は，「人間は考える葦(あし)である」という言葉で有名なパスカルとフェルマーの最終定理($x^n + y^n = z^n$ を満たす整数の組 x, y, z は， $n > 2$ では存在しない)で有名なフェルマー(F. Fermat, 1601-1665)であった．(この定理は，300年後のごく最近1995年に，A.Wilesによってはじめて証明された.)

さて，何かを行うことを試行(trial or experiment)と呼ぶ．もし一回一回の試行が何の影響もない状況で独立に行えるなら，この試行を独立試行と呼ぶ．例えば，コインを飛ばして裏がでるか表がでるか試みること(coin toss)やサイコロを振ってどの目がでるか試みるようなことである．このとき，何回か試行を繰り返して，その回数を N とすると，その間に特定の事象(event) A --例えば，表が出る--の起こった回数を $N_{\{A\}}$ とする．この事象の起こった確率 $P(A)$ は，次のように与えられる:

$$P(A) = N_{\{A\}}/N. (1)$$

一方，裏の出た回数を $N_{\{B\}}$ とすると，裏の出た確率 $P(B)$ は，

$$P(B) = N_{\{B\}}/N. (2)$$

コインの場合，裏と表が同時に現われることはないから，表裏が同時に現われた回数 $N_{\{AB\}}$ はゼロである，すなわち

$$N_{\{AB\}} = 0. \quad (3)$$

したがって、その確率は、

$$P(AB) = 0. \quad (4)$$

同様に、裏も表も出ないという可能性もないから、裏も表も出ない回数 $N_{\{E\}}$ はゼロ、すなわち、

$$N_{\{E\}} = 0. \quad (5)$$

したがって、その確率は、

$$P(E) = 0. \quad (6)$$

ここでは当然、以下の関係が成り立つ：

$$N_{\{A\}} + N_{\{B\}} + N_{\{AB\}} + N_{\{E\}} = N, \quad (7)$$

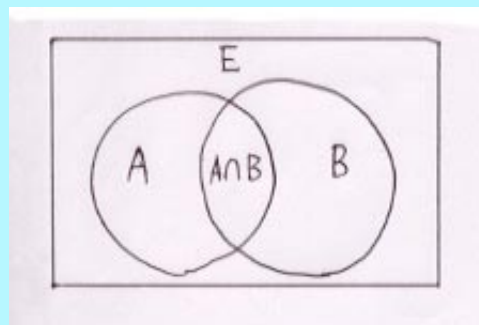
$$P(A) + P(B) + P(AB) + P(E) = 1. \quad (8)$$

一般に、試行の回数 N が有限のとき、確率 $P(A)$, $P(B)$, $P(AB)$, $P(E)$ の値はある値のまわりで揺らいでいる。この確率を非先見的(a posteriori)確率と呼ぶ。しかし、 N の試行の後、それぞれ一定の値に近づくと考えられる。この確率のことを、先見的(a priori)確率と呼ぶ、コインの表裏の確率がそれぞれ $1/2$ であるというのが、先見的確率のことである。

<集合と確率>

以上の確率論の基礎は、集合論(set theory)とうまく結び付く。というのも、物事が生じる事象が、集合の概念を満たすからである。

一般には、 $N_{\{AB\}}$ と $N_{\{E\}}$ はゼロでない。そこで、 A が起こった事象を集合 A 、 B が起こった事象を集合 B とし、 A と B が共に起こった事象を集合 $A \cap B$ と呼ぶと、 A と B がどちらも起こらない事象 E は集合 E とすることができる。この状況は、以下のようにいわゆるベン図に書くことができる。



したがって、この集合の考え方をを使うと、 A と B が共に起こる確率 $P(A \cap B)$ は、次のように書くことができる：

$$P(A \cap B) = P(A) \cdot P(B). \quad (9)$$

同様に考えて、 A と B のいずれかが生じる確率 $P(A \cup B)$ は以下のように書ける：

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). \quad (10)$$

ここで、

$$P(A) = (N_{\{A\}} + N_{\{A \cap B\}}) / N, \quad (11)$$

$$P(B) = (N_{\{B\}} + N_{\{A \cap B\}}) / N, \quad (12)$$

$$P(A \cap B) = N_{\{A \cap B\}} / N. \quad (13)$$

では、Bが起こってからAが起こる確率はどうなるだろうか？この確率を $P(A|B)$ と書こう。これは、次のように定義される：

$$P(A|B) = N_{\{A \cap B\}} / (N_{\{B\}} + N_{\{A \cap B\}}). \quad (14)$$

同様に、Aが起こってからBが起こる確率 $P(B|A)$ は、

$$P(B|A) = N_{\{A \cap B\}} / (N_{\{A\}} + N_{\{A \cap B\}}). \quad (15)$$

(11)-(15)より、以下の関係を得る：

$$P(A \cap B) = P(A)P(B|A) = P(B)P(A|B). \quad (16)$$

もし2つの出来事がお互いに相反する(mutually exclusive)なら、つまり両方同時に起こることがなければ、

$$P(A \cap B) = P(AB) = 0, \quad (17)$$

このとき、(10)から

$$P(A \cup B) = P(A) + P(B). \quad (18)$$

もし事象AとBが、統計的に独立である(statistically independent)なら、つまり

$$P(A \cap B) = P(A)P(B), \quad (19)$$

このとき、(16)から

$$\begin{aligned} P(A|B) &= P(A), \\ P(B|A) &= P(B). \end{aligned} \quad (20)$$



賭博者の友だち

ブレーズ・パスカルは、下のピエール・ド・フェルマとともに確率論をつくり上げたが、ある賭博師から、ダイス・ゲームを中断したとき賭け金をどのように分ければよいかときかれたことが、研究の動機だった。



片手間の数学者

ピエール・ド・フェルマはパスカルと共同で確率論をつくり上げ、さらに整数論の研究にも貢献し、微分法発達の推進力ともなった人である。フェルマの本職は地方議員であり、数学はほんの片手間仕事であった。

<平均と期待値>

さて、平均と期待値を考えて見よう。N個の数の組(例えば、N人のテストの点数)、 $x_{\{1\}}, x_{\{2\}}, \dots, x_{\{N\}} (>0)$ があるとしよう。このとき、これらの平均(average or mean)は、以下のように定義される:

$$\langle x \rangle_{\{s\}} = (x_{\{1\}} + x_{\{2\}} + \dots + x_{\{N\}}) / N. \quad (21)$$

これを相加(算術)平均(arithmetic mean)と呼ぶ。一方、相乗(幾何)平均(geometric mean)は、

$$\langle x \rangle_{\{p\}} = (x_{\{1\}} x_{\{2\}} \dots x_{\{N\}})^{1/N}, \quad (22)$$

調和平均(harmonic mean)は、

$$\langle x \rangle_{\{h\}} = N / (1/x_{\{1\}} + 1/x_{\{2\}} + \dots + 1/x_{\{N\}}). \quad (23)$$

これらをさらに一般化して、r次の平均を定義することができる:

$$\langle x \rangle_{\{r\}} = [(1/N) \sum_{j=1}^N \{x_{\{j\}}\}^r]^{1/r}. \quad (24)$$

この記法を使うと、次を得る:

$$\begin{aligned} \langle x \rangle_{\{1\}} &= \langle x \rangle_{\{s\}}, \\ \langle x \rangle_{\{0\}} &= \langle x \rangle_{\{p\}}, \\ \lim_{r \rightarrow 0} \langle x \rangle_{\{r\}} &= \langle x \rangle_{\{p\}}, \\ \langle x \rangle_{\{-1\}} &= \langle x \rangle_{\{h\}}. \end{aligned} \quad (25)$$

今度は、期待値(expectation value)を考えてみよう。N個の数の組、 $x_{\{1\}}, x_{\{2\}}, \dots, x_{\{N\}}$ があり [この集合のことをアンサンブル(ensemble)とよぶ]、そのそれぞれが起こる確率を、 $P(x_{\{1\}}), P(x_{\{2\}}), \dots, P(x_{\{N\}})$ としよう。もしこれらのどれかが同時に起こらないとすると、

$$P(x_{\{1\}}) + P(x_{\{2\}}) + \dots + P(x_{\{N\}}) = 1. \quad (26)$$

このとき、 x の期待値は、以下のように定義される：

$$\langle x \rangle = \sum_{j=1}^N x_{\{j\}} P(x_{\{j\}}). \quad (27)$$

ここで、もしすべての事象が等確率(equal probability)で起こるとすると、すなわち、

$$P(x_{\{1\}}) = P(x_{\{2\}}) = \dots = P(x_{\{N\}}) = P, \quad (28)$$

これを(27)へ代入すると、

$$P = 1/N. \quad (29)$$

(29)を(27)へ代入すると、通常の数平均(17)を得る。したがって、(27)を一般化すると、 r 次の期待値を以下のように定義することができる：

$$\langle x \rangle_{\{r\}} = \left[\sum_{j=1}^N x_{\{j\}}^r P(x_{\{j\}}) \right]^{1/r}. \quad (30)$$

このように、いわゆる平均と期待値はまったく等価なものである。

<確率と濃度>

さて、ある物事が全体の事象の中でどれほどの割合で起こるか考えるのが確率であった。しかし、これに似て、ある物(例えば、粒子)が全体の物に占める割合を考えることができる。

もしある系(システム, system)の中に、N個の粒子があり、その内Aという粒子が $N_{\{A\}}$ 個存在するとき、この粒子Aを1体積当りに見出す確率(density) $\rho(A)$ は、

$$\rho(A) = N_{\{A\}} / (NV) \quad (31)$$

である。これを粒子の密度あるいは濃度(density)と呼ぶ。このとき、この系の中の微小体積 dV の領域の中に存在するAの数は、

$$dN_{\{A\}}(dV) = dV N_{\{A\}} / (NV) \quad (32)$$

である。したがって、全体積中にあるAの割合は、

$$dN_{\{A\}}(dV) / N = \rho(A) dV. \quad (33)$$

<時間平均と空間平均とエルゴードの定理>

一般に、自然科学で扱う物理量(physical quantity)や人間社会で扱う観測量(observable)は、時間的に変動していることが多い。例えば、経済の株価などは、時々刻々と変動している。そのため、観測量の時間平均という概念も極めて重要な意味を持つ。そこでまず、この時間平均を考えてみよう。

ある時間変動する観測量を $x(t)$ としよう。この観測量の、ある時刻 $T_{\{0\}} = 0$ から次の時刻 T までの時間平均を次のように定義する：

$$\langle x \rangle_{\{T\}} = \frac{1}{T} \int_0^T x(t) dt. \quad (34)$$

一方，前セクションまでに述べた平均をアンサンブル平均(ensemble average)と呼ぶ，そして，時間平均と違うことを意識的に区別して， $\langle x \rangle$ と書くのである．

さてここに，物理学で非常に重要な定理--エルゴードの定理(ergodic theorem)--がある．それは， T の時間平均 $\langle x \rangle$ とアンサンブル平均 $\langle x \rangle$ が等しい:

$$\lim_{T \rightarrow \infty} \langle x \rangle_{\{T\}} = \langle x \rangle. \quad (35)$$

この等式は多くの物理学系で成り立つのであるけれども，その完全な証明は未だになされていない．

<二項分布>

さて，この辺で，より具体的に確率を求めて行こう．今，表(a)裏(b)だけあるコイン投げを考えてみよう．そこで，もし

$$aaba\dots b \quad (36)$$

のように，表がM，裏がN-Mだけあるとするなら，このときの場合の数は，2項展開:

$$(p + q)^N = \sum_{M=0}^N \frac{N!}{M!(N-M)!} p^M q^{N-M}, \quad (37)$$

の係数を使って，

$$\frac{N!}{(N-M)! M!} = \frac{N!}{M!} \quad (38)$$

と書ける．同様にして，その確率P(M)は，

$$p + q = 1 \quad (39)$$

を(37)に代入して，

$$P(M) = \frac{N!}{(N-M)! M!} p^M (1-p)^{N-M}, \quad (40)$$

$$P(M) = 1. \quad (41)$$

$$M=1$$

ここで、 p は a の出やすさを表わす確率である。この(40)の分布関数を2項分布(binomial distribution)と呼ぶ。

<ガウス分布>

この2項分布は、 N でどのような性質を持つだろうか？比較的理解しやすい2つの極限がある。今、 $N \gg M \gg 1$, $Np \gg 1$ の場合を考えると、(40)は以下のように書ける：

$$\begin{aligned} P(M) &= [1/(2^{-n})^{1/2}] \times \\ & (M/N)^{-M-1/2} [(N-M)/N]^{-N+M-1/2} p^M (1-p)^{N-M} \\ &= [1/(2^{-n})^{1/2}] \exp[-(M+1/2) \ln(M/N) \\ & - (N-M+1/2) \ln((N-M)/N) + M \ln p \\ & + (N-M) \ln(1-p)]. \quad (42) \end{aligned}$$

ここで、次のスターリング(Stirling)の公式を使った: $n \gg 1$ に対して、

$$n! \approx (2\pi n)^{-1/2} (n/e)^n [1 + 1/(12n) + 1/(288n^2) + \dots]. \quad (43)$$

(42)で、

$$M = Np + \delta, \quad \delta \ll Np \quad (44)$$

のように置くと、 $P(M)$ は次のように書ける：

$$P(M) = [1/(2\pi Np(1-p))^{1/2}] \exp[-\delta^2/(2Np(1-p))], \quad (45)$$

$$= [Np(1-p)]^{-1/2}. \quad (46)$$

(45)は確率を表わすから、当然次の規格化条件を満たさなくてはならない:

$$P(\delta) d\delta = 1. \quad (47)$$

この分布をガウス分布(Gaussian distribution)あるいは正規分布(normal distribution)と呼ぶ。はこの分布関数(distribution function)の幅に関する尺度(measure)を意味し、標準偏差(standard deviation)と呼ぶ。このガウス分布関数のことを、単にガウシアン(Gaussian)と呼んだり、誤差関数(error function)と呼ぶことがある。

この関数をなぜ誤差関数と呼ぶのか考えてみよう。今、 k 個の観測量、 $x_{\{1\}}, x_{\{2\}}, \dots, x_{\{k\}}$ があるとき、これらはそれぞれ真の平均値 $\langle x \rangle (= 0)$ から多少ずれている。このずれを表わす量 δ^2 :

$$\delta^2 = (x_{\{1\}} - \langle x \rangle)^2 / \langle x^2 \rangle + (x_{\{2\}} - \langle x \rangle)^2 / \langle x^2 \rangle + \dots + (x_{\{k\}} - \langle x \rangle)^2 / \langle x^2 \rangle. \quad (48)$$

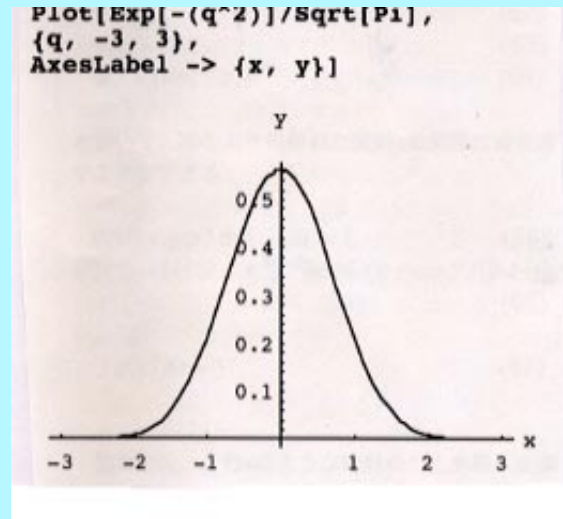
この量を $\langle x \rangle$ の関数と見て、これを最小にする $\langle x \rangle$ を求めると、

$$d\delta^2 / d\langle x \rangle = [x_{\{1\}} + x_{\{2\}} + \dots + x_{\{k\}} - k\langle x \rangle] / \langle x^2 \rangle = 0 \quad (49)$$

として、 k 個の量の平均の定義式を得る：

$$\langle \rangle = (_1 + _2 + \dots + _k) / k. \quad (50)$$

そして、(48)から $\exp[-x^2/2]$ を作ると、ガウス分布の形が得られる。以上の理由から、ガウス分布関数を誤差関数と呼ぶのである。



<ポアソン分布>

もう一つの極限は、 $N \rightarrow \infty$, $p \rightarrow 0$ で、

$$Np = s = \text{有限} \quad (51)$$

となる場合である。つまり、 a の起こる割合が b が起こる割合より極めて小さく、起こりにくい場合である。ガウス分布の場合と同様に、 $M \ll N$ として、(40)を考えると、

$$\frac{N!}{(N-M)!} = N(N-1)\dots(N-M+1) \approx N^M, \quad (52)$$

$$(1-p)^{N-M} \approx (1-p)^N (1-p)^{s/p} \approx e^{-s}. \quad (53)$$

これらを(40)に使うと、

$$P(M) = \frac{1}{M!} s^M e^{-s}. \quad (54)$$

ここで、

$$P(M) = 1. \quad (55)$$

この分布のことをポアソン分布(Poisson distribution)と呼ぶ。この分布の標準偏差は、

$$= [np(1-p)]^{1/2} \approx (np)^{1/2} = s. \quad (56)$$

(54)のポアソン分布は、 $M = s$ のところで極大を持ち、それより大きな M ですぐに減少する。これを示

そう、まず、Stirlingの公式から、

$$\ln(M!) \approx M \ln M - M. \quad (57)$$

したがって、(54)から、

$$P(M) = e^{-\ln M!} e^{M \ln s} e^{-s} = e^{M \ln s - \ln M!} e^{-s} \\ \approx \exp[M \ln s - M \ln M + M - s]. \quad (58)$$

ここで、

$$f(M) = M \ln s - M \ln M + M - s \quad (59)$$

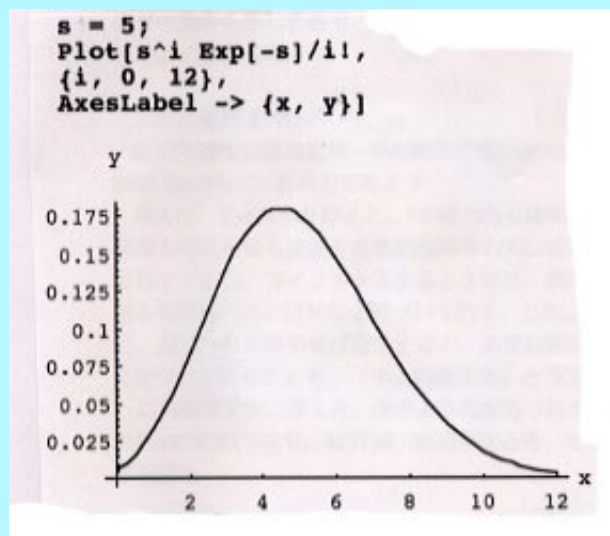
と置くと、極大は、

$$df(M)/dM = \ln s - \ln M = 0 \quad (60)$$

で得られるから、

$$M = s \quad (61)$$

のところにある。



<確率変数>

2項分布の場合は、N回の試行のうちM回の表が現われる場合の数は、(40)のようにMの離散的な (discrete)関数と考えることができる。一方、ガウス分布の場合は、N の極限をとるとき、

$$x = M, \quad (62)$$

$$\langle x \rangle = Np, \quad (63)$$

$$\sigma^2 = [Np(1-p)]^{1/2}. \quad (64)$$

と置いて、次のように連続な変数xの関数に書き直すことができる：

$$P(M) = p(x) dx, \quad dx = 1, \quad (65)$$

$$p(x) = [1/(2\sigma^2)^{1/2}] \exp[-(x - \langle x \rangle)^2 / (2\sigma^2)], \quad (66)$$

$$p(x) dx = 1. \quad (67)$$

数学では、この x のような変数のことを確率変数(probabilistic variable)と呼ぶ。一方、 $p(x)$ のような関数を確率分布関数と呼ぶ。この考え方を使って、確率的に変動し得る観測量や物理量のことを自然科学では、確率変数と考えるのである。そして、この確率分布関数を求めることが自然科学の問題を解くことになるのである。

<平均値と高次モーメント>

この確率分布関数 $p(x)$ を使うと、 x の期待値 $\langle x \rangle$ は以下のように定義できる：

$$\langle x \rangle =$$

$$\int x p(x) dx. \quad (68)$$

同様に、確率変数 x の関数 $f(x)$ の期待値 $\langle f(x) \rangle$ は、

$$\langle f(x) \rangle =$$

$$\int f(x) p(x) dx. \quad (69)$$

特に、 $f(x) = x^k$ のとき、その分布関数 $p(x)$ の k 次モーメント(the k th moment)と呼び、次のように定義される：

$$\langle x^k \rangle =$$

$$\int x^k p(x) dx. \quad (70)$$

<標準偏差と分散>

例えば、 $p(x)$ が(66)のガウス分布のとき、期待値 $\langle x \rangle$ は、

$$\langle x \rangle = [1 / (2 \pi \sigma^2)^{1/2}] \times$$

$$\int dx x \exp[-(x - x_0)^2 / (2 \sigma^2)] = x_0. \quad (71)$$

平均 $\langle x \rangle$ からの平均 2 乗偏差(mean square deviation)は、

$$\langle (x - x_0)^2 \rangle = [1 / (2 \pi \sigma^2)^{1/2}] \times$$

$$\int dx (x - x_0)^2 \exp[-(x - x_0)^2 / (2 \sigma^2)] = \sigma^2. \quad (72)$$

ここで、(71)を使うと、次のような 平均 2 乗偏差と標準偏差の関係を得る：

$$\sigma^2 = \langle (x - x_0)^2 \rangle = \langle (x - \langle x \rangle)^2 \rangle$$

$$= \langle x^2 \rangle - 2\langle x \rangle^2 + \langle x \rangle^2$$

$$= \langle x^2 \rangle - \langle x \rangle^2. \quad (73)$$

このように、標準偏差は、1次と2次のモーメントを使って書ける。標準偏差は、統計の確率分布の広がり具合を表わすので、これを分散(dispersion)と呼ぶこともある。

<中心極限定理>

統計で最も重要な定理--中心極限定理(central limit theorem)--を考えてみよう。

例えば、サイコロを振ると、1の目の出る確率は、何度も何度も振るほど、先見的な確率1/6に近づいて行く。また、コインをトスするときには、表の出る確率は1/2に限りなく近づいて行く。このように、測定の数を増やせば増やすほど、先見的確率に近づいて行くことを、「中心極限定理」と呼ぶ。

これを数学的に書くと、次のようになる：確率変数 x の測定(つまり、試行)を n 回行ったとき、その平均値を

$$A_{\{n\}} = (x_{\{1\}} + x_{\{2\}} + \dots + x_{\{n\}})/n \quad (74)$$

のように書くとしよう。このとき、真の期待値 $\langle x \rangle$ と $A_{\{n\}}$ の差:

$$A_{\{n\}} - \langle x \rangle \quad (75)$$

は、 n で、限りなくゼロに近づく。この証明は、数学的により高度な知識を必要とするので、ここでは省略する。

<多変数ガウス分布関数>

我々が実際に扱う自然現象の問題では、問題に関わる確率変数がたった1つであるとは限らない。多くの問題では、多数の確率変数が必要である。この場合、我々は(66)を拡張して、次のような多変数確率分布関数を考える:

$$p(x_{\{1\}}, x_{\{2\}}, \dots, x_{\{N\}}) = 1/[(2\pi)^{N/2} \sigma_{\{1\}} \sigma_{\{2\}} \dots \sigma_{\{N\}}] \times \exp(-\sum_{j=1}^N x_{\{j\}}^2/2\sigma_{\{j\}}^2), \quad (76)$$

ここで、

$$\sigma^2 = \sum_{j=1}^N (x_{\{j\}} - \langle x_{\{j\}} \rangle)^2 / \sigma_{\{j\}}^2, \quad (77)$$

$$\dots \int dx_{\{1\}} dx_{\{2\}} \dots dx_{\{N\}} p(x_{\{1\}}, x_{\{2\}}, \dots, x_{\{N\}}) = 1. \quad (78)$$

この分布関数は、それぞれの確率変数に対して独立であるので、次のような関係を満たす:

$$p(x_{\{1\}}, x_{\{2\}}, \dots, x_{\{N\}}) = p(x_{\{1\}})p(x_{\{2\}}) \dots p(x_{\{N\}}), \quad (79)$$

$$\langle x_{\{j\}} \rangle = \langle x_{\{j\}} \rangle, \quad (80)$$

$$\langle (x_{\{j\}} - \langle x_{\{j\}} \rangle)^2 \rangle = \sigma_{\{j\}}^2, \quad (81)$$

$$\langle x_{\{i\}} x_{\{j\}} \rangle = 0, \quad i \neq j. \quad (82)$$

さらに、より一般の場合として、確率変数が独立でない場合の分布関数がある:

$$p(x_{\{1\}}, x_{\{2\}}, \dots, x_{\{N\}}) = A \exp[-N \sum_{i < j} a_{\{ij\}} x_{\{i\}} x_{\{j\}}], \quad (83)$$

ここでAは $p(x_{\{1\}}, x_{\{2\}}, \dots, x_{\{N\}})$ が(78)を満たすようにとる定数である($A = \sqrt{N/2} / [\det(a_{\{ij\}})]^{1/2}$)。そして、以下の関係を満足する：

$$\langle x_{\{j\}}^2 \rangle = 2a_{\{jj\}}, \quad (84)$$

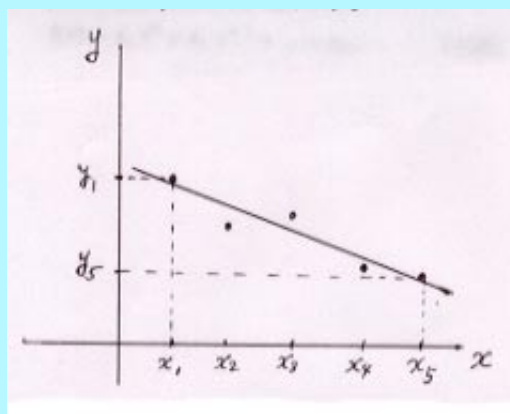
$$\langle x_{\{i\}} x_{\{j\}} \rangle = a_{\{ij\}}, \quad i \neq j. \quad (85)$$

この最後の(85)のような関係を相関関係(correlation relation)と呼ぶ。(85)は、確率変数 $x_{\{i\}}$ と確率変数 $x_{\{j\}}$ が独立ではなく、ある種の関係があることを示している。その関係の度合の強さを $a_{\{ij\}}$ が表わしている。これを相関係数(correlation coefficient)と呼ぶ。

<最小2乗法と χ^2 法>

この辺で、統計学の方法を実際に自然科学の実際的な問題に応用することを考えてみよう。

まず、実験結果をどのようにグラフで近似するか考えてみよう。実際の測定では、実験結果が下図のように、ばらつきを持ってプロット(plot)されることが多い。そのため、それらの測定データのどこに合わせて線や曲線を入れたいかすぐには分からないことがしばしばある。こういう場合に、最適な線を入れることのできる方法がある。それが、最小2乗法(least square method)と χ^2 法(カイ2乗法)と呼ばれる方法である。



今、 $x_{\{1\}}, x_{\{2\}}, \dots, x_{\{n\}}$ の点に対して、次の n 個の測定データがあったとしよう：

$$y_{\{1\}}, y_{\{2\}}, \dots, y_{\{n\}}. \quad (86)$$

これらは、今見出したい真の関数関係：

$$y = f(x) \quad (87)$$

のまわりをばらつきを持って揺らいでいる。したがって、この真の関数と測定値の間の誤差を見積もるために、ガウス関数のところで学んだような、次の量を定義する：

$$\chi^2 = [y_{\{1\}} - f(x_{\{1\}})]^2 / \sigma_{\{1\}}^2 + [y_{\{2\}} - f(x_{\{2\}})]^2 / \sigma_{\{2\}}^2 + \dots + [y_{\{n\}} - f(x_{\{n\}})]^2 / \sigma_{\{n\}}^2, \quad (88)$$

ここで、 $\sigma_{\{1\}}, \sigma_{\{2\}}, \dots, \sigma_{\{n\}}$ は各測定のときのばらつきの度合、つまり分散を表わす。この誤差

関数(88)が最小になるように、うまく $f(x)$ の形を決定する。この方法を χ^2 法 (Chi-square fitting) と呼ぶ、特に、すべての分散が等しいとき：

$$\sigma_1^2 = \dots = \sigma_n^2 = \sigma^2 \quad (89)$$

最小 2 乗法と呼ぶ。

例として、

$$f(x) = a x + b \quad (90)$$

のような直線で、測定データをフィットできる場合を考えてみよう。これを(88)へ代入すると、次式を得る：

$$\begin{aligned} \chi^2 &= [y_1 - (a x_1 + b)]^2 / \sigma_1^2 \\ &+ [y_2 - (a x_2 + b)]^2 / \sigma_2^2 + \dots \\ &+ [y_n - (a x_n + b)]^2 / \sigma_n^2. \quad (91) \end{aligned}$$

これを最小にするように、 a と b を選ばばよいから、次の条件を満たせばよい：

$$\frac{\partial \chi^2}{\partial a} = 0, \quad (92)$$

$$\frac{\partial \chi^2}{\partial b} = 0. \quad (93)$$

(92), (93) から、

$$\begin{aligned} \frac{\partial \chi^2}{\partial a} &= -[x_1 (y_1 - a x_1 - b) / \sigma_1^2 + x_2 (y_2 - a x_2 - b) / \sigma_2^2 + \dots + x_n (y_n - a x_n - b) / \sigma_n^2] \\ &= 0. \quad (94) \end{aligned}$$

$$\begin{aligned} \frac{\partial \chi^2}{\partial b} &= -[(y_1 - a x_1 - b) / \sigma_1^2 + (y_2 - a x_2 - b) / \sigma_2^2 + \dots + (y_n - a x_n - b) / \sigma_n^2] \\ &= 0. \quad (95) \end{aligned}$$

(94), (95) は以下のように書ける：

$$a S_{xx} + b S_x = S_{xy}, \quad (96)$$

$$a S_x + b S = S_y, \quad (97)$$

ここで、

$$S = 1 / \sigma_1^2 + 1 / \sigma_2^2 + \dots + 1 / \sigma_n^2, \quad (98)$$

$$S_x = x_1 / \sigma_1^2 + x_2 / \sigma_2^2 + \dots + x_n / \sigma_n^2, \quad (99)$$

$$S_y = y_1 / \sigma_1^2 + y_2 / \sigma_2^2 + \dots + y_n / \sigma_n^2, \quad (100)$$

$$S_{xx} = x_1^2 / \sigma_1^2 + x_2^2 / \sigma_2^2 + \dots + x_n^2 / \sigma_n^2. \quad (101)$$

$$S_{xy} = x_1 y_1 / \sigma_1^2 + x_2 y_2 / \sigma_2^2 + \dots + x_n y_n / \sigma_n^2. \quad (102)$$

(96), (97) を a, b に対して解くと、

$$a = (SS_{xy} - S_x S_y) / (SS_{xx} - S_x^2), \quad (103)$$

$$b = (S_{xx}S_y - S_xS_{xy}) / \quad , (104)$$

$$= SS_{xx} - (S_x)^2. (105)$$

このようにして、実験データに最もフィットする直線を見つけることができる。当然、この方法は、直線の場合だけでなく、もっと一般の関数：

$$f(x) = a_0 x^M + a_1 x^{M-1} + \dots + a_M \quad (106)$$

の場合にも適用できる。この最小2乗法や²法の考え方は、非常に重要かつパワフルで幅広い応用がある。

<乱数とモンテカルロ法>

この講義の最後に、まったく異なるタイプの積分の方法--モンテカルロ法(Monte Carlo method)-
-を考えてみよう。

まず、この講義の第一回るとき、数列:

$$x_{\{0\}}, x_{\{1\}}, \dots, x_{\{n\}}, \dots \quad (107)$$

にはさまざまなものがあるということを学んだ。この数列の中で、次に現われる数がいつも予測できない数(unpredictable number)である場合、この数列を乱数(random number)と呼ぶ。

もっとも単純な乱数の作り方は、以下のようにする。a, b, mを正の整数、 $x_{\{0\}}$ を与えられた数(seedと呼ぶ)とし、 $x_{\{n\}}$ が与えられると、その次の数 $x_{\{n+1\}}$ が次式：

$$x_{\{n+1\}} = a x_{\{n\}} + b \text{ mod } m. \quad (108)$$

で与えられるようにする。ここで、 $x \text{ mod } y$ の意味は、xをyで割った余り(residue)のことである。つまり、

$$x/y = q \dots r, \text{ or } x = q y + r, \quad (109)$$

$$r = x \text{ mod } y. \quad (110)$$

このとき、区間[0, 1]の間にあるn番目の乱数 $r_{\{n\}}$ は、

$$r_{\{n\}} = x_{\{n\}} / m \quad (111)$$

で与えられる。このようにして、乱数列 $r_{\{n\}}$ が得られると、簡単には積分できないような複雑な関数の積分をこの乱数を使って、近似的に求めることが可能となる。この方法のことを、モンテカルロ法による積分と呼ぶ。

例えば、次のような積分を行うことを考えよう：

$$\int_a^b f(x) dx. \quad (112)$$

今、積分領域[a, b]の中から勝手にN個の点、 $x_{\{1\}}, x_{\{2\}}, \dots, x_{\{N\}}$ をでたらめに取り出したとする。このような試みをランダムサンプリング(random sampling)と呼ぶ。すると、(112)は次の和と近似的に等しくなる：

$$\int_a^b f(x) dx \approx (b - a) (1/N) \sum_{i=1}^N f(x_{\{i\}})$$

$$a$$

$$N$$

$$f(x_{\{k\}}) \cdot (113)$$

$$k=1$$

このようにして、どのような複雑な積分も乱数を見つけることによって実行できるのである。しかし、この積分の精度は、どれほど今選び出した乱数が本当の乱数に近いかという、乱数の精度に依存する。(111)で得られる乱数は、 m に依存する周期性を持つので、 m が大きくなると近似は悪くなる。そのため、この自然界にある乱数を使って積分をもとめることもある。このような乱数を自然乱数と呼ぶ。

野球のピッチャーは、投球を打者に予測されないため、次にどこに投げるか乱数表を使って決めるといのはよく知られている。経済においても、株価の変動の予測を乱数の統計や確率論を使って行っている。このように、乱数の統計的独立性を使った方法は、科学や実社会のさまざまな分野ですでにたくさん利用されているのである。

| | | | | | | | |
|-------|-------|-------|-------|-------|-------|-------|-------|
| 43585 | 20270 | 74558 | 48961 | 90052 | 02750 | 82718 | 27982 |
| 84735 | 32061 | 47876 | 42152 | 89344 | 82877 | 44440 | 61944 |
| 85449 | 66261 | 38104 | 76120 | 66105 | 86843 | 17467 | 79969 |
| 21904 | 22894 | 46802 | 68360 | 67676 | 37401 | 50290 | 46941 |
| 06203 | 10840 | 07664 | 84061 | 78870 | 19046 | 94038 | 74214 |
| 46986 | 88540 | 26772 | 51091 | 60122 | 13542 | 29098 | 02527 |
| 70325 | 54459 | 74210 | 33550 | 67053 | 03497 | 00764 | 59007 |
| 82482 | 85693 | 34742 | 79244 | 54316 | 59097 | 05238 | 71302 |
| 34342 | 01936 | 91700 | 87950 | 36445 | 27181 | 94249 | 35572 |
| 78982 | 10013 | 72214 | 98454 | 63763 | 75478 | 24327 | 74597 |
| 16513 | 04318 | 44844 | 62677 | 52651 | 92644 | 60732 | 82781 |
| 81253 | 49676 | 62672 | 77020 | 33251 | 77045 | 66312 | 20038 |
| 14165 | 91983 | 19943 | 51068 | 33249 | 54613 | 76240 | 99180 |
| 70411 | 30598 | 83133 | 74098 | 05019 | 92651 | 23968 | 39257 |
| 93900 | 73882 | 25113 | 59388 | 43088 | 23301 | 32577 | 52791 |

乱数表
乱数は、無作為抽出による調査の道具として使われる。乱数を考え出すには、人間はあまりにも遠慮が過ぎる。1列に、同じ数字をつけて3回取り上げる人はいないだろう。しかしほんとうにでたための数列だったら、そんなこともありうる。ランド研究所では電子ランダムを使ひ100万個の数字をでたために並べて、上の左のような乱数表をつくった。右側はそのページの一部。

<参考文献>

- 1) ライフ/人間と科学シリーズ, 「数の世界」, (Time Life, 1973).
- 2) W. H. Press et al., Numerical Recipes, (Cambridge University, Cambridge, 1986).

<Home Work Set #13>

- 1) a) 1から6の目のあるサイコロを振って得られる数値の期待値を求めよ。(5点)
b) A, B, Cが統計的に独立なとき, 次のベイズの定理(Bayes' theorem):
 $P(B|C)/P(C|A) = [P(B)/P(C)][P(A|B)/P(A|C)]$ を証明せよ。(5点)
- 2) (21), (22), (23)で $N = 2$ として,
 $\langle x_{\{s\}} \rangle \langle x_{\{g\}} \rangle \langle x_{\{h\}} \rangle$ を証明せよ。(5点)
- 3) (42)から(45)を導け。(5点)
- 4) ポアソン分布(54)で $s = 2$ と置き, $M = 0$ から $M = 10$ までの分布を与えよ。(5点)
- 5) $\langle x_{\{3\}} \rangle \langle (x - \langle x \rangle)^3 \rangle$ を(73)のようにモーメント $\langle x^3 \rangle$, $\langle x^2 \rangle$, $\langle x \rangle$ で表現せよ。(5点)

- 6) a) 標準偏差の意味を述べよ。(5点)
 b) 日本の学校でいう「偏差値」と統計学の標準偏差の関係を述べよ。(5点)
 c) パチンコ屋へ行って、良く出るパチンコ台を探したい。でたらめに何台か見つけるのと、どこか1台を起点に横に何台か見つけるのと、どちらが良く出るパチンコ台を見つけやすいか？考察せよ。(5点)

7) 今、3点の測定点 $(x, y) = (1, 5), (2, 4.5), (3, 6.5)$ がある。最小2乗法により、この3点に最もフィットする直線を見つけよ。(5点)

8) a) (108)と(111)を使って、乱数を10個生成せよ。ただし、 $x_{\{0\}} = 1, a = 111, b = 55, m = 41$ とする。(5点)

b) a)の10個の乱数を使い、次の積分をモンテカルロ法により求めよ。(5点)

$$\int_0^1 4(1-x^2)^{1/2} dx.$$

<Home Work Set #12の解答>

1) 楕円(6)の面積を扇型の面積を求める方法で求めよ。(10点)

答え)

楕円の方程式(6)は、

$$x = a \cos t, y = b \sin t$$

で与えられる。一方、局座標表示では、

$$x = r \cos \theta, y = r \sin \theta$$

で与えられる。したがって、

$$r^2 = a^2 \cos^2 t + b^2 \sin^2 t.$$

楕円の面積 S の1/4は、

$$\begin{aligned} S/4 &= \int_0^{\pi/2} \int_0^{r(t)} r dr dt \\ &= \int_0^{\pi/2} \left(\frac{r^2}{2} \right) \Big|_0^{r(t)} dt \\ &= \int_0^{\pi/2} \frac{1}{2} (a^2 \cos^2 t + b^2 \sin^2 t) dt. \end{aligned}$$

$\tan \theta = (b/a) \tan t$ であるから、

$$\left(\frac{1}{\cos^2 \theta} \right) d\theta = (b/a) \left(\frac{1}{\cos^2 t} \right) dt.$$

また、

$$\frac{1}{\cos^2 t} = 1 + \tan^2 t = 1 + (b/a)^2 \tan^2 t \\ = (a^2 \cos^2 t + b^2 \sin^2 t) / a^2 \cos^2 t.$$

$$(d/dt) = (b/a)(\cos^2 t / \cos^2 t) \\ = (b/a) a^2 / (a^2 \cos^2 t + b^2 \sin^2 t) \\ = ab / (a^2 \cos^2 t + b^2 \sin^2 t).$$

これを積分に代入すると，

$$S/4 = (ab/2) \int_0^{\pi/2} dt = ab/4.$$

したがって，楕円の面積， $s = ab$ が得られた．

2) 楕円の方程式(6)は， $x = a \cos t$ ， $y = b \sin t$ で与えられることがわかっている．これを使って，楕円の周囲の長さを与える積分を求めよ．積分を計算する必要はない．(10点)

答え)

$ds^2 = dx^2 + dy^2$ であるから，

$$s = \int_0^{2\pi} ds = \int_0^{2\pi} \sqrt{(dx/dt)^2 + (dy/dt)^2} dt = 4 \int_0^{\pi/2} \sqrt{(dx/dt)^2 + (dy/dt)^2} dt.$$

これに

$$dx/dt = -a \sin t, \quad dy/dt = b \cos t$$

を代入すると，求める積分が得られる：

$$s = 4 \int_0^{\pi/2} \sqrt{(a^2 \sin^2 t + b^2 \cos^2 t)^{1/2}} dt.$$

さらに， $\cos^2 t = 1 - \sin^2 t$ を上式に代入すると，

$$s/(4a) = \int_0^{\pi/2} \sqrt{(1 - k^2 \sin^2 t)^{1/2}} dt = K(k),$$

ここで，

$$k^2 = (b^2 - a^2)/a^2.$$

この形の積分を第2種の楕円積分(elliptic integral of the second kind)と呼ぶ。

3) 底面の半径R, 高さLの円錐の側面の表面積を求めよ。(10点)

答え)

円錐を次の図のように

(図省略)

$$y = (R/L)x \quad (0 < x < L)$$

で表すと, 円錐の側面の表面積Sは, 以下のように書ける:

$$\begin{aligned} S &= 2 \int_0^L y(1+(dy/dx)^2)^{1/2} dx \\ &= 2 \int_0^L (R/L) x (1+(R/L)^2)^{1/2} dx \\ &= 2 (R/L)(1+(R/L)^2)^{1/2} \int_0^L x dx \\ &= (R/L)(1+(R/L)^2)^{1/2} L^2 \\ &= R (L^2+R^2)^{1/2}. \end{aligned}$$

この結果は, 円錐の斜辺の長さ $(L^2+R^2)^{1/2}$ を半径に持つ円の面積の $R/(L^2+R^2)^{1/2}$ の割合の面積と一致する:

$$\begin{aligned} &(R/(L^2+R^2)^{1/2}) [(L^2+R^2)^{1/2}]^2 \\ &= R (L^2+R^2)^{1/2}. \end{aligned}$$

(図省略)

4) $y = f(x) = a x^2 + b x + c$ とする. $y_1 = f(x_1)$, $y_2 = f(x_2)$, $y_0 = f(x_0)$, $x_0 = (x_1 + x_2)/2$, $x_2 - x_1 = h$ とするととき, 以下の関係が成り立つことを示せ。(10点)

$$\int_{x_1}^{x_2} f(x) dx = h(y_1 + y_2 + 4y_0)/6.$$

答え)

まず直接 $f(x)$ を代入して積分すると:

$$\begin{aligned}
 & \int_{x_1}^{x_2} (a x^2 + b x + c) dx \\
 &= \left[\frac{a}{3} x^3 + \frac{b}{2} x^2 + c x \right]_{x_1}^{x_2} \\
 &= \frac{a}{3} x_2^3 + \frac{b}{2} x_2^2 + c x_2 - \left[\frac{a}{3} x_1^3 + \frac{b}{2} x_1^2 + c x_1 \right] \\
 &= \frac{a}{3}(x_2^3 - x_1^3) + \frac{b}{2}(x_2^2 - x_1^2) + c(x_2 - x_1) \\
 &= (x_2 - x_1) \left[\frac{a}{3}(x_2^2 + x_2 x_1 + x_1^2) + \frac{b}{2}(x_2 + x_1) + c \right] \\
 &= \frac{(x_2 - x_1)}{6} [2a(x_2^2 + x_2 x_1 + x_1^2) + 3b(x_2 + x_1) + 6c] \\
 &= \frac{(x_2 - x_1)}{6} [a x_2^2 + b x_2 + c + a x_1^2 + b x_1 + c + a x_2^2 + 2a x_2 x_1 + a x_1^2 + 2b(x_2 + x_1) + 4c].
 \end{aligned}$$

ここで,

$$x_0 = (x_1 + x_2)/2$$

と置くと、最後の式の最後の項は、次のように書ける：

$$\begin{aligned}
 & a x_2^2 + 2a x_2 x_1 + a x_1^2 + 2b(x_2 + x_1) + 4c \\
 &= a(x_2 + x_1)^2 + 2b(x_2 + x_1) + 4c \\
 &= 4(a x_0^2 + b x_0 + c).
 \end{aligned}$$

したがって,

$$\begin{aligned}
 h &= x_2 - x_1, \\
 y_1 &= f(x_1) = a x_1^2 + b x_1 + c, \\
 y_2 &= f(x_2) = a x_2^2 + b x_2 + c, \\
 y_0 &= f(x_0) = a x_0^2 + b x_0 + c
 \end{aligned}$$

のように置くと、結局

$$\begin{aligned}
 & \int_{x_1}^{x_2} (a x^2 + b x + c) dx \\
 &= \frac{h}{6} (y_2 + y_1 + 4y_0).
 \end{aligned}$$

5) 台形公式により, $f(x) = \exp[-x^2]$ の面積を $[-10, 10]$ の区間で数値積分せよ. ただし, $\exp[x] = e^x$, そして積分区間を20等分せよ. (10点)

答え)

I =

$$\int_{-10}^{10} \exp[-x^2] dx = 2 \int_0^{10} \exp[-x^2] dx$$

$$= 2 \left(\frac{h}{2} \right) [y_0 + 2(y_1 + y_2 + \dots + y_9) + y_{10}].$$

ここで, $[-10, 10]$ の区間を20等分しているので, $h = 1$. そして

$$y_0 = \exp[0] = e^0 = 1,$$

$$y_1 = \exp[-1] = e^{-1} = 1/e = 0.3678794,$$

$$y_2 = \exp[-2^2] = e^{-4} = 1/e^4 = 0.0183156,$$

$$y_3 = \exp[-3^2] = e^{-9} = 1/e^9 = 0.0001234,$$

$$y_4 = \exp[-4^2] = e^{-16} = 1/e^{16} = 0.000000113,$$

$$y_5 = \exp[-5^2] = e^{-25} = 1/e^{25} \approx 0, \dots,$$

$$y_{10} = \exp[-10^2] = e^{-100} = 1/e^{100} \approx 0.$$

したがって,

$$I = 1 + 2(1/e + 1/e^4 + \dots + 1/e^{81}) + 1/e^{100}$$

$$= 1 + 2(0.3678794 + 0.0183156$$

$$+ 0.000123 + 0.000000113)$$

$$= 1.772637.$$

この値は厳密な値:

$$I = \int_{-10}^{10} \exp[-x^2] dx = 2 \int_0^{10} \exp[-x^2] dx$$

$$= 1.772453851\dots$$

に4桁の精度で一致している.

[前セクション](#) [目次](#)

[ホームページ](#) [和基](#) [和子](#) [維作](#) [条蒔](#) [家族](#) [Donation](#)

「井口和基博士と家族のホームページ」
〒774-0003 徳島県阿南市畷町新はり70-3
井口和基 (C)2004