ノート - 数学・数値解析

Atsushi Shimono

平成 22 年 9 月 7 日

目 次

第I部

統計解析

第1章　確率分布

ある一つの現象に対して、その全事象は一つの空間（標本空間）に含まれる座標として表される。いま、発生した全事象に対して、標本空間上での分布を取ると、この分布を全事象の数で割ったものが、対応する座標における確率となる。そして、標本空間に確率の次元を加えたものを確率空間と呼び、標本空間の一つの点に対応した確率の値をあらわす確率変数はこの空間の一つの点として表される。また、標本空間に時間などの助変数と呼ばれるパラメータが入っている場合、確率変数は助変数をパラメータに取り確率過程とよばれる。

いま、確率変数が連続量である場合、ある一つの値を取る確率自体は微少量となるため使い勝手がよくない。そこで、累積確率分布関数と確率密度関数を定義する。累積確率分布関数とは、$-\infty \leq t \leq x$ の区間の確率変数を積分した値であり、確率密度関数は累積確率分布関数をその区間の上限値 x で微分した関数をいう。

1.1　基本的な統計量

まず、確率変数を扱う演算において、もっとも基本的となるものに関して述べる。

標本空間上に存在する各事象に対して、ある演算を行ったうえで全積分した値として定義される統計量がはじめに与えられる。この、確率変数の統計的平均値である確率変数に確率密度関数の重みを掛けて積分した値を、確率変数の期待値もしくは平均と呼び

$$E(x) = \int_{\forall x} xp(x)dx \tag{1.1}$$

となる。次に、１変数についてその平均と変数の差の二乗による確率変数の期待値、つまり

$$V(x) = E\left((x - E(x))^2\right) = \int_{\forall x} (x - E(x))^2 \, p(x)dx \tag{1.2}$$

を分散と言う。ここで、変数が 2 つ以上ある場合に、そのうちの任意の 2 変数を用いて、2 変数に関しての平均と変数の差を掛けたものを用いて導出した分散

$$Cov(x, y) = E\left((x - E(x))\,(y - E(y))\right) = \int_{\forall(x,y)} (x - E(x))\,(y - E(y))\, p(x, y)dxdy \tag{1.3}$$

を、2 変数 x, y の共分散という。

つぎに、モーメントを定義する。モーメントは、導出に利用する変数とベキ指数によって多種定義できる。期待値が変数の 1 乗に確率変数を乗算した積分値であるのに対し、m 次モーメントとは、変数の m 乗に確率変数を乗算したものとなり

$$E(x^m) = \int_{\forall x} x^m p(x)dx \tag{1.4}$$

と定義される。これを利用すると、期待値は 1 次モーメントといえる。そして、m 次中心モーメントとは、m 次モーメントに対し、変数を平均と変数の差に置き換えたものである。つまり

$$E\left((x - E(x))^m\right) = \int_{\forall x} (x - E(x))^m \, p(x)dx \tag{1.5}$$

である。上と同様に、分散は 2 次の中心モーメントである。なお、多次元変数の場合には、変数の m 乗を

$$x^m = {x_1}^{m_1} {x_2}^{m_2} \cdots {x_n}^{m_n} \quad , \quad m = \sum m_i \tag{1.6}$$

とする。この場合、結合モーメントもしくは、 結合中心モーメントと呼ぶ。

離散的な確率変数に対して

離散分布を持つ確率変数に関しても、連続分布と同様に平均や分散が定義される。まず、平均は

$$\bar{x} = \sum_{\forall x} x_i p(x_i) = \frac{1}{N} \sum_{\forall i} x_i \tag{1.7}$$

となり、分散は

$$E(x) = \sum_{\forall x} (x - \bar{x})^2 p(x) = \frac{1}{N-1} \sum_{\forall i} (x_i - \bar{x})^2 \tag{1.8}$$

である。なお、分散の平方根は標準偏差と呼ばれる。

ここで、分布の指標としてこれらの値を用いることを考えると、分布は 1 次モーメントに、分散は 2 次モーメントに依存するといえる。しかしながら、分布によっては 2 次モーメントが存在しないことがあり、分散は分布を特徴づける指標としては不適切といえる。この場合に利用されるロバストな[1]評価量として、平均偏差

$$AD(x) = \sum_{\forall x_i} |x_i - \bar{x}| = \frac{1}{N} \sum_{\forall i} |x_i - \bar{x}| \tag{1.9}$$

が用いられることもある。

また、より高次の統計量として、3 次モーメントによる歪度

$$SK(x) = \sum_{\forall x_i} \left(\frac{x_i - \bar{x}}{\sigma} \right)^3 p(x_i) \tag{1.10}$$

や、4 次モーメントによる尖度

$$Kur(x) = \sum_{\forall x_i} \left(\frac{x_i - \bar{x}}{\sigma} \right)^4 p(x_i) - 3 \tag{1.11}$$

も利用される。ただし、高次のモーメントになるほど、その統計量の標準偏差は非常に大きなものとなる。

数値計算の面から見ると、ここで定義した統計量は非常に単純な演算から導出されるものがほとんどであり、計算速度としてみると改善の余地はほとんどないが、逆に計算精度を必要とする場合は多少の改善の余地がある。分散を求める際の丸め誤差を減少させるために、よく利用されるのが補正 2 パス演算法[2]であり

$$Var(x) = \frac{1}{N-1} \left(\sum (x_i - \bar{x})^2 - \frac{1}{N} \left(\sum (x_i - \bar{x}) \right)^2 \right) \tag{1.12}$$

と表される。これは、平均 $\bar{x}$ を求める際の丸め誤差によるずれを補正するための項であり、第 2 項は丸め誤差がなければ 0 となるが、そうでない場合には、丸め誤差を補正する方向に働く。なお、これが 2 パス演算法と呼ばれるのは、2 回データの並びを走査しているからである。

1.2　特性関数・母関数

確率分布関数に対して定義される、助変数を持つ複素指数関数 $\exp(ixt)$ について議論する。

まず、特性関数とは、助変数を持つ複素指数関数の期待値であり、t の複素関数となる。つまり

$$E(\exp(ixt)) = M_x(t) = \int_{\forall x} \exp(ixt) p(x) dx \tag{1.13}$$

---

[1] robust : 頑強な、分布中央からはずれた値への依存率が小さい

[2] corrected 2-pass algorithm

である。これは、確率密度関数をフーリエ変換したものであり、逆にこの特性関数のフーリエ逆変換によって確率密度関数が得られる。つまり

$$p(x) = \frac{1}{2\pi}\int_{\forall x} \exp(-ixt)M_x(t)dt \tag{1.14}$$

となり、任意次元の確率分布に対しても、確率密度関数と同じように特性関数も定義される。

この特性関数の原点における微係数は、定義式からモーメントであることが分かり、この性質を利用することで、マクローリン展開形

$$M_x(t) = 1 + \sum_{n\geq 1} \frac{(it)^n}{n!E(x^n)} \tag{1.15}$$

が得られる。

同様にして、助変数を持つ確率分布関数による指数関数 $\exp(xs)$ の期待値は、助変数 s の実関数となっており、モーメント母関数と呼ばれる。これは、確率密度関数をラプラス変換したものであり

$$E\left(\exp(xs)\right) = G_x(s) = \int_{\forall x} \exp(xs)p(x)dx \tag{1.16}$$

と表される。このモーメント母関数は特性関数とほぼ同じ性質を持ち、原点における微係数はモーメントに一致し、確率変数に関しての情報がない場合でも、これら特性関数か母関数を得られる場合、確率変数のなす構造は記述可能である。

1.3 確率分布

前節での各種確率に関わる量を元として、本節では確率分布関数とその特徴をみる。確率分布関数とは、確率変数が標本値に一致する確率を表す関数であり、現象の確率的構造の理解の基本となる関数といえる。

一般に、正規分布と呼ばれる確率分布関数がもっとも代表的な連続対称分布とされ、中心極限定理を適用すると、任意の確率分布をなす標本から集めた標本であっても多数を集めた場合の分布は正規分布に従う。これは、様々な原因が影響しあう不規則分布は正規分布に従うことを示し、偶然誤差といった性質が多くの不規則現象においてあらわれる。

1.3.1 正規分布

観測・測定に含まれる誤差は、大きく系統誤差と偶然誤差に分類される。系統誤差は、観測・測定器自身の内包するエラーによるものや測定者の引き起こすぶれによるものであり、十分な調査を行うことでその量を推定し排除することは可能である。しかしながら、偶然誤差は、その値が本質的に不規則なものであり、どのようにしても排除することはできない[3]。

正規分布の誘導

偶然誤差の確率分布については以下の３つの特徴がある。

- 標本平均は真値の最尤推定量である
- 偶然誤差の分布は誤差の絶対量のみの関数であり、対称・位置不変な性質を持つ
- 標本は同一母集団から同一の方法で独立に抽出されうる

これらの条件を与えた分布を持つ偶然誤差に関して、真値 a に対して n 個の標本が x_i と言う値を持って同時に抽出されたとする。このとき、その尤度関数が同時確率密度関数に対応しており、これらの独立な実現値の集合により一つの標本平均値が求めうるので、尤度関数 Y は

$$Y(x,a) = Y(x_1-a, x_2-a, x_3-a, \cdots, x_n-a) = p(x_1-a)p(x_2-a)\cdots p(x_n-a) \tag{1.17}$$

[3] S/N 比改善の手法がある、というのは、この誤差を排除しているのではなく、誤差の分布を狭くしているだけである。

と表現される。いま、$\bar{x}$ を a の最尤推定量[4]とする場合、尤度関数の微係数は $x = \bar{x}$ において 0 になる。つまり

$$\left.\frac{\partial Y(x,a)}{\partial x}\right|_{x=\bar{x}} = 0 \tag{1.18}$$

である。これら 2 式から

$$nY(\bar{x},a)\sum \left.\frac{p'(x_i-a)}{p(x_i-a)}\right|_{\bar{x}=a} = 0 \tag{1.19}$$

となり、任意の i に対して特解を

$$\frac{p'(x_i-a)}{p(x_i-a)} = f(x_i-a) \tag{1.20}$$

の形で持つので、積分することによって

$$\log p(x_i-a) = \frac{f}{2}(x_i-a)^2 + c \tag{1.21}$$

を得る。パラメータ c, f については、確率密度関数 p の定義から、無限遠で 0 でありかつ全領域の積分値が 1 に等しいという条件を利用すると、自由度を 1 減らすことができる。いま、f は負の値でなければならないので σ を

$$f = -\sigma^{-2} \tag{1.22}$$

と定義すると

$$e^c = \left(\sqrt{2\pi}\sigma\right)^{-1} \tag{1.23}$$

と書き換えられる。よって、任意の x_i を $\bar{x}$ とすると、確率密度関数 $p(x)$ は

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-a)^2}{2\sigma^2}\right) \tag{1.24}$$

と表される。

この 2 つの変数 a, σ^2 はそれぞれ平均と分散となり、この 2 変数を係数として上式が正規確率密度関数の表式となる。つまり、正規確率密度分布は上記の 3 つの条件を満たす偶然誤差を表す分布関数であるというのが定義である。

いま、正規確率密度関数のパラメータは a, σ^2 の 2 つであるので、この分布関数を $N(a, \sigma^2)$ と略記することが多い。また、統計データの取り扱いの場面では、2 変数に対して平均の原点移動や無次元化のためのデータ処理を行うことがあり、中心化得点 $X - a \sim N(0, \sigma^2)$ や標準化得点 $(x-a)/\sigma \sim N(0,1)$ などがある。なお、一般的に統計表に記載される関数値はこの標準化得点の値が記載されている。

正規確率密度関数の性質

中心モーメント

中心化得点のモーメントは次数によって異なるが、奇数次数のモーメントについては、元の分布関数が原点対称分布であることが条件であることから

$$E(Y^{2m+1}) = 0 \tag{1.25}$$

である。また、偶数次数のモーメントに関しても

$$E\left(Y^{2m}\right) = \sigma^{2m}(2m-1)!! = E\left(|Y|^{2m}\right) \tag{1.26}$$

と原点対称分布であることから変数の絶対値に関しては不変である。最後に、変数の絶対値に対する奇数次数のモーメントについては

$$E\left(|Y|^{2m+1}\right) = \sqrt{\frac{2}{\pi}}2^m m!\sigma^{2m+1} \tag{1.27}$$

[4] この場合、条件から標本平均である。

となる。

特性関数

特性関数は、前に示したとおり確率密度関数をフーリエ変換した関数であり、標準 (化) 正規分布の場合

$$M(t) = \exp\left(-\frac{t^2}{2}\right) \tag{1.28}$$

となる。

分布同士の演算

正規確率の分布関数 N に関して、互いに独立な分布 $N_1(a_1, {\sigma_1}^2)$, $N_2(a_2, {\sigma_2}^2)$ に従う変数 x_{1i}, x_{2i} に関して、これらの和で定義される $y_i = x_{1i} + x_{2i}$ は正規分布 $N(a_1 + a_2, {\sigma_1}^2 + {\sigma_2}^2)$ に従う。つまり、正規分布に従う変数の和が従う分布は、平均・分散をそれぞれ足しあわせた正規分布である。

次に、一般の線形和に関して、正規分布の分布関数 $N_i(a_i, {\sigma_i}^2)$ を考え、そのそれぞれの変数の線形和 $y = \sum b_i x_i + c$ で与えられる変数群 y を考える。この y が従う分布は、最初に定義した分布関数による変数群が互いに独立であると言う条件の下で、平均は $\sum b_i a_i + b$、分散は $\sigma^2 = \sum {b_i}^2 {\sigma_i}^2$ となる。

誘導される (標本) 分布

ある確率変数 Y が、別の確率変数 X_i から抽出された複数の標本による多変数関数 y によって記述されるとき、Y は X_i の標本統計量と呼ばれ、この確率分布は対応する関数 y の確率分布として誘導する手法により決定できる。

この節では、このようにして標準・任意正規分布関数から誘導された代表的な標本分布に関して、その元となる関数や標本分布の性質について述べる。まず、標準正規分布関数から誘導されるものから、χ 二乗分布を取り上げ、そこから誘導される F 分布、t 分布、Maxwell 分布、Cauchy 分布などにふれた後、これら正規分布の標本分布が含まれる、より一般的な確率分布である Γ 分布、β 分布を取り上げる[5]。

χ 自乗分布

用いる関数は、互いに独立な n 個の標準正規分布の変数 x_i に対して

$$y(i) = \sum {x_i}^2 \quad , \quad x_i \sim N(0, 1) \tag{1.29}$$

となる二乗和である。この関数によって定義される分布関数を、自由度 n の χ^2 分布（$\chi^2(n)$ 分布）とよぶ。この確率密度関数は

$$p_{\chi_2}(y, n) = \frac{2^{-n/2}}{\Gamma(n/2)} \exp\left(-\frac{y}{2}\right) y^{n/2-1} \tag{1.30}$$

である。

F 分布

互いに独立な 2 つの確率変数 X, Y がそれぞれ $\chi^2(m)$, $\chi^2(n)$ の分布に従うとする。このとき、関数

$$z(i) = \frac{x_i}{m} \cdot \left(\frac{y_i}{n}\right)^{-1} \tag{1.31}$$

によって誘導される分布を、自由度 (m, n) の F 分布という。この密度関数は β 関数を用いて

$$p_F(z) = \left(\frac{m}{n}\right)^{m/2} B(m/2, n/2)^{-1} \frac{z^{m/2-1}}{(1 + mz/n)^{(m+n)/2}} \tag{1.32}$$

[5]なお、ここでは任意正規分布関数（$N(0, 1)$ に標準化されていない、$\sigma \neq 1$ の様な値を持つ分布関数）から導出される分布関数についてはふれない。これらは、一般的に統計解析に用いられることはなく、どちらかというと、考えている対象を標準化して考えることが多いからである。なお、これらの誘導された分布関数は、中心が原点でないことから、非心母数を持つ分布関数と呼ばれる。

と表される。なお、この分布は、関数の分母・分子が標本分散を意味するので、値を標本分散比と呼ぶこともある。特に、この性質から 2 つの標本集合を分散を用いて分析する理論において用いられる。

t 分布

z 分布

確率変数 X が自由度 (m, n) の F 分布に従うとき $Z = (\log X)/2$ の分布を自由度 (m, n) の z 分布（$z(m, n)$ 分布）という。この密度関数は

$$p_{z(m,n)}(x) = \frac{2(m/n)^{m/2}}{B(m/2, n/2)} \frac{e^{mx}}{(1 + me^{2x}/n)^{(m+n)/2}} \tag{1.33}$$

となる。

第2章　推定

索引