正規分布の発見

どんなに偏りのあるコインであっても，繰り返し投げ続けていくと，すべてが同じ結果，たとえば表ばかりとか裏ばかりに一方的に偏ることがなく，適度な変動をもった結果が集約して得られると想像される。n 枚のコインを投げると（1 枚のコイン投げを n 回繰り返すと），表の出る枚数（回数）は 2 項分布にしたがうことという命題は既に学んだ。この枚数（回数）をどんどん大きくしていくと，2 項分布ではあるが，ある分布に近づくことがわかる。この分布が正規分布である。ラプラス (Pierre Simon de Laplace), 1749-1827 フランスの革命期の数学者、天文学者。数学特に解析学の多くの分野に大きな業績を残したが、古典確率論の大成はその貢献の一つである。

これは古典的な中心極限定理の魁（さきがけ）となっています。中心という言葉は Polya(1920) によるものといわれ、基本的とか「中心的に重要なもの」という意味がこめられている。つまりどんな分布であってもたくさん集めて、その観測値データの算術平均をとれば、その分布は正規分布に近づくという。これを分布収束とよぶ。

中心極限定理 (Central Limit Theorem) とは

$X_1, X_2, \cdots,$ を独立同一分布で共通の平均 μ と分散 σ^2 をもつならば、観測データの和 $S_n = \sum_{i=1}^{n} X_i$ について

$$\frac{S_n - n\mu}{\sqrt{n}\sigma} \to N(0,1), \quad n \to \infty$$

あるいは観測データの算術平均 $\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i$ について

$$\frac{\overline{X_n} - n\mu}{\sigma/\sqrt{n}} \to N(0,1), \quad n \to \infty$$

この中心極限定理は確率論の中心テーマの一つといえる。多くの研究者が取り組んでいるが、とくにロシア（ソビエト連邦）では盛隆であり、拡張された優れた結果が多く発表された。

正規分布の歴史は 1730 年ドモアブル（De Moivre）の記事 Miscellanea Analytica（さまざまな解析）、1733 年 The Doctrine of Chances（偶然の原則）から始まる。さらに 1738 年の The Doctrine of Chances（偶然の原則）第 2 版に 2 項分布を近似する内容が述べられているという。その後ラプラスにより、Analytical Theory of Probabilities （確率の解析的理論）(1812) によって拡張された。

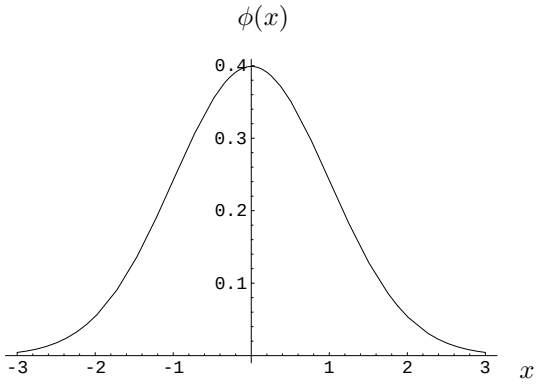
一般のコイン投げを多数回繰り返すと、その分布は正規分布に近い形でことがわかる。つまり 2 項分布の極限は正規分布である。このように 2 項分布を正規分布で近似することは、ドモアブル・ラプラスの定理とよばれる。

確率変数がパラメータ n, p の 2 項分布にしたがうならば、記号では $X \sim Bin(n,p)$ と表し, その密度関数が $P(X=x) = \binom{n}{x} p^x (1-p)^x$ であるならば、

$$P(X=x) \sim \frac{1}{\sqrt{2\pi}} \exp\left\{ -\left(\frac{x-np}{\sqrt{np(1-p)}} \right)^2 /2 \right\} \frac{1}{\sqrt{np(1-p)}}$$

(2 項分布の確率密度) ∼ (正規分布の確率密度関数)

標準正規分布の密度関数の曲線です。滑らかに変化し、ひとつ山の形をしています。



このグラフの式はつぎで与えられるものです。

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2x^2}\right) = \frac{1}{\sqrt{2\pi}}\, e^{-1/2x^2}$$

実際、数値を当てはめてみると比較をしてみます。かなり正確に当てはまることがわかる。いま 2 項確率とは $\sqrt{npq}P(X=x)$ 正規確率とは $\phi(u),\ u = \dfrac{x-np}{\sqrt{npq}},\ q = 1-p$ として計算してみます。たとえば $n = 10$, $p = 0.7$ のとき、

値	0	1	2	3	4	5	6	7	8	9	10
2 項確率	0.000	0.000	0.002	0.013	0.053	0.149	0.290	0.387	0.338	0.175	0.041
正規確率	0.000	0.000	0.001	0.009	0.047	0.154	0.314	0.399	0.314	0.154	0.047

もう少し n を大きくして $n = 30$, $p = 0.7$ のときの数値結果は

値	10	12	14	16	18	20	22	24	26	28	30
2 項確率	0.000	0.001	0.011	0.058	0.188	0.355	0.377	0.208	0.052	0.005	0.000
正規確率	0.000	0.001	0.008	0.055	0.195	0.369	0.369	0.195	0.055	0.008	0.001

証明はすこし技術的な要素が強いですが、スターリングの公式 $m! = \sqrt{2\pi m} m^m e^{-m}$ と対数関数の近似（テーラー展開）$\log(1+x) = x - \dfrac{x^2}{2} + O(x^3)$ をもちいる。$\sqrt{npq}\binom{n}{x} p^x q^{n-x} \sim \dfrac{1}{2\pi} e^{-\frac{u^2}{2}}$ の両辺の対数をとって示すことが多い。$x/n - p = u\sqrt{\dfrac{pq}{n}}$, $(n-x)/n - q = -u\sqrt{\dfrac{pq}{n}}$ であり、先の公式から

$$\log \frac{x/n}{p} = \log\left(1 + u\sqrt{\frac{q}{np}}\right) = u\sqrt{\frac{q}{np}} - \frac{u^2}{2}\frac{q}{np}$$

,

$$\log \frac{(n-x)/n}{q} = \log\left(1 - u\sqrt{\frac{p}{nq}}\right) = -u\sqrt{\frac{p}{nq}} - \frac{u^2}{2}\frac{p}{nq}$$

が得られ、より高次の n については無視することで近似される。

つぎは Pierr Simon Laplace(1749-1827) による厳密な証明である。彼は特性関数（ラプラス変換と同様）とその逆変換という技法をもちいた。つぎの変換 $\psi(t) = E(e^{itX})$ をおこなって,$n \to \infty$ とした極限関数とが正規分布のそれと一致することから、対応の一意性によって厳密に証明した。確率論では強力に使える手法である。まず 2 項分布（ベルヌイ分布 ; 1 枚のコイン投げ）の特性関数を計算し、独立な和の分布は積になるから、2 項分布の特性関数が求められる。コインの枚数 (n) を多くするから、これを n の値で評価する。微積分でのテーラー展開でもちいて、極限を計算する。この極限関数が、正規分布の特性関数に対応するものであり、対応が 1 対 1 であることから、証明される。

一方，ガウス (Karl Friedrich Gauss), 1777-1855 が正規分布について極めて大きな貢献を与えた。いわゆる中心極限定理である。ガウスはドイツの数学者、物理学者、天文学者で、幼少時より天才の誉れ高く、

数学・物理学に巨大な功績を残している。ラプラスが議論した、２項分布の極限として正規分布を導いた方法とは全く異なり，天体観測の誤差測定を解析し，変動があり得る状況から，微分方程式をたてて，正規分布を導いている。「誤差論」円錐曲線で太陽の回りを回る天体の運動理論―多くの観測結果にもっともよく合う軌道の決定―という論文（１８０９年）で密度関数を示した。

ガウスは、誤差の解析をすることで、微分方程式 $\frac{d}{dx}\phi(x) = -x\phi(x)$ を導き、その解を求める。この微分方程式は変数分離形であるから

$$\log\phi(x) = \int(-x)dx + const. = \{x \text{ の } 2 \text{ 次式 }\} = cx^2\text{の形（条件から）} = (\text{係数}) \times \frac{-1}{2}x^2$$

と求められる。

ガウスの考えかたは（観測値）＝（未知の真の値）＋（誤差分布）ととらえ、もし n 個の観測値があれば、それらを $M_i = z + v_i, i = 1, 2, \cdots n$ とおく。また誤差分布はゼロを中心として対称（正と負の両方の値をとる）で標本 n が多数であればゼロに近いとし、$\frac{M_1 + M_2 + \cdots + M_n}{n} = z$ 分布の確率は密度 $\phi(v)dv$ をもつとする。$v_i = M_i - z$ であるから、独立として積の形になるから $P = \phi(v_1)dv_1 \times \phi(v_2)dv_2 \times \cdots \times \phi(v_n)dv_n$ この P を最大にするには 対数をとって $\log P$ を最大にすればよい、すなわち $\log\phi(v_1) + \log\phi(v_2) + \cdots + \log\phi(v_n)$ を微分してゼロになるようなところを定める。

$$\frac{\partial P}{\partial z} = \frac{1}{\phi(v_1)}\frac{\partial\phi(v_1)}{\partial z_1} + \frac{1}{\phi(v_2)}\frac{\partial\phi(v_2)}{\partial z_2} + \cdots + \frac{1}{\phi(v_n)}\frac{\partial\phi(v_n)}{\partial z_n} = 0$$

合成関数の微分から $\frac{\partial\phi(v)}{\partial z} = \phi'(v)\frac{\partial v}{\partial z}$ であり、いま $\frac{\phi'(v)}{\phi(v)} = \psi(v)$ とおけば

$$\psi(v_1)\frac{\partial v_1}{\partial z_1} + \psi(v_2)\frac{\partial v_2}{\partial z_2} + \cdots + \psi(v_n)\frac{\partial v_n}{\partial z_n} = 0$$

が成り立ち、仮定からすべての i で $v_i = M_i - z$ としたから $\frac{\partial v_i}{\partial z_i} = -1$ で共通だから、

$$\psi(v_1) + \psi(v_2) + \cdots + \psi(v_n) = 0 \quad \cdots\cdots\cdots \quad (a)$$

の形に帰着される。また

$$v_1 + v_2 + \cdots + v_n = 0 \quad \cdots\cdots\cdots \quad (b)$$

である。なぜなら $v_1 + v_2 + \cdots + v_n = (M_1 - z) + (M_2 - z) + \cdots (M_n - z) = M_1 + M_2 + \cdots + M_n - nz = 0$ したがって２つの関係式 $(a), (b)$ から $\psi(x) = cv$ (c は定数) の形でなければならない。元の関数で微分がはいった式では $\frac{\phi'(v)}{\phi(v)} = cv$、$\log\phi(v) = c\frac{v^2}{2} + c'(c, c' : const)$。この c, c' は $\phi(v)$ の対称性条件から $\phi(v) = ke^{-h^2v^2}(k, h : const)$ という正規分布の形に結論として得られた。この係数の値は $k = \frac{1}{\sqrt{2\pi}}, h = \frac{1}{\sqrt{2}}$ である。

ラプラスの方法は解析の方法がかなり難しく簡単には述べられない。が、汎用性に富んでいる。これに比べればガウスの方法はシンプルであるといえる。しかし発想には天才による煌きがある。