正規密度の計算

□□□□□ □□□□□ □□□□□ □□□□□ □□□□□ □□□□□　2009 年 6 月 10 日版

1　はじめに

正規分布、とくに標準正規分布に関する確率変数 $Z \sim N(0,1)$ の密度関数

$$f_Z(z) = \phi(z) \tag{1.1}$$

ここで

$$\phi(z) = \frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{1}{2}z^2\Big) \tag{1.2}$$

は統計学、確率論では必ず記述される基本の関数である。また積分変数も習慣的に z がよく用いられるが、これは複素数ではないので、複素積分ではないことの誤解を避けたい。特性関数などでは、複素積分を使うが混同することはないと思われる。一般にこの密度関数は関数値として非負であることは容易に理解される。なぜなら指数関数とは、値が負になっても、逆数を意味するから関数が負値になることはない。が、積分値が 1 になること、あるいは

$$I = \int_0^\infty \exp\Big(-\frac{1}{2}z^2\Big)\,dz = \sqrt{\frac{\pi}{2}} \tag{1.3}$$

であることは、ある程度の基礎知識を積み重ねないとそう簡単ではない。ただし証明も多数知られている。

ここでは古典的に由緒ある De Moivre(Approximatio ad Summam Teminorum Binomu $\overline{a+b}/^n$ in Seriem expansi 1733) が１７３３年に示した方法 (H.M.Walker; Scripta Mathematica 2(4)(1934) 316-333) を紹介し、R Courant(Differential and Integral Calculus(2 Vols), Blackie 1934-6) によりガンマ関数による表現を適用すると

$$\Gamma\Big(\frac{1}{2}\Big) = \sqrt{\pi} \tag{1.4}$$

が得られることを示そう。変数変換することで、ガンマ関数 (1.4) と正規密度の計算 (1.3) とは、同値であることが示されるが、しかしガンマ関数の計算自体がそう簡単とはいえない。だが、1/2 の「語呂合わせ」が「ルート・パイ」としてシックリしているのでこれは記憶しておくと便利である。本論では簡単な三角関数の積分計算によって、この等式 (1.4) を証明する。

2　ド・モアブルってどこかで聞いたよ

ド・モアブルの公式とか定理ってのは確か $(\cos\theta + i\sin\theta)^n = \cos(n\theta) + i\sin(n\theta)$ ですよね。ちなみに i は虚数単位です。これ以外にもいろいろあります。ド・モアブル Abraham De Moivre(1667-1754) の歴史的なＣＬＴ (中心極限定理) への貢献は、現在は「二項分布の正規近似」とよばれる。ド・モアブルは 1730 年、”Miscellanea Analytica”での研究題目として出版されている。1733 年”Document of Chance”での結果を現代形で表現するとつぎのようになる：

確率変数が二項分布にしたがうとき、$X \sim \mathrm{Bin}(n,p)$ のとき、

$$P(X=x) = \binom{n}{x} p^x q^{n-x}$$

$x=0,1,2,\cdots,n$ に対して、このとき

$$\sqrt{npq}\binom{n}{x}p^xq^{n-x} \approxeq \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-np}{\sqrt{npq}}\right)^2}$$

ただし $q=1-p$ となる。

3　ド・モアブルはどう示したか？

ド・モアブルは２項分布 $N(n,p), p=1/2$ の計算からも示したが、これには James Stirling による係数評価が用いられる。A Hald; A History of Theory of Probability and Statistics and Their Application before 1750, Wiley 1990 section 24.4

もし $n=2m$ で

$$b(x)=\binom{2m}{x}\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{2m-x}=\binom{2m}{x}2^{-2m}, \qquad x=0,1,2,\cdots 2m \tag{3.5}$$

とすると

$$b(x)\simeq\frac{0.7976}{\sqrt{n}} \tag{3.6}$$

となるが、この値はどういうものか。その答えはスターリングが示した。係数の値がスターリングの公式:

$$m! \sim \sqrt{2\pi m}\, m^m e^{-m} \tag{3.7}$$

により、$0.7976 \sim \sqrt{\frac{2}{\pi}}$ であり、

$$b(m)=\binom{2m}{m}2^{-2m} \sim \frac{1}{\sqrt{\pi m}}=\sqrt{\frac{2}{\pi n}}$$

この当時は m を中心として左右対称であるから、$\{-m,-m+1,\cdots,0,1,2,\cdots,m\}$ のように変化させて計算している。$l=-m,-m+1,\cdots,0,1,2,\cdots,m$ について $(n=2m)$

$$b(m+l)\sim b(m)\exp(-2l^2/n)\sim\sqrt{\frac{2}{\pi n}}\exp(-2l^2/n) \tag{3.8}$$

分散の概念はド・モアブルには使われていないが, $\sigma^2=p(1-p)n=\frac{n}{4}$ とすると

$$b(m+l)\sim\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{l^2}{2\sigma^2}\right)=\frac{1}{\sigma}\phi\left(\frac{l}{\sigma}\right) \tag{3.9}$$

２項分布について総和の計算から

$$\begin{aligned}1=\sum_{l=-m}^{m}b(m+l) &\sim \int_{-\infty}^{\infty}\frac{1}{\sigma}\phi\left(\frac{l}{\sigma}\right)dl \\ &\sim \int_{-\infty}^{\infty}\phi(x)dx\end{aligned} \tag{3.10}$$

4　もっと簡単に

クーラント (リヒャルト・クーラント（Richard Courant, 1888 年 1 月 8 日 - 1972 年 1 月 27 日、現ポーランド領ルブリニツ生まれ、アメリカ合衆国ニューヨーク没）はドイツとアメリカの数学者。彼の書いた教科書 Methods of mathematical physics は 80 年以上後もいまだに使われている。一般書 What is Mathematics?（邦題 :『数学とは何か』）もいまだに出版されている。) によれば、

$$\int_0^{\pi/2} \sin^\alpha t dt = \frac{\sqrt{\pi}}{\alpha} \frac{\Gamma\big((1+\alpha)/2\big)}{\Gamma(\alpha/2)} \tag{4.11}$$

により、

$$I_k = \int_0^{\pi/2} \sin^k t dt \tag{4.12}$$

とおくと

$$I_k\, I_{k+1} = \frac{\pi}{2(k+1)} \tag{4.13}$$

が成り立つ。なぜならば、部分積分により自然数 n で $I_n = (n-1)(I_{n-2} - I_n)$, $I_n = \frac{n-1}{n} I_{n-2}$, $I_1 = 1$, $I_0 = \pi/2$ であるから、

$$I_{2n} = \frac{1\cdot 3\cdot 5\cdots(2n-1)}{2\cdot 4\cdot 6\cdots(2n)}\frac{\pi}{2}, \qquad I_{2n+1} = \frac{2\cdot 4\cdot 6\cdots(2n)}{1\cdot 3\cdot 5\cdots(2n+1)}$$

となる。掛け合わせると係数が消去できるから、関係式

$$I_{2n} I_{2n+1} = \cdot\frac{1}{2n+1}\frac{\pi}{2} = \frac{\pi}{4n+2} = \frac{\pi}{2(2n)+1} \tag{4.14}$$

が、つまり (4.13) 式が得られる。

また $0 \le x \le \pi/2$ において $\sin^{2n-1} x > \sin^{2n} x > \sin^{2n+1} x$ から $I_{2n-1} > I_{2n} > I_{2n+1}$ したがって

$$1 < \frac{I_{2n}}{I_{2n+1}} < \frac{I_{2n-1}}{I_{2n+1}} = \frac{2n-1}{2n} \xrightarrow[(n\to\infty)]{} 1$$

極限の値が 1 であるから、上の関係式 (4.14) から

$$\lim_n \sqrt{n} I_{2n} = \lim_n \sqrt{n} I_{2n} \sqrt{\frac{I_{2n}}{I_{2n+1}}} = \lim_n \sqrt{n}\sqrt{I_{2n} I_{2n+1}} = \lim_n \sqrt{n}\sqrt{\frac{\pi}{4n+2}} = \frac{\sqrt{\pi}}{2}$$

同様に

$$\lim_n \frac{1}{\sqrt{n}} I_{2n} = 0$$

という評価式を得る。

また

$$I_k = \frac{1}{2} B\Big(\frac{k+1}{2}, \frac{1}{2}\Big) \tag{4.15}$$

が成り立つことがわかり、$k=2$ とおけば、これから

$$\Gamma(1/2)^2 = \pi \tag{4.16}$$

を得る。なぜなら、これは (4.12) 式をもちいたが、$\sin^2 t = x$ とおき、$\sin^k t = x^k/2, \cos t = (1-x)^{1/2}$ としてベータ関数の形にできるから。

ここでベータ関数とガンマ関数のそれぞれの定義は

$$B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1} \quad (a,b>0)$$
$$\Gamma(a) = \int_0^\infty x^{a-1}e^{-x}dx \quad (a>0)$$

で、両者の関係式が

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \quad \Gamma(a+1) = a\Gamma(a), (a>0) \quad \Gamma(n+1) = n!, (n: integer), \Gamma(1) = 1$$

であることに注意する。

5　I_2 だけ

いま $k=2$ で

$$I_2 = \int_0^{\pi/2} \sin^2 t dt \tag{5.17}$$

とおき、このとき $\sin^2 t = x$ とすれば、$\sin^3 t = x^{3/2}$ $\cos t = (1-x)^{1/2}$ としてベータ関数の形にできる。$\Gamma\Big(\frac{3}{2}\Big) = \frac{1}{2}\Gamma\Big(\frac{1}{2}\Big), \Gamma(2) = \Gamma(1) = 1$ より

$$I_2 = \frac{1}{2}B\Big(\frac{3}{2}, \frac{1}{2}\Big) = \frac{1}{2}\frac{\Gamma\Big(\frac{3}{2}\Big)\Gamma\Big(\frac{1}{2}\Big)}{\Gamma(2)} = \frac{1}{4}\Gamma^2\Big(\frac{1}{2}\Big) \tag{5.18}$$

一方 $\int_0^{\pi/2} \sin^2 tdt = \int_0^{\pi/2} \cos^2 tdt$ より、

$$I_2 = \frac{1}{2}\Big(\int_0^{\pi/2} \sin^2 tdt + \int_0^{\pi/2} \cos^2 tdt\Big) = \frac{\pi}{4} \tag{5.19}$$

したがって

$$\Gamma^2\Big(\frac{1}{2}\Big) = \pi \qquad \therefore \Gamma\Big(\frac{1}{2}\Big) = \sqrt{\pi} \tag{5.20}$$

6　その他の有名な方法

I.Todhunter; A History of the Mathematical Theory of Probability from the time of Pascal to that of Laplace, reprinted NY Chelsea 1949 art.899.

（１）求めたい積分をまず変数変換して、変数 y を導入し、

$$I = \int_0^\infty \exp\Big(-\frac{z^2}{2}\Big)dz = \int_0^\infty \exp\Big(-\frac{(xy)^2}{2}\Big)ydx$$

この 2 項の積を考えると、2 重積分になり、前項の積分変数を z から x と変えれば

$$I^2 = \int_0^\infty\int_0^\infty \exp\Big(-\frac{1}{2}(x^2+1)y^2\Big)ydydx$$

ここで変数を $(1+x^2)y^2 = 2t$ とおくと、$ydy = \dfrac{1}{1+x^2}dt$ より

$$I^2 = \int_0^\infty \int_0^\infty \exp(-t)\frac{dt}{1+x^2}dx = \Big(\int_0^\infty \exp(-t)\,dt\Big)\Big(\int_0^\infty \frac{dx}{1+x^2}\Big) = 1\cdot\frac{\pi}{2}$$

が得られる。

（２）Jacob Karl Franz Strum(1803-1855) 積を考えることで、2 重積分の形にし、これに対して極座標変換を用いる方法で、微分積分での演習問題では頻出である。比較的短く示される。直交座標と極座標の変換には

$$x = r\cos\theta,\ y = r\sin\theta \quad dxdy = r\,drd\theta$$

であるから、

$$I^2 = \int_0^\infty \int_0^\infty \exp\Big(-\frac{1}{2}(x^2+y^2)\Big)\,dxdy = \int_0^{\pi/2}\int_0^\infty \exp\Big(-\frac{1}{2}r^2\Big)\,rdrd\theta = \int_0^{\pi/2} d\theta\cdot\int_0^\infty e^{-r^2/2}\,rdr = \frac{\pi}{2}$$

（３）E.T.Wittaker and G N Watson; ”A Course of Modern Analysis, Cambridge Univ Press 1902,1915,1920,1927”. 西暦年は重版。

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin\pi z}$$

において $z = 1/2$ とおけばよい。