2022 年 07 月 29 日

正規確率変数の一次式

新潟工科大学 基礎教育・教養系 竹野茂治

1 はじめに

本稿では、講義で証明を省略した、正規分布に従う独立な確率変数の一次式が正規分布に従う、という定理の証明を行う。また、それに必要な e^{-x^2} の定積分についても、講義では証明を省略したが、本稿で紹介しておく。

なお、統計や確率の専門的な書物では、通常特性関数や積率母関数などを使って計算、証明をするようであるが ([2], [3])、本稿では、計算は少し大変だが、より原始的な方法で計算する。

また、本稿では現代的な公理的確率論ではなく、古典的確率論の範疇で考える。

2 命題と簡略化

証明すべき事柄は以下の通りである。

命題 1

確率変数 $x_1, \ldots, x_n$ が独立で、$x_j \sim N(\mu_j, \sigma_j^2)$ $(1 \leq j \leq n)$ であるとき、定数 a_j $(\neq 0)$, b に対し、

$$y = \sum_{j=1}^{n} a_j x_j + b \sim N\left(\sum_{j=1}^{n} a_j \mu_j + b, \sum_{j=1}^{n} a_j^2 \sigma_j^2\right) \tag{1}$$

これをいくつかの段階に分けて考える。

命題 2

確率変数 $x_1, \ldots, x_n$ が独立で、関数 ϕ_j により $y_j = \phi_j(x_j)$ というあらたな確率変数が作られるとき、$y_1, \ldots, y_n$ も $((x_1, \ldots, x_n)$ の n 次元確率分布の元で$)$ 独立となる。

この命題の「あらたな確率変数が作られるとき」という表現については、連続確率変数の場合は任意の関数であらたな連続確率変数が作られるとは限らないために用いている。例えば、$\phi_j(x)$ が x の 1 次式、多項式などであれば問題はない。詳しくは、[1] を参照のこと。

また、「$(x_1, \ldots, x_n)$ の n 次元確率分布の元で」という表現については、「$y_1, \ldots, y_n$」の独立性は本来 $(y_1, \ldots, y_n)$ に関する n 次元確率分布のとり方によって決まることであるが、それを「$(y_1, \ldots, y_n)$ の n 次元確率分布を $(x_1, \ldots, x_n)$ の n 次元確率分布から自然に決まるものとする」、すなわち、

$$
\begin{aligned}
&\mathrm{Prob}\{(y_1, \ldots, y_n) \in B_1 \times \cdots \times B_n\} \\
&= \mathrm{Prob}\{(x_1, \ldots, x_n) \in \phi_1^{-1}(B_1) \times \cdots \times \phi_n^{-1}(B_n)\} \qquad (2)
\end{aligned}
$$

とする場合にこの命題が成り立つ、ということを意味する。n 次元確率分布とその独立性については、詳しくは、[1] を参照のこと。

命題 2 の証明

$y_1, \ldots, y_n$ が独立であることは、任意の B_j に対し

$$
\mathrm{Prob}\{(y_1, \ldots, y_n) \in B_1 \times \cdots \times B_n\} = \mathrm{Prob}\{y_1 \in B_1\} \times \cdots \times \mathrm{Prob}\{y_n \in B_n\} \qquad (3)
$$

が成り立つことを意味する。(2)、および $x_1, \ldots, x_n$ が独立であるという仮定により、(3) の左辺は

$$
\begin{aligned}
&\mathrm{Prob}\{(y_1, \ldots, y_n) \in B_1 \times \cdots \times B_n\} \\
&= \mathrm{Prob}\{(x_1, \ldots, x_n) \in \phi_1^{-1}(B_1) \times \cdots \times \phi_n^{-1}(B_n)\} \\
&= \mathrm{Prob}\{x_1 \in \phi_1^{-1}(B_1)\} \times \cdots \times \mathrm{Prob}\{x_n \in \phi_n^{-1}(B_n)\} \\
&= \mathrm{Prob}\{y_1 \in B_1\} \times \cdots \times \mathrm{Prob}\{y_n \in B_n\}
\end{aligned}
$$

となるので、(3) が成立する。■

命題 2 の正規分布に従う確率変数 x_j に対して 標準化 $u_j = (x_j - \mu_j)/\sigma_j \sim N(0, 1)$

を取れば、命題 2 より $u_1, \ldots, u_n$ は独立で、y は、

$$y = \sum_{j=1}^{n} a_j x_j + b = \sum_{j=1}^{n} a_j \sigma_j u_j + c, \quad c = \sum_{j=1}^{n} a_j \mu_j + b \tag{4}$$

となる。

補題 3

$u \sim N(0,1)$ ならば $v = -u \sim N(0,1)$

証明

$N(0,1)$ の密度関数

$$f_0(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \tag{5}$$

は偶関数なので、v の分布関数 $\mathrm{Prob}\{v \leq t\}$ は、

$$\begin{aligned} \mathrm{Prob}\{v \leq t\} &= \mathrm{Prob}\{u \geq -t\} = 1 - \mathrm{Prob}\{u < -t\} = 1 - \int_{-\infty}^{-t} f_0(u)du \\ &= 1 - \int_{t}^{\infty} f_0(u)du = \int_{-\infty}^{t} f_0(u)du \end{aligned}$$

となって $N(0,1)$ の分布関数に一致する。■

よって、u_j に対して

$$v_j = \begin{cases} u_j & (a_j > 0 \text{ のとき}) \\ -u_j & (a_j < 0 \text{ のとき}) \end{cases} \tag{6}$$

とすれば、補題 3 により $v_j \sim N(0,1)$ で、命題 2 より $v_1, \ldots, v_n$ も独立であり、また $a_j u_j = |a_j| v_j$ となるので、(4) は、

$$y = \sum_{j=1}^{n} d_j v_j + c, \quad c = \sum_{j=1}^{n} a_j \mu_j + b, \quad d_j = |a_j| \sigma_j > 0 \tag{7}$$

と書ける。つまり、命題 1 を、$\mu_j = 0$, $\sigma_j = 1$, $a_j > 0$ の形に帰着できることになる。

もしこの (7) に対して命題 1 が成立することが言えれば、

$$y \sim N\left(c, \sum_{j=1}^{n} d_j^2\right) = N\left(\sum_{j=1}^{n} a_j\mu_j + b, \sum_{j=1}^{n} a_j^2\sigma_j^2\right)$$

となり、命題 1 が得られることがわかる。よって後は、命題 1 を、$\mu_j = 0$, $\sigma_j = 1$, $a_j > 0$ の元で示せばよい。

3 積分に関する補題

補題 4

$$\int_{\mathrm{R}} e^{-x^2} dx = \sqrt{\pi} \tag{8}$$

証明

e^{-x^2} は偶関数なので、

$$\int_0^{\infty} e^{-x^2} dx = \lim_{K\to\infty} \int_0^K e^{-x^2} dx = \frac{\sqrt{\pi}}{2} \tag{9}$$

を示せばよい。それには、二重積分の極座標変換を用いる。

$$I_K = \int_0^K e^{-x^2} dx$$

とすると、

$$I_K^2 = \int_0^K e^{-x^2} dx \int_0^K e^{-y^2} dy = \int_0^K dx \int_0^K e^{-x^2-y^2} dy$$

は、正方形 $D_K = [0, K] \times [0, K]$ 上の二重積分と見ることができる。D_K は、原点中心半径 K の 1/4 円 C_K を含み、原点中心半径 $\sqrt{2}K$ の 1/4 円 $C_{\sqrt{2},K}$ に含まれるので、

$$\iint_{C_K} e^{-x^2-y^2} dxdy \leq I_K^2 \leq \iint_{C_{\sqrt{2},K}} e^{-x^2-y^2} dxdy \tag{10}$$

となる。ここで、C_K の積分を極座標変換すれば、

$$\iint_{C_K} e^{-x^2-y^2} dxdy = \int_0^{\pi/2} d\theta \int_0^K e^{-r^2} r dr = \frac{\pi}{2} \left[-\frac{1}{2} e^{-r^2} \right]_{r=0}^{r=K} = \frac{\pi}{4}(1 - e^{-K^2})$$

となるので、(10) より、

$$\frac{\pi}{4}(1 - e^{-K^2}) \leq I_K^2 \leq \frac{\pi}{4}(1 - e^{-2K^2}) \tag{11}$$

となる。ここで $K \to \infty$ とすれば、はさみうちの原理により、

$$\lim_{K \to \infty} I_K^2 = \frac{\pi}{4}$$

となるので、(9) が得られる。■

補題 5

$a > 0$, $D = b^2 - 4ac$ に対して、

$$\int_{\mathrm{R}} e^{-(ax^2+bx+c)/2} dx = \sqrt{\frac{2\pi}{a}} e^{D/(8a)} \tag{12}$$

証明

$$ax^2 + bx + c = a\left(x + \frac{b}{2a}\right)^2 - \frac{b^2}{4a} + c = a\left(x + \frac{b}{2a}\right)^2 - \frac{D}{4a}$$

より、この積分を I とすると、

$$I = e^{D/(8a)} \int_{\mathrm{R}} e^{-a(x+b/(2a))^2/2} dx$$

となるので、$\sqrt{a/2}(x+b/(2a))=t$ と置換すれば、補題 8 より

$$I = e^{D/(8a)} \int_{\mathrm{R}} e^{-t^2} dt \sqrt{\frac{2}{a}} = e^{D/(8a)} \sqrt{\frac{2\pi}{a}}$$

となる。■

4 最初の命題の証明

ここまでの準備の元、命題 1 を証明する。なお、2 節最後に述べたように、$\mu_j = 0$, $\sigma_j = 1$, $a_j > 0$ として考える。

$x_1, \ldots, x_n$ は独立なので、n 次元確率変数 $(x_1, \ldots, x_n)$ の密度関数 $f(x_1, \ldots, x_n)$ は

$$f(x_1, \ldots, x_n) = f_0(x_1) \cdots f_0(x_n) = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-(x_1^2+\cdots+x_2^2)/2}$$

となる。よって、$y = a_1 x_1 + \cdots + a_n x_n + b = S_n$ の分布関数 $G(y)$ は、$y = S_n \le t$ を x_n について書けば $x_n \le (t - S_{n-1})/a_n$ となるので、

$$\begin{aligned} G(t) &= \mathrm{Prob}\{y \le t\} = \mathrm{Prob}\{x_n \le (t - S_{n-1})/a_n\} \\ &= \int_{\mathrm{R}} f_0(x_1) dx_1 \cdots \int_{\mathrm{R}} f_0(x_{n-1}) dx_{n-1} \int_{-\infty}^{(t-S_{n-1})/a_n} f_0(x_n) dx_n \end{aligned}$$

となり、y の密度関数 $g(y) = G'(y)$ は、

$$\begin{aligned} g(t) &= G'(t) = \int_{\mathrm{R}} f_0(x_1) dx_1 \cdots \int_{\mathrm{R}} f_0(x_{n-1}) f_0\left(\frac{t - S_{n-1}}{a_n}\right) \frac{1}{a_n} dx_{n-1} \\ &= \left(\frac{1}{\sqrt{2\pi}}\right)^n \frac{1}{a_n} \int_{\mathrm{R}} dx_1 \cdots \int_{\mathrm{R}} e^{-T_{n-1}/2 - (t - S_{n-1})^2/(2a_n^2)} dx_{n-1} \end{aligned} \tag{13}$$

となる。ここで $T_n = x_1^2 + \cdots x_n^2$ とする。

この指数の部分は、x_j にする 2 次式で、$a_n^2 = p^2$ と書くことにすれば、それを x_{n-1} について整理すると、

$$-\frac{T_{n-1}}{2} - \frac{(t - S_{n-1})^2}{2p^2} = -\frac{T_{n-2} + x_{n-1}^2}{2} - \frac{(t - S_{n-2} - a_{n-1} x_{n-1})^2}{2p^2}$$

$$= -\frac{1}{2}\left\{\left(1+\frac{a_{n-1}^2}{p^2}\right)x_{n-1}^2 - \frac{2a_{n-1}}{p^2}(t-S_{n-2})x_{n-1} + \frac{(t-S_{n-2})^2}{p^2} + T_{n-2}\right\}$$

これに、補題 5 を用いれば、

$$\begin{aligned}
a &= \frac{p^2+a_{n-1}^2}{p^2}, \\
\frac{D}{4} &= \frac{a_{n-1}^2(t-S_{n-2})^2}{p^4} - \left(1+\frac{a_{n-1}^2}{p^2}\right)\left(\frac{(t-S_{n-2})^2}{p^2}+T_{n-2}\right) \\
&= -\frac{p^2+a_{n-1}^2}{p^2}T_{n-2} - \frac{(t-S_{n-2})^2}{p^2}, \\
\frac{D}{8a} &= -\frac{1}{2}T_{n-2} - \frac{(t-S_{n-2})^2}{p^2+a_{n-1}^2}
\end{aligned}$$

より、x_{n-1} での積分の結果は、

$$\sqrt{\frac{2\pi}{a}}\,e^{D/(8a)} = \sqrt{\frac{2\pi p^2}{p^2+a_{n-1}^2}}\,e^{-\{T_{n-2}+(t-S_{n-2})^2/(p^2+a_{n-1}^2)\}/2} \tag{14}$$

となる。これと、積分前の

$$e^{-\{T_{n-1}+(t-S_{n-1})^2/p^2\}/2}$$

を比較すると、全体が $\sqrt{2\pi p^2/(p^2+a_{n-1}^2)}$ 倍され、指数部分は x_{n-1} の項が消え、代わりに p^2 に a_{n-1}^2 が追加されることがわかる。

よって、(14) を x_{n-2} で積分すると、その結果は

$$\begin{aligned}
&\sqrt{\frac{2\pi p^2}{p^2+a_{n-1}^2}}\sqrt{\frac{2\pi p^2+a_{n-1}^2}{p^2+a_{n-1}^2+a_{n-2}^2}}\,e^{-\{T_{n-3}+(t-S_{n-3})^2/(p^2+a_{n-1}^2+a_{n-2}^2)\}/2} \\
=\ &\sqrt{\frac{(2\pi)^2p^2}{p^2+a_{n-1}^2+a_{n-2}^2}}\,e^{-\{T_{n-3}+(t-S_{n-3})^2/(p^2+a_{n-1}^2+a_{n-2}^2)\}/2}
\end{aligned}$$

となる。

これを続けていけば、最後の x_1 での積分の結果 $g(t)$ は、$\vec{a}=(a_1,\ldots,a_n)$ とすれば、$p^2=a_n^2$ より以下のようになる。

$$g(t) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \frac{1}{a_n}\sqrt{\frac{(2\pi)^{n-1}p^2}{p^2+a_{n-1}^2+\cdots+a_1^2}}\,e^{-(t-b)^2/(2(p^2+a_{n-1}^2+\cdots+a_1^2))}$$

$$= \frac{1}{\sqrt{2\pi}|\vec{a}|}\,e^{-(t-b)^2/(2|\vec{a}|)}$$

これは、y の密度関数 $g(y)$ が正規分布 $N(b,|\vec{a}|^2)=N(b,\sum a_j^2)$ の密度関数に等しいことを意味し、これで $\mu_j=0,\sigma_j=1$ の場合の命題 1 が示されたことになる。

参考文献

[1] 竹野茂治「多次元確率分布と独立性」(2022)
http://takeno.iee.niit.ac.jp/~shige/math/lecture/graduate/data/ndimrandvar1.pdf

[2] 水原昂廣、宇野力「例題中心 確率・統計入門 (改訂版)」、学術図書出版社 (2001)

[3] 西尾真喜子「確率論」、実教出版 (1978)