カイ2乗分布再考
(Chi Square Distribution Revisited)

緑川章一*

1 数学的準備 (Mathematical Preliminaries)

積分 $\displaystyle\int_0^\infty e^{-\alpha x}x^{s-1}\,dx$ は、α が実数ならば、$t=\alpha x$ と変数変換をおこなうと、

$$\int_0^\infty e^{-\alpha x}x^{s-1}\,dx = \frac{1}{\alpha^s}\int_0^\infty e^{-t}t^{s-1}\,dt.$$

ところが、ガンマ関数 $\Gamma(\alpha)$ の定義より、

$$\Gamma(s) = \int_0^\infty e^{-t}t^{s-1}\,dt$$

である。ゆえに、

$$\int_0^\infty e^{-\alpha x}x^{s-1}\,dx = \frac{\Gamma(s)}{\alpha^s} \tag{1}$$

が得られる。

ここで、(1) 式の x を複素数に拡張し、閉曲線 $ABCB'A'c$ についての積分をおこなうと、閉曲線の内部で、関数は正則なので、

$$\int_A^B e^{-\alpha x}x^{s-1}\,dx = \int_{A'}^{B'} e^{-\alpha z}z^{s-1}\,dz + \int_C - \int_c$$

ここで、$z=e^{i\varphi}x$ と x が実数となるようにおいて、さらに、$OA,\ OA'\to 0,\quad OB,\ OB'\to\infty$ の極限をとると、$\displaystyle\int_C \to 0,\ \int_c \to 0$ だから、

$$\int_0^\infty e^{-\alpha x}x^{s-1}\,dx = e^{is\varphi}\int_0^\infty e^{-\alpha e^{i\phi}x}x^{s-1}\,dx$$

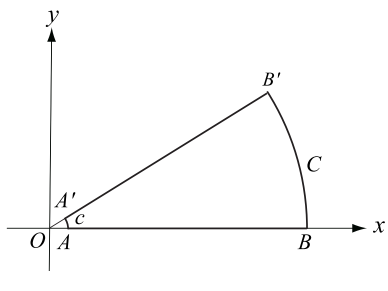

Figure 1:

*Shoichi Midorikawa

この式の左辺に (1) 式を用いると、

$$e^{is\varphi}\int_0^\infty e^{-\alpha e^{i\varphi}x}x^{s-1}\,dx=\frac{\Gamma(s)}{\alpha^s}$$

を得る。さらに、両辺を $e^{is\varphi}$ で割ると、

$$\int_0^\infty e^{-\alpha e^{i\varphi}x}x^{s-1}\,dx=\frac{\Gamma(s)}{(\alpha e^{i\varphi})^s}$$

となる。ここで、$\alpha e^{i\varphi}=p+iq$ とおくと、

$$\int_0^\infty e^{-(p+iq)x}x^{s-1}\,dx=\frac{\Gamma(s)}{(p+iq)^s} \tag{2}$$

を得る。これは、(1) 式において、α を $p+iq$ で置き換えたものと形式的には同じである。

(2) 式の両辺に e^{iqy} を掛けて、q について $-\infty$ から ∞ まで積分すると、

$$\int_0^\infty dx\,e^{-px}x^{s-1}\int_{-\infty}^\infty e^{iq(y-x)}\,dq=\Gamma(s)\int_{-\infty}^\infty\frac{e^{iqy}}{(p+iq)^s}\,dq. \tag{3}$$

ところで、

$$\int_{-\infty}^\infty e^{iq(y-x)}\,dq=2\pi\delta(y-x)$$

とデルタ関数を用いて表されるので、左辺の積分は実行できて、

$$2\pi e^{-py}y^{s-1}=\Gamma(s)\int_{-\infty}^\infty\frac{e^{iqy}}{(p+iq)^s}\,dq.$$

となる。これを整理して書き直すと、

$$\int_{-\infty}^\infty\frac{e^{iqy}}{(p+iq)^s}\,dq=\frac{2\pi}{\Gamma(s)}e^{-py}y^{s-1} \tag{4}$$

となる。

2 χ^2(カイ 2 乗) 分布 (Chi Square Distribution)

標準正規分布は、

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

で与えられる。ここで、 $x=u^2$ と置くと、確率変数 x についての確率密度関数 $T_1(x)$ は、

$$\begin{aligned}T_1(x)&=\int_{-\infty}^\infty\delta(x-u^2)f(u)\,du\\&=\int_{-\infty}^\infty\frac{1}{2\sqrt{x}}\left(\delta(u-\sqrt{x})+\delta(u+\sqrt{x})\right)f(u)\,du\\&=\frac{1}{2\sqrt{x}}\left(f(\sqrt{x})+f(-\sqrt{x})\right)\end{aligned}$$

と書けるので、これに (5) 式を代入すると、

$$T_1(x) = \frac{1}{\sqrt{2\pi}} e^{-x/2} x^{-1/2} \qquad (x \geq 0).$$

を得る。これを自由度 1 の χ^2(カイ 2 乗) 分布と言う。

変数 $x_1,\ x_2,\ \cdots,\ x_n$ は互いに独立で、各々自由度 1 の χ^2 分布に従うものとする。このとき、変数

$$y = x_1 + x_2 + \cdots + x_n.$$

の従う確率密度関数 $T_n(y)$ は、

$$T_n(y) = \int_0^\infty \delta(y - x_1 - x_2 - \cdots - x_n) T_1(x_1) T_1(x_2) \cdots T_1(x_n)\, dx_1 dx_2 \cdots dx_n.$$

と書ける。ここで、デルタ関数の積分表示

$$\delta(y - x_1 - x_2 - \cdots - x_n) = \frac{1}{2\pi} \int_{-\infty}^\infty e^{ik(y - x_1 - x_2 - \cdots - x_n)}\, dk$$

を用いて、

$$\begin{aligned} T_n(y) &= \frac{1}{2\pi} \int_{-\infty}^\infty dk e^{iky} \prod_{i=1}^n \int_0^\infty e^{-ikx_i} T_1(x_i)\, dx_i \\ &= \frac{1}{2\pi} \int_{-\infty}^\infty dk e^{iky} \left[\int_0^\infty e^{-ikx} T_1(x)\, dx \right]^n \end{aligned} \tag{6}$$

と書き直すことができる。さらに、 (2) 式を用いると、

$$\begin{aligned} \int_0^\infty e^{-ikx} T_1(x)\, dx &= \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\left(\frac{1}{2}+ik\right)x} x^{\frac{1}{2}-1}\, dx \\ &= \frac{1}{\sqrt{2\pi}} \frac{\Gamma\left(\frac{1}{2}\right)}{\left(\frac{1}{2}+ik\right)^{1/2}} \\ &= \frac{1}{\sqrt{2}\left(\frac{1}{2}+ik\right)^{1/2}} \end{aligned}$$

となる。ここで、関係 $\Gamma(1/2) = \sqrt{\pi}$ を使った。この結果を (6) 式に代入すると、

$$T_n(y) = \frac{1}{2\pi} \frac{1}{2^{n/2}} \int_{-\infty}^\infty \frac{e^{iky}}{\left(\frac{1}{2}+ik\right)^{n/2}}\, dk$$

となる。この積分は (4) 式を用いて行うことができて、

$$T_n(y) = \frac{1}{2^{n/2}\Gamma(n/2)} e^{-\frac{y}{2}} y^{\frac{n}{2}-1}.$$

となる。この分布を**自由度 n の χ^2 分布**と言う。

3 標本分散の標本分布 (Distribution of the Sample Variance)

確率変数 $u_1,\ u_2,\ \cdots,\ u_n$ が互いに独立で、それぞれ標準正規分布 $N(0,1)$ に従うとする。このとき、標本平均

$$\bar{u} = \frac{u_1 + u_2 + \cdots + u_n}{n} \tag{7}$$

を用いて作った分散

$$x = (u_1 - \bar{u})^2 + (u_2 - \bar{u})^2 + \cdots + (u_n - \bar{u})^2 \tag{8}$$

の分布 $f_X(x)$ を求めよう。これは、形式的には、

$$\begin{aligned} f_X(x) &= \int \delta\left(x - \sum_{i=1}^{n}(u_i - \bar{u})^2\right) \\ &\quad \times \delta\left(\bar{u} - \sum_{i=1}^{n} u_i/n\right) \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-u_i^2/2}\, du_i\, d\bar{u} \end{aligned} \tag{9}$$

と書ける。ここで余分な変数 $\bar{u}$ を導入したので、拘束条件 (7) を表すためのデルタ関数 $\delta\left(\bar{u} - \sum_{i=1}^{n} u_i/n\right)$ が必要となることに注意しよう。

ところで、

$$\sum_{i=1}^{n}(u_i - \bar{u})^2 = \sum_{i=1}^{n} u_i^2 - n\bar{u}^2$$

と書き直すことができる。さらに、デルタ関数の積分表示を用いると、

$$\begin{aligned} \delta\left(x - \sum_{i=1}^{n}(u_i - \bar{u})^2\right) &= \delta\left(x - \sum_{i=1}^{n} u_i^2 + n\bar{u}^2\right) \\ &= \frac{1}{2\pi} \int e^{ip\left(x - \sum_{i=1}^{n} u_i^2 + n\bar{u}^2\right)}\, dp \end{aligned} \tag{10}$$

となる。もう 1 つのデルタ関数も、

$$\delta\left(\bar{u} - \frac{\sum_{i=1}^{n} u_i}{n}\right) = \frac{1}{2\pi} \int e^{iq\left(\bar{u} - \sum_{i=1}^{n} u_i/n\right)}\, dq \tag{11}$$

と積分表示で表して、これら (10) 式と (11) 式を (9) 式に代入すると、

$$f_X(x) = \frac{1}{(2\pi)^2} \int e^{ipx} \prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{(1+2ip)}{2}u_i^2 - iqu_i/n}\, du_i\right) e^{i(pn\bar{u}^2 + q\bar{u})}\, d\bar{u}dpdq \tag{12}$$

となる。

(12) 式のカッコ内の積分は実行できて、

$$\begin{aligned}
&\frac{1}{\sqrt{2\pi}}\int e^{-\frac{(1+2ip)}{2}u_i^2-iqu_i/n}\,du_i \\
&= \frac{1}{\sqrt{2\pi}}\int \exp\left\{-\frac{(1+2ip)}{2}\left(u_i+i\frac{q}{n(1+2ip)}\right)^2-\frac{q^2}{2n^2(1+2ip)}\right\}du_i \\
&= \frac{1}{(1+2ip)^{1/2}}\,\exp\left\{-\frac{q^2}{2n^2(1+2ip)}\right\} \qquad (13)
\end{aligned}$$

を得る。この (13) 式を (12) 式に代入すると、

$$f_X(x)=\frac{1}{(2\pi)^2}\int\frac{e^{ipx}}{(1+2ip)^{n/2}}\,\exp\left\{-\frac{q^2}{2n(1+2ip)}+ipn\bar{u}^2+iq\bar{u}\right\}d\bar{u}dpdq \qquad (14)$$

となる。

さらに、指数関数の中を変形して、

$$-\frac{q^2}{2n(1+2ip)}+ipn\bar{u}^2+iq\bar{u}=-\frac{1}{2n(1+2ip)}\Big[q-in(1+2ip)\bar{u}\Big]^2-\frac{n}{2}\bar{u}^2 \qquad (15)$$

q についての積分をおこなうと、

$$\int\exp\left\{-\frac{1}{2n(1+2ip)}\Big[q-in(1+2ip)\bar{u}\Big]^2\right\}dq=\sqrt{\frac{2\pi}{n(1+2ip)}} \qquad (16)$$

となるので、(14) 式は、

$$f_X(x)=\frac{1}{(2\pi)^{3/2}}\int\frac{\sqrt{n}\,e^{ipx}}{(1+2ip)^{(n-1)/2}}\,e^{-n\bar{u}^2/2}\,d\bar{u}dp \qquad (17)$$

次に、$\bar{u}$ についての積分をおこなうと、

$$\int e^{-n\bar{u}^2/2}\,d\bar{u}=\sqrt{\frac{2\pi}{n}}$$

だから、(17) 式は、

$$\begin{aligned}
f_X(x) &= \frac{1}{2\pi}\int\frac{e^{ipx}}{(1+2ip)^{(n-1)/2}}\,dp \\
&= \frac{1}{2\pi}\frac{1}{2^{(n-1)/2}}\int\frac{e^{ipx}}{\left(\frac{1}{2}+ip\right)^{(n-1)/2}}\,dp \qquad (18)
\end{aligned}$$

最後に、(4) 式を用いると、

$$f_X(x)=\frac{1}{2^{(n-1)/2}\Gamma\left(\frac{n-1}{2}\right)}e^{-\frac{x}{2}}x^{\frac{n-1}{2}-1} \qquad (19)$$

を得る。これは、**自由度が $\boldsymbol{n-1}$ の $\boldsymbol{\chi^2}$ 分布** $T_{n-1}(x)$ である。