

2 次形式統計量の分布の高次漸近展開

Higher Order Asymptotic Expansions for the Distribution
of Quadratic Form in Normal Variates

統計数理研究所 仁木 直人 (Naoto Niki)
(Inst. Statist. Math.) 小西 貞則 (Sadanori Konishi)

Higher order asymptotic expansions for the distribution function and the percentiles of the distribution of positive definite quadratic form in normal variates are obtained by using a computer algebra system. The resulting formulas guarantee accuracy almost up to fourth decimal place if the distribution is not very skew. The normalizing transformation investigated by Jensen and Solomon (1972) is reconsidered based on the discussion by Niki and Konishi (1985) about the effects of transformation in higher order asymptotic expansions.

§ 0. はじめに

本稿で紹介するのは、確率分布の正規近似の精密化とでも言うべき Edgeworth 展開および Cornish-Fisher 展開の、正值 2 次形式統計量への応用である。このような確率分布の漸近展開においても、良い近似を得ようとすれば、有理係数多項式の計算を大量にこなす必要があり、展開の次数を上げるにつれて、項の数および係数の大きさが指数的に増大する傾向がある。また、実際に求めた式を用いて数値計算をする場合にも、安定した計算が可能な段階まで数式処理システムを用いて変形を行わねばならないことが多い。

§ 1. 2 次形式統計量

確率変数列 $x_1, x_2, \dots, x_n$ は、互いに独立に、標準正規分布 (分布関数を Φ で、確率密度関数を ϕ で表わす) に従うとする。

いま、有限で正の値をもつ定数列

$$\underline{d} = \{d_1, d_2, \dots, d_n\}, \underline{c} = \{c_1, c_2, \dots, c_n\}$$

により、

$$Q_n = Q_n(\underline{d}, \underline{c}) = \sum_{j=1}^n c_j (x_j + d_j)^2$$

を作り、この分布を考える。一般に、 n 変量正規分布に従う確率ベクトルの正值 2 次形式の分布は、適当に線型変換することにより、 Q_n の分布に帰着することができる。

2 次形式統計量 Q_n の r 次のキュムラント θ_r は、

$$\theta_r = \sum_{j=1}^n c_j^r (1 + r a_j^2)$$

とおくと、

$$x_r = 2^{r-1} (r-1)! \theta_r$$

と書ける。

$n \rightarrow \infty$ により、 $\theta_1 \rightarrow \infty$ とするとき、

$$w_r = \theta_r / \theta_1 = O(1)$$

と仮定すれば、 Q_n は正規分布に法則収束する。この場合、

$$X_n = \sqrt{\theta_1} (Q_n - \theta_1) / \sqrt{(2\theta_1\theta_2)}$$

の分布を、 θ_1 に関して Φ の周りで漸近展開する ($1/\sqrt{\theta_1}$ のべき級数に展開) ことにより、求める Q_n の分布の近似値を得ることができよう。

θ_1 が大きい値の場合はここまでで充分であるが、そうでない場合には、このような直接の漸近展開では充分な近似を得ることができない。そこで、適当な狭義の単調関数で統計量 Q_n を変換し、正規分布への収束を速めてから漸近展開を行うことを考える。

§ 2. 正值 2 次形式における正規化変換

高次の漸近展開を行うための正規化変換については、Niki & Konishi (1984) に一般論が述べられている。簡単にいえば、3 次のキウムラントの漸近展開の主項 ($-1/2$ 次の項) を 0 とするような『対称化変換』でなければならない、ということである。

正值 2 次形式について正規化変換を求めると、『べき乗変換』が得られる。これは、2 次形式の最も簡単な場合であるカイ 2 乗変量に対して非常に有効な Wilson-Hilferty の方法 (1/3 乗変換) を含んでいる。

べき乗変換された統計量

$$R_n = \sqrt{\theta_1} \{ (Q_n / \theta_1)^h - 1 \} / \sqrt{(2h^2 w_2)}$$

の 3 次のキウムラントの $\theta_1^{-1/2}$ の係数を 0 とするためには、

$$h = 1 - (2\theta_1\theta_3) / (3\theta_2^2) = 1 - (2w_3) / (3w_2^2)$$

とすればよいことがわかる。ここで、2 次形式 Q_n がカイ 2 乗変量となるような条件、すなわち、

$$\underline{d} = \{0, 0, \dots, 0\}, \quad \underline{c} = \{1, 1, \dots, 1\}$$

に対しては、 $h = 1/3$ となって Wilson-Hilferty の変換に一致すること、また、 n を固定した上で $\|\underline{d}\| \rightarrow \infty$ とするような条件に対しては、 $h \rightarrow 1/2$ と、非心カイ 2 乗変量に用いられる Fisher の平方根変換に一致すること、および、もともと対称に近い分布については $h = 1$ (すなわち何もしない) に近いことを注意しておく。

§ 3. 分布関数の漸近展開 (Edgeworth 展開)

変換された統計量の分布の漸近展開については、

$$Y_n = R_n + (\theta_3 \sqrt{2}) / (3\theta_1\theta_2 \sqrt{\theta_2})$$

とし、

$$\Pr \{Y_n < x\} \sim \Phi(x) - \phi(x) \sum_{j=1}^{\infty} a_j \theta_1^{-j/2}$$

と表わすと、 a_j ($j = 1, 2, 3, 4$) は次のようになる。ここに、 $H_j(x)$ は j 次のエルミート多項式を示す。

$$a_1 = 0$$

$$a_2 = w_2^{-3} \left\{ H_3(x) \left(\frac{1}{2} w_4 w_2 - \frac{20}{27} w_3^2 + \frac{2}{9} w_3 w_2^2 \right) + H_1(x) w_3 \left(-\frac{2}{3} w_3 + \frac{2}{3} w_2^2 \right) \right\}$$

$$a_3 = \sqrt{2} w_2^{-9/2} \left\{ H_4(x) \left(\frac{2}{5} w_5 w_2^2 - \frac{4}{3} w_4 w_3 w_2 + \frac{76}{81} w_3^3 + \frac{1}{9} w_3^2 w_2^2 - \frac{1}{9} w_3 w_2^4 \right) \right\}$$

$$\begin{aligned}
& +H_2(x)w_3\left(-2w_4w_2+\frac{184}{81}w_3^2+\frac{4}{9}w_3w_2^2-\frac{2}{3}w_2^4\right)+w_3\left(\frac{2}{9}w_3^2+\frac{1}{9}w_3w_2^2-\frac{1}{3}w_2^4\right)\Big\} \\
a_4 = & w_2^{-6}\left\{H_7(x)\left(\frac{1}{8}w_4w_2^2-\frac{10}{27}w_4w_3w_2+\frac{1}{9}w_4w_3w_2^3+\frac{200}{729}w_3^4-\frac{40}{243}w_3^3w_2+\frac{2}{81}w_3^2w_2^4\right)\right. \\
& +H_5(x)\left(\frac{2}{3}w_6w_2^3-\frac{8}{3}w_5w_3w_2^2+\frac{7}{3}w_4w_3^2w_2+\frac{5}{3}w_4w_3w_2^3+\frac{56}{405}w_3^4-\frac{1016}{405}w_3^3w_2+\frac{32}{135}w_3^2w_2^4\right. \\
& \left.+\frac{2}{15}w_3w_2^6\right) \\
& +H_3(x)w_3\left(-\frac{16}{3}w_5w_2^2+\frac{104}{9}w_4w_3w_2+\frac{16}{3}w_4w_2^3-\frac{1106}{243}w_3^3-\frac{76}{9}w_3^2w_2^2+\frac{2}{27}w_3w_2^4+\frac{4}{3}w_2^6\right) \\
& \left.+H_1(x)w_3\left(\frac{28}{9}w_4w_3w_2+\frac{8}{3}w_4w_2^3-\frac{560}{243}w_3^3-\frac{392}{81}w_3^2w_2^2-\frac{20}{27}w_3w_2^4+2w_2^6\right)\right\}
\end{aligned}$$

§ 4. Cornish-Fisher逆展開

分布のパーセントを求めるためのCornish-Fisher逆展開は、次のように与えられる。

いま、 q_α を Q_n の 100α パーセント点とし、

$$q_\alpha = \theta_1 \{h\theta_1^{-1}x_\alpha \sqrt{(2\theta_2)} + 1 + h(h-1)\theta_1^{-2}\theta_2\}^{1/h}$$

とおくと、 x_α は標準正規分布の 100α パーセント点 u_α を用いて、

$$x_\alpha \sim u_\alpha + \sum_{j=1}^{\infty} b_j \theta_1^{-j/2}$$

と展開される。係数 b_j ($j=1, 2, 3, 4$) は次のような値である。

$$b_1 = 0$$

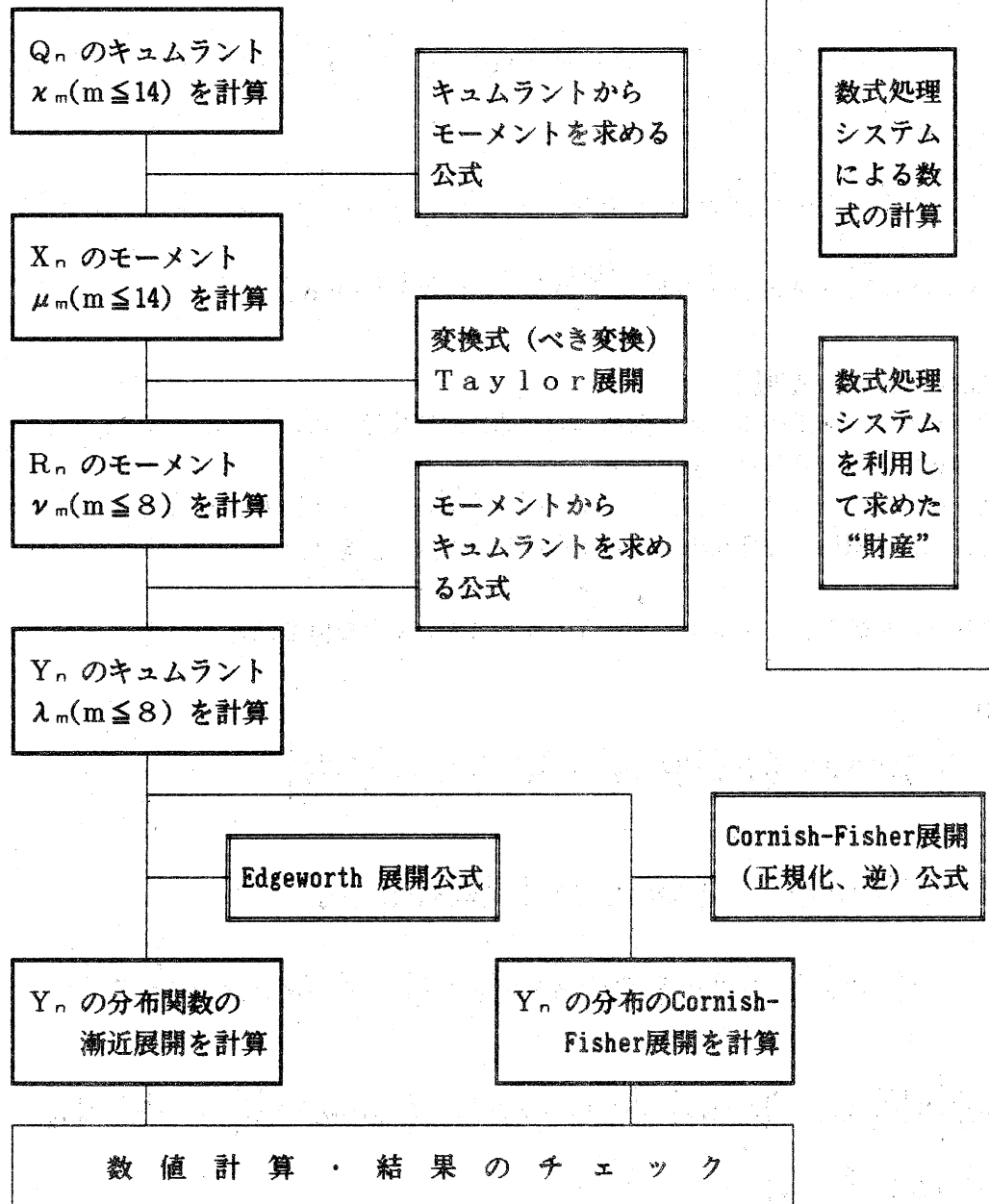
$$b_2 = w_2^{-3} \left\{ u_\alpha^3 \left(\frac{1}{2}w_4w_2 - \frac{20}{27}w_3^2 + \frac{2}{9}w_3w_2^2 \right) + u_\alpha \left(-\frac{3}{2}w_4w_2 + \frac{14}{9}w_3^2 \right) \right\}$$

$$\begin{aligned}
b_3 = & \sqrt{2}w_2^{-9/2} \left\{ u_\alpha^4 \left(\frac{2}{5}w_5w_2^2 - \frac{4}{3}w_4w_3w_2 + \frac{76}{81}w_3^3 + \frac{1}{9}w_3^2w_2^2 - \frac{1}{9}w_3w_2^4 \right) \right. \\
& \left. + u_\alpha^2 \left(-\frac{12}{5}w_5w_2^2 + 6w_4w_3w_2 - \frac{272}{81}w_3^3 - \frac{2}{9}w_3^2w_2^2 \right) + \left(\frac{6}{5}w_5w_2^2 - 2w_4w_3w_2 + \frac{62}{81}w_3^3 \right) \right\}
\end{aligned}$$

$$\begin{aligned}
b_4 = & w_2^{-6} \left\{ u_\alpha^5 \left(\frac{2}{3}w_6w_2^3 - \frac{8}{3}w_5w_3w_2^2 - \frac{9}{8}w_4^2w_2^2 + 6w_4w_3^2w_2 + \frac{1}{3}w_4w_3w_2^3 - \frac{1144}{405}w_3^4 - \frac{52}{135}w_3^3w_2^2 \right. \right. \\
& \left. - \frac{2}{15}w_3^2w_2^4 + \frac{2}{15}w_3w_2^6 \right) \\
& + u_\alpha^3 \left(-\frac{20}{3}w_6w_2^3 + \frac{64}{3}w_5w_3w_2^2 + 9w_4^2w_2^2 - \frac{367}{9}w_4w_3^2w_2 - w_4w_3w_2^3 + \frac{4144}{243}w_3^4 + \frac{20}{27}w_3^3w_2^2 \right. \\
& \left. + \frac{8}{27}w_3^2w_2^4 \right) \\
& \left. + u_\alpha \left(10w_6w_2^3 - 24w_5w_3w_2^2 - \frac{87}{8}w_4^2w_2^2 + \frac{113}{3}w_4w_3^2w_2 - \frac{350}{27}w_3^4 + \frac{4}{27}w_3^3w_2^2 \right) \right\}
\end{aligned}$$

§ 5. 数式処理

前記の式は、数式処理システムであるREDUCE 3.1を利用して、さらに θi^3 のオーダー ($j = 6$) の項まで求めている。その結果、 Q_n の分布が大きく歪んではいない場合、 θi^2 のオーダーまでの式で約3桁、 θi^3 のオーダーの式を用いれば約4桁の精度で分布関数の近似値を得ることができた。処理の概略は、次の図のようにまとめることができよう。



参考文献

- Jensen & Solomon (1972). A Gaussian approximation to the distribution of a definite quadratic form, J. Amer. Statist. Assoc., 67.
- Niki & Konishi (1984). Effects of transformations in higher order asymptotic expansions, Res. Memo. No. 284, Inst. Statist. Math.