Some Results on Univariate and Multivariate Cornish-Fisher Expansion: Algebraic Properties and Validity
Author(s): Akimichi Takemura and Kei Takeuchi
Source: *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)*, Vol. 50, No. 1 (Feb., 1988), pp. 111-136
Published by: Indian Statistical Institute
Stable URL: http://www.jstor.org/stable/25050684

Sankhyā : The Indian Journal of Statistics
1988, Volume **50**, Series A, Pt. 1, pp. 111-136.

SOME RESULTS ON UNIVARIATE AND MULTIVARIATE CORNISH-FISHER EXPANSION: ALGEBRAIC PROPERTIES AND VALIDITY

By AKIMICHI TAKEMURA and KEI TAKEUCHI
University of Tokyo, Japan

SUMMARY. We give a proof of the fact that the degree of the polynomial of order $n^{-k/2}$ in Cornish-Fisher expansion is $k+1$. The result is generalized to multivariate Cornish-Fisher expansion. Furthermore we establish validity of Cornish-Fisher expansion.

1. Introduction

Let X_n, $n = 1, 2, \ldots$, be a sequence of continuous random variables such that $X_n \to N(0, 1)$ in distribution as $n \to \infty$. Let F_n denote the distribution function of X_n. Starting from expansion of the cumulant generating function, F_n is often approximated by formal Edgeworth series :

$$F_n(x) \sim \Phi(x) - \phi(x) \sum_{i=1}^{\infty} g_i(x) n^{-i/2}, \qquad \ldots \quad (1.1)$$

where $\Phi'(x) = \phi(x)$ is the standard normal density and g_i is a polynomial in x. Sometimes it is useful to derive from (1.1) expansion of quantiles :

$$\Phi^{-1}(F_n(x)) \sim x + \sum_{i=1}^{\infty} A_{i+1}(x)\, n^{-i/2}, \qquad \ldots \quad (1.2)$$

$$F_n^{-1}(\Phi(u)) \sim u + \sum_{i=1}^{\infty} B_{i+1}(u)\, n^{-i/2}, \qquad \ldots \quad (1.3)$$

where A_{i+1}, B_{i+1} are polynomials. These expansions are called Cornish-Fisher expansions (Cornish and Fisher, 1937, Fisher and Cornish, 1960). Takeuchi (1975) discusses these expansions in detail with many examples. A general recursive algorithm for obtaining (1.2) and (1.3) from (1.1) is discussed by Hill and Davis (1968).

In this article we clarify some questions on Cornish-Fisher expansions which have been puzzling us for some time.

AMS (1980) *subject classification :* 62E20.

Key words and phrases : Cornish-Fisher expansion, validity.

The first problem concerns the algebraic properties of polynomials A_i, B_i. From explicit forms of the polynomials A_i, B_i, it is apparent that $\deg A_i = i$ and $\deg B_i = i$, $i = 2, 3, \ldots$, whereas in the Edgeworth expansion (1.1) $\deg g_i = 3i-1$. Actually we observe that many terms cancel in converting (1.1) into (1.2) or (1.3). This is by no means obvious and there does not seem to be a published proof of this fact. In section 2 we give a proof that $\deg A_i = i$ and $\deg B_i = i$. Takeuchi (1978) gave a generalization of Cornish-Fisher expansion to the multivariate case. In section 3 we prove analogous results on the polynomials of the multivariate case.

This result on the degree of polynomials in Cornish-Fisher expansion has a rather close connection with combinatorial result of James (James, 1955, 1958 ; James and Mayne, 1962) concerning the order of cumulants of a transformed random variable. In section 4 we show that James' result is almost an immediate consequence of our result on Cornish-Fisher expansion.

Another question we address is the interpretation of Cornish-Fisher expansion and its validity. As long as we regard Cornish-Fisher expansion as expansion of quantiles, there is conceptual difficulty in generalizing it to the multivariate case, since quantile function is not well defined in the multivariate case. However if we regard Cornish-Fisher expansion as approximating the distribtion of X_n by the distribution of a polynomial of a standard normal random variable U :

$$X_n \sim U + \sum_{i=1}^{\infty} B_{i+1}(U) n^{-i/2}, \qquad \cdots \quad (1.4)$$

then there is no conceptual difficulty. (1.4) means that the distribution of both sides agree up to an appropriate order of $n^{-1/2}$). Let $\boldsymbol{X} = \boldsymbol{X}_n = (X_1, \ldots, X_p)$ be a random vector such that $\boldsymbol{X}_n \to N(\mathbf{0}, \boldsymbol{\Sigma})$ in distribtion as $n \to \infty$. Corresponding to (1.4) Takeuchi (1978) defined a multivariate Cornish-Fisher expansion of $\boldsymbol{X}_n$ in the following "triangular" form :

$$\begin{aligned}
X_1 &\sim U_1 + \frac{1}{n^{1/2}} B_{1,2}(U_1) + \frac{1}{n} B_{1,3}(U_1) + \cdots, \\
X_2 &\sim U_2 + \frac{1}{n^{1/2}} B_{2,2}(U_1, U_2) + \frac{1}{n} B_{2,3}(U_1, U_2) + \cdots, \qquad \cdots \quad (1.5) \\
&\cdots \\
X_p &\sim U_p + \frac{1}{n^{1/2}} B_{p,2}(U_1, \ldots, U_p) + \frac{1}{n} B_{p,3}(U_1, \ldots, U_p) + \cdots,
\end{aligned}$$

where $B_{j,k}$ are polynomials and $\boldsymbol{U} = (U_1, \ldots, U_p)$ is distributed according to $N(\mathbf{0}, \boldsymbol{\Sigma})$. Takeuchi (1978) showed that $B_{j,k}$ can be uniquely determined by requiring that the characteristic function of the right hand side of (1.5) agree with the characteristic function of $\boldsymbol{X}_n$ when they are expanded in powers of $n^{-1/2}$. In Section 5 we regard Cornish-Fisher expansion as approximation of the distribution and establish its validity. More precisely we establish the validity of Cornish-Fisher expansion in terms of variation distance between the distributions of X_n(or $\boldsymbol{X}_n$) and of the right hand side of (1.4) (or (1.5)).

In this article X_n's are assumed to be continuous random variables. When X_n's are lattice random variables, Cornish-Fisher expansion has to be defined more carefully. This will be discussed in a subsequent paper.

2. Univariate Cornish-Fisher expansion

Let $\kappa_r(X_n)$ denote the r-th cumulant of X_n. We assume that the cumulants have asymptotic expansions of the form :

$$\kappa_r(X_n) \sim n^{r/2} \sum_{j=2r-2}^{\infty} \gamma_{r,j}\, n^{-j/2}, \qquad r \geqslant 1, \qquad \ldots \quad (2.1)$$

where $\gamma_{1,0} = \gamma_{1,1} = 0$, $\gamma_{2,2} = 1$. This is a standard assumption (see Withers (1982)). In particular if X_n is a normalized sum of i.i.d. random variables, then only the first term of the summation of (2.1) is nonzero. Now let the formal Cornish-Fisher expansion of X_n in terms of standard normal random variable U as

$$X_n \sim U + \frac{1}{n^{1/2}} B_2(U) + \frac{1}{n} B_3(U) + \ldots . \qquad \ldots \quad (2.2)$$

As discussed in Introduction, (2.2) means that the distribution of X_n is approximated by the distribution of the right hand side polynomial in the standard normal variable U up to an appropriate order of $n^{-1/2}$. The main result of this section is

Theorem 2.1 : *Degree of the polynomial B_k is k. More precisely, the degree of B_k does not exceed k and the coefficient of u^k in $B_k(u)$ is a nonzero polynomial in $\gamma_{r,j}$, $j-r \leqslant k-1$.*

Converse of Theorem 2.1 is also useful.

Theorem 2.2 : *Let a random variable X_n be formally defined by the right hand side of* (2.2), *where U is a standard normal random variable and B_k is a polynomial of degree k. Then the r-th order cumulant of X_n is of the order* $O(n^{-(r-2)/2})$, $r \geqslant 3$.

A 1–15

We have a corresponding result for the alternative type of Cornish-Fisher expansion where U is expressed as a series in X_n by inverting (2.2). Let

$$U \sim X_n + \frac{1}{n^{1/2}} A_2(X_n) + \frac{1}{n} A_3(X_n) + \cdots . \qquad \cdots \quad (2.3)$$

Then we have

Theorem 2.3 : *deg* $A_k = k$.

The rest of this section is devoted for proving Theorem 2.1 and Theorem 2.2. Theorem 2.3 can be proved entirely analogously as Theorem 2.1, or it can be derived from Theorem 2.1 using (A2.3) of Witheres (1983) which gives the relation between (2.2) and (2.3). Our discussion in this section is purely formal and algebraic. However under standard regularity conditions our results can be justified. This will be discussed in Section 5.

In the course of our proof we obtain results which may be of independent interest. Our method seems to be very useful in proving algebraic properties involved in calculation of moments and cumulants. In particular our operator Δ in (2.11) is useful in recursive calculation of moments and cumulants.

Proofs given below might seem somewhat difficult to follow without working out a few simple cases. As soon as one writes out several examples, ideas behind the proofs should become obvious.

For notational convenience let

$$\begin{aligned} \beta_r &= \kappa_r (X_n)/r\,!, & r &\neq 2 \\ &= (\kappa_r (X_n)-1)/r\,!, & r &= 2. \end{aligned} \qquad \cdots \quad (2.4)$$

Note that

$$\beta_r = O(n^{-(r-2)/2}), \qquad \text{for} \quad r > 2. \qquad \cdots \quad (2.5)$$

Then the cumulant generating function $\xi(t)$ of X_n can be written as

$$\xi(t) = -\frac{t^2}{2} + \sum_{j>0} \beta_j (\mathrm{it})^j. \qquad \cdots \quad (2.6)$$

Exponentiating (2.6), the characteristic function $\psi(t)$ is given as

$$\psi(t) = e^{-t^2/2}\left(1 + C + \frac{C^2}{2} + \cdots\right), \qquad \cdots \quad (2.7)$$

where $C = \Sigma\beta_j\,(it)^j$. Taking the inverse Fourier transform of (2.7) the density $f(x)$ of X_n is written as

$$\begin{aligned} f(x) &= \phi(x)\left[1 + (\beta_1 H_1 + \beta_2 H_2 + \ldots) + \frac{1}{2}(\beta_1^2 H_2 + 2\beta_1\beta_2 H_3 + \ldots) + \ldots\right] \\ &= \phi(x)\left(1 + E_1 + \frac{1}{2} E_2 + \frac{1}{3!} E_3 + \cdots\right), \end{aligned} \qquad \cdots \quad (2.8)$$

where $H_i = H_i(x)$ is the i-th Hermite polynomial and

$$E_1 = \sum_{i>0} \beta_i H_i,$$

$$E_2 = \sum_{i,j>0} \beta_i \beta_j H_{i+j}, \qquad \ldots \quad (2.9)$$

$$\ldots$$

$$E_k = \sum_{i(1),\ldots,i(k)>0} \beta_{i(1)} \beta_{i(2)} \ldots \beta_{i(k)} H_{i(1)+\cdots+i(k)}.$$

with $i(1), \ldots, i(k)$ running independently in the summation. (2.8) and (2.9) give the explicit expansion of the density function of X_n. The use of the notation $i(1), i(2), \ldots,$ for more than one indices makes our notation consistent with the multivariate generalization discussed in the next section.

Now the key idea in our proof is to consider the logarithm of the density function (2.8). We will obtain an explicit expression of $\log f(x)$ in terms of Hermite polynomials. We will then show that systematic cancelling of terms occurs in $\log f(x)$. Now

$$\log f(x) = \log \phi(x) + \log \left(1 + E_1 + \frac{1}{2} E_2 + \ldots\right)$$

$$= \log \phi(x) + E_1 + \frac{1}{2}(E_2 - E_1^2)$$

$$+ \frac{1}{3!}(E_3 - 3E_2E_1 + 2E_1^3)$$

$$+ \ldots$$

$$= \log \phi(x) + Q_1 + \frac{1}{2} Q_2 + \frac{1}{3!} O_3 + \ldots, \qquad \ldots \quad (2.10)$$

where

$$Q_k = k! \sum_{\substack{l \\ s_1,\ldots,s_l}} \frac{(-1)^{l-1}}{l} \frac{E_{s_1} E_{s_2} \ldots E_{s_l}}{s_1!\, s_2! \ldots s_l!}$$

and $(s_1, \ldots, s_l)$ runs over all *unordered* partitions of k.

Now we define a symbolic linear operator Δ operating on products of E_s's as

$$\Delta(E_{s_1} E_{s_2} \ldots E_{s_l})$$

$$= E_{s_1+1} E_{s_2} \ldots E_{s_l} + E_{s_1} E_{s_2+1} E_{s_3} \ldots E_{s_l} + \ldots + E_{s_1} \ldots E_{s_{l-1}} E_{s_l+1}$$

$$-l\, E_1 E_{s_1} E_{s_2} \ldots E_{s_l}. \qquad \ldots \quad (2.11)$$

For a sum of products of E_s's, Δ is defined by linearity. Then we have

Lemma 2.1 :

$$\Delta Q_k = Q_{k+1}. \qquad \cdots \quad (2.12)$$

Defining $Q_1 = \Delta 1$, (2.9) can then be written as

Lemma 2.2 :

$$\log f(x) = \log \phi(x) + \sum_{k=1}^{\infty} \frac{1}{k!} \Delta^k 1. \qquad \cdots \quad (2.13)$$

Proof of Lemma 2.1.

$$\Delta Q_k = \Sigma \frac{k!(-1)^{l-1}}{l s_1! \ldots s_l!} (E_{s_1+1} E_{s_2} \ldots E_{s_l} + \ldots + E_{s_1} \ldots E_{s_{l-1}} E_{s_l+1}$$
$$-lE_1 E_{s_1} E_{s_2} \ldots E_{s_l}). \qquad \cdots \quad (2.14)$$

Consider the coefficient of the term $E_{t_1} E_{t_2} \ldots E_{t_m}$ in (2.14).

Case 1 : $t_1 \geqslant 2, \ldots, t_m \geqslant 2$.

The term arises from m places corresponding to m-tuples $(t_1-1, t_2, \ldots, t_m)$, $\ldots, (t_1, t_2, \ldots, t_m-1)$. Hence the coefficient is

$$\frac{k!(-1)^{m-1}}{m t_1! \ldots t_m!}(t_1 + \ldots + t_m) = \frac{(k+1)!(-1)^{m-1}}{m t_1! \ldots t_m!}. \qquad \cdots \quad (2.15)$$

Case 2 : $t_\alpha = 1$ for some α.

Write

$$\frac{k!(-1)^{l-1}}{l s_1! \ldots s_l!}(-l E_1 E_{s_1} \ldots E_{s_l})$$
$$= \frac{k!(-1)^l}{(l+1) s_1! \ldots s_l!}(E_1 E_{s_1} \ldots E_{s_l} + E_{s_1} E_1 E_{s_2} \ldots E_{s_l}$$
$$+ \ldots + E_{s_1} \ldots E_{s_l} E_1). \qquad \cdots \quad (2.16)$$

Then the coefficient is seen to be

$$\frac{k!(-1)^{m-1}}{m t_1! \ldots t_m!}\left(\sum_{t_\alpha \geqslant 2} t_\alpha\right) + \frac{k!(-1)^{m-1}}{m t_1! \ldots t_m!}\left(\sum_{t_\alpha = 1} t_\alpha\right)$$
$$= \frac{k!(-1)^{m-1}}{m t_1! \ldots t_m!}(t_1 + \ldots + t_m) = \frac{(k+1)!(-1)^{m-1}}{m t_1! \ldots t_m!}. \qquad \cdots \quad (2.17)$$

Q.E.D.

Next, we define a symbolic linear operator Δ_j operating on products of Hermite polynomials by

$$\Delta_j(H_{i(1)} \ldots H_{i(l)}) = H_{i(1)+j} H_{i(2)} \ldots H_{i(l)} + H_{i(1)} H_{i(2)+j} \ldots H_{i(l)}$$
$$+ \ldots + H_{i(1)} \ldots H_{i(l-1)} H_{i(l)+j} - l H_j H_{i(1)} \ldots H_{i(l)}. \qquad \cdots \quad (2.18)$$

For a sum of products of Hermite polynomials Δ_j is defined by linearity.

Remark 2.1 : In (2.18) H_i*'s* are regarded as independent symbols rather than polynomials. For a while we will consider polynomials in the variables $H_1, H_2, \ldots$. Δ_j is a linear operator operating on polynomials in the variables $H_1, H_2, \ldots$. For example, as a polynomial in x we have the equality

$$H_i(x)\, H_j(x) = H_{i+j}(x) + ij\, H_{i+j-2}(x) + \ldots .$$

However as polynomials in $H_1, H_2, \ldots$, two sides are considered to be different. Actually

$$\Delta_k(H_i H_j) \neq \Delta_k(H_{i+j}) + ij\, \Delta_k(H_{i+j-2}) + \ldots .$$

From (2.9) we see that a product $E_{s_1} \ldots E_{s_l}$ of E_s*'s* is a polynomial in $H_1, H_2, \ldots$. We now prove

Lemma 2.3 : $\Delta = \sum_{i>0} \beta_i \Delta_i$, ... (2.19)

i.e.

$$\Delta(E_{s_1} \ldots E_{s_l}) = \sum_i \beta_i \Delta_i(E_{s_1} \ldots E_{s_l}),$$

where both sides are considered as polynomials in $H_1, H_2, \ldots$.

Proof : Note

$$\Delta(E_{s_1} \ldots E_{s_l}) = (E_{s_1+1} - E_1 E_{s_1}) E_{s_2} \ldots E_{s_l}$$
$$+ \ldots + E_{s_1} \ldots E_{s_{l-1}}(E_{s_l+1} - E_1 E_{s_l}) \qquad \ldots \quad (2.20)$$

and

$$\Delta_j(H_{i(1)} \ldots H_{i(l)}) = (H_{i(1)+j} - H_j H_{i(1)}) H_{i(2)} \ldots H_{i(l)}$$
$$+ \ldots + H_{i(1)} \ldots H_{i(l-1)}\, (H_{i(l)+j} - H_j H_{i(l)}). \qquad \ldots \quad (2.21)$$

Now consider $\beta_i \Delta_i(E_{s_1} \ldots E_{s_l})$. From (2.9) a general term of $E_{s_1} \ldots E_{s_l}$ can be written as

$$(\beta_{j(1,1)} \ldots \beta_{j(1,s_1)}) \ldots (\beta_{j(l,1)} \ldots \beta_{j(l,s_l)})$$
$$\times H_{j(1,*)} \ldots H_{j(l,*)}, \qquad \ldots \quad (2.22)$$

where

$$j(1, *) = j(1, 1) + \ldots + j(1, s_1),$$
$$\ldots$$
$$j(l, *) = j(l, 1) + \ldots + j(l, s_l).$$

Operating Δ_i on (2.22) and summing up similar terms we obtain

$$\beta_i\Delta_i(E_{s_1}\cdots E_{s_l}) \qquad \cdots \quad (2.23)$$

$$= \beta_i \sum_{j(1,1),\ldots,j(1,s_1)} \prod_{\alpha} \beta_{j(1,\alpha)} (H_{j(1,*)+i} - H_i H_{j(1,*)}) E_{s_2} \cdots E_{s_l}$$

$$+ \cdots +$$

$$+ \beta_i E_{s_1} \cdots E_{s_l} \sum_{j(l,1)\ldots,j(l,s_l)} \prod_{\alpha} \beta_{j(l,\alpha)} (H_{j(l,*)+i} - H_i H_{j(l,*)}).$$

Then adding (2.23) with respect to i we obtain form (2.9)

$$(\Sigma\, \beta_i\Delta_i)E_{s_1}\cdots E_{s_l} = (E_{s_1+1} - E_1E_{s_1})E_{s_2}\cdots E_{s_l}$$

$$+ \cdots + E_{s_1}\cdots E_{s_{l-1}}(E_{s_l+1} - E_1E_{s_l})$$

$$= \Delta\,(E_{s_1}\cdots E_{s_l}). \qquad \text{Q.E.D.}$$

Lemma 2.4 : $\Delta_i\Delta_j = \Delta_j\Delta_i$. $\qquad \cdots \quad (2.24)$

Proof is easy and omitted.

From Lemma 2.2—Lemma 2.4 we have

Lemma 2.5 :

$$\log f(x) = \log\phi(x) + \sum_{k=1}^{\infty} \frac{1}{k\,!} \left(\sum_{i>0} \beta_i\Delta_i\right)^k 1$$

$$= \log\phi(x) + \sum_k \frac{1}{k\,!} \sum_{i(1),\ldots,i(k)} \beta_{i(1)}\cdots\beta_{i(k)}\Delta_{i(1)}\cdots\Delta_{i(k)}\, 1$$

$$= \log\phi(x) + \sum_{j(1)\ldots,i(m)}^{m} \frac{\beta_1^{j(1)}\cdots\beta_m^{j(m)}}{j(1)\,!\cdots j(m)\,!}\Delta_1^{j(1)}\cdots\Delta_m^{j(m)}\, 1, \qquad \cdots \quad (2.25)$$

where on the second line of (2.25) $i(1), \ldots, i(k)$ *run independently whereas on the third line only distinct terms are counted.*

If we can obtain explicit expression for $\Delta_{i(1)}\,\Delta_{i(2)}\cdots\Delta_{i(l)}1$ in terms of Hermite polynomials, then we can obtain explicit expression for $\log f(x)$ from (2.25). The following lemma gives an explicit expression for $\Delta_{i(1)}\,\Delta_{i(2)}\cdots\Delta_{i(l)}\,1$.

Lemma 2.6 : *Let* $(1, \ldots, l)$ *be grouped into* m *mutually disjoint nonempty subsets* $J_1, \ldots, J_m$: $(1, \ldots, l) = J_1 \bigcup \cdots \bigcup J_m$ *and* $J_s \bigcap J_t = \phi$ *if* $s \neq t$. *Let* $j(1) = \sum_{s\varepsilon J_1} i(s), \ldots, j(m) = \sum_{s\varepsilon J_m} i(s)$. *Then*

$$\Delta_{i(1)}\cdots\Delta_{i(l)}1 = \Sigma(-1)^{m-1}(m-1)\,!\,H_{j(1)}H_{j(2)}\cdots H_{j(m)}, \qquad \cdots \quad (2.26)$$

where the summation is over all different groupings of $(1, \ldots, l)$

For example by direct calculation we readily verify

$$\Delta_i\Delta_j\Delta_k 1 = H_{i+j+k}-H_iH_{j+k}-H_jH_{i+k}-H_kH_{i+j}+2\,H_iH_jH_k,$$

$$\begin{aligned}\Delta_i\Delta_j\Delta_k\Delta_h 1 = {} & H_{i+j+k+h}-H_iH_{j+k+h}-H_jH_{i+k+h}-H_kH_{i+j+h}-H_hH_{i+j+k} \\ & -H_{i+j}H_{k+h}-H_{i+k}H_{j+h}-H_{i+h}H_{j+k}+2H_iH_jH_{k+h} \\ & +2H_{i+j}H_kH_h+2H_iH_kH_{j+h}+2H_{i+k}H_jH_h+2H_iH_hH_{j+k} \\ & +2H_{i+h}H_jH_k-6\,H_iH_jH_kH_h. \qquad \ldots \quad (2.27)\end{aligned}$$

Proof of Lemma 2.6 : Proof is by induction. If $l=1$ then $\Delta_i 1 = H_i$ which satisfies (2.26). For induction assume that (2.26) holds for $l=t$. Then for $l=t+1$

$$\begin{aligned}& \Delta_{i(1)}\ldots\Delta_{i(t)}\Delta_{i(t+1)}1 = \Delta_{i(t+1)}\Delta_{i(1)}\ldots\Delta_{i(t)}1 \\ & = \Sigma\,\{(-1)^{m-1}(m-1)\,!\,(H_{j(1)+i(t+1)}H_{j(2)}\ldots H_{j(m)}+\ldots \\ & \quad +H_{j(1)}\ldots H_{j(m-1)}H_{j(m)+i(t+1)})+(-1)^m m\,!\,H_{i(t+1)}H_{j(1)}\ldots H_{j(m)}\} \\ & = \Sigma\,(-1)^{m-1}(m-1)\,!\,H_{j'(1)}H_{j'(2)}\ldots H_{j'(m)},\end{aligned}$$

where on the extreme right hand side the summation is over all different groupings of $(1, \ldots, t+1)$. Q.E.D.

Remark 2.2 : Our expression (2.26) is closely related to a general expression of joint cumulants of l random variables in terms of their joint moments. See Section 2.3 of Brillinger (1981) or Section 1 of Leonov and Shiryaev (1959).

Lemma 2.5 and Lemma 2.6 give an explicit expression of $\log f(x)$.

Up to this point we have been considering $H_1, H_2, \ldots$ as independent symbols. Now we again consider $H_i = H_i(x)$ as polynomials in x and investigate properties of $\Delta_{i(1)}\ldots\Delta_{i(l)}1$ as a polynomial in x. In particular we are interested in the degree of $\Delta_{i(1)}\ldots\Delta_{i(l)}1$ as a polynomial in x.

Let $D = d/dx$ be the differential operator. Then

Lemma 2.7 : For $i(1) > 1, \ldots, i(l) > 1$

$$\begin{aligned}D(\Delta_{i(1)}\ldots\Delta_{i(l)}1) = {} & i(1)\,\Delta_{i(1)-1}\Delta_{i(2)}\ldots\Delta_{i(l)}1 \\ & + i(2)\,\Delta_{i(1)}\Delta_{i(2)-1}\ldots\Delta_{i(l)}1 \\ & + \ldots \\ & + i(l)\Delta_{i(1)}\Delta_{i(2)}\ldots\Delta_{i(l)-1}1. \qquad \ldots \quad (2.28)\end{aligned}$$

Proof: Using the relation $DH(x) = i\,H_{i-1}(x)$ we have

$$D(\Sigma\,(-1)^{m-1}(m-1)\,!\,H_{j(1)}\cdots H_{j(m)})$$

$$= \Sigma\,(-1)^{m-1}(m-1)\,!\,D(H_{j(1)}\cdots H_{j(m)})$$

$$= \Sigma\,(-1)^{m-1}(m-1)\,!\,(j(1)\,H_{j(1)-1}H_{j(2)}\cdots H_{j(m)}+\cdots$$

$$+j(m)\,H_{j(1)}\cdots H_{j(m-1)}H_{j(m)-1}). \qquad \cdots \quad (2.29)$$

Recall that $j(1),\dots,j(m)$ are sums of $i(s)$'s. Each $i(s)$ appears exactly once in the summand of the extreme right hand side of (2.29). Consider, for example, terms involving $i(1)$. Then we see that the sum of these terms is over different groupings of $(1,\dots,l)$ with $i(1)$ replaced by $i(1)-1$. Therefore summing up terms involving $i(1),\dots,i(l)$ separately we obtain the lemma.
Q.E.D.

This lemma does not hold if $i(s)=1$ for some s, because in Lemma 2.6 all subscripts have to be positive. Now we consider this case. Because of Lemma 2.4 we assume $i(1)=\dots=i(h)=1<i(h+1),\dots,i(l)$ without loss of generality. Δ_1's operating from the left show the following property.

Lemma 2.8 : $$\Delta_1\Delta_{i(2)}\cdots\Delta_{i(l)}1 = -D(\Delta_{i(2)}\cdots\Delta_{i(l)}1). \qquad \cdots \quad (2.30)$$

Proof: By linearity it suffices to show that

$$\Delta_1(H_{j(1)}\cdots H_{j(m)}) = -D(H_{j(1)}\cdots H_{j(m)}). \qquad \cdots \quad (2.31)$$

The left hand side can be written as

$$(H_{j(1)+1}-H_1H_{j(1)})\,H_{j(2)}\cdots H_{j(m)}+\cdots$$

$$+H_{j(1)}\cdots H_{j(m-1)}\,(H_{j(m)+1}-H_1H_{j(m)}). \qquad \cdots \quad (2.32)$$

Now consider (2.32) as polynomials in x. Then $H_1(x)=x$ and furthermore $H_{j+1}(x) = xH_j(x)-D\,H_j(x)$. Therefore (2.31) is

$$(-DH_{j(1)})H_{j(2)}\cdots H_{j(m)}+\cdots+H_{j(1)}\cdots H_{j(m-1)}\,(-DH_{j(m)})$$

$$= -D(H_{j(1)}\cdots H_{j(m)}). \qquad \text{Q.E.D.}$$

We can now employ mathematical induction to obtain the degree of $\Delta_{i(1)}\cdots\Delta_{i(l)}1$ as a polynomial in x, which we state as a theorem.

Theorem 2.4 : $\deg(\Delta_{i(1)} \dots \Delta_{i(l)} 1) = i(1)+ \dots +i(l)-2(l-1)$. ... (2.33)

Proof : Proof is by induction. $\deg(\Delta_i 1) = \deg(H_i(x)) = i$, which coincides with (2.33). First we want to prove that $\deg(\Delta_{i(1)} \dots \Delta_{i(l)} 1) \leqslant i(1) + \dots +i(l)-2(l-1)$. But this is easily shown by induction on l and $i(1) + \dots +i(l)$ in view of Lemma 2.7 and Lemma 2.8.

To complete the proof we have to check that the leading coefficient (the coefficient of $x^{i(1)+ \dots +i(l)-2(l-1)}$) is nonzero. Let $c(i(1), \dots, i(l))$ denote the leading coefficient. Then again by induction it can be explicitly evaluated as ($l \geqslant 2$)

$$c(i(1), \dots, i(l)) = (-1)^{l-1} i(1) \dots i(l) \prod_{t=0}^{l-3} (i(1)+ \dots +i(l)-l-t). \qquad \dots \quad (2.34)$$

For $l = 2$, $\Delta_i \Delta_j 1 = H_{i+j} - H_i H_j = -ij\, H_{i+j-2} + \dots$. Therfore (2.34) holds. Now we check Lemma 2.7 and Lemma 2.8 for induction. Lemma 2.7 implies

$$(\Sigma\, i(s)-2(l-1))\, c(i(1), \dots, i(l)) = i(1)\, c(i(1)-1, i(2), \dots, i(l)) + \dots + i(l)\, c(i(1), \dots, i(l-1), i(l)-1).$$

Clearly (2.34) satisfies this relation. Lemma 2.8 implies

$$c(1, i(2), \dots, i(l)) = -\Big(\sum_{s=2}^{l} i(s)-2(l-2) \Big) c(i(2), \dots, i(l)).$$

This also is satisfied by (2.34). Now note that (2.34) is nonzero as long as $i(1)+ \dots +i(l)-2(l-1) \geqslant 0$. This completes the proof. Q.E.D.

Now consider the order of n in $\beta_{i(1)} \dots \beta_{i(l)}$. By (2.5) it is

$$-\sum_{s=1}^{l} \frac{i(s)-2}{2} = -\frac{1}{2} [(i(1)+ \dots +i(l)-2(l-1))-2]$$

$$= -\frac{1}{2} (\deg(\Delta_{i(1)} \dots \Delta_{i(l)})-2) \qquad \dots \quad (2.35)$$

if $i(1) > 2, \dots, i(l) > 2$. If $i(s) \leqslant 2$ for some s, then the order of n is smaller than the right hand side of (2.35). Therefore as a corollary to Theorem 2.4 we obtain

Corollary 2.1 : *Write* $\log f(x)$ *as*

$$\log f(x) = -\log(2\pi)/2 - \frac{x^2}{2} + \frac{1}{n^{1/2}}\, g_3^*(x) + \frac{1}{n}\, g_4^*(x) + \dots . \qquad \dots \quad (2.36)$$

Then $g_k^*(x)$ *is a polynomial of degree* k *in* x.

Compared to the Edgeworth expansion of the density, we see that systematic cancelling of terms has occurred, thus reducing the degree of polynomials. It might be helpful to give more explicit expression of (2.36) for the case of sum of i.i.d. random variables. Writing $\gamma_{r,2r-2}=\gamma_r$, $r=3, 4, 5, 6$, the expansion of $\log f(x)$ up to the order n^{-2} is

$$\begin{aligned}\log f(x) &= \log\phi(x)+\frac{\gamma_3}{6n^{1/2}}(x^3-3x)\\ &+\frac{\gamma_4}{24n}(x^4-6x^2+3)-\frac{\gamma_3^2}{24n}(3x^4-12x^2+5)\\ &+\frac{\gamma_5}{120n^{3/2}}(x^5-10x^3+15x)-\frac{\gamma_3\gamma_4}{12n^{3/2}}(x^5-7x^3+8x)\\ &+\frac{\gamma_3^3}{24n^{3/2}}(3x^5-16x^3+15x)+\frac{\gamma_6}{720n^2}(x^6-15x^4+45x^2-15)\\ &-\frac{\gamma_5\gamma_3}{48n^2}(x^6-11x^4+25x^2-7)-\frac{\gamma_4{}^2}{144n^2}(2x^6-21x^4+48x^2-12)\\ &+\frac{\gamma_4\gamma_3^2}{48n^2}(7x^6-59x^4+109x^2-25)\\ &-\frac{\gamma_3^4}{48n^2}(7x^6-48x^4+75x^2-15)+o(n^{-2}). \qquad \dots \quad (2.37)\end{aligned}$$

Now we are ready to prove Theorem 2.1.

Proof of Theorem 2.1 : Let the formal Cornish-Fisher expansion of X_n be given as (2.2). Let $f(x)$ be the density of X_n. Then the logarithm of the density of U is given as (ignoring the constant)

$$\begin{aligned}-\frac{u^2}{2} &= \log f\left(u+\frac{1}{n^{1/2}}B_2(u)+\dots\right)+\log\left(1+\frac{1}{n^{1/2}}B_2'(u)+\dots\right)\\ &= -\frac{1}{2}\left(u+\frac{1}{n^{1/2}}B_2(u)+\dots\right)^2+\frac{1}{n^{1/2}}g_3^*\left(u+\frac{1}{n^{1/2}}B_2(u)+\dots\right)\\ &+\dots+\log\left(1+\frac{1}{n^{1/2}}B_2'(u)+\dots\right). \qquad \dots \quad (2.38)\end{aligned}$$

On the right hand side of (2.38) all terms have to vanish except for $-u^2/2$.

First consider terms of order $n^{-1/2}$. Then we have

$$-u\, B_2(u)+g_3^*(u)+B_2'(u) = 0.$$

Since deg $g_3^* = 3$ we obtain deg $B_2 = 2$.

For general $n^{-l/2}$ we argue by induction. For induction assume deg $B_k \leqslant k$, $k = 2, \ldots, l$. Let $B_1(u) = u$ for notational convenience. Note that terms involving $B_{l+2}, B_{l+3}, \ldots$ are of smaller order than $n^{-l/2}$ and can be ignored. There are only two terms of order $n^{-l/2}$ involving $B_{l+1}(u)$, i.e.,

$$-u\, B_{l+1}(u)+B_{l+1}'(u). \qquad \ldots \quad (2.39)$$

Now consider a general term involving $B_1, \ldots, B_l$. Omitting irrelevant coefficients r-th degree term in $g_q^*(r \leqslant q)$ is of the following form :

$$n^{-(q-2)/2}\, n^{-(\Sigma_{s=1}^r j(s)-r)/2}\, B_{j(1)}(u) \ldots B_{j(r)}(u), \qquad \ldots \quad (2.40)$$

whose degree is $d = \Sigma_{s=1}^r j(s)$. Supposing that (2.40) is of order $n^{-l/2}$ we have $q-2+d-r = l$. Therefore the degree d is

$$d = l+2+r-q \leqslant l+2. \qquad \ldots \quad (2.41)$$

Finally consider a general term arising from logarithm of the Jacobian :

$$n^{-(\Sigma j(s)-r)/2}\, B_{j(1)}'(u) \ldots B_{j(r)}'(u), \qquad \ldots \quad (2.42)$$

whose degree is $\Sigma\, j(s)-r$. Hence using the same notation as above we have $d = l < l+2$. We see that as far as highest degree in u is concerned, terms from the Jacobian can be ignored.

Combining (2.39)—(2.42) we have shown that

$$\begin{aligned} 0 = {} & -u\, B_{l+1}(u)+B_{l+1}''(u) \\ & + \text{polynomial in } u \text{ of degree not exceeding } l+2. \end{aligned} \qquad \ldots \quad (2.43)$$

Therefore deg $B_{l+1} \leqslant l+1$.

We now want to show that the coefficient of u^{l+1} in $B_{l+1}(u)$ is a nonzero polynomial in $\gamma_{r,j}$, $j-r \leqslant l$. From the above induction argument, it is clear that the coefficient is a polynomial in $\gamma_{r,j}$, $j-r \leqslant l$. To show that it is nonzero, consider the term involving $\gamma_{l+2,2l+2}$. There is only one term of order $n^{-\ /2}$ arising in the following form :

$$g_{l+2}^*(x) = \frac{\gamma_{l+2,\,2l+2}}{(l+2)\,!}\, H_{l+2}(x)+\text{terms not involving } \gamma_{l+2,\,2l+2}. \qquad \ldots \quad (2.44)$$

Hence from (2.39) it follows that

$$B_{l+1}(u) = \frac{\gamma_{l+2,2l+2}}{(l+2)\,!} H_{l+1}(u) + \text{terms not involving } \gamma_{l+2,2l+2}. \qquad \dots \quad (2.45)$$

Q.E.D.

This completes the proof.

Remark 2.3 : In (2.43) the last term is given in terms of $B_2, \dots, B_l$. Therefore using (2.43) we can recursively calculate $B_2 B_3, \dots$. In section 5, where validity of Cornish-Fisher expansion is discussed, we assume that $B_2, B_3 \dots$ are determined in this way in order to justify our derivation above.

We finally give a proof of Theorem 2.2. Since details of the proof are similar to that of Theorem 2.1 we only discuss main points of the proof.

Proof of Theorem 2.2 : From (2.2) we can formally calculate the moments of X_n and then expand the density of X_n in Edgeworth series. The logarithm of the density can be expressed as (2.36) where g_3^*, g_4^*, ..., are certain polynomials in x. Then analogous to the proof of Theorem 2.1 we can show that $\deg g_k^* = k$. Now writing the logarithm of the density in powers of x we have

$$\log f(x) = -\log(2\pi)/2 - \frac{x^2}{2} + c_{0,n} + c_{1,n}x + c_{2,n}x^2 + \dots, \qquad \dots \quad (2.46)$$

where

$$c_{j,n} = O(n^{-1/2}), \qquad j = 0, 1, 2,$$

$$c_{j,n} = O(n^{-(j-2)/2}), \qquad j \geqslant 3.$$

Now note that obtaining cumulant generating function from logarithm of the density function is algebraically entirely analogous to the inverse operation in view of the symmetry between Fourier transform and inverse Fourier transform. Therefore analogus to Corollary 2.1 we can show that the cumulant generating function of X_n can be written as

$$-\frac{t^2}{2} + \frac{1}{n^{1/2}} \tilde{g}_3(\mathrm{it}) + \frac{1}{n} \tilde{g}_4(\mathrm{it}) + \dots, \qquad \dots \quad (2.47)$$

where $\tilde{g}_r$ is a polynomial of degree r. This implies the theorem. Q.E.D.

3. Multivariate cornish-fisher expansion

The development of the previous section can be generalized to the multivariate case in a straightforward manner.

Let $\kappa_{r_1, \dots, r_p}(\boldsymbol{X}_n)$ denote the $(r_1, \dots, r_p)$—joint cumulant of $\boldsymbol{X}_n = (X_1, \dots, X_p)$. We assume that

$$\kappa_{\boldsymbol{r}}(\boldsymbol{X}_n) = n^{|\boldsymbol{r}|/2} \sum_{s=2|\boldsymbol{r}|-2}^{\infty} \gamma_{\boldsymbol{r};s} n^{-s/2}, \qquad \dots \quad (3.1)$$

where $r = (r_1, \ldots, r_p)$ and $|r| = r_1 + \ldots + r_p$. Further assume that (*i*) if $|r| = 1$ then $\gamma_{r;0} = \gamma_{r;1} = 0$, (*ii*) if $|r| = 2$ such that $r_\alpha > 0$ and $r_\delta > 0$ some α and δ and other components are zero, then $\gamma_{r;2} = \sigma_{\alpha,\delta}$ where $\sigma_{\alpha,\delta}$ is (α, δ)— element of a positive definite matrix $\mathbf{\Sigma}$. Now consider the multivariate Cornish-Fisher expansion discussed in Introduction :

$$X_1 \sim U_1 + \frac{1}{n^{1/2}} B_{1,2}(U_1) + \frac{1}{n} B_{1,3}(U_1) + \ldots,$$

$$X_2 \sim U_2 + \frac{1}{n^{1/2}} B_{2,2}(U_1, U_2) + \frac{1}{n} B_{2;3}(U_1, U_2) + \ldots,$$

$$\ldots$$

$$X_p \sim U_p + \frac{1}{n^{1/2}} B_{p,2}(U_1, \ldots, U_p) + \frac{1}{n} B_{p,3}(U_1, \ldots, U_p) + \ldots, \qquad \ldots \quad (3.2)$$

where $B_{j,k}$ are polynomials. Generalizing Theorem 2.1 and Theorem 2.2 we have

Theorem 3.1 : *Degree of the polynomial* $B_{j,k}$ *is* k, $j = 1, \ldots, p$.

Theorem 3.2 : *Let random variables* $X_1, \ldots, X_p$ *be formally defined by the right hand side of* (3.2) *where* $(U_1, \ldots, U_p)$ *has a multivariate normal distribution with mean* $\mathbf{0}$ *and covariance matrix* $\mathbf{\Sigma}$ *and* $B_{j,k}$ *is a polynomial of degree* k, $j = 1, \ldots, p$. *Then the r-the order joint cumulant of* $X_1, \ldots, X_p$ *is of the order* $O(n^{-(r-2)/2})$, $r \geqslant 3$.

The proof of the previous section can be applied with only a few appropriate changes in notation. Therefore we first point out the differences in notation and then discuss modifications in the proof.

In this section $i, j, \ldots$ denote multiple indices with p components and the components are denoted with subscripts : $i = (i_1, \ldots, i_p)$, $j = (j_1, \ldots, j_p)$, etc. Sum of the components are denoted by the absolute value sign : $|i| = i_1 + i_2 + \ldots + i_p$. $i > 0$ means that at least one component of i is positive. Sum of two multiple indices is defined componentwise : $(i+j)_1 = i_1 + j_1, \ldots, (i+j)_p = i_p + j_p$. For $\boldsymbol{a} = (a_1, \ldots, a_p)$ $\boldsymbol{a}^i$ stands for $a_1^{i_1} \ldots a_p^{i_p}$. $i\,!$ stands for $i_1\,! \ldots i_p\,!$.

Let

$$\begin{aligned} \beta_j &= \kappa_j(\boldsymbol{X}_n)/j\,!, & |j| &\neq 2, \\ &= (\kappa_j(X_n) - \gamma_{j;2})/j\,!, & |j| &= 2. \end{aligned} \qquad \ldots \quad (3.3)$$

Then the cumulant generating function $\xi(t)$ of X_n can be written as (2.6).

Now let

$$\phi(\boldsymbol{x}) = \phi(\boldsymbol{x}\,;\mathbf{\Sigma}) = \frac{1}{(2\pi)^{p/2}|\mathbf{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2}\,\boldsymbol{x}'\mathbf{\Sigma}^{-1}\boldsymbol{x}\right) \qquad \ldots \quad (3.4)$$

and define multivariate Hermite polynomial $H_i = H_i(\boldsymbol{x} ; \boldsymbol{\Sigma})$ by

$$H_i(\boldsymbol{x} ; \boldsymbol{\Sigma}) \phi(\boldsymbol{x} ; \boldsymbol{\Sigma}) = \left(-\frac{\partial}{\partial x_1}\right)^{i_1} \cdots \left(-\frac{\partial}{\partial x_p}\right)^{i_p} \phi(\boldsymbol{x}; \boldsymbol{\Sigma}). \qquad \dots \quad (3.5)$$

Then H_i is a polynomial in $x_1, \dots, x_p$ of degree $|i|$. Properties of multivariate Hermite polynomials are discussed in Appendix.

With these and other obvious notational changes, Lemma 2.1—Lemma 2.6 of the previous section hold for the multivariate case.

As in the univariate case we want to prove that the degree of $\Delta_{i(1)} \dots \Delta_{i(l)}1$ is given by $|i(1)| + \dots + |i(l)| - 2(l-1)$. For this, Lemma 2.7 and Lemma 2.8 need the following modifications.

As shown in Appendix, $H_i(\boldsymbol{x} ; \boldsymbol{\Sigma})$ can be more easily expressed in terms of $\boldsymbol{y} = (y_1, \dots, y_p)' = \boldsymbol{\Sigma}^{-1}\boldsymbol{x}$. Let

$$D_\alpha = \frac{\partial}{\partial x_\alpha}, \qquad \alpha = 1, \dots, p, \qquad \dots \quad (3.6)$$

and

$$\widetilde{D}_\alpha = \frac{\partial}{\partial y_\alpha} = \sum_{\delta=1}^{p} \sigma_{\alpha\delta} D_\delta, \qquad \alpha = 1, \dots, p, \qquad \dots \quad (3.7)$$

be differential operators with respect to x_α and y_α, respectively. Let $e_\alpha = (0, \dots 0, 1, 0, \dots, 0)$ denote the multiple index whose α-th component is 1 and other components are 0. Then Lemma 2.7 is generalized as follows.

Lemma 3.1 : *For* $|i(1)| > 1, \dots, |i(l)| > 1$,

$$\widetilde{D}_\alpha(\Delta_{i(1)} \dots \Delta_{i(l)}1) = i_\alpha(1)\, \Delta_{i(1)-e_\alpha} \Delta_{i(2)} \dots \Delta_{i(l)}1 + \dots + i_{\alpha(l)}\, \Delta_{i(1)}\Delta_{i(2)} \dots \Delta_{i(l)-e_\alpha}1. \qquad \dots \quad (3.8)$$

Proof is the same as in the univariate case in view of (10) of Appendix. Note that Lemma 3.1 is valid even if some of $i_\alpha(1), \dots, i_\alpha(l)$ are zeros.

For the case where $|i(1)| = 1$ such that $i(1) = e_\alpha$ for some α we have the following generalization of Lemma 2.8.

Lemma 3.2 : $\Delta_{e_\alpha}\Delta_{i(2)} \dots \Delta_{i(l)}1 = -D_\alpha(\Delta_{i(2)} \dots \Delta_{i(l)}1)$. ... (3.9)

Proof is again the same as in the univariate case in view of (14) of Appendix.

Now with the same argument as in the univariate case we obtain

Theorem 3.3 : $\deg(\Delta_{i(1)} \dots \Delta_{i(l)}1) = |i(1)| + \dots + |i(l)| - 2(l-1)$. ... (3.10)

In the univariate case leading coefficient was evaluated as (2.34). In the multivariate case there are many "leading terms" and it seems difficult to evaluate them explicitly. However we at least know that the leading terms do not simultaneously vanish, because the sum of the leading terms is reduced to the leading term of the univariate case by setting $y \equiv y_1 \equiv \dots \equiv y_p$ and $\sigma^{\alpha\delta} \equiv 1$ where $\sigma^{\alpha\delta}$ is the (α, δ)—element of $\boldsymbol{\Sigma}^{-1}$. See (3.11) and the discussion at the end of Appendix.

In the multivariate case it becomes extremely tedious to write down explicit expressions. For example, using symbolic manipulation language, we obtained the following expansion of the logarithm of the bivariate density for the case of sum of i.i.d. variables up to the order n^{-1}. For simplicity consider the case :

$$\boldsymbol{\Sigma}^{-1} = \begin{pmatrix} 1 & \tau \\ \tau & 1 \end{pmatrix}.$$

Writing $\gamma_{r;\,2|r|-2} = \gamma_r$ we have

$$\log f(\boldsymbol{x}) = \log \phi(\boldsymbol{x}; \boldsymbol{\Sigma}) + \frac{C_1}{n^{1/2}} + \frac{C_2}{n}, \qquad \dots \quad (3.11)$$

where

$$C_1 = \frac{\gamma_{3,0}}{6}(y_1{}^3 - 3y_1) + \frac{\gamma_{2,1}}{2}(y_1^2 y_2 - 2\tau y_1 - y_2) + \text{symmetric terms},$$

$$\begin{aligned} C_2 = {} & \frac{\gamma_{4,0}}{24}(y_1^4 - 6y_1^2 + 3) + \frac{\gamma_{3,1}}{6}(y_1^3 y_2 - 3\tau y_1^2 - 3y_1 y_2 + 3\tau) \\ & - \frac{\gamma_{3,0}^2}{24}(3y_1^4 - 12y_1^2 + 5) - \frac{\gamma_{3,0}\gamma_{2,1}}{4}(\tau y_1^4 + 2y_1^3 y_2 - 8\tau y_1^2 - 4y_1 y_2 + 5\tau) \\ & - \frac{\gamma_{3,0}\gamma_{1,2}}{4}(2\tau y_1^3 y_2 + y_1^2 y_2^2 - 4\tau^2 y_1^2 - y_1^2 - 6\tau y_1 y_2 - y_2^2 + 4\tau^2 + 1) \\ & - \frac{\gamma_{3,0}\gamma_{0,3}}{12}(3\tau y_1^2 y_2^2 - 3\tau y_1^2 - 6\tau^2 y_1 y_2 - 3\tau y_2^2 + 2\tau^3 + 3\tau) \\ & + \frac{\gamma_{2,2}}{4}(y_1^2 y_2^2 - y_1^2 - 4\tau y_1 y_2 - y_2{}^2 + 2\tau^2 + 1) \\ & - \frac{\gamma_{2,1}^2}{8}(y_1^4 + 4\tau y_1^3 y_2 + 4y_1^2 y_2^2 - 8\tau^2 y_1^2 - 6y_1^2 - 20\tau y_1 y_2 - 2y_2^2 + 12\tau^2 + 3) \\ & - \frac{\gamma_{2,1}\gamma_{1,2}}{4}(2y_1^3 y_2 + 5\tau y_1^2 y_2^2 - 7\tau y_1^2 + 2y_1 y_2^3 - 14\tau^2 y_1 y_2 \\ & - 8y_1 y_2 - 7\tau y_2^2 + 6\tau^3 + 9\tau) + \text{symmetric terms}, \end{aligned} \qquad \dots \quad (3.12)$$

where $(y_1, y_2)' = \mathbf{\Sigma}^{-1}(x_1, x_2)$ and "symmetric terms" are terms which can be obtained by interchanging the roles of y_1 and y_2. (3.11) checks with Theorem 3.3. Explicit expressions of Cornish-Fisher expansion for the bivariate case are discussed in Takeuchi (1978).

Once Theorem 3.3 is established, Theorem 3.1 and Theorem 3.2 can be proved in the same way as in the univariate case. Hence we omit the rest of the proof.

4. Application

In this section we discuss one important application of Theorem 2.1 and Theorem 2.2. James (1955, 1958) and James and Mayne (1962) proved the following fact :

Theorem 4.1 : (James) : *Let x be a random variable with the r-th cumulant κ_{rx}. Let a new random variable y be defined by*

$$y = c_0+c_1x+c_2x^2+ \dots, \qquad \dots \quad (4.1)$$

where c_0, c_1, ... are constants. Let κ_{ry} denote the r-th cumulant of y (obtained formally from the cumulants of x). If $\kappa_{rx} = O(\nu^{-r+1})$ then $\kappa_{ry} = O(\nu^{-r+1})$ where ν is some 'large' number.

This theorem has been used in an essential way in establishing validity of Edgeworth expansion (Bhattacharya and Ghosh (1978)). James' proof was combinatorial and very complicated. Based on our Theorem 2.1 and Theorem 2.2 we can give a very simple proof of Theorem 4.1.

Assume $E(x) = 0$ and $c_0 = 0$ without loss of generality. Furthermore for simplicity let $\kappa_{2x} = \nu^{-1}$. Now normalizing x define $\tilde{x} = \nu^{1/2} x$. Then the r-th cumulant of $\tilde{x}$ is $O(\nu^{-(r-2)/2})$. With $n = \nu$ this is the situation considered in (2.1). Hence by Theorem 2.1 the Cornish-Fisher expansion of $\tilde{x}$ is given as

$$\tilde{x} = u+ \frac{1}{\nu^{1/2}} B_2(u)+\frac{1}{\nu} B_3(u)+ \dots, \qquad \dots \quad (4.2)$$

where $B_k(u)$ is a k-th degree polynomial in u. Then

$$\begin{aligned} y &= c_1x+c_2x^2+ \dots \\ &= c_1\nu^{-1/2}\tilde{x}+c_2\nu^{-1}\tilde{x}^2 + \dots . \qquad \dots \quad (4.3) \end{aligned}$$

Hence

$$\begin{aligned} \nu^{1/2}y &= c_1\tilde{x}+c_2\nu^{-1/2}\tilde{x}^2+ \dots \\ &= c_1(u+\nu^{-1/2}B_2(u)+\nu^{-1}B_3(u)+ \dots) \\ &\quad +c_2\nu^{-1/2}(u+\nu^{-1/2}B_2(u)+\nu^{-1}B_3(u)+ \dots)^2 \qquad \dots \quad (4.4) \\ &\quad + \dots . \end{aligned}$$

Now consider a general term of the right hand side of (4.4). With $B_1(u) = u$ it is of the form

$$\nu^{-(r-1)/2}\,\nu^{-(\Sigma_{s=1}^{r} j(s)-r)/2} B_{j(1)} \ldots B_{j(r)} \qquad \ldots \quad (4.5)$$

$$= \nu^{-(d-1)/2} B_{j(1)} \ldots B_{j(r)},$$

where $d = \Sigma_{s=1}^{r} j(s) = \deg(B_{j(1)} \ldots B_{j(r)})$. Therefore term of order $\nu^{-k/2}$ is a polynomial of degree $k+1$ in u. Hence now by Theorem 2.2 the r-th cumulant of $\nu^{1/2}y$ is of order $O(\nu^{-(r-2)/2})$. This implies Theorem 4.1.

5. Validity of Cornish-Fisher expansion

In the previous sections we studied formal Cornish-Fisher expansions. In this section we establish the validity of Cornish-Fisher expansions under the same regularity conditions which are needed for the validity of Edgeworth expansions. The validity of Edgeworth expansions has been well established (see Bhattacharya and Ghosh, 1978). On the other hand the validity of Cornish-Fisher expansions does not seem to be appropriately discussed in literature. We will show that when Edgeworth expansion is valid, then the corresponding Cornish-Fisher expansion is valid as well.

For simplicity of notation we only discuss one-dimensional case. However the argument for the multivariate case is entirely similar and Theorem 5.1 below holds for the multivariate case as well.

Mathematically there can be many forms of validity corresponding to different notions of convergence in probability theory. As discussed in Section 1 Cornish-Fisher expansion is usually considered to be expansion of percentiles. This corresponds to the pointwise convergence of quantile function. Here we prefer to consider Cornish-Fisher expansion as approximation of the distribution of X_n. We discuss convergence of distributions in terms of variation norm, since it seems to be more natural and convenient.

Let F_n be the distribution function of random variable X_n which is asymptotically normally distributed. Let $\hat{F}_{n,k}$ denote the approximation of F_n based on the Edgeworth expansion up to the order $n^{-k/2}$. Under suitable regularity conditions it has been established that

$$\|F_n - \hat{F}_{n,k}\| = 0(n^{-k/2}), \qquad \ldots \quad (5.1)$$

where $\| \quad \|$ denotes the variation norm of the signed measure. For example, Theorem 2(a) of Bhattacharya and Ghosh (1978) establishes (5.1) when X_n

is (a smooth function of) sum of i.i.d. continuous random variables (see Bhattacharya and Ghosh (1978) for more precise statements).

Based on Cornish-Fisher expansion for X_n up to the order $n^{-k/2}$, define random variable $\widetilde{X}_n$ by

$$\widetilde{X}_n = U + \frac{1}{n^{1/2}} B_2(U) + \dots + \frac{1}{n^{k/2}} B_{k+1}(U), \qquad \dots \quad (5.2)$$

where U is a standard normal variable. To be consistent with our derivation in Section 2, we assume that $B_2, \dots, B_{k+1}$ are obtained from the Edgeworth expansion as in the proof of Theorem 2.1, i.e., by equating the logarithm of the density up to the order $n^{-k/2}$ (see Remark 2.3). Let $\widetilde{F}_{n,k}$ denote the distribtuion function of $\widetilde{X}_n$. Then we have the following theorem :

Theorem 5.1 : *If* $\|F_n - \hat{F}_{n,k}\| = o(n^{-k/2})$ *then*

$$\|F_n - \widetilde{F}_{n,k}\| = o(n^{-k/2}). \qquad \dots \quad (5.3)$$

As an immediate consequence of Theorem 5.1 we obtain

Corollary 5.1 : *Let g be a bounded Borel measurable function, then*

$$E(g(X_n)) = E(g(\widetilde{X}_n)) + o(n^{-k/2}). \qquad \dots \quad (5.4)$$

If g is taken to be an indicator function of an interval, then (5.4) gives an approximation of the cumulative distribution function. If g is taken to be e^{itx}, then (5.4) gives an approximation of the characteristic function.

Now we prove Theorem 5.1 in several steps.

We first consider truncation of U : $|U| \leqslant \log n$. Let $\widetilde{X}_n$ of (5.2) be regarded as a function of U : $\widetilde{X}_n = \tilde{x}_n(U)$. Then for any Borel measurable set B,

$$\widetilde{F}_{n,k}(B) = \Phi(\tilde{x}_n^{-1}(B)). \qquad \dots \quad (5.5)$$

Now corresponding to the truncation $|U| \leqslant \log n$, define a (subprobability) measure $\widetilde{G}_{n,k}$ by

$$\widetilde{G}_{n,k}(B) = \Phi(\tilde{x}_n^{-1}(B) \cap (-\log n, \log n)). \qquad \dots \quad (5.6)$$

Then

Lemma 5.1 : *For any positive l*

$$\|\widetilde{F}_{n,k} - \widetilde{G}_{n,k}\| = o(n^{-l}). \qquad \dots \quad (5.7)$$

Proof: Immediate from

$$\|\widetilde{F}_{n,k}-\widetilde{G}_{n,k}\| = \Phi((-\log n, \log n)^c)$$

and

$$\Phi((-\log n, \log n)^c) = o(n^{-l})$$

for any positive l. Q.E.D.

Let

$$u_n = \tilde{x}_n(\log n), \qquad l_n = \tilde{x}_n(-\log n). \qquad \dots \quad (5.8)$$

Note that

$$u_n = \log n+o(1), \quad l_n = -\log n+o(1). \qquad \dots \quad (5.9)$$

Now consider trunction of $\hat{F}_{n,k}$ and define a (signed) measure $\hat{G}_{n,k}$ as

$$\hat{G}_{n,k}(B) = \hat{F}_{n,k}(B \cap (l_n, u_n)). \qquad \dots \quad (5.10)$$

Then

Lemma 5.2 : *For any positive* l

$$\|\hat{F}_{n,k}-\hat{G}_{n,k}\| = o(n^{-l}). \qquad \dots \quad (5.11)$$

Proof: For any Borel set B

$$\hat{F}_{n,k}(B)-\hat{G}_{n,k}(B) = \hat{F}_{n,k}(B \cap (l_n, u_n)^c).$$

Hence

$$\|\hat{F}_{n,k}-\hat{G}_{n,k}\| = \int_{x\leqslant l_n} |\hat{F}_{n,k}(dx) + \int_{x\geqslant u_n} |\hat{F}_{n,k}(dx)|. \qquad \dots \quad (5.12)$$

Now

$$|\hat{F}_{n,k}(dx)| = \phi(x)|P_{n,k}(x)|dx, \qquad \dots \quad (5.13)$$

where $P_{n,k}(x)$ is a polynomial in x whose coefficients converge to 0 as $n \to \infty$. Recall that for any positive m, $\int_{|x|\geqslant \log n} |x|^m \phi(x)\, dx$ converges to 0 faster than any negative power of n. Therefore considering (5.9) we see that the right hand side of (5.12) converges to 0 faster than any negative power of n. Q.E.D.

The last lemma we need is

Lemma 5.3 : $\|\hat{G}_{n,k}-\widetilde{G}_{n,k}\| = o(n^{-k/2})$. $\qquad \dots \quad (5.14)$

Proof: $\tilde{x}_n'(u) = 1 + \frac{1}{n^{1/2}} B_2'(u)+ \dots + \frac{1}{n^{k/2}} B_{k+1}'(u)$.

Therefore for all sufficiently large n, the function $\tilde{x}_n$ is monotone in the range $|u| \leqslant \log n$. Then the inverse transform from $\tilde{x}_n$ to u is uniquely defined in the range (l_n, u_n). Let $u(x)$ denote this inverse map.

Then

$$\|\hat{G}_{n,k}-\widetilde{G}_{n,k}\| = \int_{l_n}^{u_n} |\hat{G}_{n,k}(dx)-\tilde{G}_{n,k}(dx)|$$

$$= \int_{l_n}^{u_n} \left| \phi(x)P_{n,k}(x)-\phi(u(x)) \frac{du}{dx} \right| dx$$

$$= \int_{-\log n}^{\log n} \left| \phi(\tilde{x}_n(u))P_{n,k}(\tilde{x}_n(u)) \frac{d\tilde{x}_n}{du} - \phi(u) \right| du. \qquad \dots \quad (5.15)$$

Now $B_2, \dots, B_{k+2}$ are defined in such a way that

$$R_n = R_n(u) = \log \phi(\tilde{x}_n(u)) + \log P_{n,k}(\tilde{x}_n(u)) + \log(d\tilde{x}_n/du) - \log \phi(u)$$
$$= o(n^{-k/2})$$

for each u. Now bounding the remainder terms in the Taylor expansion of $\log(1+(P_{n,k}-1))$ and $\log(1+((d\tilde{x}_n/du)-1))$ it can be easily shown that there exists some positive c and m such that

$$\sup_{|u| \leqslant \log n} |R_n(u)| \leqslant c \frac{(\log n)^m}{n^{(k+1)/2}}, \qquad \dots \quad (5.16)$$

for all sufficiently large n. Hence

$$\phi(\tilde{x}_n(u))P_{n,k}(\tilde{x}_n(u)) \frac{d\tilde{x}_n}{du} = \phi(u)\, e^{R_n(u)} = \phi(u)+R_n^*(u) \qquad \dots \quad (5.17)$$

where for some c'

$$\sup_{|u| \leqslant \log u} |R_n^*(u)| \leqslant c' \frac{(\log n)^m}{n^{(k+1)/2}} \qquad \dots \quad (5.18)$$

for all sufficiently large n. Then the right hand side of (5.15) is smaller than or equal to

$$\int_{-\log n}^{\log n} |R_n^*(u)|\, du \leqslant 2c' \frac{(\log n)^{m+1}}{n^{(k+1)/2}} = o(n^{-k/2}). \qquad \dots \quad (5.19)$$

Q.E.D.

Proof of Theorem 5.1 : Immediate from Lemmas 5.1—5.3 by triangular inequality. Q.E.D.

Appendix

Multivariate Hermite polynomials. Let multivariate Hermite polynomial be defined by (3.5). Let $\boldsymbol{y} = \boldsymbol{\Sigma}^{-1}\boldsymbol{x}$ and the differential operators $D_{\boldsymbol{\alpha}}$ and $\widetilde{D}_{\boldsymbol{\alpha}}$ be defined as (3.6) and (3.7).

Now we define a polynomial $\widetilde{H}_i(\boldsymbol{x}\,;\boldsymbol{\Sigma})$ by

$$\widetilde{H}_i(\boldsymbol{x}\,;\boldsymbol{\Sigma})\,\phi(\boldsymbol{x}\,;\boldsymbol{\Sigma}) = (-\widetilde{D}_1)^{i_1} \dots (-\widetilde{D}_p)^{i_p}\,\phi(\boldsymbol{x}\,;\boldsymbol{\Sigma}). \qquad \dots \quad (\text{A.1})$$

$\widetilde{H}_i(\boldsymbol{x}\,;\boldsymbol{\Sigma})$ is also a polynomial in $x_1, \dots, x_p$ with degree $|\,i\,|$. Amari and Kumon (1983) introduced $\widetilde{H}_i(\boldsymbol{x}\,;\boldsymbol{\Sigma})$ as *tensorial Hermite polynomial.* Here we prefer to call it *dual Hermite polynomial* because of Lemma A1 below.

To investigate properties of these polynomials we consider their generating functions. For H_i the generating function is simply given by

$$\Sigma\,\frac{\boldsymbol{t}^i}{i\,!}\,H_i(\boldsymbol{x}\,;\boldsymbol{\Sigma}) = \phi(\boldsymbol{x}-\boldsymbol{t}\,;\boldsymbol{\Sigma})/\phi(\boldsymbol{x}\,;\boldsymbol{\Sigma})$$

$$= \exp(\boldsymbol{t}'\boldsymbol{\Sigma}^{-1}\boldsymbol{x} - \frac{1}{2}\,\boldsymbol{t}'\boldsymbol{\Sigma}^{-1}\boldsymbol{t}). \qquad \dots \quad (\text{A.2})$$

Now by definition of $\widetilde{D}_{\boldsymbol{\alpha}}$

$$\left[\frac{\partial}{\partial y_{\boldsymbol{\alpha}}}\,f(\boldsymbol{\Sigma y})\right]_{\boldsymbol{y}=\boldsymbol{\Sigma}^{-1}\boldsymbol{x}} = \widetilde{D}_{\boldsymbol{\alpha}}\,f(x) \qquad \dots \quad (\text{A.3})$$

for any function f. Therefore

$$\left[\left(-\frac{\partial}{\partial y_1}\right)^{i_1} \dots \left(-\frac{\partial}{\partial y_p}\right)^{i_p} \exp\left(-\frac{1}{2}\,\boldsymbol{y}'\boldsymbol{\Sigma y}\right)\right]_{\boldsymbol{y}=\boldsymbol{\Sigma}^{-1}\boldsymbol{x}}$$

$$= \widetilde{H}_i(\boldsymbol{x}\,;\boldsymbol{\Sigma})\,\exp\left(-\frac{1}{2}\,\boldsymbol{x}'\boldsymbol{\Sigma}^{-1}\boldsymbol{x}\right). \qquad \dots \quad (\text{A.4})$$

Summing (4) up we obtain

$$\Sigma\,\frac{\boldsymbol{t}^i}{i\,!}\,\widetilde{H}_i(\boldsymbol{x}\,;\boldsymbol{\Sigma})\,\exp\left(-\frac{1}{2}\,\boldsymbol{x}'\boldsymbol{\Sigma}^{-1}\boldsymbol{x}\right)$$

$$= \left[\exp\left\{-\frac{1}{2}\,(\boldsymbol{y}-\boldsymbol{t})'\,\boldsymbol{\Sigma}(\boldsymbol{y}-\boldsymbol{t})\right\}\right]_{\boldsymbol{y}=\boldsymbol{\Sigma}^{-1}\boldsymbol{x}}$$

$$= \exp\left(-\frac{1}{2}\,\boldsymbol{x}'\boldsymbol{\Sigma}^{-1}\boldsymbol{x} + \boldsymbol{x}\boldsymbol{t}' - \frac{1}{2}\,\boldsymbol{t}'\boldsymbol{\Sigma t}\right). \qquad \dots \quad (\text{A.5})$$

Therefore the generating function for $\tilde{H}_i$'s is

$$\Sigma \frac{\boldsymbol{t}^i}{i\,!} \tilde{H}_i(\boldsymbol{x}\,;\boldsymbol{\Sigma}) = \exp\left(\boldsymbol{x}'\boldsymbol{t} - \frac{1}{2}\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}\right). \qquad \text{(A.6)}$$

Comparing (A.3) and (A.6) we obtain the relation between H_i and $\tilde{H}_i$ as

$$H_i(\boldsymbol{\Sigma}\boldsymbol{y}\,;\boldsymbol{\Sigma}) = \tilde{H}_i(\boldsymbol{y}\,;\boldsymbol{\Sigma}^{-1}). \qquad \text{(A.7)}$$

As a polynomial $\tilde{H}_i$ has simpler form than H_i. Actually as discussed later explicit expression of $\tilde{H}_i$ can be written down. (A.7) shows that expression of H_i in terms of $\boldsymbol{y}$ can be obtained from $\tilde{H}_i$ by substituting $\boldsymbol{\Sigma}^{-1}$ for $\boldsymbol{\Sigma}$. In proving (3.8) of Section 3 we needed derivatives of H_i. Now differentiate (A.6) with respect to x_α. Then

$$\Sigma \frac{\boldsymbol{t}^i}{i\,!} \frac{\partial}{\partial x_\alpha} \tilde{H}_i(\boldsymbol{x};\boldsymbol{\Sigma}) = t_\alpha \Sigma \frac{\boldsymbol{t}^i}{i\,!} \tilde{H}_i(\boldsymbol{x}\,;\boldsymbol{\Sigma}). \qquad \text{(A.8)}$$

Hence we obtain

$$\frac{\partial}{\partial x_\alpha} \tilde{H}_i(\boldsymbol{x};\boldsymbol{\Sigma}) = i_\alpha \tilde{H}_{i-e_\alpha}(\boldsymbol{x}\,;\boldsymbol{\Sigma}). \qquad \text{(A.9)}$$

Then by (A.3) and (A.7) we have

$$\tilde{D}_\alpha H_i(\boldsymbol{x}\,;\boldsymbol{\Sigma}) = i_\alpha H_{i-e_\alpha}(\boldsymbol{x}\,;\boldsymbol{\Sigma}). \qquad \text{(A.10)}$$

Note that (A.9) and (A.10) are valid even if $i_\alpha = 0$.

Now differentiate (A.6) with respect to t_α. Then

$$\Sigma \frac{t^i}{i\,!} \tilde{H}_{i+e_\alpha}(\boldsymbol{x}\,;\boldsymbol{\Sigma}) = (x_\alpha - \sum_\delta \sigma_{\alpha\delta} t_\delta) \Sigma \frac{t^i}{i\,!} \tilde{H}_i(\boldsymbol{x}\,;\boldsymbol{\Sigma}). \qquad \text{(A.11)}$$

Therefore

$$\begin{aligned} \tilde{H}_{i+e_\alpha}(\boldsymbol{x}\,;\boldsymbol{\Sigma}) &= x_\alpha \tilde{H}_i(\boldsymbol{x}\,;\boldsymbol{\Sigma}) - \sum_\delta \sigma_{\alpha\delta}\, i_\delta \tilde{H}_{i-e_\delta}(\boldsymbol{x}\,;\boldsymbol{\Sigma}) \\ &= x_\alpha \tilde{H}_i(\boldsymbol{x}\,;\boldsymbol{\Sigma}) - \sum_\delta \sigma_{\alpha\delta} \frac{\partial}{\partial x_\delta} \tilde{H}_i(\boldsymbol{x}\,;\boldsymbol{\Sigma}) \\ &= \tilde{H}_{e_\alpha}(\boldsymbol{x}\,;\boldsymbol{\Sigma})\, \tilde{H}_i(\boldsymbol{x}\,;\boldsymbol{\Sigma}) - \sum_\delta \sigma_{\alpha\delta} \frac{\partial}{\partial x_\delta} \tilde{H}_i(\boldsymbol{x}\,;\boldsymbol{\Sigma}). \qquad \text{(A.12)} \end{aligned}$$

This gives a recurrence relation for computing $\widetilde{H}_{i+e_\alpha}$. Now substituting (A.7) into (A.12) and using (A.10) we have

$$H_{i+e_\alpha}(\boldsymbol{x};\boldsymbol{\Sigma})-H_{e_\alpha}(\boldsymbol{x};\boldsymbol{\Sigma})\,H_i(\boldsymbol{x};\boldsymbol{\Sigma}) = -\sum_\delta \sigma^{\alpha\delta}\widetilde{D}_\delta H_i(\boldsymbol{x};\boldsymbol{\Sigma}), \quad \ldots \quad (A.13)$$

where $\sigma^{\alpha\delta}$ is (α, δ)—element of Σ^{-1}. However $\Sigma_\delta \sigma^{\alpha\delta}\widetilde{D}_\delta = D_\alpha$. Hence

$$H_{i+e_\alpha}(\boldsymbol{x};\boldsymbol{\Sigma})-H_{e_\alpha}(\boldsymbol{x};\boldsymbol{\Sigma})\,H_i(\boldsymbol{x};\boldsymbol{\Sigma}) = -D_\alpha H_i(\boldsymbol{x};\boldsymbol{\Sigma}). \quad \ldots \quad (A.14)$$

This was needed for the proof of (3.9) of Section 3.

Next, we discuss mutual orthogonality of $\{H_i(\boldsymbol{x};\boldsymbol{\Sigma})\}$ and $\{\widetilde{H}_i(\boldsymbol{x};\boldsymbol{\Sigma})\}$ with respect to $\phi(\boldsymbol{x};\boldsymbol{\Sigma})$.

Lemma A1 :

$$\int H_i(\boldsymbol{x};\boldsymbol{\Sigma})\,\widetilde{H}_i(\boldsymbol{x};\boldsymbol{\Sigma})\,\phi(x;\boldsymbol{\Sigma})\,dx = \begin{cases} 0 & \textit{if } i \neq j, \\ i\,! & \textit{if } i = j. \end{cases} \quad \ldots \quad (A.15)$$

Proof : From (2) and (6) we have

$$\int \sum_i \sum_j \frac{t^i}{i!} H_i(\boldsymbol{x};\boldsymbol{\Sigma}) \frac{s^j}{j\,!} \widetilde{H}_j(\boldsymbol{x};\boldsymbol{\Sigma})\,\phi(\boldsymbol{x};\boldsymbol{\Sigma})\,dx$$

$$= \exp(\boldsymbol{t}'\mathbf{s}) \int \frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{t}-\boldsymbol{\Sigma}\mathbf{s})'\boldsymbol{\Sigma}^{-1}(x-t-\boldsymbol{\Sigma}\mathbf{s})\right\} dx$$

$$= \exp(\boldsymbol{t}'\mathbf{s})$$

$$= \sum_i \frac{\boldsymbol{t}^i\mathbf{s}^i}{i\,!}\,. \qquad \text{Q.E.D.}$$

Finally we discuss explicit expression for $\widetilde{H}_i(\boldsymbol{x};\boldsymbol{\Sigma})$. First consider multiple index $i = (i_1, \ldots, i_p)$ such that $i_\alpha, \alpha = 1, \ldots, p$, are either 0 or 1. For example consider $i = (1^k) = (1, \ldots, 1, 0, \ldots, 0)$, where k first components of i are 1. Then

Lemma A2 :

$$\widetilde{H}_{(1^k)}(\boldsymbol{x};\boldsymbol{\Sigma}) = x_1x_2 \ldots x_k - \Sigma\sigma_{\alpha_1\alpha_2}x_{\alpha_3}\ldots x_{\alpha_k} + \Sigma\,\sigma_{\alpha_1\alpha_2}\sigma_{\alpha_3\alpha_4}x_{\alpha_5}\ldots x_{\alpha_k} - \ldots$$

$$= \Sigma(-1)^m\sigma_{\alpha_1\alpha_2}\ldots\sigma_{\alpha_{2m-1}\alpha_{2m}}x_{\alpha_{2m+1}}\ldots x_{\alpha_k}, \quad \ldots \quad (A.16)$$

where $(\alpha_1, \ldots, \alpha_k)$ is a permutation of $(1, \ldots, k)$ and only distinct terms are counted on the right hand side of (A.16).

Proof : Obvious by inspection of the term $t_1t_2 \ldots t_k$ in the expansion of $\exp\{x_1t_1+\ldots+x_pt_p-(t_1t_2\sigma_{12}+\ldots+t_{p-1}t_p\sigma_{p-1,p})-\frac{1}{2}(t_1^2\sigma_{11}+\ldots+t_p^2\,\sigma_{pp})\}$. Q.E.D.

For i such that $i_\alpha = 0$ or 1, $\alpha = 1, \ldots, p$, $\widetilde{H}_i$ can be obtained from (A.16) by appropriate substitution of indices.

Now consider the case where $i_1 > 0, \ldots, i_k > 0$, $i_{k+1} = \ldots = i_p = 0$. $\widetilde{H}_i$ can be expressed as follows. Let $l = i_1 + \ldots + i_k$ and consider degenerate multivariate normal random vector $\tilde{\boldsymbol{x}}$ such that $\tilde{x}_1 = \ldots = \tilde{x}_{i_1} = x_1$, $\tilde{x}_{i_1+1} = \ldots = \tilde{x}_{i_1+i_2} = x_2, \ldots$. Then

$$\widetilde{H}_i(\boldsymbol{x};\boldsymbol{\Sigma}) = \widetilde{H}_{(1^l)}(\tilde{\boldsymbol{x}};\widetilde{\boldsymbol{\Sigma}}), \qquad \ldots \quad (A.17)$$

where $\widetilde{\boldsymbol{\Sigma}}$ is the convariance matrix of $\tilde{\boldsymbol{x}}$ (so that $\tilde{\boldsymbol{\sigma}}_{11} = \tilde{\boldsymbol{\sigma}}_{12} = \ldots = \tilde{\boldsymbol{\sigma}}_{i_1 i_1}$, etc.). (A.17) can be proved by considering differentiation of $\exp\{\boldsymbol{x}'\boldsymbol{t} - (1/2)\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}\}$ with respect to $\boldsymbol{t}$. For the extreme case where $x_1 = \ldots = x_p$ and $\sigma_{\alpha\delta} \equiv 1$, $\widetilde{H}_i(\boldsymbol{x}\,;\,\boldsymbol{\Sigma})$ reduces to the $|i|$—th univariate Hermite polynomial. This was mentioned at the end of Section 3.

References

Amari, S. and Kumon, M. (1983): Differential geometry of Edgeworth expansions in curved exponential family. *Ann. Inst. Statist. Math.*, **35**, 1-24.

Bhattacharya, R. N. and Ghosh, J. K. (1978): On the validity of the formal Edgeworth expansion. *Ann. Statist.*, **6**, 434-451.

Brillinger, D. R. (1981): *Time Series*, Expanded edition. Holden-Day, San Francisco.

Cornish, E. A. and Fisher, R. A. (1937): Moments and cumulants in the specification of distributions. *Revue de l'Institut Internat. de Statist.* **5**, 307-322.

Fisher, R. A. and Cornish, E. A. (1960): The percentile points of distributions having known cumulants. *Technometrics*, **2**, 209-226.

Hill, G. W. and Davis, A. W. (1968): Generalized asymptotic expansions of Cornish-Fisher type. *Ann. Math. Statist.*, **39**, 1264-1273.

James, G. S. (1955): Cumulants of a transformed variate. *Biometrika*, **42**, 529-531.

——— (1958): On moments and cumulants of systems of statistics. *Sankhyā*, **20**, 1-30.

James, G. S. and Mayne, A. J. (1962): Cumulants of functions of random variables. *Sānkhya* Ser. A, **24**, 47-54.

Leonov, V. P. and Shiryaev, A. N. (1959): On a method of calculation of semi-invariants. *Theor. Prob. Appl.*, **4**, 319-329.

Takeuchi, K. (1975): Approximation of probability distributions. (In Japanese). *Kyoiku Shuppan*, Tokyo.

——— (1978): A multivariate generalization of Cornish-Fisher expansion and its applications. (In Japanese). *Keizaigaku Ronshu*, **44**, No. 2, 1-12.

Withers, C. S. (1982): Second order inference for asymptotically normal random variables. *Sankhyā*, Ser. B, **44**, 19-27.

Withers, C. S. (1983): Accurate confidence intervals for distributions with one parameters. *Ann. Inst. Statist. Math.*, **35**, 49-61.

Paper received: *February, 1986.*
Revised: *June, 1987.*