

Article

Not peer-reviewed version

Edgeworth Expansions When the Parameter Dimension Increases with Sample Size

[Christopher Stroude Withers](#) *

Posted Date: 20 November 2025

doi: 10.20944/preprints202511.1584.v1

Keywords: standard estimate; increasing number of unknown parameters; extended Edgeworth expansions



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Edgeworth Expansions When the Parameter Dimension Increases with Sample Size

Christopher Stroude Withers

DSIR and Industrial Research Ltd., 101 Allington Road, Wellington 6012, New Zealand; kit.withers@gmail.com

Abstract

Suppose that we have a statistical model with $q = q_n$ unknown parameters, $w = (w_1, \dots, w_q)'$, estimated by $\hat{w}$, based on a sample of size n . Suppose also, that we have Edgeworth expansions for the density and distribution of $X_n = n^{1/2}(\hat{w} - w)$. How fast can $q = q_n$ increase with n for the three main Edgeworth expansions to remain valid? We show that it is sufficient that $q_n = o(n^{1/6})$, if the estimate $\hat{w}$ is a standard estimate. That is, $E \hat{w} \rightarrow w$ as $n \rightarrow \infty$, and for $r \geq 1$, its r th order cumulants have magnitude n^{1-r} and can be expanded in powers of n^{-1} . This very large class of estimates has a huge range of potential applications. When $\hat{w} = t(\bar{X})$ for $t: R^q \rightarrow R^p$ a smooth function of a sample mean $\bar{X}$ from a distribution on R^q , and $p_n q_n = pq \rightarrow \infty$ as $n \rightarrow \infty$, I show that the Edgeworth expansions for $\hat{w}$ remain valid if $q_n^8 p_n^6 = o(n)$. For example, this holds for fixed $p = p_n$ if $q_n = o(n^{1/8})$. I also give a method that greatly reduces the number of terms needed for the 2nd and 3rd order terms in the Edgeworth expansions, that is, for the 1st and 2nd order corrections to the Central Limit Theorems (CLTs).

Keywords: standard estimate; increasing number of unknown parameters; extended Edgeworth expansions

1. Introduction and Summary

Suppose that we have a estimate $\hat{w}$ of an unknown parameter $w \in R^q$ of a statistical model, and that

$$\text{as } n \rightarrow \infty, X_n = n^{1/2}(\hat{w} - w) \xrightarrow{L} X \sim \mathcal{N}_q(0, V), \tag{1}$$

the multivariate normal on R^q , with density and distribution

$$\phi_V(x) = (2\pi)^{-q/2}(\det V)^{-1/2} \exp(-x'V^{-1}x/2), \Phi_V(x) = \int_{-\infty}^x \phi_V(x)dx. \tag{2}$$

Here we assume that $\hat{w}$ is a *standard estimate*. That is, $E \hat{w} \rightarrow w$ as $n \rightarrow \infty$, and for $r \geq 1$, its r th order cumulants have magnitude n^{1-r} and can be expanded in powers of n^{-1} . The class of standard estimates includes smooth functions of sample means or empirical distributions, based on one or more random samples, or on samples from a stationary time series. So it has a huge range of potential applications.

For $\hat{w}$ non-lattice, the density and distribution of X_n of (1) can be expanded in powers of $n^{-1/2}$ about those of X of (1). These are the Edgeworth expansions. To be self-contained, Section 2 summarises these expansions.

How fast can $q = q_n$ increase with n , for these expansions to hold? Section 3 shows that they hold if $q_n = o(n^{1/6})$.

Section 4 gives a theorem that reduces the number of terms needed for 2nd and 3rd order Edgeworth expansions. For example, if $q = 3$, it reduces the number of terms needed for 2nd and 3rd order Edgeworth expansions by 57% and 94%. This % reduction increases with q . If $q = 4$ or 5, it

reduces the number of terms by 65% or 69% for the 2nd order Edgeworth expansions, and by 98% for the 3rd order Edgeworth expansions.

Section 5 considers the case when $\hat{w} = t(\bar{X})$ for $t : R^q \rightarrow R^p$ a smooth function of a sample mean $\bar{X}$ from a distribution on R^q with finite moments. When $p_n q_n = pq \rightarrow \infty$ as $n \rightarrow \infty$, it shows the Edgeworth expansions for $\hat{w}$ remain valid if $q_n^8 p_n^6 = o(n)$. For example, this holds for fixed $p = p_n$ if $q_n = o(n^{1/8})$. These may be the 1st CLTs or Edgeworth expansions when, not 1, but 2 parameters are allowed to increase to ∞ with n .

Earlier work done when dimension $q = q_n$ increases with sample size n , has been mainly for sample means, including some CLTs and a 2nd order Edgeworth-type expansion. [10,11] showed asymptotic normality for M-estimators with q_n regression parameters when q_n^2/n is large. [3] gave a CLT for M-estimates in a linear regression model of dimension q_n when $q_n/n \rightarrow a \in (0, 1)$. Remarkably, (8) of [2] gave a CLT for $\hat{w}$ the sample mean of *bounded* random vectors that holds for $q_n = O(\exp(an^c))$ if $c < 1/7$. (Substitute into their condition to confirm.) (1.4) of [7], appears to allow for $q_n = O(n^c)$ if $c < 1/3$ to suffice for a CLT for the sample mean, when $X_1, \dots, X_n$ are log-concave, quoting [4]. It will be interesting to see if this bound can be extended to a broader class of estimates than sample means, and if the log-concave condition can be removed.

Section 2.1 of the very recent paper [7] considers the 2nd order Edgeworth expansion for the distribution of a standardized sample mean. His Theorem 2.1 gives conditions for this which hold when $q_n = O(n^c)$ for any c , or even when $q_n = O(\exp(bn^c))$ and $c < 1/3$. The bounds (1.4) and (2.4) of [7] give a 2nd order Edgeworth expansion for a sample mean, that allows for q_n of magnitude n^c if $c < 1/3$, a remarkable result.

[8] considered sampling from vectors, and investigated the simultaneous estimation of the marginal distributions for large q_n . [5] considered the asymptotic distributions of the canonical correlations between $X_1 \in R^p$ and $X_2 \in R^q$ with $p \leq q$. They derived asymptotic distributions of the canonical correlations when p is fixed, $q = q_n \rightarrow \infty$, and $q_n/n \rightarrow a \in [0, 1)$, as $n \rightarrow \infty$. It assumes that X_1 , and X_2 have a joint *normal* distribution. [9] gave a CLT for the sample mean for large q_n . [6] gave a number of results for large dimensions. [1] and [7] considered the validity and accuracy of Edgeworth expansions of $\max_{i=1}^q X_n$ for large q when X_n is a standardized sample mean.

2. Multivariate Edgeworth Expansions

Suppose that $\hat{w}$ is a *standard estimate* of $w \in R^q$ with respect to n . (n is typically the sample size.) That is, $E \hat{w} \rightarrow w$ as $n \rightarrow \infty$, where we use E for expected value, and for $r \geq 1$ and $1 \leq i_1, \dots, i_r \leq q$, the r th order cumulants of $\hat{w} = (\hat{w}_1, \dots, \hat{w}_q)'$ can be expanded as

$$\bar{k}^{1-r} = k^{i_1 \dots i_r} = \kappa(\hat{w}_{i_1}, \dots, \hat{w}_{i_r}) \approx \sum_{d=r-1}^{\infty} n^{-d} \bar{k}_d^{1-r}, \text{ where } \bar{k}_d^{1-r} = k_d^{i_1 \dots i_r}, \quad (3)$$

$\approx$ indicates an asymptotic expansion, and the *cumulant coefficients* $\bar{k}_d^{1-r}$ may depend on n but are bounded as $n \rightarrow \infty$. So *the bar replaces each i_k by k* . For example, $\bar{k}_0^1 = \bar{w}_1 = w_{i_1}$ and $\bar{k}_1^{12} = k_1^{i_1 i_2}$. I reserve i_k for this bar notation to avoid double subscripts. So, (1) holds with $V = (\bar{k}_1^{12})$, $q \times q$. V may depend on n , but I assume that $\det V$ is bounded away from 0.

Let $\bar{P}_r^{1-k} = P_r^{i_1 \dots i_k}$ be the k th Edgeworth coefficient of order r for $\hat{w}$.

These are Bell polynomials in the cumulant coefficients, $\bar{k}_d^{1-r}$ of (3), defined and given in [14] for $1 \leq r \leq 3$. Their importance lies in their central role in the Edgeworth expansions of X_n of (1): see (8) and (14) below.

The $\bar{P}_r^{1-k}$ needed for $r = 1, 2, 3$ are given in (19)–(21) of [14]:

$$\bar{P}_1^1 = \bar{k}_1^1, \bar{P}_1^{1-3} = \bar{k}_2^{1-3}/6, \bar{P}_2^{12} = \bar{k}_2^{12}/2 + \bar{k}_1^1 \bar{k}_1^2/2, \quad (4)$$

$$\bar{P}_2^{1-4} = \bar{k}_3^{1-4}/4! + \mathcal{S} \bar{k}_1^4 \bar{k}_2^{1-3}/6, \bar{P}_2^{1-6} = \mathcal{S} \bar{k}_2^{1-3} \bar{k}_2^{4-6}/72,$$

$$\bar{P}_3^1 = \bar{k}_2^1, \bar{P}_3^{1-3} = \bar{k}_3^{1-3}/6 + \mathcal{S} \bar{k}_2^{12} \bar{k}_1^3/2 + \bar{k}_1^1 \bar{k}_1^2 \bar{k}_1^3/6,$$

$$\bar{P}_3^{1-5} = \bar{k}_4^{1-5}/5! + S_1/24 + S_2/12 + S_3/12$$

$$\text{where } S_1 = \mathcal{S} \bar{k}_3^{1-4} \bar{k}_1^5, S_2 = \mathcal{S} \bar{k}_2^{12} \bar{k}_2^{345}, S_3 = \mathcal{S} \bar{k}_1^1 \bar{k}_1^2 \bar{k}_2^{345},$$

$$\bar{P}_3^{1-7} = \mathcal{S} \bar{k}_2^{123} \bar{k}_3^{4-7}/144 + \mathcal{S} \bar{k}_2^{123} \bar{k}_2^{4-6} \bar{k}_1^7/72,$$

$$\bar{P}_3^{1-9} = \mathcal{S} \bar{k}_2^{1-3} \bar{k}_2^{4-6} \bar{k}_2^{7-9}/6^4, \quad (5)$$

where $\mathcal{S}$ is the operator $\mathcal{S}$ that symmetrizes $\bar{f}^{1-k}$ over $i_1, \dots, i_k$.

Set $P(A)$ = Probability that A is true. By [15], or [14], for $\hat{w}$ non-lattice, the density and distribution of X_n can be expanded as

$$p_{X_n}(x) \approx \sum_{r=0}^{\infty} n^{-r/2} p_r(x), \quad P(X_n \leq x) \approx \sum_{r=0}^{\infty} n^{-r/2} P_r(x), \quad x \in R^q, \quad (6)$$

where $p_0(x) = \phi_V(x)$, $P_0(x) = \Phi_V(x)$, of (2), and for $r \geq 1$,

$$p_r(x)/\phi_V(x) = \sum_{k=1}^{3r} [\bar{p}_{rk} : k-r \text{ even}] = \bar{p}_r(x) \text{ say, } \bar{p}_{rk} = \bar{P}_r^{1-k} \bar{H}^{1-k}, \quad (7)$$

$$P_r(x) = \sum_{k=1}^{3r} [P_{rk}(x) : k-r \text{ even}], \quad P_{rk}(x) = \bar{P}_r^{1-k} \bar{H}_*^{1-k}, \quad (8)$$

$$\text{and } \bar{H}^{1-k} = \bar{H}^{1-k}(x, V) = H^{i_1 \dots i_k} = \phi_V(x)^{-1} \bar{O}^{1-k} \phi_V(x),$$

$$\bar{H}_*^{1-k} = \bar{H}_*^{1-k}(x, V) = \bar{O}^{1-k} \Phi_V(x) = \int_{-\infty}^x \bar{H}^{1-k} \phi_V(x) dx, \quad (9)$$

$$\bar{O}^{1-k} = (-\bar{\partial}_1) \dots (-\bar{\partial}_k), \quad \bar{\partial}_k = \partial_{i_k}, \quad \partial_i = \partial/\partial x_i.$$

$\bar{H}^{1-k}(x, V) = \bar{H}^{1-k}$ and $\bar{H}_*^{1-k}(x, V) = \bar{H}_*^{1-k}$ are the multivariate Hermite polynomial, and the integrated multivariate Hermite polynomial. By [12],

$$\bar{H}^{1-k} = E (\bar{y}_1 + \sqrt{-1} \bar{Y}_1) \dots (\bar{y}_k + \sqrt{-1} \bar{Y}_k)$$

$$\text{where } y = V^{-1}x, \quad Y = V^{-1}X \sim \mathcal{N}_q(0, V^{-1}), \quad \bar{y}_k = y_{i_k}, \text{ and } \bar{Y}_k = Y_{i_k}.$$

I use the tensor summation convention: repetition of $i_1, \dots, i_k$ in $\bar{p}_{rk}$ of (7) and $P_{rk}(x)$ of (8) implies their implicit summation over their range, $1, \dots, q$. [14] gave $\bar{H}^{1-k}$ explicitly for $k \leq 6$, and for $k \leq 9$ when $q = 2$.

$$\text{Set } \bar{\mu}^{1-2k} = E \bar{Y}_1 \dots \bar{Y}_{2k} = \sum^{1.3 \dots (2k-1)} \bar{V}^{12} \dots \bar{V}^{2k-1, 2k},$$

where $\sum^N \bar{f}^{1-k}$ sums $\bar{f}^{1-k}$ over all N permutations of $i_1, \dots, i_k$ giving distinct values. For example,

$$\bar{\mu}^{1-4} = V^{12}V^{34} + V^{13}V^{24} + V^{14}V^{23}, \quad \bar{H}^1 = \bar{y}_1, \quad \bar{H}^{12} = \bar{y}_1 \bar{y}_2 - \bar{V}^{12},$$

$$\bar{H}^{1-3} = \bar{y}_1 \bar{y}_2 \bar{y}_3 - \sum^3 \bar{y}_1 \bar{V}^{23} = \bar{y}_1 \bar{y}_2 \bar{y}_3 - \bar{y}_1 \bar{V}^{23} - \bar{y}_2 \bar{V}^{13} - \bar{y}_3 \bar{V}^{12},$$

$$\bar{H}_*^1 = \bar{f}^1, \quad \bar{H}_*^{12} = \bar{f}^{12} - \bar{V}^{12} \Phi_V(x), \quad \bar{H}_*^{1-3} = \bar{f}^{123} - \sum^3 \bar{f}^1 \bar{V}^{23}, \quad \text{where}$$

$$\bar{f}^{1-k} = \bar{f}^{1-k}(x, V) = E \bar{Y}_1 \dots \bar{Y}_k I(X \leq x) = \bar{V}^{1, k+1} \dots \bar{V}^{k, 2k} \bar{M}^{k+1-2k}, \quad (10)$$

$$\text{and } \bar{M}^{a-b} = \bar{M}^{a-b}(x, V) = E \bar{X}_1 \dots \bar{X}_k I(X \leq x) = \int_{-\infty}^x \bar{x}_a \dots \bar{x}_b \phi_V(x) dx, \quad (11)$$

for $\bar{x}_a = x_{i_a}$. So the repeated $i_{k+1}, \dots, i_{2k}$ in (10) implies their repeated summation over $1, \dots, q$. As x lies in R^q , $\int_{-\infty}^x dx$ in (9) and (11), stands for $\int_{-\infty}^{x_1} dx_1 \cdots \int_{-\infty}^{x_q} dx_q$. (6) with the $\bar{P}_r^{1-k}$ of [14], give the Edgeworth expansions for the density and distribution of X_n of (1) to $O(n^{-2})$. $\tilde{p}_{rk}$ and P_{rk} each have q^k terms, but many are duplicates as $\bar{P}_r^{1-k}$ is symmetric in $i_1, \dots, i_k$. This is exploited in Section 4 to greatly reduce the number of terms in (8) and (14) below.

By (7), the density of X_n relative to its asymptotic value is

$$p_{X_n}(x)/\phi_V(x) \approx 1 + \sum_{r=1}^{\infty} n^{-r/2} \tilde{p}_r(x) = 1 + n^{-1/2} \tilde{p}_1(x) + O(n^{-1}),$$

for $x \in R^q$. For measurable $C \subset R^q$,

$$P(X_n \in C) \approx \sum_{r=0}^{\infty} n^{-r/2} P_r(C), \text{ where } P_0(C) = \Phi_V(C), \text{ and for } r \geq 1, \quad (12)$$

$$P_r(C) = E p_r(X) I(X \in C) = \int_C p_r(x) \phi_V(x) dx = \sum_{k=1}^{3r} [P_{rk}(C) : k-r \text{ even}], \quad (13)$$

$$P_{rk}(C) = E \tilde{p}_{rk}(X) I(X \in C) = \int_C \tilde{p}_{rk}(x) \phi_V(x) dx = \bar{P}_r^{1-k} \bar{H}_*^{1-k}(C), \quad (14)$$

$$\text{and } \bar{H}_*^{1-k}(C) = E \bar{H}^{1-k}(X, V) I(X \in C) = \int_C \bar{H}^{1-k} \phi_V(x) dx.$$

This paper focuses on the three Edgeworth expansions, (6) and (12), when $q \rightarrow \infty$ as $n \rightarrow \infty$.

If $-C = C$, then for r odd, $P_{rk}(C) = P_r(C) = 0$, so that

$$P(X_n \in C) \approx \sum_{r=0}^{\infty} n^{-r} P_{2r}(C) = \Phi_V(C) + n^{-1} P_2(C) + O(n^{-2}).$$

Examples 3 and 4 of [14] gave $P_2(C)$ for

$C = \{x : x'V^{-1}x \leq u\}$, and $C = \{x : |(V^{-1/2}x)_j| \leq u_j, j = 1, \dots, q\}$.

The main take-away here is that for $x \in R^q, C \subset R^q, \tilde{p}_r(x), p_r(x), P_r(C)$ of (7), (8) and (12), and $s = 1, 2, \dots$,

$$p_{X_n}(x)/\phi_V(x) = \sum_{r=0}^{s-1} n^{-r/2} \tilde{p}_r(x) + O(n^{-s/2}),$$

$$P(X_n \leq x) = \sum_{r=0}^{s-1} n^{-r/2} P_r(x) + O(n^{-s/2}),$$

$$P(X_n \in C) = \sum_{r=0}^{s-1} n^{-r/2} P_r(C) + O(n^{-s/2}),$$

where, for example, by (4),

$$\tilde{p}_1(x) = p_1(x)/\phi_V(x) = \sum_{k=1,3} \tilde{p}_{1k}, \tilde{p}_{11} = \bar{k}_1^1 \bar{H}^1, \tilde{p}_{13} = \bar{k}_2^{1-3} \bar{H}^{1-3}/6,$$

$$P_1(x) = \sum_{k=1,3} P_{1k}(x), P_{11}(x) = \bar{k}_1^1 \bar{H}_*^1, P_{13}(x) = \bar{k}_2^{1-3} \bar{H}_*^{1-3}/6,$$

$$P_1(C) = \sum_{k=1,3} P_{1k}(C), P_{11}(C) = \bar{k}_1^1 \bar{H}_*^1(C), P_{13}(C) = \bar{k}_2^{1-3} \bar{H}_*^{1-3}(C)/6.$$

These asymptotic expansions generally diverge, as normal moments and Hermite polynomials increase very rapidly with their degree.

3. The Case $q = q_n \rightarrow \infty$ as $n \rightarrow \infty$

Theorem 1. Let $\hat{w}$ be a non-lattice estimate of $w \in R^q$, satisfying $E \hat{w} \rightarrow w$, and (3). Set $X_n = n^{1/2}(\hat{w} - w)$. Take $s \geq 1$. Suppose that as $n \rightarrow \infty$,

$$q = q_n \rightarrow \infty, \text{ and } v_n = n^{-1/2} q_n^3 \rightarrow 0, \text{ that is, } q_n = o(n^{1/6}).$$

Then, for $\tilde{p}_r(x)$ of (7), $P_r(x)$ of (8), and $P_r(C)$ of (13),

$$p_{X_n}(x)/\phi_V(x) = \sum_{r=0}^{s-1} n^{-r/2} \tilde{p}_r(x) + O(v_n^s), \quad (15)$$

$$P(X_n \leq x) = \sum_{r=0}^{s-1} n^{-r/2} P_r(x) + O(v_n^s), \quad (16)$$

$$P(X_n \in C) = \sum_{r=0}^{s-1} n^{-r/2} P_r(C) + O(v_n^s). \quad (17)$$

PROOF Set $Q = q^2$. $\tilde{p}_{rk}(x)$ and $P_{rk}(x)$ of (8), and $\tilde{p}_{rk}(C)$ of (14), each have q^k terms for $q = q_n$. So $p_r(x)$ of (7), $P_r(x)$ of (8), and $P_r(C)$ of (12), each have N_r terms, where

$$\text{for } r \text{ odd, } N_r = q + q^3 + q^5 + \cdots + q^{3r} = q(Q^{(3r+1)/2} - 1)/(Q - 1), \quad (18)$$

$$\text{for } r \text{ even, } N_r = q^2 + q^4 + \cdots + q^{3r} = Q(Q^{3r/2} - 1)/(Q - 1). \quad (19)$$

$$\text{So, } N_r = O(q^{3r}) \text{ as } n \rightarrow \infty.$$

So $p_r(x)$, $P_r(x)$ and $P_r(C)$ each have magnitude q_n^{3r} as $n \rightarrow \infty$.

So $n^{-s/2} (p_s(x), P_s(x), P_s(C))$ has magnitude $n^{-s/2} q_n^{3s} = v_n^s$. The theorem follows. $\square$

Example 1. Let $\hat{w}$ be the sample mean from a distribution on R^q with finite cross cumulants $\bar{\kappa}^{1-r}$, $r \geq 1$. Then $E \hat{w} = w$, and only the leading coefficient in (3) is non-zero. As in Example 2 of [14], by (4)–(5), the non-zero Edgeworth coefficients P_r^{1-k} needed for the three 4th order Edgeworth expansions of $\hat{w}$ with $V = (\bar{\kappa}^{12})$, are

$$\begin{aligned} \bar{P}_1^{1-3} &= \bar{\kappa}^{1-3}/3!, \quad \bar{P}_2^{1-4} = \bar{\kappa}^{1-4}/4!, \quad \bar{P}_2^{1-6} = S \bar{\kappa}^{1-3} \bar{\kappa}^{4-6}/72, \\ \bar{P}_3^{1-5} &= \bar{\kappa}^{1-5}/5!, \quad \bar{P}_3^{1-7} = S \bar{\kappa}^{123} \bar{\kappa}^{4-7}/144, \quad \bar{P}_3^{1-9} = S \bar{\kappa}^{1-3} \bar{\kappa}^{4-6} \bar{\kappa}^{7-9}/6^4. \end{aligned}$$

Substitution gives $(p_s(x), P_s(x), P_s(C))$ for $s = 1, 2, 3$. For example, as $q_n = q \rightarrow \infty$, the coefficients of $n^{-1/2}$ in the 2nd terms in the Edgeworth expansions for $p_{X_n}(x)/\phi_V(x)$, $P(X_n \leq x)$, and $P(X_n \in C)$, are

$$\begin{aligned} \tilde{p}_1(x) &= \tilde{p}_{13} = \bar{\kappa}^{1-3} \bar{H}^{1-3}/6 = O(q_n^3), \\ P_1(x) &= P_{13}(x) = \bar{\kappa}^{1-3} \bar{H}_*^{1-3}/6 = O(q_n^3), \\ P_1(C) &= P_{13}(C) = \bar{\kappa}^{1-3} \bar{H}_*^{1-3}(C)/6 = O(q_n^3). \end{aligned}$$

Example 2. Suppose that $X_1, X_2, \dots, X_n$ are independent random vectors in R^q , with mean $\hat{w} = \bar{X}$, and that X_j has finite cross-cumulants

$$\bar{\kappa}_j^{1-r} = \kappa(X_{ji_1}, \dots, X_{ji_r}) \text{ for } r \geq 1 \text{ and } i_1, \dots, i_r \text{ in } 1, \dots, q.$$

Then for $r \geq 1$,

$$\kappa(\hat{w}_{i_1}, \dots, \hat{w}_{i_r}) = n^{1-r} \bar{\kappa}^{1-r} \text{ where } \bar{\kappa}^{1-r} = n^{-1} \sum_{j=1}^n \bar{\kappa}_j^{1-r}.$$

Suppose also, that $\{\bar{\kappa}^{1-r}\}$ are bounded in n , and that for $V = (\bar{\kappa}^{12})$, $\det V$ is bounded away from 0, as n increases. Then the Edgeworth coefficients needed for the three 4th order Edgeworth expansions for $\hat{w}$, are given by Example 3.1.

4. Further Reduction of Terms

Our next theorem gives a way to reduce the number of terms in $\tilde{p}_r(x)$ of (7), $P_r(x)$ of (8), and $P_r(C)$ of (14), from q^k , to $L_k = q_k [1 + O(q^{-1})]$ as $q \rightarrow \infty$, where

$$q_k = \binom{q}{k} = q(q-1) \dots (q-k+1)/k!.$$

As we do not use q_n for q in this section, there is no ambiguity with this different use of q_k . k

$\tilde{p}_{rk}(x)$, $P_{rk}(x)$ and $P_{rk}(C)$ have k summations over $1, \dots, q$, and so have q^k terms. But many are duplicates as $\tilde{P}_r^{1-k}$ is symmetric in $i_1, \dots, i_k$. Set $\binom{k}{a \dots b} = k!/a! \dots b!$, the multinomial coefficient. For example $\binom{3}{111} = 6$.

Theorem 2. Let $T^{i_1 \dots i_k} \in R$ be symmetric in $i_1, \dots, i_k$.

$$\begin{aligned} \text{Set } T^{i_1^2 i_2} &= T^{i_1 i_1 i_2} \text{ and so on, and } u_k = \sum_{i_1, \dots, i_k=1}^q T^{i_1 \dots i_k}. \\ \text{Then, } u_1 &= \sum_{i_1=1}^q T^{i_1}, \quad u_2 = \sum_{i_1=1}^q T^{i_1^2} + 2! \sum_{i_1 > i_2} T^{i_1 i_2}, \\ u_3 &= \sum_{i_1=1}^q T^{i_1^3} + \binom{3}{1} \sum_{i_1=1}^{q(q-1)} T^{i_1^2 i_2} + 3! \sum_{i_1 > i_2 > i_3} T^{i_1 i_2 i_3}, \\ u_4 &= \sum_{i_1=1}^q T^{i_1^4} + \binom{4}{1} \sum_{i_1=1}^{q(q-1)} T^{i_1^3 i_2} + \binom{4}{2} \sum_{i_1 > i_2} T^{i_1^2 i_2^2} + \binom{4}{211} \sum_{i_1 > i_2} T^{i_1^2 i_2 i_3} \\ &\quad + 4! \sum_{i_1 > i_2 > i_3 > i_4} T^{i_1 i_2 i_3 i_4}, \\ u_5 &= \sum_{i_1=1}^q T^{i_1^5} + \binom{5}{1} \sum_{i_1=1}^{q(q-1)} T^{i_1^4 i_2} + \binom{5}{2} \sum_{i_1=1}^{q(q-1)} T^{i_1^3 i_2^2} + \binom{5}{221} \sum_{i_1 > i_2} T^{i_1^2 i_2^2 i_3} \\ &\quad + \binom{5}{2111} \sum_{i_2 > i_3 > i_4} T^{i_1^2 i_2 i_3 i_4} + 5! \sum_{i_1 > i_2 > i_3 > i_4 > i_5} T^{i_1 i_2 i_3 i_4 i_5}, \\ u_6 &= \sum_{i_1=1}^q T^{i_1^6} + \binom{6}{1} \sum_{i_1=1}^{q(q-1)} T^{i_1^5 i_2} + \binom{6}{2} \sum_{i_1=1}^{q(q-1)} T^{i_1^4 i_2^2} + \binom{6}{3} \sum_{i_1 > i_2} T^{i_1^3 i_2^3} \\ &\quad + \binom{6}{411} \sum_{i_2 > i_3} T^{i_1^4 i_2 i_3} + \binom{6}{321} \sum_{i_1 > i_2} T^{i_1^3 i_2^2 i_3} + \binom{6}{222} \sum_{i_1 > i_2 > i_3} T^{i_1^2 i_2^2 i_3^2} \\ &\quad + \binom{6}{3111} \sum_{i_2 > i_3 > i_4} T^{i_1^3 i_2 i_3 i_4} + \binom{6}{2211} \sum_{i_1 > i_2, i_3 > i_4} T^{i_1^2 i_2^2 i_3 i_4} \\ &\quad + \binom{6}{21111} \sum_{i_2 > i_3 > i_4 > i_5} T^{i_1^2 i_2 i_3 i_4 i_5} + 6! \sum_{i_1 > \dots > i_6} T^{i_1 \dots i_6}, \end{aligned}$$

where the sums are for distinct i_j . This reduces the number of terms in u_k from q^k to L_{qk} where

$$\begin{aligned} L_{q1} &= q, \quad L_{q2} = q + q_2, \quad L_{q3} = q + 2q_2 + q_3, \\ L_{q4} &= q + 3q_2 + 3q_3 + q_4, \quad L_{q5} = q + 4q_2 + 5q_3 + 4q_4 + q_5, \\ L_{q6} &= q + 5q_2 + 10q_3 + 14q_4 + 5q_5 + q_6. \end{aligned}$$

For example, $(q, k) = (5, 3) \Rightarrow L_{qk}/q^k = 0.28$; that is, Theorem 4.1 reduces the number of terms to 28%.

As a check, $q^k = \sum_{j=1}^k S(k, j) (q)_j$ where $(q)_j = j!q_j = q(q-1)\dots(q-j+1)$, and $S(k, j)$ is the Stirling number of the 2nd kind tabled on p310 of Comtet (1974). For example, counting the number of terms to calculate in u_4 above gives $q^4 = q + (4+3)(q)_2 + 6(q)_3 + (q)_4$.

Taking

$$T^{i_1\dots i_k} = \bar{P}_r^{1-k} \bar{f}^{1-k}, \text{ where } \bar{f}^{1-k} = \bar{H}^{1-k}, \text{ or } \bar{H}_*^{1-k}, \text{ or } \bar{H}_*^{1-k}(C), \quad (20)$$

and tensor summation is *not* used, gives $u_k = \tilde{p}_{rk}$ of (7), or $u_k = P_{rk}(x)$ of (8), or $u_k = P_{rk}(C)$ of (14). For example,

$$\begin{aligned} \tilde{p}_{r1} &= \sum_{i_1=1}^q P_r^{i_1} H^{i_1}, P_{r1}(x) = \sum_{i_1=1}^q P_r^{i_1} H_*^{i_1}, P_{r1}(C) = \sum_{i_1=1}^q P_r^{i_1} H_*^{i_1}(C), \\ \tilde{p}_{r2} &= \sum_{i_1=1}^q P_r^{i_1 i_1} H^{i_1 i_1} + 2 \sum_{i_1 > i_2}^{q_2} P_r^{i_1 i_2} H^{i_1 i_2}, P_{r2}(x) = \sum_{i_1=1}^q P_r^{i_1 i_1} H_*^{i_1 i_1} + 2 \sum_{i_1 > i_2}^{q_2} P_r^{i_1 i_2} H_*^{i_1 i_2}, \\ P_{r2}(C) &= \sum_{i_1=1}^q P_r^{i_1 i_1} H_*^{i_1 i_1}(C) + 2 \sum_{i_1 > i_2}^{q_2} P_r^{i_1 i_2} H_*^{i_1 i_2}(C), \\ \tilde{p}_{r3} &= \sum_{i_1=1}^q P_r^{i_1 i_1 i_1} H^{i_1 i_1 i_1} + 3 \sum_{i_1 \neq i_2}^{q(q-1)} P_r^{i_1 i_1 i_2} H^{i_1 i_1 i_2} + 6 \sum_{i_1 > i_2 > i_3}^{q_3} P_r^{i_1 i_2 i_3} H^{i_1 i_2 i_3}, \\ P_{r3}(x) &= \sum_{i_1=1}^q P_r^{i_1 i_1 i_1} H_*^{i_1 i_1 i_1} + 3 \sum_{i_1 \neq i_2}^{q(q-1)} P_r^{i_1 i_1 i_2} H_*^{i_1 i_1 i_2} + 6 \sum_{i_1 > i_2 > i_3}^{q_3} P_r^{i_1 i_2 i_3} H_*^{i_1 i_2 i_3}, \\ P_{r3}(C) &= \sum_{i_1=1}^q P_r^{i_1 i_1 i_1} H_*^{i_1 i_1 i_1}(C) + 3 \sum_{i_1 \neq i_2}^{q(q-1)} P_r^{i_1 i_1 i_2} H_*^{i_1 i_1 i_2}(C) + 6 \sum_{i_1 > i_2 > i_3}^{q_3} P_r^{i_1 i_2 i_3} H_*^{i_1 i_2 i_3}(C). \end{aligned}$$

This shows that we can replace q^k in the derivation of Theorem 3.1 by L_{qk} . So for $r = 1, 2$, the number of terms, N_r of (18) and (19), in $p_r(x)$ of (7), $P_r(x)$ of (8), and $P_r(C)$ of (12), can be reduced to N'_r where

$$N'_1 = L_{q1} + L_{q3} = 2q + 2q_2 + q_3, \quad (21)$$

$$N'_2 = L_{q2} + L_{q4} + L_{q6} = 3q + 10q_2 + 10q_3 + 15q_4 + 5q_5 + q_6. \quad (22)$$

So, $q = 1 \Rightarrow N_1 = 2, N'_1 = 2, N_2 = 3, N'_2 = 3$,

$q = 2 \Rightarrow N_1 = 10, N'_1 = 6 = .6 N_1, N_2 = 84, N'_2 = 16 = .19 N_2$,

$q = 3 \Rightarrow N_1 = 30, N'_1 = 13 = .43 N_1, N_2 = 819, N'_2 = 49 = .06 N_2$,

$q = 4 \Rightarrow N_1 = 68, N'_1 = 24 = .35 N_1, N_2 = 4368, N'_2 = 127 = .02 N_2$,

$q = 5 \Rightarrow N_1 = 130, N'_1 = 40 = .31 N_1, N_2 = 16275, N'_2 = 295 = .02 N_2$,

where the values of N'_1/N_1 and N'_2/N_2 are approximate.

For example, if $q = 3$, Theorem 4.1 reduces the number of terms needed for 2nd and 3rd order Edgeworth expansions by 57% and 94%. If $q = 4$ or 5, Theorem 4.1 reduces the number of terms to calculate by 65% or 69% for the 2nd order Edgeworth expansions, and by 98% for the 3rd order Edgeworth expansions.

When $q = 2$, this reduction of terms was used in Section 4 of [14].

Example 3. In Example 3.1, using the expression for u_3 in Theorem 4.1,

$$\begin{aligned} 6\tilde{p}_1(x) &= \sum_{i_1=1}^q \kappa^{i_1^3} H^{i_1^3} + 3 \sum_{i_1=1}^{q(q-1)} \kappa^{i_1^2 i_2} H^{i_1^2 i_2} + 6 \sum_{i_1 > i_2 > i_3}^{q_3} \kappa^{i_1 i_2 i_3} H^{i_1 i_2 i_3}, \\ 6P_1(x) &= \sum_{i_1=1}^q \kappa^{i_1^3} H_*^{i_1^3} + 3 \sum_{i_1=1}^{q(q-1)} \kappa^{i_1^2 i_2} H_*^{i_1^2 i_2} + 6 \sum_{i_1 > i_2 > i_3}^{q_3} \kappa^{i_1 i_2 i_3} H_*^{i_1 i_2 i_3}, \\ 6P_1(C) &= \sum_{i_1=1}^q \kappa^{i_1^3} H_*^{i_1^3}(C) + 3 \sum_{i_1=1}^{q(q-1)} \kappa^{i_1^2 i_2} H_*^{i_1^2 i_2}(C) + 6 \sum_{i_1 > i_2 > i_3}^{q_3} \kappa^{i_1 i_2 i_3} H_*^{i_1 i_2 i_3}(C). \end{aligned}$$

As $\bar{P}_1^1 = \bar{P}_2^{12} = 0$, L_{q1} in (21), and L_{q2} in (22) need to be deleted.

5. Functions of a Vector Sample Mean

Let $\bar{X}$ be the sample mean from a non-lattice distribution on R^q with mean μ , and finite cross cumulants $\bar{\kappa}^{1-r} = \kappa^{j_1 \dots j_r}$, $r \geq 1$. Let $t(\cdot) : R^q \rightarrow R^p$ be a function with i_k th component $\bar{t}^k(\cdot) = t^{i_k}(\cdot)$ having finite derivatives at μ ,

$$\bar{t}_{r-s}^k = \partial_{j_r} \dots \partial_{j_s} t^k(\mu), \text{ where } \partial_j = \partial / \partial \mu_j, \quad s \leq r.$$

So we expand the bar convention used earlier using $j_1, j_2, \dots$ in $1, \dots, p$ as well as $i_1, i_2, \dots$ in $1, \dots, q$ as before. Here we have a notation dilemma. We chose $t(\cdot) : R^q \rightarrow R^p$, as this allows us to keep the notation $\bar{\kappa}_d^{1-r} = \kappa_d^{j_1 \dots j_r}$ and $\bar{P}_r^{1-k} = P_r^{i_1 \dots i_k}$, as used earlier. However, now $\hat{w} \in R^p$ not R^q , so that implicit summation in, say $\bar{P}_r^{1-k} = P_r^{i_1 \dots i_k}$, as used earlier. So now, implicit summation in $u_k = \bar{P}_r^{1-k} \bar{H}^{1-k}$ is over $i_1, \dots, i_k$ in $1, \dots, p$, not in $1, \dots, q$, and this u_k has p^k terms, reducible to L_{pk} using Theorem 4.1.

If I had chosen $t(\cdot) : R^p \rightarrow R^q$, then $\bar{\kappa}_d^{1-r}$ and $\bar{P}_r^{1-k}$ would have had to be reinterpreted as $\kappa_d^{j_1 \dots j_r}$ and $P_r^{j_1 \dots j_k}$, which would likely be confusing.

Let us use $\sum f^{i_1 \dots i_k} \sim p^k$, to mean $\sum f^{i_1 \dots i_k}$ has magnitude p^k from summing $i_1, \dots, i_k$ over $1, \dots, p$.

$$\text{Set } \sum \bar{t}_{a-b}^1 \bar{t}_{c-d}^2 = \bar{t}_{a-b}^1 \bar{t}_{c-d}^2 + \bar{t}_{a-b}^2 \bar{t}_{c-d}^1.$$

More generally, for $\pi_1, \dots, \pi_r$ any partition of $1, \dots, k$, let $\sum^N \bar{t}_{\pi_1}^1 \dots \bar{t}_{\pi_r}^r$ denote the sum of $t_{\pi_1}^{b_1} \dots t_{\pi_r}^{b_r}$ over all N permutations $b_1, \dots, b_r$ of $j_1, \dots, j_r$, giving distinct values. For example,

$$\sum \bar{t}_{13}^1 \bar{t}_{24}^2 \bar{t}_4^3 = \bar{t}_{13}^1 \bar{t}_{24}^2 \bar{t}_4^3 + \bar{t}_{13}^2 \bar{t}_{24}^1 \bar{t}_4^3 + \bar{t}_{13}^3 \bar{t}_{24}^2 \bar{t}_4^1.$$

I now give the cumulant coefficients, $\bar{\kappa}_d^{1-r}$ of (3), for $\hat{w} = t(\bar{X})$, and track their magnitude in p from summing $i_1, \dots, i_k$ over $1, \dots, p$.

Theorem 3. For $\hat{w} = t(\bar{X}) : R^q \rightarrow R^p$, (3) holds, where the cumulant coefficients, $\bar{\kappa}_d^{1-r} = \kappa_d^{i_1 \dots i_r}$, needed for the 3rd order Edgeworth expansions of order $s = 1, 2, 3, 4$, are given by

$$\text{For } s = 1 : \bar{\kappa}_1^{12} = V_{i_1 i_2} = \bar{t}_1^1 \bar{t}_2^2 \bar{\kappa}^{12} = \sum_{j_1, j_2=1}^q t_{j_1}^{i_1} \bar{t}_{j_2}^{i_2} \kappa^{j_1 j_2} \sim q^2,$$

$$\text{For } s = 2 : \bar{\kappa}_1^1 = \bar{t}_{12}^1 \bar{\kappa}^{12} / 2 \sim q^2,$$

$$\bar{\kappa}_2^{1-3} = \bar{t}_1^1 \bar{t}_2^2 \bar{t}_3^3 \bar{\kappa}^{1-3} + T_{1-4}^{1-3} \bar{\kappa}^{12} \bar{\kappa}^{34} \sim q^2 + q^4 \sim q^4, \text{ where } T_{1-4}^{1-3} = \sum \bar{t}_{13}^1 \bar{t}_2^2 \bar{t}_4^3.$$

$$\text{For } s = 3 : \bar{\kappa}_2^{12} = T_{1-3}^{12} \bar{\kappa}^{1-3} / 2 + T_{1-4}^{12} \bar{\kappa}^{12} \bar{\kappa}^{34} / 2 \sim q^3 + q^4 \sim q^4, \text{ where}$$

$$T_{1-3}^{12} = \sum \bar{t}_{12}^1 \bar{t}_3^2, \quad T_{1-4}^{12} = \sum \bar{t}_{1-3}^1 \bar{t}_4^2 + \bar{t}_{13}^1 \bar{t}_{24}^2,$$

$$\bar{k}_3^{1-4} = \bar{t}_1^1 \bar{t}_2^2 \bar{t}_3^3 \bar{t}_4^4 \bar{k}^{1-4} + T_{1-5}^{1-4} \bar{k}^{1-3} \bar{k}^{45} + T_{1-6}^{1-4} \bar{k}^{12} \bar{k}^{34} \bar{k}^{56} \sim q^4 + q^5 + q^6 \sim q^6,$$

$$\text{where } T_{1-5}^{1-4} = \sum_{12} \bar{t}_1^1 \bar{t}_2^2 \bar{t}_3^3 \bar{t}_4^4, \quad T_{1-6}^{1-4} = \sum_{135} \bar{t}_1^1 \bar{t}_2^2 \bar{t}_3^3 \bar{t}_4^4 + \sum_{125} \bar{t}_1^1 \bar{t}_2^2 \bar{t}_3^3 \bar{t}_4^4.$$

$$\text{For } s = 4: \bar{k}_2^1 = \bar{t}_{1-3}^1 \bar{k}^{1-3} / 6 + \bar{t}_{1-4}^1 \bar{k}^{12} \bar{k}^{34} / 8 \sim q^3 + q^4 \sim q^4,$$

$$\bar{k}_3^{1-3} = T_{1-4}^{1-3} \bar{k}^{1-4} / 2 + T_{1-5}^{1-3} \bar{k}^{1-3} \bar{k}^{45} + T_{1-6}^{1-3} \bar{k}^{12} \bar{k}^{34} \bar{k}^{56} \sim q^4 + q^5 + q^6 \sim q^6,$$

$$\text{where } T_{1-4}^{1-3} = \sum_{13} \bar{t}_1^1 \bar{t}_2^2 \bar{t}_3^3,$$

$$T_{1-5}^{1-3} = \sum_{124} \bar{t}_1^1 \bar{t}_2^2 \bar{t}_3^3 / 2 + \sum_{145} \bar{t}_1^1 \bar{t}_2^2 \bar{t}_3^3 / 2 + \sum_{12} \bar{t}_1^1 \bar{t}_2^2 \bar{t}_3^3 / 2 + \sum_{14} \bar{t}_1^1 \bar{t}_2^2 \bar{t}_3^3,$$

$$T_{1-6}^{1-3} = \sum_{1235} \bar{t}_1^1 \bar{t}_2^2 \bar{t}_3^3 / 2 + \sum_{1-3} \bar{t}_1^1 \bar{t}_2^2 \bar{t}_3^3 + \sum_{135} \bar{t}_1^1 \bar{t}_2^2 \bar{t}_3^3 / 2 + \bar{t}_{13}^1 \bar{t}_{25}^2 \bar{t}_{46}^3.$$

$$\bar{k}_4^{1-5} = \bar{t}_1^1 \dots \bar{t}_5^5 \bar{k}^{1-5} + T_{1-6}^{1-5} \bar{k}^{1-4} \bar{k}^{56} + U_{1-6}^{1-5} \bar{k}^{1-3} \bar{k}^{4-6} \\ + T_{1-7}^{1-5} \bar{k}^{1-3} \bar{k}^{45} \bar{k}^{67} + T_{1-8}^{1-5} \bar{k}^{12} \bar{k}^{34} \bar{k}^{56} \bar{k}^{78} \sim p^5 + \dots + p^8 \sim p^8,$$

$$\text{where } T_{1-6}^{1-5} = \sum_{15} \bar{t}_1^1 \bar{t}_2^2 \bar{t}_3^3 \bar{t}_4^4 \bar{t}_5^5, \quad U_{1-6}^{1-5} = \sum_{14} \bar{t}_1^1 \bar{t}_2^2 \bar{t}_3^3 \bar{t}_4^4 \bar{t}_5^5,$$

$$T_{1-7}^{1-5} = \sum_{146} \bar{t}_1^1 \bar{t}_2^2 \bar{t}_3^3 \bar{t}_4^4 \bar{t}_5^5 + \sum_{14} \bar{t}_1^1 \bar{t}_2^2 \bar{t}_3^3 \bar{t}_4^4 \bar{t}_5^5 + \sum_{14} \bar{t}_1^1 \bar{t}_2^2 \bar{t}_3^3 \bar{t}_4^4 \bar{t}_5^5,$$

$$T_{1-8}^{1-5} = \sum_{1357} \bar{t}_1^1 \bar{t}_2^2 \bar{t}_3^3 \bar{t}_4^4 \bar{t}_5^5 / 5 + \sum_{135} \bar{t}_1^1 \bar{t}_2^2 \bar{t}_3^3 \bar{t}_4^4 \bar{t}_5^5 + \sum_{13} \bar{t}_1^1 \bar{t}_2^2 \bar{t}_3^3 \bar{t}_4^4 \bar{t}_5^5,$$

PROOF This is a special case of Theorem 2 of [13] with $\bar{k}_d^{1-r}$ replaced by $\bar{k}^{1-r} I(d = r - 1)$. $\square$

Lemma 1. For $\bar{f}^{1-k} = f^{i_1 \dots i_k} \in R$, and $r = 1, 2, 3$,

$$\bar{P}_r^{1-k} \sim q^{r+k}. \text{ So, } \bar{P}_r^{1-k} \bar{f}^{1-k} \sim q^{r+k} p^k.$$

$$\text{Set } P_r = \sum_{k=1}^{3r} [\bar{P}_r^{1-k} \bar{f}^{1-k} : k - r \text{ even}]. \text{ Then } P_r \sim q^{4r} p^{3r}. \quad (23)$$

PROOF Use Theorem 5.1 to check that for each of the Edgeworth coefficients given by (4)–(5), $\bar{P}_r^{1-k} \sim q^{r+k}$. The dominant term in P_r is for $k = 3r$. So, (23) holds. $\square$

Theorem 4. Set $\hat{w} = t(\bar{X}) : R^q \rightarrow R^p$. Suppose that $\hat{w}$ is non-lattice, that $E \hat{w} \rightarrow w$, and that (3) holds.. Suppose that as $n \rightarrow \infty$, $p_n q_n = pq \rightarrow \infty$, and that

$$v_n = n^{-1/2} q_n^4 p_n^3 \rightarrow 0. \text{ That is, } q_n^8 p_n^6 = o(n). \quad (24)$$

Then (15)–(17) hold for $s = 1, 2, 3, 4$ with v_n of (24).

PROOF For P_r of (23), $n^{-r/2} P_r \sim v_n^r$. Now take $\bar{f}^{1-k}$ of (20). $\square$ For example, (24) holds for fixed $p = p_n$ if $q_n = o(n^{1/8})$, and for fixed $q = q_n$ if $p_n = o(n^{1/6})$.

We now apply Theorem 4.1 to the components of Theorem 5.1.

Theorem 5.

$$\begin{aligned}
&\text{For } s = 1 : \bar{k}_1^{12} = V_{i_1 i_2} \sim L_{q2}. \\
&\text{For } s = 2 : \bar{k}_1^1 \sim L_{q2}, \text{ reduced from } q^2, \\
&\bar{k}_2^{1-3} \sim L_{q3} + 3L_{q4}, \text{ reduced from } q^3 + 3q^4. \\
&\text{For } s = 3 : \bar{k}_2^{12} \sim 2L_{q3} + 3L_{q4}, \text{ reduced from } 2q^3 + 3q^4, \\
&\bar{k}_3^{1-4} \sim L_{q4} + 12L_{q5} + 16L_{q6}, \text{ reduced from } q^4 + 12q^5 + 16q^6. \\
&\text{For } s = 4 : \bar{k}_2^1 \sim L_{q3} + L_{q4}, \text{ reduced from } q^3 + q^4, \\
&\bar{k}_3^{1-3} \sim 3L_{q4} + 18L_{q5} + 16L_{q6}, \text{ reduced from } 3q^4 + 18q^5 + 16q^6, \\
&\bar{k}_4^{1-5} \sim L_{q5} + (20 + 15)L_{q6} + 150L_{q7} + 125L_{q8}, \\
&\text{reduced from } q^5 + (20 + 15)q^6 + 150q^7 + 125q^8.
\end{aligned}$$

Let us write these as

$$\bar{k}_d^{1-r} \sim N_{rd}.$$

For example, $L_{q2}/q^2 = (q+1)/2q \rightarrow 1/2$ as $q \rightarrow \infty$, and for $q = 4$, the number of calculations needed for $\bar{k}_2^{1-3}$ is reduced by the factor $N_{32}/(q^3 + 3q^4) = (L_{q3} + 3L_{q4})/(q^3 + 3q^4) = 3/26$.

One can now work out similar results for $T^{i_1 \dots i_k}$ of (20), and so for the terms of the Edgeworth expansions, $\tilde{p}_r(x)$, $P_r(x)$ and $P_r(C)$. For example,

$$\begin{aligned}
\tilde{p}_{11} &= \bar{P}_1^1 \bar{H}^1 = \bar{k}_1^1 \bar{H}^1 \sim pN_{11} = pL_{q2}, \\
\tilde{p}_{13} &= \bar{P}_1^{1-3} \bar{H}^{1-3} \sim k_2^{1-3} \bar{H}^{1-3} \sim L_{p3} N_{32}, \\
&\Rightarrow \tilde{p}_1(x) \sim pN_{11} + L_{p3} N_{32} = pL_{q2} + L_{p3}(L_{q3} + 3L_{q4}),
\end{aligned}$$

and $T^{i_1 i_2} = P_2^{i_1 i_2} H^{i_1 i_2}$ without implicit summation,

$$\tilde{p}_{22} = \bar{P}_2^{12} \bar{H}^{12} = \sum_{i_1=1}^p T^{i_1^2} + 2 \sum_{i_1 > i_2}^{p_2} T^{i_1 i_2} \sim pL_{q2} + p_2 L_{q0}.$$

6. Conclusions

Let $\hat{w}$ be a standard estimate of an unknown $w \in R^q$. That is, $\hat{w}$ is a consistent estimate, and for $r \geq 1$, its r th order cumulants have magnitude n^{1-r} , and can be expanded in powers of n^{-1} . This is a very large class of estimates. It includes functions of sample moments and empirical distributions, samples of independent but not identically distributed random vectors, and samples from stationary series. Then by §2, for fixed q , and non-lattice estimates, Edgeworth-type expansions hold for the density and distribution of $X_n = n^{1/2}(\hat{w} - w)$, in terms of the Edgeworth coefficients given in [14] for $s \leq 4$. For $s \geq 1$, their first s terms of the three Edgeworth expansions, give the density and distribution of X_n , and $P(X_n \in C)$, to $O(n^{-s/2})$ as $n \rightarrow \infty$. Theorem 3.1 shows that this remains true when $q_n = q \rightarrow \infty$, if the remainder, $O(n^{-s/2})$, is replaced by $O(v_n^s)$, where $v_n = n^{-1/2} q_n^3$. So the three Edgeworth expansions hold when $v_n \rightarrow \infty$, that is, when $q_n = o(n^{1/6})$.

Theorem 4.1 gives formulas that dramatically reduce the number of terms needed by 2nd and 3rd order Edgeworth-type expansions, that is, for 1st and 2nd order corrections to the CLT.

When $\hat{w} = t(\bar{X}) : R^q \rightarrow R^p$, and $p_n q_n = pq \rightarrow \infty$, Theorem 5.2 shows that the three 4th order Edgeworth expansions hold if $q_n^8 p_n^6 = o(n)$. For example, this holds for fixed $p = p_n$ if $q_n = o(n^{1/8})$.

7. Discussion

Theorem 3.1, showed that the three Edgeworth expansions considered here, hold for standard estimates when $q_n = o(n^{1/6})$. Is this result optimal? It is certainly not optimal for a sample mean. For, as noted in Section 1, Theorem 2.1 of the recent paper [7] gives conditions for a 2nd order

Edgeworth expansion for the distribution of a sample mean, when $q_n = O(n^c)$ for any c , or even when $q_n = O(\exp(bn^c))$ and $c < 1/3$.

It will be interesting to see if such weak conditions on q_n can be extended to a wide class of estimates, such as standard estimates.

Theorem 5.2 showed that the three 4th order Edgeworth expansions hold for $\hat{w} = t(\bar{X}) : R^q \rightarrow R^p$, when $q_n^8 p_n^6 = o(n)$. It will be interesting to see how much this condition can be weakened. It should not be hard to extend Theorem 5.2 to $\hat{w} = T(F_n) \in R^q$, where $F_n(x)$ is the empirical distribution of a random sample of size n from a distribution $F(x)$ on R^p .

One can also extend this to $\hat{w}$ a function of K independent sample means, $\hat{w} = t(\bar{X}_1, \dots, \bar{X}_K) : R^{q_1} \times \dots \times R^{q_K} \rightarrow R^p$.

The method employed here should also be able to extend many of the results in the references from a sample mean to a standard estimate.

References

1. Chernozhukov, V., Chetverikov, D., and Kato, K. (2013) Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors. *Ann. Stat.*, 41 (6), 2786–2819.
2. Chernozhukov, V., Chetverikov, D., and Kato, K. (2017). Central limit theorems and bootstrap in high dimensions. *Ann. Probab.*, 45 (4), 2309–2352. arXiv:1412.3661.
3. Donoho, D. and Montanari, A. (2016) High dimensional robust m-estimation: Asymptotic variance via approximate message passing. *Probability Theory and Related Fields*, Springer.
4. Fang, X. and Koike, Y. (2021). High-dimensional central limit theorems by Stein's method. *Ann. Appl. Probab.* 31, 1660–1686.
5. Fujikoshi, Y. and Sakurai, T. (2009) High-dimensional asymptotic expansions for the distributions of canonical correlations. *Journal of Multivariate Analysis*, 100 (1), 231–242.
6. Fujikoshi, Y., Ulyanov, V. V., and Shimizu, R. (2011). *Multivariate statistics: High-dimensional and large-sample approximations*. John Wiley and Sons.
7. Koike, Yuta (2025). High-dimensional bootstrap and asymptotic expansion. arXiv preprint arXiv:2404.05006
8. Kosorok, M. and Ma, S. (2007). Marginal asymptotics for the large p , small n paradigm: With application to microarray data. *Ann. Statist.*, 35, 1456–1486. MR2351093.
9. Kuelbs, J. and Vidyashankar, A.N. (2010). Asymptotic inference for high-dimensional data. *Annals of Statistics*, 38 (2) 836–869.
10. Portnoy, S. (1984). Asymptotic behavior of M-estimators of p regression parameters when p^2/n is large. I. Consistency. *Ann. Statist.*, 12, 1298–1309. MR0760690
11. Portnoy, S. (1985). Asymptotic behavior of M-estimators of p regression parameters when p^2/n is large. II: normal approximation, *Ann. Statist.*, 13, 1403–1417.
12. Withers, C.S. (2000) A simple expression for the multivariate Hermite polynomials. *Stat. Prob. Lett.*, 47, 165–169.
13. Withers, C.S. (2024) 5th order multivariate Edgeworth expansions for parametric estimates. *Mathematics*, 2024, 12 (6), 905, *Advances in Applied Probability and Statistical Inference*. doi: 10.3390/math12060905 <https://www.mdpi.com/2227-7390/12/6/905/pdf>
14. Withers, C.S. (2025) Edgeworth coefficients for standard multivariate estimates. *New Perspectives in Mathematical Statistics*, 2nd Edition. *Axioms*. Page 5 has 2 typos. k_1^1 in (19) should be $\bar{k}_1^1$. Delete /12 is S_3 , 2 lines before (21).
15. Withers, C.S. and Nadarajah, S. (2010) Tilted Edgeworth expansions for asymptotically normal vectors. *Annals of the Institute of Statistical Mathematics*, 62 (6), 1113–1142. DOI:10.1007/s10463-008-0206-0.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.