Behaviormetrika
Vol.39, No.2, 2012, 211–241

ASYMPTOTIC CUMULANTS OF FUNCTIONS OF MULTINOMIAL SAMPLE PROPORTIONS WITH ADJUSTMENT FOR EMPTY CELLS

Haruhiko Ogasawara*

Asymptotic cumulants of functions of multinomial sample proportions with and without studentization up to the fourth order are derived, where observed proportions are possibly added by some quantities. Some of the asymptotic cumulants of non-studentized estimators are invariant with respect to the added quantities used. On the other hand, most of the asymptotic cumulants for studentized estimators are the same as those for the estimators without the added quantities when the estimator of the asymptotic variance of the non-studentized estimator is appropriately constructed to avoid the problem of sampling zeroes or empty cells. Especially, when the quantities of order $O(1/n)$ are used, all the asymptotic cumulants of the studentized estimators up to the fourth order are the same as those for the estimators without the added quantities. A numerical example using the log odds-ratio and Yule's coefficients is illustrated.

1. Introduction

Among various statistical applications of the multinomial distribution, one of the simplest problems is the inference of a binomial proportion π. The standard Wald confidence interval (CI) for the binomial proportion based on the studentized sample proportion is easily constructed and widespread in introductory textbooks of statistics. For instance, the endpoints of the two-sided Wald CI with the asymptotic confidence level 0.95 based on the normal approximation are $p \pm 1.96\{p(1-p)/n\}^{1/2}$, where p is the sample proportion with $\mathrm{E}(p) = \pi$, and n is the sample size. On the other hand, it is known that the CI has poor properties with comparison to e.g, the score test given by Wilson (1927) (see Ghosh, 1979; Agresti & Coull, 1998; Agresti & Caffo, 2000; Brown, Cai & DasGupta, 2001, 2002; Cai, 2005). While the so-called exact test by Clopper and Pearson (1934) is available, it is only "exact" in the sense that the corresponding CI has at least the nominal coverage with poor average length of the CI (see Agresti & Coull, 1998).

For the problems of estimating general functions of multinomial proportions, the corresponding exact test may be constructed at least in principle. However, we expect the similar poor behavior not to say the excessive computation possibly required especially when the number of multinomial categories is large. In these general cases, one of the natural methods of estimation more accurate than the usual normal approximation may be to use the asymptotic expansions of the studentized estimator i.e., the Wald statistic.

Note that the normal approximation uses only the estimator $n^{-1}p(1-p)$ of the

Key Words and Phrases: multinomial distribution, sampling zeroes, studentization, asymptotic expansion, odds ratio, Yule's coefficients

* Department of Information and Management Science, Otaru University of Commerce, 3–5–21, Midori, Otaru 047–8501, Japan.
Mail Address: hogasa@res.otaru-uc.ac.jp

variance $n^{-1}\pi(1-\pi)$ of p. On the other hand, the asymptotic expansions for the distributions of p, its functions, and their studentized versions (the Wald statistics) e.g., $n^{1/2}(p-\pi)/\{p(1-p)\}^{1/2}$ use the asymptotic cumulants as well as the asymptotic variances. When the asymptotic bias and skewness are used in addition to the asymptotic variance, a typical approximation to the distributions is the single-term Edgeworth expansion. Further, the two-term Edgeworth expansion using the asymptotic kurtosis and the higher-order asymptotic variance is also used. For the review of the asymptotic expansions for the estimators associated with the multinomial distribution, see Subsection A.1 of the appendix.

While the probabilities in the binomial and multinomial distributions are of direct interest in many cases, their ratios frequently with logarithmic transformation have been used in various fields. However, the log odds or logit is undefined when a population/sample proportion is zero, which yields the non-existence of finite moments of e.g., $\ln\{p/(1-p)\}$. On the other hand, their asymptotic moments are well defined.

The above problem in samples corresponds to the case of an empty cell in the binomial distribution. Instead of using the unbiased sample proportion p, Haldane (1956) and Anscombe (1956) independently proposed that 0.5 is added to each cell frequency, which gives the modified estimator $\ln\{(p+0.5n^{-1})/(1-p+0.5n^{-1})\}$. The constant 0.5 is employed to remove the asymptotic bias of the sample logit. Hitchcock (1962) proposed 0.25 for the added constant in the case of logistic regression. Recently, Bonett and Price (2007) proposed the constant 0.1 in order to give more accurate interval estimation of a function of the multinomial proportions with 4 categories in the 2×2 table based on simulations.

The similar problem also occurs for the variance estimators of sample proportions. For instance, the usual Wald statistic $n^{1/2}(p-\pi)/\{p(1-p)\}^{1/2}$ is undefined in the case of an empty cell. Similarly, the variance estimator $\{np(1-p)\}^{-1}$ of the sample logit is undefined in the case of an empty cell as well as the estimator $\ln\{p/(1-p)\}$. However, the reciprocal of the variance estimator is always defined giving a well defined Wald statistic $\{np(1-p)\}^{1/2}[\ln\{(p+cn^{-1})/(1-p+cn^{-1})\}-\ln\{\pi/(1-\pi)\}]$ with $c=O(1)>0$. Haldane (1956) proposed $(np+1)^{-1}+\{n(1-p)+1\}^{-1}$ instead of $\{np(1-p)\}^{-1}$, which gives relatively unbiased variance estimator when c=0.5 is used. For various variance estimators in the case of the binomial distribution, see Gart and Zweifel (1967).

In this paper, the two-term local Edgeworth expansions up to $O(n^{-1})$ for the studentized estimators of functions of multinomial proportions are given as well as those for the non-studentized estimators, where the sample proportions are possibly added some quantities of orders $O(n^{-1/2})$, $O(n^{-1})$ or $O(n^{-3/2})$. Note that the case of $O(n^{-1})$ is frequently used in practice as explained above. This paper is organized as follows. In Section 2, the vector of sample and population proportions with added constants of different orders with respect to n are defined. The Bayes estimators of multinomial proportions using the Dirichlet prior are addressed. Section 3 gives the asymptotic cumulants for a non-studentized function of sample proportions up to the fourth order with the higher-order asymptotic variance. These results can be used to see the asymptotic properties of the point estimators. In Section 4, the similar results of the corresponding studentized estimators are derived, which are required for interval estimation asymptotically more accurate than that given by the usual normal approximation. In Section 5, an example will be given using the log odds-ratio, Yule's (1900, 1912) coefficients, and their variations. Section 6 gives some remarks associated with the numerical example.

2. Estimators using added constants

As mentioned in Section 1, the difficulties associated with sampling zeroes or empty cells guide us to define estimators of functions of multinomial proportions such that the sample proportions are added by some quantities possibly different from cell to cell with different orders in terms of n. Let

$$\mathbf{p}=(p_1,\ldots,p_r)' \quad \text{and} \quad \boldsymbol{\pi}=(\pi_1,\ldots,\pi_r)' \tag{2.1}$$

be the vectors of sample and population proportions for a multinomial distribution, respectively, where $\mathrm{E}(\mathbf{p})=\boldsymbol{\pi}$ and r is the number of the multinomial categories. Let

$$\mathbf{p}_n=\mathbf{p}+n^{-1/2}\mathbf{b}+n^{-1}\mathbf{c}+n^{-3/2}\mathbf{d} \quad \text{and} \quad \boldsymbol{\pi}_n=\boldsymbol{\pi}+n^{-1/2}\mathbf{b}+n^{-1}\mathbf{c}+n^{-3/2}\mathbf{d}, \tag{2.2}$$

where $\mathbf{b}$, $\mathbf{c}$ and $\mathbf{d}$ are $r\times 1$ known constant vectors of order $O(1)$. Then, $\mathbf{p}_n$ is the vector of modified sample proportions with $\boldsymbol{\pi}_n$ being its population counterpart. When the adjustment to obtain the unit sum of the modified proportions, yielding say $\mathbf{p}_n^*$, is employed, $\mathbf{p}_n^*-\mathbf{p}$ becomes stochastic rather than fixed. However, in this paper, (2.2) is adopted, which can be used in the cases of the odds and its functions irrespective of the adjustment.

Note that $\mathbf{p}_n$ and $\boldsymbol{\pi}_n$ with the constant terms of different orders are introduced for generality and comparison. In practice, mostly the constant term $n^{-1}\mathbf{c}$ seems to have been used. For instance, the case of Haldane (1956) and Anscombe (1956) mentioned earlier gives $\mathbf{p}_n=\mathbf{p}+n^{-1}\mathbf{c}$ and $\boldsymbol{\pi}_n=\boldsymbol{\pi}+n^{-1}\mathbf{c}$ with $\mathbf{c}=(0.5,0.5)'$. Agresti (2007, p.154) stated that "adding a very small constant (such as 10^{-8}) is adequate for ensuring convergence" when we have the problem of non-convergence due to infinite estimates in computation (for a similar treatment, see also Clogg & Eliason, 1987, p.13). This can be seen as adding $n^{-3/2}\mathbf{d}$ or a higher-order constant term rather than $n^{-1}\mathbf{c}$. Although the introduction of the term $n^{-1/2}\mathbf{b}$ is primarily for comparison, some of the simulated results using $n^{-1/2}\mathbf{b}$ in Section 5 are reasonable.

When some prior information is available for the multinomial proportions, the Bayesian method of inference is also used. It is known that a conjugate prior of the multinomial distribution is the Dirichlet distribution with the density proportional to $\prod_{i=1}^{r}\pi_i^{\beta_i-1}$, where $\beta_i>0$ $(i=1,\ldots,r)$ are the concentration parameters (for the distribution, see e.g., Agresti, 2002, Section 15.2.2; Press, 2005, Section 6.3). After n counts are observed, the prior parameters are updated as the posterior ones β_i+np_i $(i=1,\ldots,r)$. The β_i's used by the Bayesian method can be seen as pseudo-counts added to observed np_i's. When no information is available, the non-informative uniform prior with $\beta_i = 1$ $(i = 1, \ldots, r)$ is used, which gives the updated parameters or counts $1 + np_i$ $(i = 1, \ldots, r)$. The non-informative Jeffreys prior with the square root of the determinant of the Fisher information matrix giving $\beta_i = 0.5$ $(i = 1, \ldots, r)$ for the proportions is also often used, and shows good results in many cases (Gart, 1966; Brown et al., 2001). It is known that if we use the maximum a posterior (MAP) estimator or Bayes modal estimator with the Jeffreys prior for a parameter in a distribution of the exponential family with canonical form, the asymptotic bias of the MAP estimator becomes zero (Firth, 1993; recall the case of Haldane, 1956 and Anscombe, 1956).

When we use the expected a posteriori (EAP) estimator $(\beta_i + np_i)/(\sum_{i=1}^{r} \beta_i + n)$ for the probability of the i-th category, a formal correspondence of the numerators of these Bayes estimators and (2.2) with $\mathbf{p}_n = \mathbf{p} + n^{-1}\mathbf{c}$ and $\mathbf{c} = (\beta_1, \ldots, \beta_r)'$ is found. However, it is to be noted that the constant terms in (2.2) are assumed to be given empirically or theoretically by considering e.g., the reduction of the asymptotic bias (recall the case of $\mathbf{c} = (0.5, 0.5)'$) without using the Bayesian estimation. An advantage of using the non-informative priors as well as informative ones by the Bayes method is that the difficulties due to empty cells can be avoided (see e.g., Galindo-Garre, Vermunt & Bergsma, 2004; Galindo-Garre & Vermunt, 2006) as long as the EAP estimators for proportions are used. On the other hand, if we use the MAP estimator $(\beta_i + np_i - 1)/(\sum_{i=1}^{r} \beta_i + n - r)$, this estimator with $(\beta_i - 1)/(\sum_{i=1}^{r} \beta_i + n - r)$ for an empty cell remains to be zero when $\beta_i \leq 1$ (including the Jeffreys prior with $\beta_i = 0.5$) is adopted. Since this looks contradictory and gives an unpleasant feeling, some explanation is required. First, note that the binomial distribution with the usual parameter π is not in canonical form. So, the MAP estimator of π with the Jeffreys prior does not retain the unbiasedness of the usual sample proportion. On the other hand, the MAP estimator of the logit or $\log\{\pi/(1-\pi)\}$, which is of canonical form, has the Fisher information $\pi(1-\pi)$. In this case, the Jeffreys prior$\{\pi(1-\pi)\}^{1/2}$ gives $\beta_i = 3/2$ and the positive pseudocount of $\beta_i - 1 = 0.5$ for an empty cell (compare the Fisher information $\{\pi(1-\pi)\}^{-1}$ and $\beta_i - 1 = -0.5$ for the empty cell in the case of the usual proportion parameter π). For the review of the Bayes methods for categorical data, see Leonard and Hsu (1994) and Agresti (2002, Section 15.2.3).

Let $\theta = \theta(\cdot)$ be a parameter of interest, which is a function of multinomial proportions. Define

$$\hat{\theta} = \theta(\mathbf{p}), \quad \hat{\theta}_n = \theta(\mathbf{p}_n), \quad \theta_0 = \theta(\boldsymbol{\pi}) \quad \text{and} \quad \theta_n = \theta(\boldsymbol{\pi}_n), \tag{2.3}$$

where $\hat{\theta}$ is a special case of $\hat{\theta}_n$ when the constants are null. The function $\theta(\cdot)$ is assumed to be differentiable any number of times required with respect to the argument in a domain including $\boldsymbol{\pi}$ and $\boldsymbol{\pi}_n$. Since there is a linear dependency among the elements of $\mathbf{p}$, $\theta(\cdot)$ is seen also as a function of $r-1$ elements when necessary. In this case, r is interpreted as the number of the categories minus 1 with some adjustment for associated expressions in the remaining part of this paper.

Let

$$\mathbf{u} = n^{1/2}(\mathbf{p} - \boldsymbol{\pi}), \quad v = n^{1/2}(\hat{\theta}_n - \theta_0) \quad \text{and} \quad w = n^{1/2}(\hat{\theta}_n - \theta_n), \tag{2.4}$$

where $v = w + n^{1/2}(\theta_n - \theta_0)$. Since $\theta_n - \theta_0$ is obtained from (2.2) and (2.3) using the Taylor series expansion when necessary, w with $\mathbf{u}$ is mainly used rather than v in the following. Assume that w is expanded as

$$w = \frac{\partial \theta_n}{\partial \boldsymbol{\pi}'}\mathbf{u} + \frac{n^{-1/2}}{2}\frac{\partial^2 \theta_n}{(\partial \boldsymbol{\pi}')^{<2>}}\mathbf{u}^{<2>} + \frac{n^{-1}}{6}\frac{\partial^3 \theta_n}{(\partial \boldsymbol{\pi}')^{<3>}}\mathbf{u}^{<3>} + O_p(n^{-3/2}), \tag{2.5}$$

where

$$\frac{\partial \theta_n}{\partial \boldsymbol{\pi}'} \equiv \left.\frac{\partial \theta(\mathbf{p}_n)}{\partial \mathbf{p}'}\right|_{\mathbf{p}=\boldsymbol{\pi}} = \left.\frac{\partial \theta(\mathbf{p}_n)}{\partial \mathbf{p}_n'}\right|_{\mathbf{p}_n=\boldsymbol{\pi}_n} \left.\frac{\partial \mathbf{p}_n}{\partial \mathbf{p}'}\right|_{\mathbf{p}=\boldsymbol{\pi}} = \left.\frac{\partial \theta(\mathbf{p}_n)}{\partial \mathbf{p}_n'}\right|_{\mathbf{p}_n=\boldsymbol{\pi}_n} \tag{2.6}$$

for simplicity of notation with other expressions of partial derivatives defined similarly, and $\mathbf{u}^{<k>} = \mathbf{u}\otimes\cdots\otimes\mathbf{u}$ (k times of $\mathbf{u}$) is the k-fold Kronecker product of $\mathbf{u}$.

Expand the non-stochastic $\partial\theta_n/\partial\boldsymbol{\pi}$ by

$$\frac{\partial \theta_n}{\partial \boldsymbol{\pi}} = \frac{\partial \theta}{\partial \boldsymbol{\pi}_A} + n^{-1/2}\frac{\partial \theta}{\partial \boldsymbol{\pi}_B} + n^{-1}\frac{\partial \theta}{\partial \boldsymbol{\pi}_C} + O(n^{-3/2}), \tag{2.7}$$

where $\partial\theta/\partial\boldsymbol{\pi}_A \equiv \partial\theta(\mathbf{p})/\partial\mathbf{p}|_{\mathbf{p}=\boldsymbol{\pi}}$ i.e., the partial derivative of $\hat{\theta}$ (recall (2.3)) with respect to $\mathbf{p}$ evaluated at $\mathbf{p}=\boldsymbol{\pi}$; $n^{-1/2}\partial\theta/\partial\boldsymbol{\pi}_B$ and $n^{-1}\partial\theta/\partial\boldsymbol{\pi}_C$ are the vectors collected according to the orders in terms of n in the residual of $\partial\theta_n/\partial\boldsymbol{\pi}$ after eliminating $\partial\theta/\partial\boldsymbol{\pi}_A$. The notations using partial derivatives with subscripts A, B and C are employed for convenience and discrimination.

Using similar notations, the remaining partial derivatives in (2.5) are expanded by

$$\begin{aligned}\frac{\partial^2 \theta_n}{(\partial \boldsymbol{\pi})^{<2>}} &= \frac{\partial^2 \theta}{(\partial \boldsymbol{\pi}_A)^{<2>}} + n^{-1/2}\frac{\partial^2 \theta}{(\partial \boldsymbol{\pi}_B)^{<2>}} + O(n^{-1}),\\ \frac{\partial^3 \theta_n}{(\partial \boldsymbol{\pi})^{<3>}} &= \frac{\partial^3 \theta}{(\partial \boldsymbol{\pi}_A)^{<3>}} + O(n^{-1/2}).\end{aligned} \tag{2.8}$$

From (2.5) through (2.8), w becomes

$$\begin{aligned} w &= \frac{\partial \theta}{\partial \boldsymbol{\pi}_A'}\mathbf{u} + n^{-1/2}\left\{\frac{\partial \theta}{\partial \boldsymbol{\pi}_B'}\mathbf{u} + \frac{1}{2}\frac{\partial^2 \theta}{(\partial \boldsymbol{\pi}_A')^{<2>}}\mathbf{u}^{<2>}\right\},\\ &+ n^{-1}\left\{\frac{\partial \theta}{\partial \boldsymbol{\pi}_C'}\mathbf{u} + \frac{1}{2}\frac{\partial^2 \theta}{(\partial \boldsymbol{\pi}_B')^{<2>}}\mathbf{u}^{<2>} + \frac{1}{6}\frac{\partial^3 \theta}{(\partial \boldsymbol{\pi}_A')^{<3>}}\mathbf{u}^{<3>}\right\} + O_p(n^{-3/2}). \end{aligned} \tag{2.9}$$

As addressed earlier, θ_n is expanded about θ_0 using the above notations:

$$\begin{aligned} \theta_n &= \theta_0 + \frac{\partial \theta}{\partial \boldsymbol{\pi}_A'}(n^{-1/2}\mathbf{b} + n^{-1}\mathbf{c} + n^{-3/2}\mathbf{d}) + \frac{1}{2}\frac{\partial^2 \theta}{(\partial \boldsymbol{\pi}_A')^{<2>}}(n^{-1/2}\mathbf{b} + n^{-1}\mathbf{c})^{<2>}\\ &+ \frac{1}{6}\frac{\partial^3 \theta}{(\partial \boldsymbol{\pi}_A')^{<3>}}(n^{-1/2}\mathbf{b})^{<3>} + O(n^{-2})\\ &= \theta_0 + n^{-1/2}\frac{\partial \theta}{\partial \boldsymbol{\pi}_A'}\mathbf{b} + n^{-1}\left\{\frac{\partial \theta}{\partial \boldsymbol{\pi}_A'}\mathbf{c} + \frac{1}{2}\frac{\partial^2 \theta}{(\partial \boldsymbol{\pi}_A')^{<2>}}\mathbf{b}^{<2>}\right\}\\ &+ n^{-3/2}\left\{\frac{\partial \theta}{\partial \boldsymbol{\pi}_A'}\mathbf{d} + \frac{\partial^2 \theta}{(\partial \boldsymbol{\pi}_A')^{<2>}}(\mathbf{b}\otimes\mathbf{c}) + \frac{1}{6}\frac{\partial^3 \theta}{(\partial \boldsymbol{\pi}_A')^{<3>}}\mathbf{b}^{<3>}\right\} + O(n^{-2}) \end{aligned}$$

$$\equiv \theta_0 + n^{-1/2}\theta_{0X} + n^{-1}\theta_{0Y} + n^{-3/2}\theta_{0Z} + O(n^{-2}), \tag{2.10}$$

where θ_{0X}, θ_{0Y} and θ_{0Z} can be used to restore θ_0 from θ_n, when necessary, depending on the orders desired.

3. Asymptotic cumulants for non-studentized estimators

In this section, the asymptotic cumulants of w are derived up to the fourth order, which are denoted by α_i $(i=1,\ldots,4)$, $\alpha_{1a}, \alpha_{2a}, \alpha_{2b}$ and α_{3a}, where α_{ia} and α_{ib} stand for the higher-order terms for the i-th asymptotic cumulant with α_{ib} more higher than α_{ia}. Note that α_i, α_{ia} and α_{ib} are the terms of order $O(1)$ in the asymptotic cumulants obtained after multiplication of appropriate powers of n.

As addressed earlier, the definition of $\mathbf{p}_n$ in (2.2) was adopted for generality and comparison. In practice, however, mostly $\mathbf{p}_n = \mathbf{p} + n^{-1}\mathbf{c}$ seems to be used. Define Methods A through D depending on the constants in $\mathbf{p}_n$ used as follows:

$$\text{Method A: } \mathbf{p}_n = \mathbf{p}, \quad \text{Method B: } \mathbf{p}_n = \mathbf{p} + n^{-1/2}\mathbf{b}, \quad \text{Method C: } \mathbf{p}_n = \mathbf{p} + n^{-1}\mathbf{c},$$
$$\text{and Method D: } \mathbf{p}_n = \mathbf{p} + n^{-3/2}\mathbf{d}. \tag{3.1}$$

Note that Methods A and D give the same asymptotic expansions of the estimators up to $O(n^{-1})$.

Let $\kappa_i(\cdot)$ denote the i-th cumulant of the variable in parentheses. Then, the first cumulant of w is given from (2.9) as

$$\begin{aligned}\kappa_1(w) &= \mathrm{E}(w) = n^{-1/2}\frac{1}{2}\mathrm{tr}\left(\frac{\partial^2\theta}{\partial\boldsymbol{\pi}_A\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\right) + n^{-1}\frac{1}{2}\mathrm{tr}\left(\frac{\partial^2\theta}{\partial\boldsymbol{\pi}_B\partial\boldsymbol{\pi}'_B}\boldsymbol{\Sigma}\right) + O(n^{-3/2}) \\ &\equiv n^{-1/2}\alpha_1 + n^{-1}\alpha_{1a} + O(n^{-3/2}) = n^{-1/2}\alpha_1^{(A)} + n^{-1}\alpha_{1a} + O(n^{-3/2}),\end{aligned} \tag{3.2}$$

where $(\boldsymbol{\Sigma})_{ij} = \mathrm{E}(u_i u_j) = \delta_{ij}\pi_i - \pi_i\pi_j$ $(i,j=1,\ldots,r)$, $(\cdot)_{ij}$ is the (i,j)th element of the matrix in parentheses, δ_{ij} denotes the Kronecker delta, and $\alpha_1^{(A)}$ is α_1 given by Method A with no adjustment of counts for cells. For v in (2.4),

$$\begin{aligned}\kappa_1(v) &= \mathrm{E}(v) = \kappa_1(w) + n^{1/2}(\theta_n - \theta_0) \\ &= \theta_{0X} + n^{-1/2}(\alpha_1 + \theta_{0Y}) + n^{-1}(\alpha_{1a} + \theta_{0Z}) + O(n^{-3/2}) \\ &\equiv \theta_{0X} + n^{-1/2}\alpha_{v1} + n^{-1}\alpha_{v1a} + O(n^{-3/2}).\end{aligned} \tag{3.3}$$

The remaining cumulants up to the fourth order are as follows:

$$\begin{aligned}\kappa_2(v) = \kappa_2(w) &= \alpha_2 + n^{-1/2}\alpha_{2a} + n^{-1}\alpha_{2b} + O(n^{-3/2}) \\ &= \alpha_2^{(A)} + n^{-1/2}\alpha_{2a} + n^{-1}\alpha_{2b} + O(n^{-3/2}), \\ \kappa_3(v) = \kappa_3(w) &= n^{-1/2}\alpha_3 + n^{-1}\alpha_{3a} + O(n^{-3/2}) \\ &= n^{-1/2}\alpha_3^{(A)} + n^{-1}\alpha_{3a} + O(n^{-3/2}), \\ \kappa_4(v) = \kappa_4(w) &= n^{-1}\alpha_4 + O(n^{-3/2})\end{aligned}$$

$$= n^{-1}\alpha_4^{(A)} + O(n^{-3/2}), \tag{3.4}$$

where $\alpha_i^{(A)}$ $(i=2,3,4)$ (and similar expressions given later) are defined similarly to $\alpha_1^{(A)}$. The results of (3.4) are given from (2.9) and the cumulants of $\mathbf{u}$. The actual expressions of the results will be shown in Subsection A.2 of the appendix. From the results,

Theorem 1 *The asymptotic cumulants of w and v defined by (2.4) up to the fourth order are given from (3.2), (3.3), and (3.4) with (2.10), (A.1), (A.2) and (A.3). Among the cumulants, α_i $(i=1,\ldots,4)$ are the same over Methods A to D, and $\alpha_{1a}=\alpha_{2a}=\alpha_{3a}=0$ by Methods A, C and D. Generally, α_{1a},α_{2a} and α_{3a} by Method B are not zero. The higher-order asymptotic variance α_{2b} by Method C is generally different from α_{2b} common to Methods A and D, and α_{2b} by Method B.*

Note that α_{2b} by Method C depends on the quantity $n^{-1}\mathbf{c}$ used and that the common α_{2b} by Methods A and D does not depend on the quantity $n^{-3/2}\mathbf{d}$. The results of Theorem 1 are new in that so far only those of Method A without added constant terms are known (Ogasawara, 2009, Section 3). The asymptotic cumulants of Theorem 1 gives the asymptotic properties of the point estimators and will be used in Theorems 2 and 3. The new results of the higher-order asymptotic standard errors, which are different among Methods A (D), B and C, will be numerically illustrated in Section 5 with comparison to the usual asymptotic standard errors common to Methods A to D.

Next, the two-term local Edgeworth expansion of w and v up to order $O(n^{-1})$ are derived. Let $\mathrm{i}=\sqrt{-1}$. Then, from the definition of the cumulants, the characteristic function for w is

$$\begin{aligned}
&\exp\left\{\sum_{j=1}^{\infty}(\mathrm{i}t)^j\kappa_j(w)/j!\right\}\\
&= \exp\left(-\frac{1}{2}\alpha_2 t^2\right)\exp\left\{n^{-1/2}\alpha_1\mathrm{i}t + \frac{n^{-1/2}}{2}\alpha_{2a}(\mathrm{i}t)^2 + \frac{n^{-1/2}}{6}\alpha_3(\mathrm{i}t)^3\right.\\
&\quad\left. + n^{-1}\alpha_{1a}\mathrm{i}t + \frac{n^{-1}}{2}\alpha_{2b}(\mathrm{i}t)^2 + \frac{n^{-1}}{6}\alpha_{3a}(\mathrm{i}t)^3 + \frac{n^{-1}}{24}\alpha_4(\mathrm{i}t)^4\right\} + O(n^{-3/2})\\
&= \exp\left(-\frac{1}{2}\alpha_2 t^2\right)\left[1 + n^{-1/2}\left\{\alpha_1\mathrm{i}t + \frac{1}{2}\alpha_{2a}(\mathrm{i}t)^2 + \frac{1}{6}\alpha_3(\mathrm{i}t)^3\right\}\right.\\
&\quad + n^{-1}\left\{\alpha_{1a}\mathrm{i}t + \frac{1}{2}(\alpha_{2b}+\alpha_1^2)(\mathrm{i}t)^2 + \left(\frac{\alpha_{3a}}{6} + \frac{\alpha_1\alpha_{2a}}{2}\right)(\mathrm{i}t)^3\right.\\
&\quad\left.\left. + \left(\frac{\alpha_4}{24} + \frac{\alpha_1\alpha_3}{6} + \frac{\alpha_{2a}^2}{8}\right)(\mathrm{i}t)^4 + \frac{\alpha_{2a}\alpha_3}{12}(\mathrm{i}t)^5 + \frac{\alpha_3^2}{72}(\mathrm{i}t)^6\right\}\right] + O(n^{-3/2}).
\end{aligned} \tag{3.5}$$

Inverting (3.5) formally, the following is obtained

Theorem 2 *The approximate density of $w/\alpha_2^{1/2}$ at x using the local Edgeworth expansion is given by*

$$\begin{aligned} f\left(\frac{w}{\alpha_2^{1/2}}=x\right) &= \left[1+n^{-1/2}\left\{\frac{\alpha_1 x}{\alpha_2^{1/2}}+\frac{\alpha_{2a}}{2\alpha_2}(x^2-1)+\frac{\alpha_3}{6\alpha_2^{3/2}}(x^3-3x)\right\}\right. \\ &\quad +n^{-1}\left\{\frac{\alpha_{1a}x}{\alpha_2^{1/2}}+\frac{1}{2}(\alpha_{2b}+\alpha_1^2)\frac{x^2-1}{\alpha_2}+\left(\frac{\alpha_{3a}}{6}+\frac{\alpha_1\alpha_{2a}}{2}\right)\frac{x^3-3x}{\alpha_2^{3/2}}\right. \\ &\quad +\left(\frac{\alpha_4}{24}+\frac{\alpha_1\alpha_3}{6}+\frac{\alpha_{2a}^2}{8}\right)\frac{x^4-6x^2+3}{\alpha_2^2}+\frac{\alpha_{2a}\alpha_3}{12\alpha_2^{5/2}}(x^5-10x^3+15x) \\ &\quad \left.\left.+\frac{\alpha_3^2}{72\alpha_2^3}(x^6-15x^4+45x^2-15)\right\}\right]\phi(x)+O(n^{-3/2}). \end{aligned} \tag{3.6}$$

where $\phi(x)$ is the standard normal density at x. The distribution function is

$$\begin{aligned} \Pr\left(\frac{w}{\alpha_2^{1/2}}\le x\right) &= \Phi(x)-n^{-1/2}\left\{\frac{\alpha_1}{\alpha_2^{1/2}}+\frac{\alpha_{2a}}{2\alpha_2}x+\frac{\alpha_3}{6\alpha_2^{3/2}}(x^2-1)\right\}\phi(x) \\ &\quad -n^{-1}\left\{\frac{\alpha_{1a}}{\alpha_2^{1/2}}+\frac{1}{2}(\alpha_{2b}+\alpha_1^2)\frac{x}{\alpha_2}+\left(\frac{\alpha_{3a}}{6}+\frac{\alpha_1\alpha_{2a}}{2}\right)\frac{x^2-1}{\alpha_2^{3/2}}\right. \\ &\quad +\left(\frac{\alpha_4}{24}+\frac{\alpha_1\alpha_3}{6}+\frac{\alpha_{2a}^2}{8}\right)\frac{x^3-3x}{\alpha_2^2}+\frac{\alpha_{2a}\alpha_3}{12\alpha_2^{5/2}}(x^4-6x^2+3) \\ &\quad \left.+\frac{\alpha_3^2}{72\alpha_2^3}(x^5-10x^3+15x)\right\}\phi(x)+O(n^{-3/2}), \end{aligned} \tag{3.7}$$

where $\Phi(x)=\int_{-\infty}^{x}\phi(x^*)dx^*$. The corresponding results for v are given by replacing w,α_1 and α_{1a} in (3.6) and (3.7) with $v-\theta_{0X}$, α_{v1} and α_{v1a}, respectively.

Theorem 2 gives the approximate distribution of the estimator, which is asymptotically more accurate than that given by the normal approximation using only the first terms $O(1)$ on the right-hand sides of (3.6) and (3.7). The approximations using the terms up to $O(n^{-1/2})$ and $O(n^{-1})$ in (3.6) and (3.7) are the single- and two-term Edgeworth expansions, respectively. The results are new since so far only those by Method A are available (Ogasawara, 2009, Section 3).

4. Asymptotic cumulants for studentized estimators

The studentized estimator t using $\hat{\theta}_n$ is given from $w/\alpha_2^{1/2}$, where α_2 is replaced by its sample counterpart. Recall $\alpha_2=(\partial\theta/\partial\boldsymbol{\pi}'_A)\boldsymbol{\Sigma}\partial\theta/\partial\boldsymbol{\pi}_A$. A natural estimator of $\boldsymbol{\Sigma}$ in α_2 is given by $\hat{\boldsymbol{\Sigma}}=\mathrm{diag}(\mathbf{p})-\mathbf{pp}'$, where $\mathrm{diag}(\mathbf{p})$ is the diagonal matrix whose diagonal elements are the same as those of $\mathbf{p}$. However, $\hat{\boldsymbol{\Sigma}}=0$ occurs, which yields the non-existence of finite moments of t. A similar difficulty occurs for the estimator of $\partial\theta/\partial\boldsymbol{\pi}_A$ when $\partial\theta(\mathbf{p}^*)/\partial\mathbf{p}^*|_{\mathbf{p}^*=\mathbf{p}}$ (e.g., $\hat{\theta}=\ln\{p/(1-p)\}$) is used. Let

$$\mathbf{p}_l\equiv(\mathbf{p}+n^{-3/2}\mathbf{1}_{(r)})/(1+n^{-3/2}r)=\mathbf{p}+O_p(n^{-3/2}), \tag{4.1}$$

where $\mathbf{1}_{(r)}=(1,\ldots,1)'$ (r times of 1). Then, in order to avoid the difficulties, the following estimator $\hat{\alpha}_{2n}$ of α_2 is used in this paper:

$$\hat{\alpha}_{2n} = \left.\frac{\partial\theta(\mathbf{p}^*)}{\partial\mathbf{p}^{*\prime}}\right|_{\mathbf{p}^*=\mathbf{p}_n} \{\mathrm{diag}(\mathbf{p}_l) - \mathbf{p}_l\mathbf{p}_l'\} \left.\frac{\partial\theta(\mathbf{p}^*)}{\partial\mathbf{p}^*}\right|_{\mathbf{p}^*=\mathbf{p}_n}. \tag{4.2}$$

The vector $\mathbf{p}_l$ rather than $\mathbf{p}_n$ is used in $\hat{\mathbf{\Sigma}}$ because of its simplicity requiring no additional expansion and its similarity to the natural estimator $\mathbf{p}$ of $\boldsymbol{\pi}$. Using (4.2), define

$$t = n^{1/2}(\hat{\theta}_n - \theta_n)/\hat{\alpha}_{2n}^{1/2}. \tag{4.3}$$

Recalling (2.6) and (2.7), the population counterpart of $\hat{\alpha}_{2n}$ is defined as

$$\begin{aligned}
\alpha_{2n} &= \frac{\partial\theta_n}{\partial\boldsymbol{\pi}'}\mathbf{\Sigma}\frac{\partial\theta_n}{\partial\boldsymbol{\pi}'} + O(n^{-3/2}) \\
&= \left(\frac{\partial\theta}{\partial\boldsymbol{\pi}_A'} + n^{-1/2}\frac{\partial\theta}{\partial\boldsymbol{\pi}_B'} + n^{-1}\frac{\partial\theta}{\partial\boldsymbol{\pi}_C'}\right)\mathbf{\Sigma}\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}_A} + n^{-1/2}\frac{\partial\theta}{\partial\boldsymbol{\pi}_B} + n^{-1}\frac{\partial\theta}{\partial\boldsymbol{\pi}_C}\right) \\
&\quad + O(n^{-3/2}) \\
&= \frac{\partial\theta}{\partial\boldsymbol{\pi}_A'}\mathbf{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A} + n^{-1/2}2\frac{\partial\theta}{\partial\boldsymbol{\pi}_B'}\mathbf{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A} + n^{-1}\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}_B'}\mathbf{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_B} + 2\frac{\partial\theta}{\partial\boldsymbol{\pi}_C'}\mathbf{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A}\right) \\
&\quad + O(n^{-3/2}) \\
&\equiv \alpha_2 + n^{-1/2}\alpha_{2X} + n^{-1}\alpha_{2Y} + O(n^{-3/2}),
\end{aligned} \tag{4.4}$$

where we find that $\alpha_{2X} = \alpha_{2a}$. From the above definitions, t is expressed as

$$\begin{aligned}
t &= \mathbf{g}_{(11)}'\mathbf{u} + n^{-1/2}\{\mathbf{g}_{(21)}'\mathbf{u} + \mathbf{g}_{(22)}'\mathbf{u}^{<2>}\} \\
&\quad + n^{-1}\{\mathbf{g}_{(31)}'\mathbf{u} + \mathbf{g}_{(32)}'\mathbf{u}^{<2>} + \mathbf{g}_{(33)}'\mathbf{u}^{<3>}\} + O_p(n^{-3/2}),
\end{aligned} \tag{4.5}$$

where the vectors $\mathbf{g}_{(\cdot)}$ are of order $O(1)$ and will be given in (A.5) of the appendix. From (A.5), we find that $\partial\theta/\partial\boldsymbol{\pi}_C$ and α_{2Y} for Methods B and C are included in $\mathbf{g}_{(31)}$ of (4.5) and the other $\mathbf{g}_{(\cdot)}$'s do not include the quantities specific to Method C.

The asymptotic cumulants α_i' $(i=1,\ldots,4)$, $\alpha_{1a}', \alpha_{2a}', \alpha_{2b}'$ and α_{3a}' of t are given from (4.5) and (A.5) as

$$\kappa_1(t) = n^{-1/2}\alpha_1' + n^{-1}\alpha_{1a}' + O(n^{-3/2}), \quad \alpha_1' = \alpha_1^{(A)\prime} = \alpha_2^{-1/2}\alpha_1 - \frac{\alpha_2^{-3/2}}{2}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A'}\mathbf{\Sigma}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_A},$$

$$\alpha_{1a}' = \alpha_2^{-1/2}\alpha_{1a} - \frac{1}{2}\alpha_2^{-3/2}\alpha_{2X}\alpha_1 - \sum_{(A,B)}^{2}\frac{\alpha_2^{-3/2}}{2}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A'}\mathbf{\Sigma}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_B} + \frac{3}{4}\alpha_2^{-5/2}\alpha_{2X}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A'}\mathbf{\Sigma}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_A},$$

$$\kappa_2(t) = 1 + n^{-1}\alpha_{2b}' + O(n^{-3/2}), \quad \alpha_2' = \alpha_2^{(A)\prime} = 1, \quad \alpha_{2a}' = \alpha_{2a}^{(A)\prime} = 0, \quad \alpha_{2b}' = \alpha_{2b}^{(A)\prime},$$

$$\kappa_3(t) = n^{-1/2}\alpha_3' + n^{-1}\alpha_{3a}' + O(n^{-3/2}), \quad \alpha_3' = \alpha_3^{(A)\prime} = \alpha_2^{-3/2}\alpha_3 - 3\alpha_2^{-3/2}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A'}\mathbf{\Sigma}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_A},$$

$$\begin{aligned}
\alpha_{3a}' &= \alpha_2^{-3/2}\alpha_{3a} - 3\alpha_2^{-3/2}\sum_{(A,B)}^{2}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A'}\mathbf{\Sigma}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_B} + \alpha_2^{-5/2}\alpha_{2X}\left\{-\frac{3}{2}\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}_A'}\right)^{<3>}\boldsymbol{\sigma}^{<3>}\right. \\
&\quad \left. - \frac{9}{2}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A'}\mathbf{\Sigma}\frac{\partial^2\theta}{\partial\boldsymbol{\pi}_A\partial\boldsymbol{\pi}_A'}\mathbf{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A} + \frac{9}{2}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A'}\mathbf{\Sigma}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_A}\right\} + O(n^{-3/2}),
\end{aligned}$$

$$\kappa_4(t) = n^{-1}\alpha'_4 + O(n^{-3/2}), \quad \alpha'_4 = \alpha_4^{(A)'}. \tag{4.6}$$

The remaining actual expressions of the asymptotic cumulants and the derivations will be given in (A.6) to (A.9) of the appendix. From the above results with associated derivations in the appendix,

Theorem 3 *The asymptotic cumulants of t up to the fourth order are given by (4.6) with (A.6) through (A.9). The asymptotic cumulants α'_i $(i=1,\ldots,4)$, α'_{2a} $(=0)$ and α'_{2b} are the same over those by Methods A to D, and are invariant with respect to the added quantities in (2.2). Generally, α'_{1a} and α'_{3a} given by Method B are not zero, while those by Methods A, C and D are zero.*

It is of interest to see that α'_{2b} is the same over the four methods while α_{2b}'s were partially different from method to method, and that $\alpha'_{2a}=0$ over the four methods while generally $\alpha_{2a} \neq 0$ by Method B. These properties come from the use of $\partial\theta(\mathbf{p}^*)/\partial\mathbf{p}^*|_{\mathbf{p}^*=\mathbf{p}_n}$ in the estimator of α_2 (see (4.2)). If we use the usual one $\partial\theta(\mathbf{p}^*)/\partial\mathbf{p}^*|_{\mathbf{p}^*=\mathbf{p}}$ instead, the equality does not generally hold. The local Edgeworth expansions of the distributions of the studentized estimators are given from Theorem 2 by replacing w, α_i $(i=1,\ldots,4)$, $\alpha_{1a},\alpha_{2a},\alpha_{2b}$ and α_{3a} with t, α'_i $(i=1,\ldots,4)$, $\alpha'_{1a},\alpha'_{2a},\alpha'_{2b}$ and α'_{3a}, respectively.

The results of Theorem 3 are new in that so far only those by Method A are known (Ogasawara, 2009, 2010). For interval estimation of the parameter more accurate than that given by the normal approximation, the asymptotic cumulants derived by Theorem 3 are required. The accuracy of the asymptotic cumulants and the asymptotic expansions of the distribution of the estimator will be illustrated using simulations in the following section.

5. A numerical example

In this section, the results of the previous sections are illustrated using the log odds-ratio, Yule's coefficients and their generalization. For two possibly correlated variables with dichotomous realizations, a 2×2 contingency table is obtained, where the unbiased sample proportions in the four cells are denoted by p_{11},p_{12},p_{21} and p_{22} whose population counterparts or probabilities are $\pi_{11},\pi_{12},\pi_{21}$ and π_{22}. The subscripts 1 and 2 indicate two categories for each dichotomous variable. This situation is also described by the multinomial distribution with four categories i.e., $\mathbf{p}=(p_{11},p_{12},p_{21},p_{22})'$ and $\boldsymbol{\pi}=(\pi_{11},\pi_{12},\pi_{21},\pi_{22})'$ using the notations used earlier with some adjustment for subscripts.

Let $\hat{\omega}=p_{11}p_{22}/(p_{12}p_{21})$ be the sample odds ratio whose population counterpart is $\omega=\pi_{11}\pi_{22}/(\pi_{12}\pi_{21})$. Yule (1900, 1912) developed the following coefficients with the range $[-1,1]$:

$$\hat{Q} = \tanh\frac{\ln\hat{\omega}}{2} = \frac{\hat{\omega}-1}{\hat{\omega}+1}, \quad \hat{Y} = \tanh\frac{\ln\hat{\omega}}{4} = \frac{\hat{\omega}^{1/2}-1}{\hat{\omega}^{1/2}+1}. \tag{5.1}$$

Coefficient $\hat{Q}$ was used as an approximation to the tetrachoric correlation while $\hat{Y}$ was

Table 1: Quantities added uniformly to all cell proportions

Methods	B		C1		C2		D
Quantities	$0.1n^{-1/2}$		$0.5n^{-1}$		$0.1n^{-1}$		$n^{-3/2}$
$n=25$	0.02	$=$	0.02	$>$	0.004	$<$	0.008
$n=100$	0.01	$>$	0.005	$>$	0.001	$=$	0.001
$n=400$	0.005	$>$	0.00125	$>$	0.00025	$>$	0.000125

used as an approximation to ϕ coefficient or a sample product-moment correlation for dichotomous variables (Bonett & Price, 2007, p.444). Digby's (1983) $\hat{H}$ is defined by

$$\hat{H}=\frac{\hat{\omega}^{3/4}-1}{\hat{\omega}^{3/4}+1}, \tag{5.2}$$

which was developed in order to improve the approximation to the tetrachoric correlation. It is obvious that Yule's coefficients can be generalized to $\hat{G}^*\equiv(\hat{\omega}^{c*}-1)/(\hat{\omega}^{c*}+1)$, where c^* is usually a fixed constant while there are some cases of stochastic c^* (Bonett & Price, 2005, 2007). Digby's constant $c^*=0.75$ is close to Edwards' (1957) $c^*=\widehat{\pi}/4$, where $\widehat{\pi}\doteq 3.14159$, which was derived by Fisher's transformation $\tanh^{-1}\rho=(1/2)\ln\{(1+\rho)/(1-\rho)\}$ of the tetrachoric correlation ρ and Sheppard's Theorem (Stuart & Ort, 1994, Section 15.10). Note that Fisher's transformation of $\hat{G}^*$ yields $(c*/2)\ln\hat{\omega}$. That is, when c^* is a fixed constant, the distributions of Fisher's transformations of $\hat{Q},\hat{Y},\hat{H}$ and $\hat{G}^*$ are the same except for the known scale and that their studentized versions are exactly the same. So, in this section, we deal with only $\ln\hat{\omega}$ in addition to $\hat{Q},\hat{Y}$ and $\hat{H}$.

The problem of sampling zeroes or empty cells also occur for the coefficients and $\ln\hat{\omega}$. Note that when $p_{11}p_{22}=0$ or $p_{12}p_{21}=0$, $\hat{G}^*$ can be reasonably defined as -1 or 1, respectively while when e.g., $p_{11}=p_{12}=0$, $\hat{G}^*$ is undefined. In order to avoid the problem, Method B with $\mathbf{b}=0.1\times\mathbf{1}_{(4)}$, Method C with $\mathbf{c}=0.5\times\mathbf{1}_{(4)}$ and $\mathbf{c}=0.1\times\mathbf{1}_{(4)}$ (named as Methods C1 and C2, respectively) and Method D with $\mathbf{d}=\mathbf{1}_{(4)}$ are used for illustration with sample sizes, n=25, 100 and 400. The relative sizes of the added quantities among Methods B, C and D vary with n, which are summarized in Table 1. When $n=25$, the added quantities for Methods B and C1 are the same and the smallest is given by Method C2. When $n=100$, the added quantities for C2 and D are the same. When $n=400$, the added quantities are all different. That is, when $n=25$ the true distribution of the estimators are the same over Methods B and C1 while their asymptotic ones may be different.

Four cases of $\boldsymbol{\pi}$ or population proportions were constructed as

$$\text{Case 1: } \boldsymbol{\pi}=(0.1,0.1,0.1,0.7)',\quad \text{Case 2: } \boldsymbol{\pi}=(0.1,0.2,0.3,0.4)',$$
$$\text{Case 3: } \boldsymbol{\pi}=(0.3,0.1,0.2,0.4)' \text{ and } \text{Case 4: } \boldsymbol{\pi}=(0.4,0.1,0.1,0.4)'.$$

The population values of $\ln\omega$, Q, Y and H are shown in the last column under θ_0 in Table 2. Tables 2, 3A, 4 and 5 give the asymptotic cumulants (see Theorem 3) and the corresponding accurate values, derived by the multinomial distributions, for

Table 2: Asymptotic biases (α_1') and the corresponding accurate ones for studentized estimators

Case			$n^{1/2}\times$ accurate biases										
			$n=25$			$n=100$			$n=400$				
No		α_1'	B=C1	>C2	<D	B	>C1	>C2=D	B	>C1	>C2	>D	θ_0
	$\ln\omega$	.49	.49	.21	.34	.46	.47	.47	.47	.48	.48	.48	1.946
1	Q	2.59	4.51	8.20	6.49	2.76	2.94	3.11	2.55	2.66	2.69	2.70	.750
	H	1.80	2.52	3.09	2.99	1.83	1.93	2.02	1.75	1.82	1.84	1.85	.623
	Y	1.12	1.34	1.20	1.35	1.10	1.15	1.19	1.08	1.12	1.14	1.14	.451
	$\ln\omega$	−.09	−.18	.18	.03	−.11	−.09	−.06	−.10	−.09	−.08	−.08	−.405
2	Q	−.54	−.97	−1.14	−1.11	−.60	−.63	−.65	−.54	−.56	−.56	−.57	−.200
	H	−.35	−.57	−.39	−.50	−.38	−.38	−.38	−.35	−.35	−.35	−.35	−.151
	Y	−.20	−.34	−.03	−.18	−.23	−.22	−.20	−.21	−.21	−.20	−.20	−.101
	$\ln\omega$	.44	.50	.08	.25	.44	.43	.41	.44	.43	.43	.43	1.792
3	Q	2.07	3.38	4.32	4.03	2.20	2.29	2.37	2.06	2.12	2.13	2.13	.714
	H	1.44	2.00	1.96	2.07	1.48	1.53	1.57	1.42	1.46	1.47	1.47	.586
	Y	.92	1.14	.81	.99	.93	.94	.95	.91	.92	.93	.93	.420
	$\ln\omega$	.75	.82	.15	.44	.73	.72	.70	.73	.74	.74	.74	2.773
4	Q	2.96	5.37	9.04	7.30	3.22	3.38	3.53	2.94	3.04	3.07	3.07	.882
	H	2.21	3.24	3.58	3.59	2.29	2.39	2.47	2.17	2.24	2.26	2.27	.778
	Y	1.50	1.88	1.42	1.71	1.50	1.54	1.57	1.46	1.51	1.52	1.52	.600

Note. By Methods B ($0.1n^{-1/2}$), C1 ($0.5n^{-1}$), C2 ($0.1n^{-1}$) or D ($n^{-3/2}$), the quantity in parentheses is added uniformly to all cell proportions. The equality/inequality signs indicate the relative sizes of these values when n is given.

Table 3A: Asymptotic and accurate standard errors for studentized estimators

Case		n=25				n=100				n=400				
			SE*				SE*				SE*			
No		HASE*	B=C1	>C2	<D	HASE*	B	>C1	>C2=D	HASE*	B	>C1	>C2	>D
	$\ln\omega$	.977	1.048	.775	.876	.994	1.002	.994	.986	.999	.999	.998	.998	.998
1	Q	1.666	3.264	9.948	5.779	1.202	1.277	1.313	1.348	1.054	1.055	1.058	1.059	1.060
	H	1.389	1.890	2.504	2.198	1.110	1.136	1.144	1.148	1.029	1.029	1.030	1.030	1.030
	Y	1.168	1.323	1.088	1.210	1.044	1.055	1.052	1.047	1.011	1.012	1.011	1.011	1.011
	$\ln\omega$	.972	1.023	.884	.928	.993	.999	.994	.989	.998	.999	.998	.998	.998
2	Q	1.347	1.788	1.958	1.922	1.097	1.119	1.125	1.130	1.025	1.026	1.026	1.026	1.026
	H	1.196	1.369	1.259	1.321	1.052	1.063	1.062	1.060	1.013	1.014	1.014	1.014	1.014
	Y	1.076	1.155	1.008	1.066	1.020	1.026	1.023	1.019	1.005	1.005	1.005	1.005	1.005
	$\ln\omega$	.982	1.026	.895	.937	.995	1.001	.996	.991	.999	1.000	.999	.999	.999
3	Q	1.483	2.483	4.997	3.448	1.140	1.182	1.196	1.207	1.037	1.038	1.039	1.039	1.039
	H	1.276	1.595	1.630	1.634	1.076	1.091	1.093	1.093	1.019	1.020	1.020	1.020	1.020
	Y	1.116	1.219	1.056	1.130	1.030	1.037	1.035	1.031	1.008	1.008	1.008	1.008	1.008
	$\ln\omega$	.989	1.041	.848	.915	.997	1.003	.997	.991	.999	1.000	.999	.999	.999
4	Q	1.662	3.318	9.819	5.741	1.200	1.285	1.320	1.351	1.054	1.055	1.058	1.059	1.059
	H	1.405	1.950	2.601	2.268	1.115	1.145	1.153	1.156	1.030	1.030	1.031	1.032	1.032
	Y	1.186	1.346	1.145	1.252	1.050	1.059	1.058	1.054	1.013	1.013	1.013	1.013	1.013

Note. B, C1, C2 and D: as before. HASE*=$(1+n^{-1}\alpha_{2b}')^{1/2}$; SE*: accurate standard error.

Table 3B: Asymptotic and accurate ratios of standard errors for non-studentized estimators

		n = 100						
		B		C1		C2	D	C2 D
Case No		$\frac{\text{HASE}}{\text{ASE}}$	$\frac{\text{SE}}{\text{ASE}}$	$\frac{\text{HASE}}{\text{ASE}}$	$\frac{\text{SE}}{\text{ASE}}$	$\frac{\text{HASE}}{\text{ASE}}$	$\frac{\text{HASE}}{\text{ASE}}$	$\frac{\text{SE}}{\text{ASE}}$
	$\ln\omega$	.985	.969	1.019	1.023	1.056	1.065	1.074
1	Q	1.077	1.058	1.090	1.092	1.101	1.104	1.124
	H	1.006	.993	1.028	1.030	1.050	1.055	1.064
	Y	.979	.967	1.008	1.011	1.038	1.046	1.050
	$\ln\omega$	.986	.977	1.012	1.015	1.039	1.045	1.049
2	Q	.936	.936	.961	.964	.985	.992	.987
	H	.956	.951	.981	.982	1.007	1.013	1.009
	Y	.972	.964	.998	.999	1.024	1.031	1.029
	$\ln\omega$	.986	.977	1.012	1.015	1.039	1.045	1.049
3	Q	1.026	1.021	1.029	1.025	1.030	1.030	1.029
	H	.988	.985	1.000	1.000	1.011	1.014	1.012
	Y	.977	.972	.997	.998	1.017	1.022	1.020
	$\ln\omega$	.985	.972	1.016	1.020	1.049	1.057	1.065
4	Q	1.098	1.096	1.079	1.078	1.061	1.057	1.062
	H	1.021	1.018	1.022	1.021	1.021	1.021	1.022
	Y	.983	.979	1.001	1.001	1.016	1.020	1.020

Note. B, C1, C2 and D: as before. HASE = $(n^{-1}\alpha_2 + n^{-3/2}\alpha_{2a} + n^{-2}\alpha_{2b})^{1/2}$; ASE = $(n^{-1}\alpha_2)^{1/2}$; SE: accurate standard error.

Table 4: Asymptotic third cumulants (α_3') and the corresponding accurate ones for studentized estimators

			$n^{1/2}\times$ accurate third cumulants									
Case			n = 25			n = 100			n = 400			
No		α_3'	B=C1	>C2	<D	B	>C1	>C2=D	B	>C1	>C2	>D
	$\ln\omega$	−.56	−.14	−.76	−.69	−.34	−.47	−.60	−.47	−.54	−.56	−.56
1	Q	12.06	1.2e3	8.1e4	1.2e4	43.28	63.78	161	14.67	15.48	15.72	15.75
	H	7.30	109	700	295	14.43	16.47	16.59	8.03	8.41	8.51	8.53
	Y	3.24	13.18	8.99	11.44	4.57	4.69	4.51	3.35	3.47	3.50	3.50
	$\ln\omega$	.74	−.01	1.13	1.00	.39	.67	.93	.59	.73	.78	.78
2	Q	−2.00	−34	−21	−43	−6.34	−7.78	−8.68	−2.38	−2.43	−2.44	−2.44
	H	−.81	−7.22	−1.53	−4.23	−2.16	−2.07	−1.64	−.99	−.94	−.93	−.93
	Y	.05	−1.63	1.20	.42	−.50	−.23	.09	−.09	.02	.05	.05
	$\ln\omega$	−1.13	−.25	−1.81	−1.63	−.74	−1.06	−1.34	−.95	−1.12	−1.17	−1.18
3	Q	8.65	612	3.1e4	4.7e3	23.83	29.25	33.43	10.13	10.42	10.51	10.52
	H	4.89	57	183	99	8.91	9.26	8.71	5.38	5.47	5.49	5.49
	Y	1.74	8.05	2.04	4.26	2.77	2.53	2.19	1.92	1.86	1.85	1.85
	$\ln\omega$	−1.50	−.36	−2.42	−2.16	−.95	−1.38	−1.78	−1.26	−1.48	−1.55	−1.55
4	Q	11.74	1.1e3	6.5e4	9.8e3	47.53	68.75	170	14.60	15.17	15.33	15.35
	H	7.25	121	719	306	16.33	18.13	17.44	8.25	8.47	8.54	8.54
	Y	3.00	15.80	9.27	12.19	5.07	4.89	4.34	3.29	3.28	3.27	3.27

Note. B, C1, C2 and D: as before. $x\text{e}y = x \times 10^y$.

Table 5: Asymptotic fourth cumulants (α_4') and the corresponding accurate ones for studentized estimators

Case			$n^{1/2}\times$ accurate fourth cumulants									
			$n=25$			$n=100$			$n=400$			
No		α_4'	B=C1	>C2	<D	B	>C1	>C2=D	B	>C1	>C2	>D
	$\ln\omega$	−16.4	−5.1	−4.9	−10.8	−4.8	−13.0	−19.9	−12.8	−16.2	−17.2	−17.3
1	Q	307	2.2e5	8.0e7	5.8e6	5.4e3	2.0e4	2.4e6	426	467	479	481
	H	128	6.8e3	1.2e5	3.4e4	621	921	1.0e3	152	163	166	166
	Y	30.0	272	242	254	89.1	85.8	53.9	34.7	34.4	34.3	34.3
	$\ln\omega$	−13.0	−4.3	−5.6	−8.8	−4.9	−10.3	−14.0	−10.3	−12.6	−13.3	−13.4
2	Q	76.3	8.4e3	1.5e6	1.1e5	564	953	1.2e3	92.4	95.3	96.0	96.1
	H	34.8	435	580	422	120	123	78	40.0	39.6	39.5	39.4
	Y	7.4	47.9	4.2	10.6	27.4	18.8	8.2	9.9	8.2	7.7	7.6
	$\ln\omega$	−8.6	−3.3	−1.7	−5.1	−2.3	−6.7	−9.5	−6.9	−8.4	−8.9	−8.9
3	Q	180	1.7e5	5.4e7	3.9e6	2.1e3	4.0e3	5.9e3	234	246	249	249
	H	73.0	3.8e3	4.7e4	1.4e4	322	367	238	85.7	87.7	88.2	88.3
	Y	16.8	147	60.1	86.6	53.5	42.4	23.2	19.9	18.6	18.2	18.1
	$\ln\omega$	−6.0	−3.5	3.2	−1.7	.7	−4.4	−7.7	−4.9	−6.1	−6.4	−6.4
4	Q	288	1.7e5	4.8e7	3.7e6	6.7e3	2.2e4	2.5e6	412	443	452	453
	H	126	6.9e3	1.1e5	3.1e4	807	1.1e3	1.2e3	156	164	166	167
	Y	32.5	306	282	286	115	106	61	37.8	37.5	37.4	37.4

Note. B, C1, C2 and D: as before. $xey=x\times10^y$.

studentized estimators of $\ln\omega$, Q,H and Y. Table 3B is for the non-studentized estimators corresponding to Table 3A (see Theorem 1). The partial derivatives required for computation (see (2.7) and (2.8)) will be given in Subsection A.4 of the appendix.

Table 2 shows the results of bias. Note that the accurate bias for a studentized estimator is its expectation. When n=25, the accurate absolute bias for Q by Method C2 is much larger than the corresponding asymptotic value. But, this is not necessarily so for other coefficients. The tendency of the true values approaching the corresponding asymptotic ones as monotonic functions of an added quantity is observed in the table. The effect of Fisher's transformation is well seen for $\ln\hat{\omega}$ in the reduction of the biases.

Table 3A gives the common higher-order asymptotic standard errors (HASE* $=(1+n^{-1}\alpha_{2b}')^{1/2}$) and the corresponding accurate standard errors (SE*). Note that the usual asymptotic standard error (ASE*) is 1. When n = 25, the results show that many of the SE*s are far from their unit ASE*s while they are partially explained by the HASE*s. The decreasing tendency of the discrepancies by increasing the added quantity is observed. Method B was employed mainly for comparison. However, in Table 3A, Method B seems to give reasonable results in that many of the associated SE*s are relatively closer to the corresponding HASE*s. Of course, exactly the same results can also be obtained by increasing the constant for Method C.

While the values of HASE* in Table 3A are common to all the methods and are invariant with respect to the constants used, they are different for non-studentized estimators. Table 3B shows the HASEs ($=(n^{-1}\alpha_2+n^{-3/2}\alpha_{2a}++n^{-2}\alpha_{2b})^{1/2}$) (see

Table 6: $10^5\times$ root mean square errors of approximate distribution functions for studentized estimators ($n = 100$)

Case		B			C1			C2=D		
No		N*	E1	E2	N*	E1	E2	N*	E1	E2
	$\ln\omega$	952	218	162	984	206	157	1013	227	175
1	Q	3849	1553	1316	4057	1606	770	4237	1672	756
	H	2685	932	947	2826	943	355	2946	966	364
	Y	1709	491	274	1780	465	188	1841	445	190
	$\ln\omega$	312	129	91	331	118	77	362	143	92
2	Q	1173	861	191	1215	885	193	1249	905	197
	H	746	509	254	753	505	107	758	499	99
	Y	457	261	813	451	233	84	442	208	74
	$\ln\omega$	974	157	98	1003	134	97	1032	149	111
3	Q	3157	1124	437	3273	1158	465	3370	1196	465
	H	2226	679	250	2296	686	223	2356	697	222
	Y	1486	360	878	1520	339	118	1545	322	104
	$\ln\omega$	1595	474	427	1651	440	423	1695	435	419
4	Q	4481	1671	1032	4668	1736	882	4833	1811	883
	H	3377	1103	565	3512	1124	575	3626	1157	574
	Y	2395	706	1497	2474	684	457	2541	671	439

Note. B, C1, C2 and D: as before. N*: normal approximation; E1: the single-term Edgeworth expansion; E2: the two-term Edgeworth expansion.

Theorem 1) and the actual standard errors (SEs) as the ratios to the corresponding ASEs ($= (n^{-1}\alpha_2)^{1/2}$) when $n = 100$. It is found that the HASEs well explain the differences of the ratios SE/ASE from 1, and that the ratios HASE/ASE and SE/ASE by Method B tend to be smaller than the remaining ones but are not necessarily closer to 1.

Table 4 shows the results of skewness, whose relative sizes seem to parallel those of biases in Table 2. That is, when the added quantity is larger, the absolute skewness tends to be smaller and closer to the corresponding asymptotic value. The normalizing effect in the reduction of skewness for $\ln\hat{\omega}$ is also observed. Table 5 gives the results of kurtosis. The monotone tendency with respect to the added quantity is also found.

As an application of the asymptotic cumulants of the studentized estimator, the approximate cumulative distribution functions are constructed by three methods: the usual normal approximation (N*), the single-term Edgeworth expansion up to order $O(n^{-1/2})$ (E1), and the two-term Edgeworth expansion up to order $O(n^{-1})$ (E2) (for E1 and E2, see Theorem 2 and the associated description). Table 6 shows the $10^5\times$ root mean square errors of the approximations. An error is defined by the difference of an approximate value minus the corresponding accurate one. They are evaluated at the equally spaced values from -3.99 to 4.00 by steps 0.01. The square roots of the means of the squared errors are given in the table. From the table, it is seen that the errors by E1 are much smaller than N*. While the errors by E2 are smaller than those by E1 using Methods C1, C2 and D, some of the errors by E2 are larger than those by E1 when Method B is used indicating over correction by E2. Among the four parameters, overall the errors for Q are largest while those for lnω are smallest.

6. Some remarks

As mentioned after Theorem 3 for studentized estimators, the equalities of all the asymptotic cumulants considered in this paper by Methods A, C and D are due to the use of $\mathbf{p}_n$ in the estimator $\hat{\alpha}_{2n}$. The equalities give tractable results in that we can use the same asymptotic expansion as that using $\mathbf{p}$ rather than $\mathbf{p}_n$, and that the sizes of the added quantities are irrelevant to the asymptotic results. The results by Method B in Tables 1 though 5 look promising. However, as explained earlier, the seemingly reasonable results come mainly from the relatively large constants used. When we look at Table 1 again, we see that when n=25, the values of the added quantities by Methods B and C1 are equal, and that when $n = 100$, $0.01 = 0.1n^{-1/2}$ by Method B is two times of $0.005 = 0.5n^{-1}$ by Method C1. The value 0.01 by Method B is also obtained by increasing the constant as $1.0n^{-1}$ by Method C, which is equivalent to a single additional count to all cell frequencies.

Though Bonett and Price (2007) recommended Method C2 with the constant 0.1, the results of the numerical example are not consistent with their proposal though the numerical results in this paper are limited. Agresti and Coull (1998) proposed the formula of two successes and two failures added in the cell frequencies for inference of the binomial probability. Their formula corresponds to $2n^{-1}$ by Method C in our notation when the odds ratio is used. Though the formula was developed when the asymptotic coverage is 95% with the normal deviate $z_{0.025} = 1.96$ whose rounded value is 2, the formula may be used in other situations. As illustrated above, $2n^{-1}$ by Method C is equal to $0.1n^{-1/2}$ by Method B when $n = 400$.

Considering the smallest errors for lnω among the four parameters in Table 6, ln$\hat{\omega}$ (or ln$\hat{\omega}_n$) is to be used for estimation of Q, H or Y, which is recommended by e.g., Bonett and Price (2007). For instance, the two-sided CI for lnω by Methods C and D (and lnω_n by Methods B, C and D) with the asymptotic confidence level 0.95 using the Cornish-Fisher expansion (see Hall, 1992b; Ogasawara, 2012) with the estimated cumulants $\hat{\alpha}'_1$ and $\hat{\alpha}'_3$ is given by

$$\ln\hat{\omega}_n + [\pm z_{0.025} - n^{-1/2}\{\hat{\alpha}'_1 + (\hat{\alpha}'_3/6)(z^2_{0.025} - 1)\}]n^{-1/2}\hat{\alpha}^{1/2}_{2n} \equiv \hat{\omega}* \pm z* \qquad (6.1)$$

The corresponding intervals for Q, H and Y are given by $\tanh\{(c*/2)(\omega * \pm z*)\}$ with $c^* = 1$, 0.75 and 0.5, respectively.

Acknowledgements

This work was partially supported by a Grant-in-Aid for Scientific Research from the Japanese Ministry of Education, Culture, Sports, Science and Technology No.23500341.

Appendix

A.1 A review of the asymptotic expansions associated with the multinomial distribution

In the simple case of the binomial proportion, Hall (1982) used the asymptotic expansion of the studentized proportion up to $O(n^{-1/2})$ or up to the term next beyond the normal approximation. Zhou, Tsao and Qin (2004) obtained a similar but extended result for estimating the difference of the probabilities of two independent binomial distributions. Zhou and Qin (2005) gave the corresponding results when the two binomial distributions are correlated. This problem reduces to that of estimating the contrast $\pi_{11}+\pi_{12}-(\pi_{11}+\pi_{21})=\pi_{12}-\pi_{21}$ in the multinomial distribution with four categories, where π_{ij} is the probability corresponding to the $(i,\ j)$th cell in a 2×2 contingency table. Further, Zhou, Li and Yang (2008) derived the asymptotic expansion of the distribution of the sample logit or log odds in a binomial distribution. Zhou et al. (2004, 2005, 2008) used the cubic transformation by Hall (1992a). It is to be noted that Hall's expansion is asymptotically equal to the single-term local Edgeworth expansion up to $O(n^{-1/2})$ i.e., the expansion when the sample proportions are regarded as continuous variables or when the oscillatory part of the discrete distribution is neglected. For the total expansion, see Esséen (1945, Theorem 3), Ranga Rao (1961, Theorem 4), Bikyalis (1961, Theorem 1), Yarnold (1972, Section 1), Siotani and Fujikoshi (1984, Lemma 2.1). Bhattacharya and Ranga Rao (1986, Theorem 23.1), Brown et al. (2002, Lemma 1), Götze and Ulyanov (2003, Section 1) and Staicu (2009, Appendix A).

The two-term local Edgeworth expansion up to $O(n^{-1})$ for the distribution of explicit/implicit functions of sample multinomial proportions without studentization and the corresponding single-term expansion up to $O(n^{-1/2})$ for studentized estimators are given by Ogasawara (2009). Partial justification of using the local asymptotic expansion or using only the continuous part is the relatively small property of the discrete part in the sense that the part becomes $O(n^{-3/2})$ in the case of the binomial proportion when the discrete term is integrated over its domain (see Brown et al, 2002, Theorem 6; see also Hall, 1982 for the smallness of the discrete part). The author conjectures that this property more or less conveys to the general cases of functions of multinomial proportions.

It is known that the local Edgeworth expansion becomes asymptotically equivalent to the corresponding total one when the continuity correction is employed with Sheppard's correction for cumulants (Feller, 1971, Theorem 2, p.540; Kolassa & McCullagh, 1990; Kolassa, 2006, Section 3.15). In the case of the single-term Edgeworth expansion, Sheppard's correction for associated cumulants gives unchanged ones. For the two-term Edgeworth expansion, only the variances are influenced by the correction with unchanged covariances and other cumulants (Wold, 1934, Equation (5); Kolassa, 1989, Chapter 3; Stuart & Ort, 1994, Section 3.25). It is to be noted that the correction of variances is to decrease it by the quantity $-1/(12n^2)$ rather than to increase it in the case of the binomial distribution (see e.g., Kolassa & McCullagh, 1990, p.982;

Stuart & Ort, 1994, Section 3.18). So, it is expected that using estimated variances without using Sheppard's correction typically give conservative results in e.g., interval estimation of parameters.

A.2 The asymptotic cumulants for non-studentized estimators

The second cumulants are

$$\begin{aligned}
\kappa_2(v) &= \kappa_2(w) = \mathrm{E}(w^2) - \{\mathrm{E}(w)\}^2 \\
&= \frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A} + n^{-1/2}2\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_B} + n^{-1}\left[\left\{\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\otimes\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_A)^{<2>}}\right\}n^{1/2}\mathrm{E}(\mathbf{u}^{<3>})\right. \\
&\quad + \frac{\partial\theta}{\partial\boldsymbol{\pi}'_B}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_B} + \frac{1}{4}\left\{\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_A)^{<2>}}\right\}^{<2>}\mathrm{E}(\mathbf{u}^{<4>}) - \alpha_1^2 \\
&\quad \left. + 2\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_C} + \frac{1}{3}\left\{\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\otimes\frac{\partial^3\theta}{(\partial\boldsymbol{\pi}'_A)^{<3>}}\right\}\mathrm{E}(\mathbf{u}^{<4>})\right] + O(n^{-3/2}) \\
&= \frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A} + n^{-1/2}2\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_B} + n^{-1}\left[\left\{\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\otimes\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_A)^{<2>}}\right\}\boldsymbol{\sigma}^{(3)}\right. \\
&\quad + \frac{\partial\theta}{\partial\boldsymbol{\pi}'_B}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_B} + \frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}\frac{\partial^2\theta}{\partial\boldsymbol{\pi}_A\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial^2\theta}{\partial\boldsymbol{\pi}_A\partial\boldsymbol{\pi}'_A}\right) \\
&\quad \left. + 2\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_C} + \left\{\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\otimes\frac{\partial^3\theta}{(\partial\boldsymbol{\pi}'_A)^{<3>}}\right\}(\boldsymbol{\sigma}^{(2)})^{<2>}\right] + O(n^{-3/2}) \\
&\equiv \alpha_2 + n^{-1/2}\alpha_{2a} + n^{-1}\alpha_{2b} + O(n^{-3/2}), \qquad \text{(A.1)}
\end{aligned}$$

where $\boldsymbol{\sigma}^{(2)} = \mathrm{vec}(\boldsymbol{\Sigma})$, $\mathrm{vec}(\cdot)$ is a vectorizing operator stacking the columns of a matrix, $\boldsymbol{\sigma}^{(3)} \equiv n^{1/2}\mathrm{E}(\mathbf{u}^{<3>})$ with

$$\begin{aligned}
(\boldsymbol{\sigma}^{(3)})_{(ijk)} &= n^2\mathrm{E}\{(p_i-\pi_i)(p_j-\pi_j)(p_k-\pi_k)\} \\
&= \delta_{ij}\delta_{ik}(\pi_i - 3\pi_i^2) - \{\delta_{ij}(1-\delta_{ik})\pi_i\pi_k + \delta_{ik}(1-\delta_{ij})\pi_i\pi_j + \delta_{jk}(1-\delta_{ji})\pi_j\pi_i\} \\
&\quad + 2\pi_i\pi_j\pi_k \quad (i,j,k=1,\ldots,r),
\end{aligned}$$

and $(\cdot)_{(ijk)}$ denotes an element of the vector in parentheses corresponding to π_i, π_j and π_k (see e.g., Stuart & Ort, 1994, Equation (7.18)).

The third cumulants are

$$\begin{aligned}
\kappa_3(v) &= \kappa_3(w) = \mathrm{E}[\{w-\mathrm{E}(w)\}^3] = \mathrm{E}(w^3) - 3\mathrm{E}(w^2)\mathrm{E}(w) + O(n^{-3/2}) \\
&= n^{-1/2}\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<3>}\boldsymbol{\sigma}^{(3)} + n^{-1/2}\frac{3}{2}\left\{\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<2>}\otimes\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_A)^{<2>}}\right\}\mathrm{E}(\mathbf{u}^{<4>}) \\
&\quad + n^{-1}3\left\{\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<2>}\otimes\frac{\partial\theta}{\partial\boldsymbol{\pi}'_B}\right\}\boldsymbol{\sigma}^{(3)} \\
&\quad + n^{-1}\frac{3}{2}\left\{\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<2>}\otimes\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_B)^{<2>}}\right\}\mathrm{E}(\mathbf{u}^{<4>})
\end{aligned}$$

$$
\begin{aligned}
&+n^{-1}3\left\{\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\otimes\frac{\partial\theta}{\partial\boldsymbol{\pi}'_B}\otimes\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_A)^{<2>}}\right\}\mathrm{E}(\mathbf{u}^{<4>})\\
&-3(\alpha_2+n^{-1/2}\alpha_{2a})(n^{-1/2}\alpha_1+n^{-1}\alpha_{1a})+O(n^{-3/2})\\
=\ & n^{-1/2}\left\{\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<3>}\boldsymbol{\sigma}^{(3)}+3\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial^2\theta}{\partial\boldsymbol{\pi}_A\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A}\right\}\\
&+n^{-1}\left[3\left\{\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<2>}\otimes\frac{\partial\theta}{\partial\boldsymbol{\pi}'_B}\right\}\boldsymbol{\sigma}^{(3)}+3\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial^2\theta}{\partial\boldsymbol{\pi}_B\partial\boldsymbol{\pi}'_B}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A}\right.\\
&\left.+\,6\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial^2\theta}{\partial\boldsymbol{\pi}_A\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_B}\right]+O(n^{-3/2})\\
\equiv\ & n^{-1/2}\alpha_3+n^{-1}\alpha_{3a}+O(n^{-3/2}).
\end{aligned}
\tag{A.2}
$$

The fourth cumulants become

$$
\begin{aligned}
\kappa_4(v)\ =\ & \kappa_4(w)=\mathrm{E}[\{w-\mathrm{E}(w)\}^4]-3\kappa_2(w)^2\\
=\ & \mathrm{E}(w^4)-4\mathrm{E}(w^3)\mathrm{E}(w)+6\mathrm{E}(w^2)\mathrm{E}(w)^2-3\kappa_2(w)^2+O(n^{-2})\\
=\ & \mathrm{E}(w^4)-4\{n^{-1/2}\alpha_3+n^{-1}\alpha_{3a}\\
&+3(\alpha_2+n^{-1/2}\alpha_{2a})(n^{-1/2}\alpha_1+n^{-1}\alpha_{1a})\}n^{-1/2}\alpha_1\\
&+6(\alpha_2+n^{-1/2}\alpha_{2a})n^{-1}\alpha_1^2-3(\alpha_2+n^{-1/2}\alpha_{2a}+n^{-1}\alpha_{2b})^2+O(n^{-3/2})\\
=\ & \mathrm{E}(w^4)-3\alpha_2^2-n^{-1/2}6\alpha_2\alpha_{2a}-n^{-1}(4\alpha_1\alpha_3+6\alpha_2\alpha_1^2+6\alpha_2\alpha_{2b}+3\alpha_{2a}^2)\\
&+O(n^{-3/2}),
\end{aligned}
$$

where

$$
\begin{aligned}
\mathrm{E}(w^4)\ =\ & \left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<4>}\mathrm{E}(\mathbf{u}^{<4>})+n^{-1/2}4\left\{\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<3>}\otimes\frac{\partial\theta}{\partial\boldsymbol{\pi}'_B}\right\}\mathrm{E}(\mathbf{u}^{<4>})\\
&+n^{-1}2\left\{\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<3>}\otimes\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_A)^{<2>}}\right\}n^{1/2}\mathrm{E}(\mathbf{u}^{<5>})\\
&+n^{-1}\binom{4}{2}\left[\left\{\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<2>}\otimes\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_B}\right)^{<2>}\right\}\mathrm{E}(\mathbf{u}^{<4>})\right.\\
&\left.+\frac{1}{4}\left\{\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<2>}\otimes\left(\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_A)^{<2>}}\right)^{<2>}\right\}\mathrm{E}(\mathbf{u}^{<6>})\right]\\
&+n^{-1}4\left[\left\{\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<3>}\otimes\frac{\partial\theta}{\partial\boldsymbol{\pi}'_C}\right\}\mathrm{E}(\mathbf{u}^{<4>})\right.\\
&\left.+\frac{1}{6}\left\{\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<3>}\otimes\frac{\partial^3\theta}{(\partial\boldsymbol{\pi}'_A)^{<3>}}\right\}\mathrm{E}(\mathbf{u}^{<6>})\right]+O(n^{-3/2})\\
=\ & 3\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<4>}(\boldsymbol{\sigma}^{(2)})^{<2>}+n^{-1}\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<4>}n^3\boldsymbol{\kappa}^{(4)}(\mathbf{p})\\
&+n^{-1/2}12\left\{\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<3>}\otimes\frac{\partial\theta}{\partial\boldsymbol{\pi}'_B}\right\}
\end{aligned}
$$

$$\times(\boldsymbol{\sigma}^{(2)})^{<2>}+n^{-1}2\left\{\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<3>}\otimes\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_A)^{<2>}}\right\}\sum^{10}\boldsymbol{\sigma}^{(2)}\otimes\boldsymbol{\sigma}^{(3)}$$
$$+n^{-1}\left[6\alpha_2\frac{\partial\theta}{\partial\boldsymbol{\pi}'_B}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_B}+12\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_B}\right)^2\right.$$
$$\left.+\frac{3}{2}\left\{\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<2>}\otimes\left(\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_A)^{<2>}}\right)^{<2>}\right\}\sum^{15}(\boldsymbol{\sigma}^{(2)})^{<3>}\right]$$
$$+n^{-1}\left[12\alpha_2\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_C}+\frac{2}{3}\left\{\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<3>}\otimes\frac{\partial^3\theta}{(\partial\boldsymbol{\pi}'_A)^{<3>}}\right\}\sum^{15}(\boldsymbol{\sigma}^{(2)})^{<3>}\right]$$
$$+O(n^{-3/2}),$$

where $\sum^k$ denotes the sum of k similar terms considering the permutation and combination of the multivariate moments concerned, and $\boldsymbol{\kappa}^{(4)}(\mathbf{p})$ is the $n^4\times 1$ vector of the multivariate fourth cumulants of $\mathbf{p}$, whose elements corresponding to p_i,p_j,p_k and p_l are

$$n^3\kappa_4(p_i,p_i,p_i,p_i)=\pi_i(1-\pi_i)\{1-6\pi_i(1-\pi_i)\},$$
$$n^3\kappa_4(p_i,p_i,p_i,p_j)=-\pi_i\pi_j\{1-6\pi_i(1-\pi_i)\},$$
$$n^3\kappa_4(p_i,p_i,p_j,p_j)=-\pi_i\pi_j\{(1-2\pi_i)(1-2\pi_j)+2\pi_i\pi_j\},$$
$$n^3\kappa_4(p_i,p_i,p_j,p_k)=2\pi_i\pi_j\pi_k(1-3\pi_i),\quad n^3\kappa_4(p_i,p_j,p_k,p_l)=-6\pi_i\pi_j\pi_k\pi_l,$$

$(i,j,k,l=1,\ldots,r;\ i\neq j,\ i\neq k,\ i\neq l,\ j\neq k,\ j\neq l,\ k\neq l$; see e.g., Stuart & Ort, 1994, Equation (7.18)). Noting $\alpha_2=(\partial\theta/\partial\boldsymbol{\pi}'_A)^{<2>}\boldsymbol{\sigma}^{(2)}$ and $\alpha_{2a}=2(\partial\theta/\partial\boldsymbol{\pi}'_A)\boldsymbol{\Sigma}\partial\theta/\partial\boldsymbol{\pi}_B$, it follows that

$$\kappa_4(w)=n^{-1}\left[\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<4>}n^3\boldsymbol{\kappa}^{(4)}(\mathbf{p})+2\left\{\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<3>}\otimes\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_A)^{<2>}}\right\}\sum^{10}\boldsymbol{\sigma}^{(2)}\otimes\boldsymbol{\sigma}^{(3)}\right.$$
$$+6\alpha_2\frac{\partial\theta}{\partial\boldsymbol{\pi}'_B}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_B}+\frac{3}{2}\left\{\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<2>}\otimes\left(\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_A)^{<2>}}\right)^{<2>}\right\}\sum^{15}(\boldsymbol{\sigma}^{(2)})^{<3>}$$
$$+12\alpha_2\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_C}+\frac{2}{3}\left\{\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<3>}\otimes\frac{\partial^3\theta}{(\partial\boldsymbol{\pi}'_A)^{<3>}}\right\}\sum^{15}(\boldsymbol{\sigma}^{(2)})^{<3>}$$
$$\left.-(4\alpha_1\alpha_3+6\alpha_2\alpha_1^2+6\alpha_2\alpha_{2b})\right]+O(n^{-3/2})$$
$$\equiv n^{-1}\alpha_4+O(n^{-3/2})=n^{-1}\alpha_4^{(A)}+O(n^{-3/2}).\tag{A.3}$$

The equality $\alpha_4=\alpha_4^{(A)}$ in (A.3) follows from the relationship $6\alpha_2(\partial\theta/\partial\boldsymbol{\pi}'_B)\boldsymbol{\Sigma}\partial\theta/\partial\boldsymbol{\pi}_B+12\alpha_2(\partial\theta/\partial\boldsymbol{\pi}'_A)\boldsymbol{\Sigma}\partial\theta/\partial\boldsymbol{\pi}_C-6\alpha_2\alpha_{2b}=-6\alpha_2\alpha_{2b}^{(A)}$ (see (A.1)).

A.3 The asymptotic cumulants for studentized estimators

First, using the notations similar to (2.7) and (2.8), define the partial derivatives for later use as follows:

$$\frac{\partial\alpha_{2n}}{\partial\boldsymbol{\pi}} = \frac{\partial\hat{\alpha}_{2n}}{\partial\mathbf{p}}\bigg|_{\mathbf{p}=\boldsymbol{\pi}} = \frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_A} + n^{-1/2}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_B} + O(n^{-1}),$$
$$\frac{\partial^2\alpha_{2n}}{(\partial\boldsymbol{\pi})^{<2>}} = \frac{\partial^2\hat{\alpha}_{2n}}{(\partial\mathbf{p})^{<2>}}\bigg|_{\mathbf{p}=\boldsymbol{\pi}} = \frac{\partial^2\alpha_2}{(\partial\boldsymbol{\pi}_A)^{<2>}} + O(n^{-1/2}). \tag{A.4}$$

Let $(\cdot)_k$ be the k-th element of the vector in parentheses, $\mathbf{E}_{kk}$ be the $r\times r$ matrix whose k-th diagonal element is 1 with remaining ones being 0, $\mathbf{e}_{(k)}$ be the $r\times 1$ vector with unit norm whose k-th element is 1. Then, the terms on the right-hand sides of (A.4) are

$$\begin{aligned}
\frac{\partial\alpha_2}{\partial(\boldsymbol{\pi}_A)_k} &= 2\frac{\partial^2\theta}{\partial(\boldsymbol{\pi}_A)_k\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A} + \frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\frac{\partial\boldsymbol{\Sigma}}{\partial\pi_k}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A},\\
\frac{\partial\boldsymbol{\Sigma}}{\partial\pi_k} &\equiv \frac{\partial\{\mathrm{diag}(\mathbf{p})-\mathbf{p}\mathbf{p}'\}}{\partial p_k}|_{\mathbf{p}=\boldsymbol{\pi}} = \mathbf{E}_{kk} - \mathbf{e}_{(k)}\boldsymbol{\pi}' - \pi\mathbf{e}'_{(k)},\\
\frac{\partial^2\alpha_2}{\partial(\boldsymbol{\pi}_A)_k\partial(\boldsymbol{\pi}_A)_l} &= 2\frac{\partial^3\theta}{\partial(\boldsymbol{\pi}_A)_k\partial(\boldsymbol{\pi}_A)_l\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A} + 2\sum_{(k,l)}^{2}\frac{\partial^2\theta}{\partial(\boldsymbol{\pi}_A)_k\partial\boldsymbol{\pi}'_A}\frac{\partial\boldsymbol{\Sigma}}{\partial(\boldsymbol{\pi}_A)_l}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A},\\
&\quad + 2\frac{\partial^2\theta}{\partial(\boldsymbol{\pi}_A)_k\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial^2\theta}{\partial(\boldsymbol{\pi}_A)_l\partial\boldsymbol{\pi}_A} + \frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\frac{\partial^2\boldsymbol{\Sigma}}{\partial\pi_k\partial\pi_l}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A},\\
\frac{\partial^2\boldsymbol{\Sigma}}{\partial\pi_k\partial\pi_l} &= -\mathbf{e}_{(k)}\mathbf{e}'_{(l)} - \mathbf{e}_{(l)}\mathbf{e}'_{(k)},\\
\frac{\partial\alpha_2}{\partial(\boldsymbol{\pi}_B)_k} &= \frac{\partial\alpha_{2X}}{\partial\pi_k} = \frac{\partial}{\partial\pi_k}\left(2\frac{\partial\theta}{\partial\boldsymbol{\pi}'_B}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A}\right)\\
&= 2\frac{\partial^2\theta}{\partial(\boldsymbol{\pi}_B)_k\partial\boldsymbol{\pi}'_B}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A} + 2\frac{\partial\theta}{\partial\boldsymbol{\pi}'_B}\frac{\partial\boldsymbol{\Sigma}}{\partial\pi_k}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A} + 2\frac{\partial\theta}{\partial\boldsymbol{\pi}'_B}\boldsymbol{\Sigma}\frac{\partial^2\theta}{\partial(\boldsymbol{\pi}_A)_k\partial\boldsymbol{\pi}_A}\\
&\quad (k,l=1,\ldots,r),
\end{aligned}$$

where $\sum_{(k,l)}^{2}$denotes the sum of two terms given by exchanging the subscripts k and l.

Noting

$$\begin{aligned}
\alpha_{2n}^{-1/2} &= \alpha_2^{-1/2} - \frac{1}{2}\alpha_2^{-3/2}(n^{-1/2}\alpha_{2X} + n^{-1}\alpha_{2Y}) + \frac{3}{8}\alpha_2^{-5/2}(n^{-1/2}\alpha_{2X})^2 + O(n^{-3/2})\\
&= \alpha_2^{-1/2} - \frac{n^{-1/2}}{2}\alpha_2^{-3/2}\alpha_{2X} + n^{-1}\left(-\frac{\alpha_2^{-3/2}}{2}\alpha_{2Y} + \frac{3}{8}\alpha_2^{-5/2}\alpha_{2X}^2\right) + O(n^{-3/2})
\end{aligned}$$

and $\alpha_{2n}^{-3/2} = \alpha_2^{-3/2} - n^{-1/2}(3/2)\alpha_2^{-5/2}\alpha_{2X} + O(n^{-1})$, and using (4.4) and (A.4), t becomes

$$\begin{aligned}
t &= n^{1/2}(\hat{\theta}_n - \theta_n)/\hat{\alpha}_{2n}^{1/2}\\
&= \alpha_{2n}^{-1/2}\frac{\partial\theta_n}{\partial\boldsymbol{\pi}'}\mathbf{u} + \frac{n^{-1/2}}{2}\alpha_{2n}^{-1/2}\frac{\partial^2\theta_n}{(\partial\boldsymbol{\pi}')^{<2>}}\mathbf{u}^{<2>} - \frac{n^{-1/2}}{2}\alpha_{2n}^{-3/2}\frac{\partial\theta_n}{\partial\boldsymbol{\pi}'}\mathbf{u}\frac{\partial\alpha_{2n}}{\partial\boldsymbol{\pi}'}\mathbf{u}\\
&\quad + \frac{n^{-1}}{6}\alpha_{2n}^{-1/2}\frac{\partial^3\theta_n}{(\partial\boldsymbol{\pi}')^{<3>}}\mathbf{u}^{<3>} - \frac{n^{-1}}{4}\alpha_{2n}^{-3/2}\frac{\partial^2\theta_n}{(\partial\boldsymbol{\pi}')^{<2>}}\mathbf{u}^{<2>}\frac{\partial\alpha_{2n}}{\partial\boldsymbol{\pi}'}\mathbf{u}
\end{aligned}$$

$$
\begin{aligned}
&+n^{-1}\frac{3}{8}\alpha_{2n}^{-5/2}\frac{\partial\theta_n}{\partial\boldsymbol{\pi}'}\mathbf{u}\left(\frac{\partial\alpha_{2n}}{\partial\boldsymbol{\pi}'}\right)^{<2>}\mathbf{u}^{<2>}-\frac{n^{-1}}{4}\alpha_{2n}^{-3/2}\frac{\partial\theta_n}{\partial\boldsymbol{\pi}'}\mathbf{u}\frac{\partial^2\alpha_{2n}}{(\partial\boldsymbol{\pi}')^{<2>}}\mathbf{u}^{<2>}\\
&+O_p(n^{-3/2})\\
=\;&\alpha_2^{-1/2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\mathbf{u}+n^{-1/2}\left\{\alpha_2^{-1/2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_B}\mathbf{u}-\frac{\alpha_2^{-3/2}}{2}\alpha_{2X}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\mathbf{u}+\frac{\alpha_2^{-1/2}}{2}\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_A)^{<2>}}\mathbf{u}^{<2>}\right.\\
&\left.-\frac{\alpha_2^{-3/2}}{2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\mathbf{u}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}'_A}\mathbf{u}\right\}+n^{-1}\left\{\alpha_2^{-1/2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_C}\mathbf{u}-\frac{\alpha_2^{-3/2}}{2}\alpha_{2X}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_B}\mathbf{u}\right.\\
&+\left(-\frac{\alpha_2^{-3/2}}{2}\alpha_{2Y}+\frac{3}{8}\alpha_2^{-5/2}\alpha_{2X}^2\right)\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\mathbf{u}+\frac{\alpha_2^{-1/2}}{2}\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_B)^{<2>}}\mathbf{u}^{<2>}\\
&-\frac{\alpha_2^{-3/2}}{4}\alpha_{2X}\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_A)^{<2>}}\mathbf{u}^{<2>}\\
&-\sum_{(A,B)}^{2}\frac{\alpha_2^{-3/2}}{2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\mathbf{u}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}'_B}\mathbf{u}+\frac{3}{4}\alpha_2^{-5/2}\alpha_{2X}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\mathbf{u}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}'_A}\mathbf{u}\\
&+\frac{\alpha_2^{-1/2}}{6}\frac{\partial^3\theta}{(\partial\boldsymbol{\pi}'_A)^{<3>}}\mathbf{u}^{<3>}-\frac{\alpha_2^{-3/2}}{4}\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_A)^{<2>}}\mathbf{u}^{<2>}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}'_A}\mathbf{u}\\
&\left.+\frac{3}{8}\alpha_2^{-5/2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\mathbf{u}\left(\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}'_A}\right)^{<2>}\mathbf{u}^{<2>}-\frac{\alpha_2^{-3/2}}{4}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\mathbf{u}\frac{\partial^2\alpha_2}{(\partial\boldsymbol{\pi}'_A)^{<2>}}\mathbf{u}^{<2>}\right\}+O_p(n^{-3/2})\\
=\;&\alpha_2^{-1/2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\mathbf{u}+n^{-1/2}\left[\left(\alpha_2^{-1/2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_B}-\frac{\alpha_2^{-3/2}}{2}\alpha_{2X}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)\mathbf{u}+\left\{\frac{\alpha_2^{-1/2}}{2}\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_A)^{<2>}}\right.\right.\\
&\left.\left.-\frac{\alpha_2^{-3/2}}{2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\otimes\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}'_A}\right\}\mathbf{u}^{<2>}\right]+n^{-1}\left[\left\{\alpha_2^{-1/2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_C}-\frac{\alpha_2^{-3/2}}{2}\alpha_{2X}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_B}\right.\right.\\
&\left.+\left(-\frac{\alpha_2^{-3/2}}{2}\alpha_{2Y}+\frac{3}{8}\alpha_2^{-5/2}\alpha_{2X}^2\right)\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right\}\mathbf{u}\\
&+\left\{\frac{\alpha_2^{-1/2}}{2}\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_B)^{<2>}}-\frac{\alpha_2^{-3/2}}{4}\alpha_{2X}\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_A)^{<2>}}\right.\\
&\left.-\sum_{(A,B)}^{2}\frac{\alpha_2^{-3/2}}{2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\otimes\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}'_B}+\frac{3}{4}\alpha_2^{-5/2}\alpha_{2X}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\otimes\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}'_A}\right\}\mathbf{u}^{<2>}\\
&+\left\{\frac{\alpha_2^{-1/2}}{6}\frac{\partial^3\theta}{(\partial\boldsymbol{\pi}'_A)^{<3>}}-\frac{\alpha_2^{-3/2}}{4}\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_A)^{<2>}}\otimes\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}'_A}+\frac{3}{8}\alpha_2^{-5/2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\otimes\left(\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}'_A}\right)^{<2>}\right.\\
&\left.-\frac{\alpha_2^{-3/2}}{4}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\otimes\frac{\partial^2\alpha_2}{(\partial\boldsymbol{\pi}'_A)^{<2>}}\right\}\mathbf{u}^{<3>}+O_p(n^{-3/2})\\
\equiv\;&\mathbf{g}'_{(11)}\mathbf{u}+n^{-1/2}\{\mathbf{g}'_{(21)}\mathbf{u}+\mathbf{g}'_{(22)}\mathbf{u}^{<2>}\}\\
&+n^{-1}\{\mathbf{g}'_{(31)}\mathbf{u}+\mathbf{g}'_{(32)}\mathbf{u}^{<2>}+\mathbf{g}'_{(33)}\mathbf{u}^{<3>}\}+O_p(n^{-3/2}). \qquad \text{(A.5)}
\end{aligned}
$$

The asymptotic cumulants of t are derived from (A.5) and the moments of $\mathbf{u}$. The first cumulant is

$$\begin{aligned}\kappa_1(t) &= n^{-1/2}\mathbf{g}'_{(22)}\boldsymbol{\sigma}^{(2)}+n^{-1}\mathbf{g}'_{(32)}\boldsymbol{\sigma}^{(2)}+O(n^{-3/2})\\ &= n^{-1/2}\left\{\frac{\alpha_2^{-1/2}}{2}\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_A)^{<2>}}\boldsymbol{\sigma}^{(2)}-\frac{\alpha_2^{-3/2}}{2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_A}\right\}\\ &\quad+n^{-1}\left\{\frac{\alpha_2^{-1/2}}{2}\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_B)^{<2>}}\boldsymbol{\sigma}^{(2)}-\frac{\alpha_2^{-3/2}}{4}\alpha_{2X}\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_A)^{<2>}}\boldsymbol{\sigma}^{(2)}\right.\\ &\quad\left.-\sum_{(A,B)}^{2}\frac{\alpha_2^{-3/2}}{2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_B}+\frac{3}{4}\alpha_2^{-5/2}\alpha_{2X}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_A}\right\}+O(n^{-3/2})\\ &\equiv n^{-1/2}\alpha'_1+n^{-1}\alpha'_{1a}+O(n^{-3/2})=n^{-1/2}\alpha_1^{(A)'}+n^{-1}\alpha'_{1a}+O(n^{-3/2}),\end{aligned}\tag{A.6}$$

where the alternative expression $\alpha_1^{(A)'}$ of α'_1 was given as for (3.2).

The second cumulant of t is

$$\begin{aligned}\kappa_2(t) &= \mathbf{g}_{(11)}^{<2>'}\boldsymbol{\sigma}^{(2)}+n^{-1/2}2(\mathbf{g}'_{(11)}\otimes\mathbf{g}'_{(21)})\boldsymbol{\sigma}^{(2)}\\ &\quad+n^{-1}\{2(\mathbf{g}'_{(11)}\otimes\mathbf{g}'_{(22)})n^{1/2}\mathrm{E}(\mathbf{u}^{<3>})+\mathbf{g}_{(21)}^{<2>'}\boldsymbol{\sigma}^{(2)}+\mathbf{g}_{(22)}^{<2>'}\mathrm{E}(\mathbf{u}^{<4>})-(\alpha'_1)^2\\ &\quad+2(\mathbf{g}'_{(11)}\otimes\mathbf{g}'_{(31)})\boldsymbol{\sigma}^{(2)}+2(\mathbf{g}'_{(11)}\otimes\mathbf{g}'_{(33)})\mathrm{E}(\mathbf{u}^{<4>})\}+O(n^{-3/2})\\ &= 1+n^{-1/2}2(\mathbf{g}'_{(11)}\otimes\mathbf{g}'_{(21)})\boldsymbol{\sigma}^{(2)}\\ &\quad+n^{-1}\left\{2(\mathbf{g}'_{(11)}\otimes\mathbf{g}'_{(22)})\boldsymbol{\sigma}^{(3)}+\mathbf{g}_{(21)}^{<2>'}\boldsymbol{\sigma}^{(2)}+\mathbf{g}_{(22)}^{<2>'}\sum^{3}(\boldsymbol{\sigma}^{(2)})^{<2>}\right.\\ &\quad\left.-(\alpha'_1)^2+2(\mathbf{g}'_{(11)}\otimes\mathbf{g}'_{(31)})\boldsymbol{\sigma}^{(2)}+2(\mathbf{g}'_{(11)}\otimes\mathbf{g}'_{(33)})\sum^{3}(\boldsymbol{\sigma}^{(2)})^{<2>}\right\}+O(n^{-3/2})\\ &\equiv 1+n^{-1/2}\alpha'_{2a}+n^{-1}\alpha'_{2b}+O(n^{-3/2})=1+n^{-1}\alpha_{2b}^{(A)'}+O(n^{-3/2}),\end{aligned}\tag{A.7}$$

where

$$\begin{aligned}\sum^{3}(\boldsymbol{\sigma}^{(2)})^{<2>} &= \boldsymbol{\sigma}^{(2)}\otimes\boldsymbol{\sigma}^{(2)}+\mathrm{E}[\{\mathbf{u}\otimes\mathbf{1}_{(r)}\}^{<2>}]\odot\mathrm{E}[\{\mathbf{1}_{(r)}\otimes\mathbf{u}\}^{<2>}]\\ &\quad+\mathrm{E}[\{\mathbf{u}\otimes\mathbf{1}_{(r)}^{<2>}\otimes\mathbf{u}\}]\odot\mathrm{E}[\{\mathbf{1}_{(r)}\otimes\mathbf{u}^{<2>}\otimes\mathbf{1}_{(r)}\}],\end{aligned}$$

$\odot$ denotes the Hadamard or elementwise product, and

$$\alpha'_{2a}=2(\mathbf{g}'_{(11)}\otimes\mathbf{g}'_{(21)})\boldsymbol{\sigma}^{(2)}=2\alpha_2^{-1}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_B}-\alpha_2^{-1}\alpha_{2X}=\alpha_2^{-1}\alpha_{2a}-\alpha_2^{-1}\alpha_{2X}=0.$$

The equality of α'_{2b} and $\alpha_{2b}^{(A)'}$ is derived as follows. The part of α'_{2b}, given by Methods B and C, different from those by Methods A and D is

$$\begin{aligned}&\mathbf{g}_{(21)}^{<2>'}\boldsymbol{\sigma}^{(2)}+2(\mathbf{g}'_{(11)}\otimes\mathbf{g}'_{(31)})\boldsymbol{\sigma}^{(2)}\\ &= \alpha_2^{-1}\frac{\partial\theta}{\partial\pi'_B}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_B}-\alpha_2^{-2}\alpha_{2X}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_B}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A}+\frac{\alpha_2^{-2}}{4}\alpha_{2X}^2\end{aligned}$$

$$+2\left(\alpha_2^{-1}\frac{\partial\theta}{\partial\boldsymbol{\pi}_C'}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A}-\frac{\alpha_2^{-2}}{2}\alpha_{2X}\frac{\partial\theta}{\partial\boldsymbol{\pi}_B'}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A}-\frac{\alpha_2^{-1}}{2}\alpha_{2Y}+\frac{3}{8}\alpha_2^{-2}\alpha_{2X}^2\right)$$
$$=\alpha_2^{-1}\frac{\partial\theta}{\partial\boldsymbol{\pi}_B'}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_B}-\frac{\alpha_2^{-2}}{4}\alpha_{2X}^2+2\left\{\alpha_2^{-1}\frac{\partial\theta}{\partial\boldsymbol{\pi}_C'}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A}-\frac{\alpha_2^{-2}}{4}\alpha_{2X}^2\right.$$
$$\left.-\frac{\alpha_2^{-1}}{2}\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}_B'}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_B}+2\frac{\partial\theta}{\partial\boldsymbol{\pi}_C'}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A}\right)+\frac{3}{8}\alpha_2^{-2}\alpha_{2X}^2\right\}=0$$

where $\alpha_{2X}=2\frac{\partial\theta}{\partial\boldsymbol{\pi}_B'}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A}$ and $\alpha_{2Y}=\frac{\partial\theta}{\partial\boldsymbol{\pi}_B'}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_B}+2\frac{\partial\theta}{\partial\boldsymbol{\pi}_C'}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A}$ are used.

The third cumulant of t is

$$\begin{aligned}
\kappa_3(t) &= \mathrm{E}[\{t-\mathrm{E}(t)\}^3]=\mathrm{E}(t^3)-3\mathrm{E}(t^2)\mathrm{E}(t)+O(n^{-3/2})\\
&= n^{-1/2}\left\{\mathbf{g}_{(11)}^{<3>\prime}\boldsymbol{\sigma}^{(3)}+3(\mathbf{g}_{(11)}^{<2>\prime}\otimes\mathbf{g}_{(22)}')\sum^{3}(\boldsymbol{\sigma}^{(2)})^{<2>}\right\}\\
&\quad +n^{-1}\left\{3(\mathbf{g}_{(11)}^{<2>\prime}\otimes\mathbf{g}_{(21)}')\boldsymbol{\sigma}^{(3)}+3(\mathbf{g}_{(11)}^{<2>\prime}\otimes\mathbf{g}_{(32)}')\sum^{3}(\boldsymbol{\sigma}^{(2)})^{<2>}\right.\\
&\quad \left.+6(\mathbf{g}_{(11)}'\otimes\mathbf{g}_{(21)}'\otimes\mathbf{g}_{(22)}')\sum^{3}(\boldsymbol{\sigma}^{(2)})^{<2>}\right\}\\
&\quad -3(1+n^{-1/2}\alpha_{2a}')(n^{-1/2}\alpha_1'+n^{-1}\alpha_{1a}')+O(n^{-3/2})\\
&= n^{-1/2}\left\{\mathbf{g}_{(11)}^{<3>\prime}\boldsymbol{\sigma}^{(3)}+3(\mathbf{g}_{(11)}^{<2>\prime}\otimes\mathbf{g}_{(22)}')\sum^{3}(\boldsymbol{\sigma}^{(2)})^{<2>}-3\alpha_1'\right\}\\
&\quad +n^{-1}\left\{3(\mathbf{g}_{(11)}^{<2>\prime}\otimes\mathbf{g}_{(21)}')\boldsymbol{\sigma}^{(3)}+3(\mathbf{g}_{(11)}^{<2>\prime}\otimes\mathbf{g}_{(32)}')\sum^{3}(\boldsymbol{\sigma}^{(2)})^{<2>}\right.\\
&\quad \left.+6(\mathbf{g}_{(11)}'\otimes\mathbf{g}_{(21)}'\otimes\mathbf{g}_{(22)}')\sum^{3}(\boldsymbol{\sigma}^{(2)})^{<2>}-3\alpha_{1a}'-3\alpha_{2a}'\alpha_1'\right\}+O(n^{-3/2})\\
&\equiv n^{-1/2}\alpha_3'+n^{-1}\alpha_{3a}'+O(n^{-3/2})=n^{-1/2}\alpha_3^{(A)\prime}+n^{-1}\alpha_{3a}'+O(n^{-3/2}), \qquad (A.8)
\end{aligned}$$

where

$$\begin{aligned}
\alpha_3' &= \alpha_2^{-3/2}\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}_A'}\right)^{<3>}\boldsymbol{\sigma}^{(3)}+\frac{3}{2}\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}_A'}\right)^{<2>}\otimes\left(\alpha_2^{-3/2}\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}_A')^{<2>}}-\alpha_2^{-5/2}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A'}\otimes\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_A}\right)\\
&\quad\times\sum^{3}(\boldsymbol{\sigma}^{(2)})^{<2>}-\frac{3}{2}\left(\alpha_2^{-1/2}\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}_A')^{<2>}}\boldsymbol{\sigma}^{(2)}-\alpha_2^{-3/2}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A'}\boldsymbol{\Sigma}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_A}\right)\\
&= \alpha_2^{-3/2}\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}_A'}\right)^{<3>}\boldsymbol{\sigma}^{(3)}+3\alpha_2^{-3/2}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A'}\boldsymbol{\Sigma}\frac{\partial^2\theta}{\partial\boldsymbol{\pi}_A\partial\boldsymbol{\pi}_A'}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A}-3\alpha_2^{-3/2}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A'}\boldsymbol{\Sigma}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_A}\\
&= \alpha_2^{-3/2}\alpha_3-3\alpha_2^{-3/2}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A'}\boldsymbol{\Sigma}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_A}=\alpha_3^{(A)\prime},
\end{aligned}$$

and

$$\alpha_{3a}' = 3(\mathbf{g}_{(11)}^{<2>\prime}\otimes\mathbf{g}_{(21)}')\boldsymbol{\sigma}^{(3)}+3(\mathbf{g}_{(11)}^{<2>\prime}\otimes\mathbf{g}_{(32)}')\sum^{3}(\boldsymbol{\sigma}^{(2)})^{<2>}$$

$$
\begin{aligned}
&+6(\mathbf{g}'_{(11)}\otimes\mathbf{g}'_{(21)}\otimes\mathbf{g}'_{(22)})\sum^{3}(\boldsymbol{\sigma}^{(2)})^{<2>}-3\alpha'_{1a}-3\alpha'_{2a}\alpha'_{1}\\
=\;&\underbrace{3\alpha_2^{-3/2}\left\{\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<2>}\otimes\frac{\partial\theta}{\partial\boldsymbol{\pi}'_B}\right\}\boldsymbol{\sigma}^{(3)}}_{[0]}\underbrace{-\frac{3}{2}\alpha_2^{-5/2}\alpha_{2X}\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<3>}\boldsymbol{\sigma}^{(3)}}_{[5]}\\
&+\left\{\underbrace{\frac{3}{2}\alpha_2^{-3/2}\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<2>}\otimes\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_B)^{<2>}}}_{[1]}\underbrace{-\frac{3}{4}\alpha_2^{-5/2}\alpha_{2X}\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<2>}\otimes\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_A)^{<2>}}}_{[6]}\right.\\
&\underbrace{-\frac{3}{2}\alpha_2^{-5/2}\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<2>}\otimes\sum_{(A,B)}^{2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\otimes\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}'_B}}_{[4]}\\
&\underbrace{\left.+\frac{9}{4}\alpha_2^{-7/2}\alpha_{2X}\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<3>}\otimes\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}'_A}\right\}\sum^{3}(\boldsymbol{\sigma}^{(2)})^{<2>}}_{[7]}\\
&+3\alpha_2^{-3/2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\otimes\left(\underbrace{\frac{\partial\theta}{\partial\boldsymbol{\pi}'_B}}_{[2],[3]}\underbrace{-\frac{\alpha_2^{-1}}{2}\alpha_{2X}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}}_{[6],[7]}\right)\otimes\left(\underbrace{\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_A)^{<2>}}}_{[2],[6]}\underbrace{-\alpha_2^{-1}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\otimes\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}'_A}}_{[3],[7]}\right)\\
&\times\sum^{3}(\boldsymbol{\sigma}^{(2)})^{<2>}-3\left\{\underbrace{\frac{\alpha_2^{-1/2}}{2}\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_B)^{<2>}}\boldsymbol{\sigma}^{(2)}}_{[1]}\underbrace{-\frac{\alpha_2^{-3/2}}{4}\alpha_{2X}\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_A)^{<2>}}\boldsymbol{\sigma}^{(2)}}_{[6]}\right.\\
&\left.\underbrace{-\sum_{(A,B)}^{2}\frac{\alpha_2^{-3/2}}{2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_B}}_{[4]}\underbrace{+\frac{3}{4}\alpha_2^{-5/2}\alpha_{2X}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_A}}_{[7]}\right\}\\
&-\frac{3}{2}\alpha_2^{-3/2}\left(\underbrace{2\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_B}}_{[2],[3]}\underbrace{-\alpha_{2X}}_{[6],[7]}\right)\left(\underbrace{\frac{\partial^2\theta}{(\partial\boldsymbol{\pi}'_A)^{<2>}}\boldsymbol{\sigma}^{(2)}}_{[2],[6]}\underbrace{-\alpha_2^{-1}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_A}}_{[3],[7]}\right)\\
=\;&\underbrace{3\alpha_2^{-3/2}\left\{\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<2>}\otimes\frac{\partial\theta}{\partial\boldsymbol{\pi}'_B}\right\}\boldsymbol{\sigma}^{(3)}}_{[0]}\underbrace{+3\alpha_2^{-3/2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial^2\theta}{\partial\boldsymbol{\pi}_B\partial\boldsymbol{\pi}'_B}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A}}_{[1]}\\
&\underbrace{+6\alpha_2^{-3/2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial^2\theta}{\partial\boldsymbol{\pi}_A\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_B}}_{[2]}\underbrace{-3\alpha_2^{-5/2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_B}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_A}}_{[3]}\\
&\underbrace{-3\alpha_2^{-3/2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_B}\boldsymbol{\Sigma}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_A}}_{[3]}\underbrace{-3\alpha_2^{-5/2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_B}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_A}-3\alpha_2^{-3/2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_B}}_{[4]}
\end{aligned}
$$

$$-\underset{[5]\cdots\cdots}{\underline{\frac{3}{2}\alpha_2^{-5/2}\alpha_{2X}\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<3>}\boldsymbol{\sigma}^{(3)}}}-\underset{[6]\cdots\cdots}{\underline{\frac{9}{2}\alpha_2^{-5/2}\alpha_{2X}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial^2\theta}{\partial\boldsymbol{\pi}_A\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A}}}$$

$$+\underset{[7]\cdots\cdots}{\underline{\frac{15}{2}\alpha_2^{-5/2}\alpha_{2X}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_A}}}$$

$$\begin{aligned}
&= \alpha_2^{-3/2}\alpha_{3a}-6\alpha_2^{-5/2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_B}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_A}-3\alpha_2^{-3/2}\sum_{(A,B)}^{2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_B}\\
&\quad+\alpha_2^{-5/2}\alpha_{2X}\left\{-\frac{3}{2}\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<3>}\boldsymbol{\sigma}^{(3)}-\frac{9}{2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial^2\theta}{\partial\boldsymbol{\pi}_A\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A}+\frac{15}{2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_A}\right\}\\
&= \alpha_2^{-3/2}\alpha_{3a}-3\alpha_2^{-3/2}\sum_{(A,B)}^{2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_B}\\
&\quad+\alpha_2^{-5/2}\alpha_{2X}\left\{-\frac{3}{2}\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<3>}\boldsymbol{\sigma}^{(3)}-\frac{9}{2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial^2\theta}{\partial\boldsymbol{\pi}_A\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\theta}{\partial\boldsymbol{\pi}_A}+\frac{9}{2}\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\boldsymbol{\Sigma}\frac{\partial\alpha_2}{\partial\boldsymbol{\pi}_A}\right\},
\end{aligned}$$

where the dotted underscores with numbers are for confirmation of correspondence. In the derivation of α'_{3a}, the term $-3\alpha'_{2a}\alpha'_1$ $(=0)$ was temporarily retained for convenience to derive the last expression using α_{3a}.

The fourth cumulant of t is

$$\begin{aligned}
\kappa_4(t) &= \mathrm{E}[\{t-\mathrm{E}(t)\}^4]-3\kappa_2(t)^2\\
&= \mathrm{E}(t^4)-4\mathrm{E}(t^3)\mathrm{E}(t)+6\mathrm{E}(t^2)\mathrm{E}(t)^2-3\kappa_2(t)^2+O(n^{-2})\\
&= \mathrm{E}(t^4)-4\{n^{-1/2}\alpha'_3+n^{-1}\alpha'_{3a}+3(1+n^{-1/2}\alpha'_{2a})(n^{-1/2}\alpha'_1+n^{-1}\alpha'_{1a})\}n^{-1/2}\alpha'_1\\
&\quad+6(1+n^{-1/2}\alpha'_{2a})n^{-1}(\alpha'_1)^2-3(1+n^{-1/2}\alpha'_{2a}+n^{-1}\alpha'_{2b})^2+O(n^{-3/2})\\
&= \mathrm{E}(t^4)-3-n^{-1/2}6\alpha'_{2a}-n^{-1}\{4\alpha'_1\alpha'_3+6(\alpha'_1)^2+6\alpha'_{2b}+3(\alpha'_{2a})^2\}+O(n^{-3/2}),
\end{aligned}$$

where

$$\begin{aligned}
\mathrm{E}(t^4) &= \mathbf{g}_{(11)}^{<4>\prime}\mathrm{E}(\mathbf{u}^{<4>})+n^{-1/2}4(\mathbf{g}_{(11)}^{<3>\prime}\otimes\mathbf{g}'_{(21)})\mathrm{E}(\mathbf{u}^{<4>})\\
&\quad+n^{-1}\Big\{4(\mathbf{g}_{(11)}^{<3>\prime}\otimes\mathbf{g}'_{(22)})n^{1/2}\mathrm{E}(\mathbf{u}^{<5>})\\
&\quad+\binom{4}{2}(\mathbf{g}_{(11)}^{<2>\prime}\otimes\mathbf{g}_{(21)}^{<2>\prime})\mathrm{E}(\mathbf{u}^{<4>})+\binom{4}{2}(\mathbf{g}_{(11)}^{<2>\prime}\otimes\mathbf{g}_{(22)}^{<2>\prime})\mathrm{E}(\mathbf{u}^{<6>})\\
&\quad+4(\mathbf{g}_{(11)}^{<3>\prime}\otimes\mathbf{g}'_{(31)})\mathrm{E}(\mathbf{u}^{<4>})+4(\mathbf{g}_{(11)}^{<3>\prime}\otimes\mathbf{g}'_{(33)})\mathrm{E}(\mathbf{u}^{<6>})\Big\}+O(n^{-3/2})\\
&= 3+n^{-1}\alpha_2^{-2}\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<4>}n^3\boldsymbol{\kappa}^{(4)}(\mathbf{p})+n^{-1/2}12(\mathbf{g}'_{(11)}\otimes\mathbf{g}'_{(21)})\boldsymbol{\sigma}^{(2)}\\
&\quad+n^{-1}\Bigg\{4(\mathbf{g}_{(11)}^{<3>\prime}\otimes\mathbf{g}'_{(22)})\sum^{10}\boldsymbol{\sigma}^{(2)}\otimes\boldsymbol{\sigma}^{(3)}+6\mathbf{g}_{(21)}^{<2>\prime}\boldsymbol{\sigma}^{(2)}\\
&\quad+12\{(\mathbf{g}'_{(11)}\otimes\mathbf{g}'_{(21)})\boldsymbol{\sigma}^{(2)}\}^2+6(\mathbf{g}_{(11)}^{<2>\prime}\otimes\mathbf{g}_{(22)}^{<2>\prime})\sum^{15}(\boldsymbol{\sigma}^{(2)})^{<3>}
\end{aligned}$$

$$+12(\mathbf{g}'_{(11)}\otimes\mathbf{g}'_{(31)})\boldsymbol{\sigma}^{(2)}+4(\mathbf{g}^{<3>\prime}_{(11)}\otimes\mathbf{g}'_{(33)})\sum^{15}(\boldsymbol{\sigma}^{(2)})^{<3>}\Bigg\}+O(n^{-3/2})$$

From $\alpha'_{2a}=2(\mathbf{g}'_{(11)}\otimes\mathbf{g}'_{(21)})\boldsymbol{\sigma}^{(2)}$ and $\mathbf{g}^{<2>\prime}_{(21)}\boldsymbol{\sigma}^{(2)}+2(\boldsymbol{g}'_{(11)}\otimes\mathbf{g}'_{(31)})\boldsymbol{\sigma}^{(2)}=0$ (see the derivation for $\alpha'_{2b}=\alpha^{(A)\prime}_{2b}$), the following result is obtained.

$$\begin{aligned}\kappa_4(t) &= n^{-1}\Bigg[\alpha_2^{-2}\left(\frac{\partial\theta}{\partial\boldsymbol{\pi}'_A}\right)^{<4>}n^3\boldsymbol{\kappa}^{(4)}(\mathbf{p})+4(\mathbf{g}^{<3>\prime}_{(11)}\otimes\mathbf{g}'_{(22)})\sum^{10}\boldsymbol{\sigma}^{(2)}\otimes\boldsymbol{\sigma}^{(3)}\\ &\quad+6(\mathbf{g}^{<2>\prime}_{(11)}\otimes\mathbf{g}^{<2>\prime}_{(22)})\sum^{15}(\boldsymbol{\sigma}^{(2)})^{<3>}+4(\mathbf{g}^{<3>\prime}_{(11)}\otimes\mathbf{g}'_{(33)})\sum^{15}(\boldsymbol{\sigma}^{(2)})^{<3>}\\ &\quad-\{4\alpha'_1\alpha'_3+6(\alpha'_1)^2+6\alpha'_{2b}\}\Bigg]+O(n^{-3/2})\\ &\equiv n^{-1}\alpha'_4+O(n^{-3/2})=n^{-1}\alpha^{(A)\prime}_4+O(n^{-3/2}).\end{aligned}\tag{A.9}$$

Note that (A.9) was derived without using $\alpha'_{2a}=0$. The last equation is given by the expression of (A.9) which does not involve those for other than Methods A and D.

A.4 Partial derivatives

A.4.1 The log odds-ratio

Let $p_{nii}\equiv p_{ii}+n^{-1/2}b_{ii}+n^{-1}c_{ii}+n^{-3/2}d_{ii}$ and $p_{nij}\equiv p_{ij}+n^{-1/2}b_{ij}+n^{-1}c_{ij}+n^{-3/2}d_{ij}$, where $b_{ii},b_{ij},c_{ii},c_{ij},d_{ii}$ and d_{ij} are constants, and $i\neq j$ $(i,j=1,2)$ is assumed in Subsection A.4. Let $\hat{\omega}_n\equiv p_{n11}p_{n22}/(p_{n12}p_{n21})=p_{nii}p_{njj}/(p_{nij}p_{nji})$. Then,

$$\begin{aligned}\frac{\partial\ln\hat{\omega}_n}{\partial p_{ii}} &= \frac{\partial\ln\hat{\omega}_n}{\partial p_{nii}}=\frac{1}{p_{nii}}=\frac{1}{p_{ii}}-\frac{1}{p_{ii}^2}(n^{-1/2}b_{ii}+n^{-1}c_{ii})+\frac{n^{-1}}{p_{ii}^3}b_{ii}^2+O_p(n^{-3/2})\\ &= \frac{1}{p_{ii}}-n^{-1/2}\frac{b_{ii}}{p_{ii}^2}+n^{-1}\left(\frac{b_{ii}^2}{p_{ii}^3}-\frac{c_{ii}}{p_{ii}^2}\right)+O_p(n^{-3/2}),\\ \frac{\partial\ln\hat{\omega}_n}{\partial p_{ij}} &= -\frac{1}{p_{ij}}+n^{-1/2}\frac{b_{ij}}{p_{ij}^2}+n^{-1}\left(-\frac{b_{ij}^2}{p_{ij}^3}+\frac{c_{ij}}{p_{ij}^2}\right)+O_p(n^{-3/2}),\\ \frac{\partial^2\ln\hat{\omega}_n}{\partial p_{ii}^2} &= -\frac{1}{p_{ii}^2}+n^{-1/2}\frac{2b_{ii}}{p_{ii}^3}+O_p(n^{-1}),\quad \frac{\partial^2\ln\hat{\omega}_n}{\partial p_{ij}^2}=\frac{1}{p_{ij}^2}-n^{-1/2}\frac{2b_{ij}}{p_{ij}^3}+O_p(n^{-1}),\\ \frac{\partial^3\ln\hat{\omega}_n}{\partial p_{ii}^3} &= \frac{2}{p_{ii}^3}+O_p(n^{-1/2}),\quad \frac{\partial^3\ln\hat{\omega}_n}{\partial p_{ij}^3}=-\frac{2}{p_{ij}^3}+O_p(n^{-1/2}).\end{aligned}$$

A.4.2 Yule's coefficients

Let $\hat{\theta}^*_n$ and $\hat{\theta}^*$ be the estimated generalized Yule's coefficients defined by $\hat{\theta}^*_n=\dfrac{\hat{\omega}_n^{c*}-1}{\hat{\omega}_n^{c*}+1}$ and $\hat{\theta}^*=\dfrac{\hat{\omega}^{c*}-1}{\hat{\omega}^{c*}+1}$, respectively, where $\hat{\omega}_n$ is expanded as

$$\hat{\omega}_n = \hat{\omega}+n^{-1/2}\frac{\partial\hat{\omega}}{\partial\mathbf{p}'}\mathbf{b}+n^{-1}\left(\frac{\partial\hat{\omega}}{\partial\mathbf{p}'}\mathbf{c}+\frac{1}{2}\frac{\partial^2\hat{\omega}}{(\partial\mathbf{p}')^{<2>}}\mathbf{b}^{<2>}\right)+O_p(n^{-3/2})$$

$$\equiv \hat{\omega}+n^{-1/2}\hat{\omega}_X+n^{-1}\hat{\omega}_Y+O_p(n^{-3/2})$$

with $\mathbf{b}=(b_{11},b_{12},b_{21},b_{22})'$ and $\mathbf{c}=(c_{11},c_{12},c_{21},c_{22})'$. Then, the partial derivatives of $\hat{\omega}_n$ with respect to $\mathbf{p}$ used in the asymptotic expansions are

$$\frac{\partial\hat{\omega}_n}{\partial\mathbf{p}}=\frac{\partial\hat{\omega}}{\partial\mathbf{p}}+n^{-1/2}\frac{\partial\hat{\omega}_X}{\partial\mathbf{p}}+n^{-1}\frac{\partial\hat{\omega}_Y}{\partial\mathbf{p}}+O_p(n^{-3/2}),$$
$$\frac{\partial^2\hat{\omega}_n}{(\partial\mathbf{p})^{<2>}}=\frac{\partial^2\hat{\omega}}{(\partial\mathbf{p})^{<2>}}+n^{-1/2}\frac{\partial^2\hat{\omega}_X}{(\partial\mathbf{p})^{<2>}}+O_p(n^{-1}),\quad \frac{\partial^3\hat{\omega}_n}{(\partial\mathbf{p})^{<3>}}=\frac{\partial^3\hat{\omega}}{(\partial\mathbf{p})^{<3>}}+O_p(n^{-1/2}),$$
$$\frac{\partial^k\hat{\omega}_X}{(\partial\mathbf{p})^{<k>}}=\frac{\partial^{k+1}\hat{\omega}}{(\partial\mathbf{p})^{<k>}\partial\mathbf{p}'}\mathbf{b}\ (k=1,2),\quad \frac{\partial\hat{\omega}_Y}{\partial\mathbf{p}}=\frac{\partial^2\hat{\omega}}{\partial\mathbf{p}\partial\mathbf{p}'}\mathbf{c}+\frac{1}{2}\frac{\partial^3\hat{\omega}}{\partial\mathbf{p}(\partial\mathbf{p}')^{<2>}}\mathbf{b}^{<2>},$$

where the non-zero derivatives of $\hat{\omega}$ with respect to $\mathbf{p}$ are

$$\frac{\partial\hat{\omega}}{\partial p_{ii}}=\frac{p_{jj}}{p_{ij}p_{ji}},\quad \frac{\partial\hat{\omega}}{\partial p_{ij}}=-\frac{p_{ii}p_{jj}}{p_{ij}^2p_{ji}},\quad \frac{\partial^2\hat{\omega}}{\partial p_{ii}\partial p_{jj}}=\frac{1}{p_{ij}p_{ji}},\quad \frac{\partial^2\hat{\omega}}{\partial p_{ii}\partial p_{ij}}=-\frac{p_{jj}}{p_{ij}^2p_{ji}},$$
$$\frac{\partial^2\hat{\omega}}{\partial p_{ij}^2}=\frac{2p_{ii}p_{jj}}{p_{ij}^3p_{ji}},\quad \frac{\partial^2\hat{\omega}}{\partial p_{ij}\partial p_{ji}}=\frac{p_{ii}p_{jj}}{p_{ij}^2p_{ji}^2},\quad \frac{\partial^3\hat{\omega}}{\partial p_{ii}\partial p_{jj}\partial p_{ij}}=-\frac{1}{p_{ij}^2p_{ji}},\quad \frac{\partial^3\hat{\omega}}{\partial p_{ii}\partial p_{ij}^2}=\frac{2p_{jj}}{p_{ij}^3p_{ji}},$$
$$\frac{\partial^3\hat{\omega}}{\partial p_{ii}\partial p_{ij}\partial p_{ji}}=\frac{p_{jj}}{p_{ij}^2p_{ji}^2},\quad \frac{\partial^3\hat{\omega}}{\partial p_{ij}^3}=-\frac{6p_{ii}p_{jj}}{p_{ij}^4p_{ji}},\quad \frac{\partial^3\hat{\omega}}{\partial p_{ij}^2\partial p_{ji}}=-\frac{2p_{ii}p_{jj}}{p_{ij}^3p_{ji}^2}.$$

Using the results above,

$$\begin{aligned}\frac{\partial\hat{\theta}_n^*}{\partial\mathbf{p}} &= \frac{\partial\hat{\theta}_n^*}{\partial\hat{\omega}_n}\frac{\partial\hat{\omega}_n}{\partial\mathbf{p}}=\left\{\frac{\partial\hat{\theta}^*}{\partial\hat{\omega}}+\frac{\partial^2\hat{\theta}^*}{\partial\hat{\omega}^2}(n^{-1/2}\hat{\omega}_X+n^{-1}\hat{\omega}_Y)+\frac{n^{-1}}{2}\frac{\partial^3\hat{\theta}^*}{\partial\hat{\omega}^3}\hat{\omega}_X^2\right\}\\
&\quad\times\left(\frac{\partial\hat{\omega}}{\partial\mathbf{p}}+n^{-1/2}\frac{\partial\hat{\omega}_X}{\partial\mathbf{p}}+n^{-1}\frac{\partial\hat{\omega}_Y}{\partial\mathbf{p}}\right)+O_p(n^{-3/2})\\
&= \frac{\partial\hat{\theta}^*}{\partial\hat{\omega}}\frac{\partial\hat{\omega}}{\partial\mathbf{p}}+n^{-1/2}\left(\frac{\partial^2\hat{\theta}^*}{\partial\hat{\omega}^2}\frac{\partial\hat{\omega}}{\partial\mathbf{p}}\hat{\omega}_X+\frac{\partial\hat{\theta}^*}{\partial\hat{\omega}}\frac{\partial\hat{\omega}_X}{\partial\mathbf{p}}\right)\\
&\quad+n^{-1}\left\{\frac{1}{2}\frac{\partial^3\hat{\theta}^*}{\partial\hat{\omega}^3}\frac{\partial\hat{\omega}}{\partial\mathbf{p}}\hat{\omega}_X^2+\frac{\partial^2\hat{\theta}^*}{\partial\hat{\omega}^2}\left(\frac{\partial\hat{\omega}_X}{\partial\mathbf{p}}\hat{\omega}_X+\frac{\partial\hat{\omega}}{\partial\mathbf{p}}\hat{\omega}_Y\right)+\frac{\partial\hat{\theta}^*}{\partial\hat{\omega}}\frac{\partial\hat{\omega}_Y}{\partial\mathbf{p}}\right\}\\
&\quad+O_p(n^{-3/2}),\\
\frac{\partial^2\hat{\theta}_n^*}{(\partial\mathbf{p})^{<2>}} &= \frac{\partial^2\hat{\theta}^*}{\partial\hat{\omega}^2}\left(\frac{\partial\hat{\omega}}{\partial\mathbf{p}}\right)^{<2>}+\frac{\partial\hat{\theta}^*}{\partial\hat{\omega}}\frac{\partial^2\hat{\omega}}{(\partial\mathbf{p})^{<2>}}+n^{-1/2}\left\{\frac{\partial^3\hat{\theta}^*}{\partial\hat{\omega}^3}\left(\frac{\partial\hat{\omega}}{\partial\mathbf{p}}\right)^{<2>}\hat{\omega}_X\right.\\
&\quad\left.+\frac{\partial^2\hat{\theta}^*}{\partial\hat{\omega}^2}\left(\frac{\partial^2\hat{\omega}}{(\partial\mathbf{p})^{<2>}}\hat{\omega}_X+\frac{\partial\hat{\omega}}{\partial\mathbf{p}}\otimes\frac{\partial\hat{\omega}_X}{\partial\mathbf{p}}+\frac{\partial\hat{\omega}_X}{\partial\mathbf{p}}\otimes\frac{\partial\hat{\omega}}{\partial\mathbf{p}}\right)+\frac{\partial\hat{\theta}^*}{\partial\hat{\omega}}\frac{\partial^2\hat{\omega}_X}{(\partial\mathbf{p})^{<2>}}\right\}\\
&\quad+O_p(n^{-1}),\\
\frac{\partial^3\hat{\theta}_n^*}{(\partial\mathbf{p})^{<3>}} &= \frac{\partial^3\hat{\theta}^*}{\partial\hat{\omega}^3}\left(\frac{\partial\hat{\omega}}{\partial\mathbf{p}}\right)^{<3>}+\frac{\partial^2\hat{\theta}^*}{\partial\hat{\omega}^2}\sum^3\frac{\partial^2\hat{\omega}}{(\partial\mathbf{p})^{<2>}}\otimes\frac{\partial\hat{\omega}}{\partial\mathbf{p}}+\frac{\partial\hat{\theta}^*}{\partial\hat{\omega}}\frac{\partial^3\hat{\omega}}{(\partial\mathbf{p})^{<3>}}\\
&\quad+O_p(n^{-1/2}),\end{aligned}$$

where

$$\frac{\partial\hat{\theta}^*}{\partial\hat{\omega}}=\frac{2c^*\hat{\omega}^{c*-1}}{(\hat{\omega}^{c*}+1)^2},\quad \frac{\partial^2\hat{\theta}^*}{\partial\hat{\omega}^2}=\frac{2c^*(c^*-1)\hat{\omega}^{c*-2}}{(\hat{\omega}^{c*}+1)^2}-\frac{4(c^*)^2\hat{\omega}^{2c*-2}}{(\hat{\omega}^{c*}+1)^3},$$

$$\frac{\partial^3\hat{\theta}^*}{\partial\hat{\omega}^3}=\frac{2c^*(c^*-1)(c^*-2)\hat{\omega}^{c*-3}}{(\hat{\omega}^{c*}+1)^2}-\frac{12(c^*)^2(c^*-1)\hat{\omega}^{2c*-3}}{(\hat{\omega}^{c*}+1)^3}+\frac{12(c^*)^3\hat{\omega}^{3c*-3}}{(\hat{\omega}^{c*}+1)^4}.$$

References

Agresti, A. (2002). *Categorical data analysis* (2nd ed.). Hoboken, N. J.: Wiley.

Agresti, A. (2007). *An introduction to categorical data analysis* (2nd ed.). Hoboken, N. J.: Wiley.

Agresti, A., & Caffo, B. (2000). Simple and effective confidence intervals for proportions and differences of proportions result from adding two successes and two failures. *The American Statistician*, **54**, 280–288.

Agresti, A., & Coull, B. A. (1998). Approximate is better than "exact" for interval estimation of binomial proportions. *The American Statistician*, **52**, 119–126.

Anscombe, F. J. (1956). On estimating binomial response relations. *Biometrika*, **43**, 461–464.

Bhattacharya, R. N., & Ranga Rao, R. (1986). *Normal approximation and asymptotic expansions* (reprint ed.). Malabar, FL: Krieger.

Bikyalis, A. (1969). Asymptotic expansions for distributions of sums of identically distributed independent lattice random variables. *Theory of Probability and Its Applications*, **14**, 481–489.

Bonett, D. G., & Price, R. M. (2005). Inferential methods for the tetrachoric correlation coefficient. *Journal of Educational and Behavioral Statistics*, **30**, 213–225.

Bonett, D. G., & Price, R. M. (2007). Statistical inference for generalized Yule coefficients in 2×2 contingency tables. *Sociological Methods and Research*, **35**, 429–446.

Brown, L. D., Cai, T. T., & DasGupta, A. (2001). Interval estimation for a binomial proportion. *Statistical Science*, **16**, 101–123.

Brown, L. D., Cai, T. T., & DasGupta, A. (2002). Confidence interval for a binomial proportion and asymptotic expansions. *The Annals of Statistics*, **30**, 160–201.

Cai, T. T. (2005). One-sided confidence intervals in discrete distributions. *Journal of Statistical Planning and Inference*, **131**, 63–88.

Clogg, C. C., & Eliason, S. R. (1987). Some common problems in log-linear analysis. *Sociological Methods and Research*, **16**, 8–44.

Clopper, C. J., & Pearson, E. S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, **26**, 404–413.

Digby, P. G. N. (1983). Approximating the tetrachoric correlation coefficient. *Biometrics*, **39**, 753–757.

Edwards, J. H. (1957). A note on the practical interpretation of 2 × 2 tables. *British Journal of Preventive and Social Medicine*, **11**, 73–78.

Esséen, C.-G. (1945). Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law. *Acta Mathematica*, **77**, 1–125.

Feller, W. (1971). *An introduction to probability theory and its applications* (2nd ed., Vol. II). New York: Wiley.

Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika*, **80**, 27–38.

Galindo-Garre, F., & Vermunt, J. K. (2006). Avoiding boundary estimates in latent class analysis by Bayesian posterior mode estimation. *Behaviormetrika*, **33**, 43–59.

Galindo-Garre, F., Vermunt, J. K., & Bergsma, W. P. (2004). Bayesian posterior estimation of logit parameters with small samples. *Sociological Methods and Research*, **33**, 88–117.

Gart, J. J. (1966). Alternative analyses of contingency tables. *Journal of the Royal Statistical Society*, B, **28**, 164–179.

Gart, J. J., & Zweifel, J. R. (1967). On the bias of various estimators of the logit and its variance with application to quantal bioassay. *Biometrika*, **54**, 181–187.

Ghosh, B. K. (1979). A comparison of some approximate confidence intervals for the binomial parameter. *Journal of the American Statistical Association*, **74**, 894–900.

Götze, F., & Ulyanov, V. V. (2003). *Asymptotic distribution of* χ^2*-type statistics*. Unpublished paper. http://www.math.uni-bielefeld.de/fgweb/main-e.html.

Haldane, J. B. S. (1956). The estimation and significance of the logarithm of a ratio of frequencies. *Annals of Human Genetics*, **20**, 309–311.

Hall, P. (1982). Improving the normal approximation when constructing one-sided confidence intervals for binomial or Poisson parameters. *Biometrika*, **69**, 647–652

Hall, P. (1992a). On the removal of skewness by transformation. *Journal of the Royal Statistical Society*, B, **54**, 221–228.

Hall, P. (1992b). *The bootstrap and Edgeworth expansion.* New York: Springer. Corrected printing, 1997.

Hitchcock, S. E. (1962). A note on the estimation of the parameters of the logistic function, using the minimum logit χ^2 method. *Biometrika*, **49**, 250–252.

Kolassa. J. E. (1989). *Topics in series approximations to distribution functions.* Unpublished doctoral dissertation, Department of Statistics, The University of Chicago, Chicago, IL.

Kolassa. J. E. (2006). *Series approximation methods in statistics* (3rd ed.). New York: Springer.

Kolassa. J. E., & McCullagh, P. (1990). Edgeworth series for lattice distributions. *The Annals of Statistics*, **18**, 981–985.

Leonard, T., & Hsu, J. S. J. (1994). The Bayesian analysis of categorical data: A selective review. In P. R. Freeman & A. F. M. Smith (Eds.), *Aspects of uncertainty: A tribute to D. V. Lindley* (pp.283–310). New York: Wiley.

Ogasawara, H. (2009). Asymptotic cumulants of the parameter estimators in item response theory. *Computational Statistics*, **24**, 313–331.

Ogasawara, H. (2010). Asymptotic expansions for the pivots using log-likelihood derivatives with an application in item response theory. *Journal of Multivariate Analysis*, **101**, 2149–2167.

Ogasawara, H. (2012). Cornish-Fisher expansions using sample cumulants and monotonic transformations. *Journal of Multivariate Analysis*, **103**, 1–18.

Press, S. J. (2005). *Applied multivariate analysis: Using Bayesian and frequentist methods of inference* (2nd ed.). Mineola, N. Y.: Dover.

Ranga Rao, R. (1961). On the central limit theorem in R_k. *Bulletin of the American Mathematical Society*, **67**, 359–361.

Siotani, M., & Fujikoshi, Y. (1984). Asymptotic approximations for the distributions of multinomial goodness-of-fit statistics. *Hiroshima Mathematical Journal*, **14**, 115–124.

Staicu, A.-M. (2009). Higher-order approximations for interval estimation in binomial settings. *Journal of Statistical Planning and Inference*, **139**, 3393–3404.

Stuart, A., & Ord, J. K. (1994). *Kendall's advanced theory of statistics: Distribution theory* (6th ed., Vol.1). London: Arnold.

Wilson, E. B. (1927). Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association*, **22**, 209–212.

Wold, H. (1934). Sheppard's correction formulae in several variables. *Skandinavisk Aktuarietidskrift*, **17**, 248–255.

Yarnold, J. K. (1972). Asymptotic approximations for the probability that a sum of lattice random vectors lies in a convex set. *The Annals of Mathematical Statistics*, **43**, 1566–1580.

Yule, G. U. (1900). On the association of attributes in statistics. *Philosophical Transactions of the Royal Society of London*, A, **194**, 257–319.

Yule, G. U. (1912). On the methods of measuring association between two attributes. *Journal of the Royal Statistical Society*, **75**, 579–642.

Zhou, X. H., Li, C. M., & Yang, Z. (2008). Improving interval estimation of binomial proportions. *Philosophical Transactions of the Royal Society*, A, **366**, 2405–2418.
Zhou, X. H., & Qin, G. (2005). A new confidence interval for the difference between two binomial proportions of paired data. *Journal of Statistical Planning and Inference*, **128**, 527–542.
Zhou, X. H., Tsao, M., & Qin, G. (2004). New intervals for the difference between two independent binomial proportions. *Journal of Statistical Planning and Inference*, **123**, 97–115.

(Received May 22 2012, Revised July 4 2012)