Working Paper Series, N. 9, September 2014

Towards unifying second-order theory of likelihoods and pseudolikelihoods

Nicola Lunardon
Department of Statistical Sciences
University of Padova
Italy

Abstract: Theory is developed to show that second-order distributional behaviour of pseudolikelihood ratios can be modified to resemble that of likelihood counterparts by means of a suitable adjustment. The latter is conceived to enable the Bartlett correction for pseudolikelihood ratios when inference focuses on a scalar parameter. The proposed methodology can be framed in the likelihood setting where it can be interpreted as a device to achieve second-order accurate inference that takes into an account potential erroneous model assumptions. The efficacy of the proposal is demonstrated via simulation studies.

Keywords: Bartlett correction; Composite likelihood; Model misspecification; Pseudolikelihood; Second-order asymptotics.

Contents

Department of Statistical Sciences
Via Cesare Battisti, 241
35121 Padova
Italy

tel: +39 049 8274168
fax: +39 049 8274170
http://www.stat.unipd.it

Corresponding author:
Nicola Lunardon
tel: +39 049 827 4124
lunardon@stat.unipd.it
http://homes.stat.unipd.it/
nicolalunardon/

Towards unifying second-order theory of likelihoods and pseudolikelihoods

Nicola Lunardon
Department of Statistical Sciences
University of Padova
Italy

Abstract: Theory is developed to show that second-order distributional behaviour of pseudolikelihood ratios can be modified to resemble that of likelihood counterparts by means of a suitable adjustment. The latter is conceived to enable the Bartlett correction for pseudolikelihood ratios when inference focuses on a scalar parameter. The proposed methodology can be framed in the likelihood setting where it can be interpreted as a device to achieve second-order accurate inference that takes into an account potential erroneous model assumptions. The efficacy of the proposal is demonstrated via simulation studies.

Keywords: Bartlett correction; Composite likelihood; Model misspecification; Pseudolikelihood; Second-order asymptotics.

1 Introduction

Pseudolikelihood is the heading that subsumes a wide class of inference functions conceived to conduct likelihood-like inference yet circumventing restrictive model assumptions. Typically, pseudolikelihoods and derived quantities possess only a few key properties of the likelihood counterparts. These are related to first-order asymptotics, as the consistency and the asymptotic normality of estimators, and guarantee the validity of the inferential conclusions. Nevertheless, the distributional characterisation of pseudolikelihood ratios may be different from that of likelihoods and the discrepancies may arise even at first-order (Kent, 1982). On the one hand, standard first-order distributional behaviour can be restored by means of suitable modifications, as witnessed by the substantial body of work by Rotnitzky and Jewell (1990), Chandler and Bate (2007), and Pace *et al.* (2011). On the other hand, the development of general strategies to correct for the second-order behaviour have been neglected. Contributions are usually devoted to assess the properties of specific instances of pseudolikelihoods (DiCiccio *et al.*, 1991) or to describe how close they relate to likelihoods (Mykland, 1999). Consequently, it is seldom possible to draw a direct link between second-order theory for pseudolikelihoods and likelihoods. Our endeavour is to create the breading ground to try to fill this gap by showing that second-order behaviour of pseudolikelihoods can be manipulated to resemble that of likelihoods. In particular, we prove that it is possible to create the necessary conditions to enable the Bartlett correction for pseudolikelihood ratios. The result is not only of relevance for pseudolikelihoods because it is susceptible of a clear-cut interpretation from the standpoint of likelihood theory: second-order accurate inference can be safeguarded against erroneous model assumptions.

We focus on a broad class of pseudolikelihoods that generalises and includes the likelihood, namely marginal composite likelihoods (Varin, 2008). Let $y_1, \dots, y_n$ be a sample of size n of independent and identically distributed observations from a q-dimensional random vector Y having unknown density $g(y)$. Marginal composite likelihoods may be defined by considering a parametric statistical model $\{f(y;\theta), \theta \in \Theta \subseteq \mathbb{R}, y \in \mathbb{R}^q\}$ and a set of marginal events on the sample space $\{\mathcal{E}_1, \dots, \mathcal{E}_K\}$ involving the components of y_i. If we denote the likelihood function associated to each event by $f(y_i \in \mathcal{E}_k;\theta)$, then the marginal composite log likelihood is

$$\ell(\theta) = \sum_{i=1}^{n}\sum_{k=1}^{K} w_k \log f(y_i \in \mathcal{E}_k;\theta) = \sum_{i=1}^{n} \ell(\theta; y_i), \tag{1}$$

where w_k are non-negative weights. The events $\mathcal{E}_k$ may regard subsets of components of y_i whose dimension are, for instance, 1, 2, up to q, leading to respectively the independence likelihood, the marginal pairwise likelihood, and the likelihood. This is in no way an exhaustive list and we defer the reader to Varin (2008, Sect. 2) for an overall view.

The remainder of this introduction is devoted to introduce further definitions and notation. Let $W(\theta) = 2\{\ell(\hat\theta) - \ell(\theta)\}$ be the composite log likelihood ratio for θ, with $\hat\theta = \operatorname{argmax}_\theta \ell(\theta)$ the maximum composite likelihood estimate. Denote by $\ell_j(\theta) = \partial^j \ell(\theta)/\partial\theta^j$ the j-th order derivative of the composite log likelihood. We define

$$\alpha_{rstu}(\theta) = \nu\, \mathbb{E}_g \left\{[\ell_1(\theta;Y)]^r[\ell_2(\theta;Y)]^s[\ell_3(\theta;Y)]^t[\ell_4(\theta;Y)]^u\right\}$$

along with the centred random variables

$$A_{rstu}(\theta) = \nu\, n^{-1}\sum_{i=1}^{n}[\ell_1(\theta;y_i)]^r[\ell_2(\theta;y_i)]^s[\ell_3(\theta;y_i)]^t[\ell_4(\theta;y_i)]^u - \alpha_{rstu}(\theta),$$

where r, s, t, u are non negative integers. The factor $\nu = (-1)^{(2r+s+2t+2u)!}$ switches the sign of α_{01} and A_{01} only, i.e. it ensures $\alpha_{01} > 0$. We shall adopt the shorthand $\alpha_{101}(\theta) \equiv \alpha_{1010}(\theta)$, $A_2(\theta) \equiv A_{2000}(\theta)$, and so forth, i.e. zeroes are retained when they precede an index greater or equal than 1. Further, we denote by $\kappa_j(T)$ the j-th cumulant of some random variable T.

2 Background

We give a brief review about the precise meaning of consistency of estimators and model correctness for marginal composite likelihoods (Sect. 2.1). These concepts are crucial to frame properly the differences that arise at first- and second-order between composite likelihood and likelihood ratios (Sect. 2.2) and provide the suitable environment for our developments.

2.1 Model correctness and consistency of estimators

The definition of model correctness for marginal composite likelihoods is termed to as marginal correct specification by Xu and Reid (2011), i.e. $g(y \in \mathcal{E}_k) = f(y \in \mathcal{E}_k;\theta')$ for all $k = 1, \dots, K$ and for some $\theta' \in \operatorname{int}(\Theta)$. This definition is weaker than the usual one of model correctness $g(y) = f(y;\theta_0)$, $\theta_0 \in \operatorname{int}(\Theta)$, because the latter involves q-dimensional densities.

The maximum composite likelihood estimator $\hat\theta$ is root-n consistent for the pseudo true parameter value θ^*, which is defined as the minimiser of the composite Kullback-Leibler divergence

(Varin and Vidoni, 2005)

$$\mathbb{E}_g \left\{ \sum_{k=1}^{K} w_k \left[\log g(Y \in \mathcal{E}_k) - \log f(Y \in \mathcal{E}_k; \theta) \right] \right\},$$

where $\mathbb{E}_g$ denotes expectation with respect to $g(y)$. If it holds $g(y) = f(y; \theta_0)$ and further $g(y \in \mathcal{E}_k) = f(y \in \mathcal{E}_k; \theta_0)$, all k, then we also have $\theta^* = \theta_0$; this implies that $\hat{\theta}$ converges in probability to the true parameter value even under the marginal correct specification (Xu and Reid, 2011). Nonetheless, as our results are not tied to such circumstance, we hereafter must assume that θ^* still has a meaningful scientific interpretation because it is the only quantity for which we may conduct inference. We remark in passing that when $\ell(\theta)$ is the ordinary likelihood function our setting recovers the more familiar theory of misspecified likelihoods developed by Kent (1982) and White (1982).

To ease the notation, in the sequel we drop the dependence on the parameter whenever quantities defined as functions of θ are evaluated at θ^*, e.g. $W \equiv W(\theta^*)$.

2.2 Bartlett identities and first- and second-order asymptotics for W

Bartlett identities regard expected balancing relations involving moments of likelihood derivatives and hold for the log likelihood under model correctness $g(y) = f(y; \theta_0)$ (see, e.g., Barndorff-Nielsen and Cox, 1994, pp 146-147). For our purposes it suffices to consider the first four identities only, which are respectively (reading from top to bottom and left to right)

$$\begin{array}{ll} \alpha_1(\theta_0) = 0 & \alpha_{001}(\theta_0) + 3\alpha_{11}(\theta_0) + \alpha_3(\theta_0) = 0 \\ \alpha_2(\theta_0) - \alpha_{01}(\theta_0) = 0 & \alpha_{0001}(\theta_0) + 4\alpha_{101}(\theta_0) + 3\alpha_{02}(\theta_0) + 6\alpha_{21}(\theta_0) + \alpha_4(\theta_0) = 0. \end{array}$$

Since marginal composite log likelihoods are formed by the sum of n contributions that do not necessarily originate from proper density functions, such identities, but the first, do not hold even under the marginal correct specification. The first identity is still valid regardless such condition, i.e. $\alpha_1 \equiv \alpha_1(\theta^*) = 0$, as can be deduced from Section 2.1.

Because some identities do not hold, the properties of $W(\theta)$ depart remarkably from those of the log likelihood ratio. The differences are here outlined by referring to formal Edgeworth series for the density of $n^{1/2}R(\theta)$. The latter is the signed square root of $W(\theta)$, i.e. a random variable chosen to fulfil $W = nR^2 + O_p(n^{-3/2})$. It is understood that the desired properties of W are derived from the density of $n^{1/2}R$ by using transformation rules of random variables. From the expansion of W in the Appendix 1, we have $R = R_1 + R_2 + R_3$, with $R_j = O_p(n^{-j/2})$, $j = 1, 2, 3$, where

$$\begin{aligned} R_1 &= \frac{A_1}{\alpha_{01}^{1/2}} \qquad R_2 = \frac{A_1 A_{01}}{2\alpha_{01}^{3/2}} + \frac{\alpha_{001} A_1^2}{6\alpha_{01}^{5/2}} \\ R_3 &= \frac{3A_1 A_{01}^2}{8\alpha_{01}^{5/2}} + \frac{A_1^2 A_{001}}{6\alpha_{01}^{5/2}} + \frac{5\alpha_{001} A_1^2 A_{01}}{12\alpha_{01}^{7/2}} + \frac{\alpha_{001}^2 A_1^3}{9\alpha_{01}^{9/2}} + \frac{\alpha_{0001} A_1^3}{24\alpha_{01}^{7/2}}. \end{aligned}$$

The leading terms of the cumulants of $n^{1/2}R$ are

$$\begin{array}{lll} \kappa_1(n^{1/2}R) = O(n^{-1/2}) & \kappa_2(n^{1/2}R) = \alpha_2\alpha_{01}^{-1} + O(n^{-1}) & \kappa_3(n^{1/2}R) = O(n^{-1/2}) \\ \kappa_4(n^{1/2}R) = O(n^{-1}) & \kappa_5(n^{1/2}R) = O(n^{-3/2}) & \kappa_j(n^{1/2}R) = o(n^{-2}), j \geq 6. \end{array} \tag{2}$$

First-order behaviour of W may be assessed by constructing a series for the density of $n^{1/2}R$ based on the leading term of $\kappa_2(n^{1/2}R)$. Because of the failure of the second Bartlett identity such term is not equal to 1, consequently W is not asymptotically chi-square distributed as the log likelihood ratio. It follows

$$W \xrightarrow{d} \alpha_2 \alpha_{01}^{-1} Z^2,$$

with $Z \sim N(0,1)$ (see Kent, 1982). The same first-order limiting behaviour of the log likelihood ratio may be restored by using suitable modifications to $W(\theta)$, as suggested by Rotnitzky and Jewell (1990) and Pace *et al.* (2011). When θ is scalar these adjustments coincide and result in a modified statistic of the form $W_1(\theta) = \alpha_2(\theta)^{-1}\alpha_{01}(\theta)W(\theta)$. A further adjustment is provided by Chandler and Bate (2007) and the purpose is to modify the curvature of the composite log likelihood about $\hat{\theta}$ by defining $\ell_{cb}(\theta) = \ell(\theta_{cb}(\theta))$, with $\theta_{cb}(\theta) = \hat{\theta} - (\hat{\theta} - \theta)C_1$. The associated composite log likelihood ratio $W_{cb}(\theta) = 2\{\ell_{cb}(\hat{\theta}) - \ell_{cb}(\theta)\}$ achieves the desired limit if $C_1 = \alpha_2(\theta)^{-1/2}\alpha_{01}(\theta)^{1/2}$ (Chandler and Bate, 2007, Sect. 3·2).

For the second-order properties of W, and in particular to enquire about the Bartlett correction, we need to develop a series for the density of $n^{1/2}R$ up to $O(n^{-3/2})$. If W was the log likelihood ratio, then

$$\begin{array}{lll} \kappa_1(n^{1/2}R) = O(n^{-1/2}) & \kappa_2(n^{1/2}R) = 1 + O(n^{-1}) & \kappa_3(n^{1/2}R) = O(n^{-3/2}) \\ \kappa_4(n^{1/2}R) = O(n^{-2}) & \kappa_5(n^{1/2}R) = O(n^{-3/2}) & \kappa_j(n^{1/2}R) = o(n^{-2}), j \geq 6, \end{array} \quad (3)$$

where the second, third, and fourth cumulant are different from those in (2) due to the validity of the second, third, and fourth Bartlett identities. This mean that the series for $n^{1/2}R$ can be based on $\kappa_1(n^{1/2}R)$ and $\kappa_2(n^{1/2}R)$ only. Computation of the cumulants of W leads to

$$\kappa_j(W) = 2^{j-1}(j-1)!\,[\mathbb{E}_g W]^j + O(n^{-3/2}),$$

where $2^{j-1}(j-1)!$ is the j-th cumulant of a chi-square variate with one degree of freedom. Standard properties of cumulants suggest that division of W by its expectation results in (see, e.g., Barndorff-Nielsen and Cox, 1994, Ch. 5)

$$\mathrm{P}\left\{W[\mathbb{E}_g W]^{-1} \leq c_\gamma\right\} = \gamma + O(n^{-2}),$$

where c_γ is the γ-quantile of a chi-square variate with one degree of freedom. The expectation of W admits the expansion $1 + n^{-1}b + O(n^{-2})$, where b is the Bartlett factor, provided, for instance, in Barndorff-Nielsen and Cox (1994, formula 5·30). When the composite log likelihood ratio is considered, then the required Bartlett identities are not satisfied, whereby the cumulants of its signed root do not exhibit the structure in (3), implying that it is not Bartlett-correctable. This is also the case for W_1 and W_{cb} as the adjustments do not account for the third and fourth Bartlett identities.

3 Second-order accuracy via the extended curvature adjustment

To establish our results in the present section, we assume conditions (A0)-(A7) in Xu and Reid (2011) for the consistency of $\hat{\theta}$ and conditions (A1)-(A5) in Jensen (1993, Sect. 1·1). Contextualised to our framework, the latter regard moment and smoothness conditions of composite likelihood derivatives that are necessary to ensure the validity of the Edgeworth expansion for the density of the signed root given in (4). All proofs are deferred to the Appendix 2.

3.1 Expected extended curvature adjustment and Bartlett factor

In order to account for the failure of the second, third, and fourth Bartlett identities for marginal composite likelihoods and to supply a version of W which is Bartlett-correctable, we generalise the approach by Chandler and Bate (2007) as follows. We define $\ell_e(\theta) = \ell(\theta_e(\theta))$ along with $W_e(\theta) = 2\{\ell_e(\hat{\theta}) - \ell_e(\theta)\}$, where

$$\theta_e(\theta) = \hat{\theta} - \sum_{j=1}^{3}(\hat{\theta} - \theta)^j C_j$$

provides what we term to as the extended curvature adjustment, $C_j = O(1)$, $j = 1, 2, 3$. Clearly $\hat{\theta} = \mathrm{argmax}_\theta\, \ell(\theta) = \mathrm{argmax}_\theta\, \ell_e(\theta)$. Provided the expansion of W_e in (6), we have $W_e = nR_e^2 + O_p(n^{-3/2})$, $R_e = R_{e1} + R_{e2} + R_{e3}$, with $R_{ej} = O_p(n^{-j/2}), j = 1, 2, 3$, and

$$\begin{aligned}
R_{e1} &= \frac{A_1 C_1}{\alpha_{01}^{1/2}} \qquad R_{e2} = \frac{C_1 A_1 A_{01}}{2\alpha_{01}^{3/2}} + \frac{C_2 A_1^2}{\alpha_{01}^{3/2}} + \frac{\alpha_{001} C_1^2 A_1^2}{6\alpha_{01}^{5/2}} \\
R_{e3} &= \frac{3C_1 A_1 A_{01}^2}{8\alpha_{01}^{5/2}} + \frac{C_1^2 A_1^2 A_{001}}{6\alpha_{01}^{5/2}} + \frac{3C_2 A_1^2 A_{01}}{2\alpha_{01}^{5/2}} + \frac{5\alpha_{001} C_1^2 A_1^2 A_{01}}{12\alpha_{01}^{7/2}} + \frac{\alpha_{001} C_2 A_1^3}{2\alpha_{01}^{7/2}} + \frac{\alpha_{001} C_1 C_2 A_1^3}{3\alpha_{01}^{7/2}} + \\
&+ \frac{C_3 A_1^3}{\alpha_{01}^{5/2}} - \frac{\alpha_{001}^2 C_1 A_1^3}{8\alpha_{01}^{9/2}} + \frac{\alpha_{001}^2 C_1^2 A_1^3}{4\alpha_{01}^{9/2}} - \frac{\alpha_{001}^2 C_1^3 A_1^3}{72\alpha_{01}^{9/2}} - \frac{\alpha_{0001} C_1 A_1^3}{12\alpha_{01}^{7/2}} + \frac{\alpha_{0001} C_1^2 A_1^3}{6\alpha_{01}^{7/2}} - \frac{\alpha_{0001} C_1^3 A_1^3}{24\alpha_{01}^{7/2}}.
\end{aligned} \tag{4}$$

The key idea to enable the Bartlett correction for W_e is to use the constants C_1, C_2, and C_3 to act on the cumulants of $n^{1/2}R_e$, given in (7)-(11), so that they achieve the same structure of those in (3), i.e. the ones resulting from the signed root of the log likelihood ratio. Specifically, C_1 is employed to obtain $\kappa_2(n^{1/2}R_e) = 1 + O(n^{-1})$, whereas C_2 and C_3 are tuned to downsize $\kappa_3(n^{1/2}R_e)$ and $\kappa_4(n^{1/2}R_e)$ to $O(n^{-3/2})$ and $O(n^{-2})$, respectively. In the following theorem we provide expressions for C_j, as well as that for the resulting Bartlett factor for W_e, $j = 1, 2, 3$.

Theorem 1. *Let $W_e = W(\theta_e)$ and $\theta_e \equiv \hat{\theta} - \sum_{j=1}^{3}(\hat{\theta} - \theta^*)^j C_j$, with*

$$\begin{aligned}
C_1 &= \alpha_2^{-1/2}\alpha_{01}^{1/2} \qquad C_2 = -\frac{C_1\alpha_{11}}{2\alpha_2} - \frac{C_1\alpha_3\alpha_{01}}{6\alpha_2^2} - \frac{C_1^2\alpha_{001}}{6\alpha_{01}} \\
C_3 &= -\frac{2C_2^2}{C_1} - \frac{3C_1\alpha_{11}^2}{4\alpha_2^2} - \frac{7C_2\alpha_{11}}{2\alpha_2} - \frac{C_1^2\alpha_{101}}{6\alpha_2} - \frac{C_1\alpha_{21}\alpha_{01}}{4\alpha_2^2} - \frac{C_1\alpha_{11}\alpha_3\alpha_{01}}{4\alpha_2^3} - \frac{C_2\alpha_3\alpha_{01}}{\alpha_2^2} + \\
&- \frac{C_1\alpha_4\alpha_{01}^2}{24\alpha_2^3} - \frac{C_1\alpha_{02}}{8\alpha_2} - \frac{C_1^2\alpha_3\alpha_{001}}{6\alpha_2^2} - \frac{C_2\alpha_{001}}{2\alpha_{01}} - \frac{C_1C_2\alpha_{001}}{\alpha_{01}} - \frac{3C_1^2\alpha_{11}\alpha_{001}}{4\alpha_2\alpha_{01}} + \frac{C_1\alpha_{001}^2}{8\alpha_{01}^2} + \\
&- \frac{C_1^2\alpha_{001}^2}{4\alpha_{01}^2} - \frac{C_1^3\alpha_{001}^2}{24\alpha_{01}^2} + \frac{C_1\alpha_{0001}}{12\alpha_{01}} - \frac{C_1^2\alpha_{0001}}{6\alpha_{01}} + \frac{C_1^3\alpha_{0001}}{24\alpha_{01}},
\end{aligned}$$

then $P\left\{W_e[1 + n^{-1}b_e]^{-1} \leq c_\gamma\right\} = \gamma + O(n^{-2})$, where b_e is the Bartlett factor for W_e whose expression is

$$b_e = \frac{5\alpha_3^2}{12\alpha_2^3} - \frac{\alpha_4}{4\alpha_2^2} - \frac{\alpha_{11}^2}{4\alpha_2\alpha_{01}^2} - \frac{\alpha_{21}}{2\alpha_2\alpha_{01}} + \frac{\alpha_{11}\alpha_3}{2\alpha_2^2\alpha_{01}} + \frac{\alpha_{02}}{4\alpha_{01}^2}.$$

Because the class of marginal composite likelihoods include as a special instance the likelihood, the result in Theorem 1 may be also framed in the likelihood setting. Here it can be interpreted as a device to achieve a robust Bartlett correction whenever the researcher is not confident about the validity of the required Bartlett identities or, equivalently, about the correctness of model assumptions. Note that when Bartlett identities hold, then Theorem 1 retrieves $C_1 = 1$, $C_2 = C_3 = 0$, and the Bartlett factor b_e reduces to the one of the likelihood ratio.

Should it be considered in the composite likelihood or likelihood framework, the result in Theorem 1 provides a striking description of a general-purpose adjustment to manipulate first- and second-order asymptotic properties of composite likelihood and likelihood ratios. Nevertheless, it is pointless from a practical point of view because $\theta_e(\theta)$ still depends on the unknown $g(y)$ through expected moments of likelihood derivatives.

3.2 Observed extended curvature adjustment and Bartlett factor

The statistic W_e depends on expected moments α_{rstu} in C_j, $j = 1, 2, 3$, and whenever they are replaced by their root-n consistent estimates

$$\hat{\alpha}_{rstu} = \hat{\alpha}_{rstu}(\hat{\theta}) = \nu\, n^{-1} \sum_{i=1}^{n} [\ell_1(\hat{\theta}; y_i)]^r [\ell_2(\hat{\theta}; y_i)]^s [\ell_3(\hat{\theta}; y_i)]^t [\ell_4(\hat{\theta}; y_i)]^u, \tag{5}$$

the result in Theorem 1 is struck down. A brief explanation is as follows. Let $\hat{C}_1$ be the empirical counterpart of C_1 in Theorem 1, i.e. expected moments are replaced by (5). Then it follows $\hat{C}_1 = C_1 + r_1 + r_2$, where $r_1 = O_p(n^{-1/2})$ and $r_2 = O_p(n^{-1})$ are given in (12) and (13). When $\hat{C}_1$ is plugged in R_{1e} and R_{2e} it produces disturbances of size $O_p(n^{-1})$ and $O_p(n^{-3/2})$ that modify the current expressions of R_{2e} and R_{3e}, respectively. This implies that C_2 and C_3 need to the be updated. Similarly, once a new expression for C_2 is retrieved, the estimation process gives rise to an error term that affects the expression of R_{3e}. Note that estimation of C_3 does not alter R_{3e} because the induced reminder is $O_p(n^{-2})$.

In order to cope with these difficulties, we define a revised version of $W_e(\theta)$, namely $W'_e(\theta) = 2\{\ell'_e(\hat{\theta}) - \ell'_e(\theta)\}$, where $\ell'_e(\theta) = \ell(\theta'_e(\theta))$ and $\theta'_e(\theta) = \hat{\theta} - \sum_{j=1}^{3} (\hat{\theta} - \theta)^j \hat{C}_j$. The function $W'_e(\theta)$ is suitable for practical purposes because in Theorem 2 we provide expressions for $\hat{C}_1, \hat{C}_2$, and $\hat{C}_3$ which are derived by taking into an account the estimation error of expected moments and are readily provided in terms of sample moments.

Theorem 2. *Let $W'_e = W(\theta'_e)$ and $\theta'_e \equiv \hat{\theta} - \sum_{j=1}^{3} (\hat{\theta} - \theta^*)^j \hat{C}_j$ with*

$$\begin{aligned}
\hat{C}_1 &= \hat{\alpha}_2^{-1/2} \hat{\alpha}_{01}^{1/2} \qquad \hat{C}_2 = \frac{\hat{C}_1 \hat{\alpha}_{11}}{\hat{\alpha}_2} + \frac{\hat{C}_1 \hat{\alpha}_3 \hat{\alpha}_{01}}{3\hat{\alpha}_2^2} + \frac{\hat{C}_1 \hat{\alpha}_{001}}{2\hat{\alpha}_{01}} - \frac{\hat{C}_1^2 \hat{\alpha}_{001}}{6\hat{\alpha}_{01}} \\
\hat{C}_3 &= -\frac{2\hat{C}_2^2}{\hat{C}_1} + \frac{\hat{C}_1 \hat{\alpha}_{11}^2}{\hat{\alpha}_2^2} + \frac{5\hat{C}_2 \hat{\alpha}_{11}}{2\hat{\alpha}_2} - \frac{\hat{C}_1 \hat{\alpha}_{101}}{2\hat{\alpha}_2} - \frac{\hat{C}_1 \hat{\alpha}_{21} \hat{\alpha}_{01}}{\hat{\alpha}_2^2} + \frac{7\hat{C}_1 \hat{\alpha}_{11} \hat{\alpha}_3 \hat{\alpha}_{01}}{6\hat{\alpha}_2^3} + \frac{\hat{C}_2 \hat{\alpha}_3 \hat{\alpha}_{01}}{\hat{\alpha}_2^2} + \\
&+ \frac{\hat{C}_1 \hat{\alpha}_3^2 \hat{\alpha}_{01}^2}{3\hat{\alpha}_2^4} - \frac{\hat{C}_1 \hat{\alpha}_4 \hat{\alpha}_{01}^2}{4\hat{\alpha}_2^3} - \frac{\hat{C}_1 \hat{\alpha}_{02}}{2\hat{\alpha}_2} + \frac{\hat{C}_1^2 \hat{\alpha}_3 \hat{\alpha}_{001}}{6\hat{\alpha}_2^2} + \frac{3\hat{C}_2 \hat{\alpha}_{001}}{2\hat{\alpha}_{01}} - \frac{\hat{C}_1 \hat{C}_2 \hat{\alpha}_{001}}{\hat{\alpha}_{01}} - \frac{\hat{C}_1 \hat{\alpha}_{11} \hat{\alpha}_{001}}{4\hat{\alpha}_2 \hat{\alpha}_{01}} + \\
&+ \frac{5\hat{C}_1^2 \hat{\alpha}_{11} \hat{\alpha}_{001}}{12\hat{\alpha}_2 \hat{\alpha}_{01}} - \frac{\hat{C}_1 \hat{\alpha}_{001}^2}{4\hat{\alpha}_{01}^2} + \frac{\hat{C}_1^2 \hat{\alpha}_{001}^2}{4\hat{\alpha}_{01}^2} - \frac{\hat{C}_1^3 \hat{\alpha}_{001}^2}{24\hat{\alpha}_{01}^2} - \frac{\hat{C}_1 \hat{\alpha}_{0001}}{6\hat{\alpha}_{01}} + \frac{\hat{C}_1^3 \hat{\alpha}_{0001}}{24\hat{\alpha}_{01}},
\end{aligned}$$

then $P\left\{W'_e[1+n^{-1}b'_e]^{-1} \leq c_\gamma\right\} = \gamma + O(n^{-2})$, *where* b'_e *is the Bartlett factor for* W'_e *whose expression is*

$$b'_e = \frac{\alpha_4}{2\alpha_2^2} - \frac{\alpha_3^2}{3\alpha_2^3}.$$

The result in Theorem 2 is still valid when b_e is replaced by its root-n consistent estimate $\hat{b}_e$ computed with sample moments $\hat{\alpha}_j$, $j = 2, 3, 4$. We highlight that the Bartlett factor for W'_e depends on the standardised third and fourth moments of the composite score function only. Incidentally, it is equal to that for the empirical likelihood, with the difference that standardised moments appearing in the latter are those of Y (DiCiccio *et al.*, 1991).

4 Empirical evidence

In the sequel an example dealing with marginal pairwise likelihoods is considered to assess, via Monte Carlo simulation, the coverage accuracy of confidence intervals for θ based on $W_e^b(\theta) = W_e(\theta)[1+n^{-1}b_e]^{-1}$, $W_1(\theta)$, $W_{cb}(\theta)$, $W_o(\theta)$, $W_{eo}^b(\theta) = W_{eo}(\theta)[1+n^{-1}b_e]^{-1}$. The latter two are the ordinary likelihood ratio and its robust Bartlett-corrected version computed with the extended curvature adjustment and Bartlett factor provided in Theorem 1. We also consider the following versions of the aforementioned statistics: $W_e'^b(\theta) = W'_e(\theta)[1+n^{-1}\hat{b}'_e]^{-1}$, $W'_1(\theta)$, $W'_{cb}(\theta)$, and $W_{eo}'^b(\theta) = W'_{eo}(\theta)[1+n^{-1}\hat{b}'_e]^{-1}$, where the second and third are the analogues of $W_1(\theta)$ and $W_{cb}(\theta)$ computed by replacing expected moments in the adjustments by empirical moments (5), whereas the fourth is the analogue of $W_{eo}^b(\theta)$ computed according to the quantities given in Theorem 2. For the computation of $W_e'^b(\theta)$, $W'_{cb}(\theta)$, and $W_{eo}'^b(\theta)$ we use a bias-corrected version of $\hat{C}_1$, namely $\hat{C}_1^{bc} = \hat{C}_1 - \hat{\mathbb{E}}_g[r_2]$, where $\hat{\mathbb{E}}_g[r_2]$ is the sample counterpart of $\mathbb{E}_g[r_2]$, without affecting the validity of Theorem 2. The resulting expression for $\hat{C}_1^{bc}$ is

$$\begin{aligned}\hat{C}_1^{bc} &= \hat{C}_1 - \frac{1}{n}\left[\frac{3\hat{C}_1\hat{\alpha}_4}{8\hat{\alpha}_2^2} + \frac{\hat{C}_1\hat{\alpha}_{11}^2}{\hat{\alpha}_2\hat{\alpha}_{01}^2} - \frac{\hat{C}_1\hat{\alpha}_{101}}{\hat{\alpha}_{01}^2} + \frac{3\hat{C}_1\hat{\alpha}_{21}}{4\hat{\alpha}_2\hat{\alpha}_{01}} - \frac{3\hat{C}_1\hat{\alpha}_{11}\hat{\alpha}_3}{2\hat{\alpha}_2^2\hat{\alpha}_{01}} - \frac{5\hat{C}_1\hat{\alpha}_{02}}{8\hat{\alpha}_{01}^2} + \right. \\ &\quad + \left. \frac{3\hat{C}_1\hat{\alpha}_{11}\hat{\alpha}_{001}}{4\hat{\alpha}_{01}^3} + \frac{\hat{C}_1\hat{\alpha}_3\hat{\alpha}_{001}}{4\hat{\alpha}_2\hat{\alpha}_{01}^2} - \frac{3\hat{C}_1\hat{\alpha}_2\hat{\alpha}_{001}^2}{8\hat{\alpha}_{01}^4} + \frac{\hat{C}_1\hat{\alpha}_2\hat{\alpha}_{0001}}{4\hat{\alpha}_{01}^3}\right].\end{aligned}$$

Similarly, for $W'_1(\theta)$ we adopt a bias-corrected version of the scaling factor $\hat{\alpha}_2^{-1}\hat{\alpha}_{01}$, whose expression is

$$\frac{\hat{\alpha}_{01}}{\hat{\alpha}_2} - \frac{1}{n}\left[\frac{\hat{\alpha}_{21}}{\hat{\alpha}_2^2} - \frac{4\hat{\alpha}_{11}\hat{\alpha}_3}{\hat{\alpha}_2^3} + \frac{4\hat{\alpha}_{11}^2}{\hat{\alpha}_2^2\hat{\alpha}_{01}} - \frac{2\hat{\alpha}_{101}}{\hat{\alpha}_2\hat{\alpha}_{01}} + \frac{\hat{\alpha}_4\hat{\alpha}_{01}}{\hat{\alpha}_2^3} - \frac{\hat{\alpha}_{02}}{\hat{\alpha}_2\hat{\alpha}_{01}} + \frac{\hat{\alpha}_3\hat{\alpha}_{001}}{\hat{\alpha}_2^2\hat{\alpha}_{01}} - \frac{\hat{\alpha}_{001}^2}{2\hat{\alpha}_{01}^3} - \frac{\hat{\alpha}_{0001}}{2\hat{\alpha}_{01}^2}\right]$$

The number of Monte Carlo trials is 100000 and expected moments of likelihood derivatives, needed to compute the expected extended curvature adjustment and associated Bartlett factor in Theorem 1, are approximated via an auxiliary simulation of 10000 replicates.

The `R` source code of a function that computes $W'_e(\theta)$ and $W_e'^b(\theta)$ for an arbitrary log likelihood function is available from the Author.

4.1 Marginal pairwise likelihood

Suppose that $y_1, \ldots, y_n$ is a sample from a q-dimensional normal distribution with null vector of means and covariance matrix Σ whose diagonal and off-diagonal elements are 1 and $\rho = \mathrm{cor}(Y_j, Y_k)$, $j \neq k = 1, \ldots, q$, $\rho \in (-(q-1)^{-1}, 1)$, respectively. The log likelihood and marginal pairwise log likelihood for ρ admit an analytic expression, and for the latter is (Cox and Reid, 2004)

$$\ell(\rho) = -\frac{nq(q-1)}{4}\log(1-\rho^2) - \frac{q-1+\rho}{2(1-\rho^2)}SS_W - \frac{(q-1)(1-\rho)}{2q(1-\rho^2)}SS_B,$$

where $SS_W = \sum_{i=1}^n \sum_{j=1}^q (y_{ij} - \bar{y}_i)^2$, $SS_B = q^2 \sum_{i=1}^n \bar{y}_i^2$, and $\bar{y}_i = q^{-1}\sum_{j=1}^q y_{ij}$.

Simulations are from the true model (multivariate normal) and from a misspecified model, i.e. a multivariate t_τ distribution with $\tau = 10$ degrees of freedom. In the first case, our aim is to validate the results in Section 3 for pairwise likelihoods and to assess the behaviour of the likelihood when we are too cautious and misuse the extended curvature adjustments along with the related Bartlett factors. Note that the pairwise likelihood is correctly specified, in the sense of Section 2.1. In the second case, the purpose is to assay the ability of the proposed methodology to retain the stability, also to second-order, of levels of confidence intervals against misspecification. In this case, neither the pairwise nor the likelihood are correctly specified.

We consider samples of size $n \in \{15, 30\}$ and $\rho \in \{0{\cdot}2,\ 0{\cdot}5,\ 0{\cdot}9\}$. The results for the first and second setting discussed above are in Table 1 and Table 2, respectively. For the former, we have that empirical coverages resulting from $W_o(\rho)$, $W_1(\rho)$, $W_{cb}(\rho)$, $W_e^b(\rho)$, and $W_{eo}^b(\rho)$ compare similarly and are close to the nominal levels. When adjustments are estimated, second-order accurate statistics $W_e'^b(\rho)$ and $W_{oe}'^b(\rho)$ outperforms $W_1'(\rho)$ and $W_{cb}'(\rho)$. The results for $W_{oe}^b(\rho)$ and $W_{oe}'^b(\rho)$ are slightly worse than those of $W_o(\rho)$ but still comparable, meaning that the use of the extended curvature adjustments do not harm substantially coverage accuracy. When we consider the simulation from the t_{10} distribution, we have a different picture than before. On the one hand, coverages from $W_o(\rho)$ drop dramatically, highlighting that the likelihood ratio is overwhelmed by the model misspecification. On the other hand, the expected adjustments for $W_1(\rho)$, $W_{cb}(\rho)$, $W_e^b(\rho)$, and $W_{eo}^b(\rho)$ are able to fix for the misspecification and lead to sensible coverages. Once again $W_e'^b(\rho)$ and $W_{oe}'^b(\rho)$ provide better results than $W_1'(\rho)$ and $W_{cb}'(\rho)$.

Appendix 1

Expansion of W_e and W

To obtain the expansion of W_e to $O_p(n^{-3/2})$ we need that of $\hat\theta - \theta^*$ to the same order, which may be found, for instance, in Barndorff-Nielsen and Cox (1994, p. 150), along with the first four derivatives of $\theta_e(\theta)$ and $\ell_e(\theta) = \ell(\theta_e(\theta))$. Let $\theta_{ej}(\theta) = \partial^j \theta_e(\theta)/\partial\theta^j$ and $\ell_{ej}(\theta) = \partial^j \ell_e(\theta)/\partial\theta^j$. It follows $\theta_{ej}(\theta) = \sum_{t=j}^3 (-1)^{j+1} t![(t-j)!]^{-1}(\hat\theta - \theta)^{t-j} C_j$, $\theta_{e4}(\theta) = 0$, and

$$\begin{aligned}
\ell_{e1}(\theta) &= \ell_1(\theta_e(\theta))\theta_{e1}(\theta),\\
\ell_{e2}(\theta) &= \ell_2(\theta_e(\theta))\theta_{e1}^2(\theta) + \ell_1(\theta_e(\theta))\theta_{e2}(\theta),\\
\ell_{e3}(\theta) &= \ell_3(\theta_e(\theta))\theta_{e1}^3(\theta) + 3\ell_2(\theta_e(\theta))\theta_{e1}(\theta)\theta_{e2}(\theta) + \ell_1(\theta_e(\theta))\theta_{e3}(\theta),\\
\ell_{e4}(\theta) &= \ell_4(\theta_e(\theta))\theta_{e1}^4(\theta) + 6\ell_3(\theta_e(\theta))\theta_{e1}^2(\theta)\theta_{e2}(\theta) + \ell_2(\theta_e(\theta))[3\theta_{e2}^2(\theta) + 4\theta_{e1}(\theta)\theta_{e3}(\theta)].
\end{aligned}$$

Table 1: Empirical coverage probabilities for confidence intervals for ρ when simulation is from the multivariate normal distribution. Monte Carlo standard errors for nominal levels $\{90, 95, 99\}$ per cent are $\{0{\cdot}09, 0{\cdot}07, 0{\cdot}03\}$, respectively

n	ρ	Level	W_o	W_1	W_{cb}	W_e^b	W_{eo}^b	W_1'	W_{cb}'	$W_e'^b$	$W_{eo}'^b$
15	0·2	90	89·8	90·6	89·7	90·2	90·7	85·9	85·1	87·4	90·4
		95	94·8	95·5	94·7	95·1	95·4	90·7	90·0	92·1	94·6
		99	98·9	99·1	98·7	99·0	99·1	95·5	95·3	96·4	98·0
	0·5	90	90·0	90·1	89·9	90·4	90·6	87·7	87·4	88·7	90·8
		95	95·0	95·0	94·9	95·2	95·4	92·6	92·6	93·4	94·9
		99	99·0	99·0	98·7	99·0	99·1	97·4	97·5	97·6	98·5
	0·9	90	90·1	89·9	89·7	90·8	90·7	89·7	89·1	90·7	90·7
		95	95·0	94·9	94·6	95·4	95·4	94·3	93·8	95·0	95·0
		99	99·0	98·9	98·8	99·1	99·1	98·3	98·0	98·5	98·5
30	0·2	90	89·9	90·1	89·1	89·5	90·1	88·1	87·6	89·4	90·4
		95	94·9	95·1	94·4	94·6	95·0	92·8	92·5	94·1	95·2
		99	99·0	99·0	98·7	98·8	99·0	97·2	97·1	98·1	98·9
	0·5	90	90·0	89·9	89·4	89·8	90·1	89·1	88·9	89·9	90·7
		95	95·0	95·0	94·7	94·8	95·1	93·9	93·9	94·6	95·4
		99	99·0	99·0	98·7	98·9	99·0	98·2	98·3	98·5	99·0
	0·9	90	89·9	89·7	89·4	89·9	90·0	89·8	89·5	90·6	90·7
		95	94·9	94·9	94·6	95·1	95·0	94·7	94·4	95·4	95·4
		99	99·0	99·0	98·9	99·1	99·0	98·8	98·6	99·0	99·0

The Taylor expansion of $\ell_e(\theta^*)$ about $\hat\theta$ yields

$$
\begin{aligned}
W_e &= (\theta^*-\hat\theta)^2\ell_{e2}(\hat\theta)+\frac{1}{3}(\theta^*-\hat\theta)^3\ell_{e3}(\hat\theta)+\frac{1}{12}(\theta^*-\hat\theta)^4\ell_{e4}(\hat\theta)+\ldots \qquad (6)\\
&= nC_1^2\left\{\frac{A_1^2}{\alpha_{01}}+\frac{A_1^2A_{01}}{\alpha_{01}^2}+\frac{2A_1^3C_2}{\alpha_{01}^2C_1}+\frac{A_1^3C_1\alpha_{001}}{3\alpha_{01}^3}+\frac{A_1^2A_{01}^2}{\alpha_{01}^3}+\frac{A_1^3A_{001}C_1}{3\alpha_{01}^3}+\frac{4A_1^3A_{01}C_2}{\alpha_{01}^3C_1}+\right.\\
&+\frac{A_1^4C_2^2}{\alpha_{01}^3C_1^2}+\frac{2A_1^4C_3}{\alpha_{01}^3C_1}+\frac{A_1^3A_{01}C_1\alpha_{001}}{\alpha_{01}^4}+\frac{A_1^4C_2\alpha_{001}}{\alpha_{01}^4C_1}+\frac{A_1^4C_2\alpha_{001}}{\alpha_{01}^4}-\frac{A_1^4\alpha_{001}^2}{4\alpha_{01}^5}+\frac{A_1^4C_1\alpha_{001}^2}{2\alpha_{01}^5}+\\
&-\left.\frac{A_1^4\alpha_{0001}}{6\alpha_{01}^4}+\frac{A_1^4C_1\alpha_{0001}}{3\alpha_{01}^4}-\frac{A_1^4C_1^2\alpha_{0001}}{12\alpha_{01}^4}\right\}+O_p(n^{-3/2}).
\end{aligned}
$$

The expansion for W are readily recovered from that of W_e by setting $C_1=1$ and $C_2=C_3=0$.

The signed roots R_e and R, and their cumulants

The signed root $n^{1/2}R_e$ is derived by matching the expansion (6) order by order. Write $W_e = W_{e1}+W_{e2}+W_{e3}+O_p(n^{-3/2})$, where $W_{ej}=O_p(n^{-(j-1)/2})$, $j=1,2,3$, then it suffices to solve for R_{e1}, R_{e2}, and R_{e3} the equations $R_{e1}^2=n^{-1}W_{e1}$, $2R_{e1}R_{e2}=n^{-1}W_{e2}$, and $2R_{e1}R_{e3}-R_{e2}^2=$

Table 2: Empirical coverage probabilities for confidence intervals for ρ when simulation is from the multivariate t_{10} distribution. Monte Carlo standard errors for nominal levels $\{90, 95, 99\}$ per cent are $\{0{\cdot}09, 0{\cdot}07, 0{\cdot}03\}$, respectively

n	ρ	Level	W_o	W_1	W_{cb}	W_e^b	W_{eo}^b	W_1'	W_{cb}'	$W_e'^b$	$W_{eo}'^b$
15	0·2	90	59·6	90·9	88·7	89·7	90·5	84·0	83·0	87·3	86·9
		95	68·0	95·5	93·9	94·5	95·4	88·8	88·3	91·7	91·5
		99	81·2	98·9	98·1	98·6	99·0	94·5	94·2	95·7	95·9
	0·5	90	52·3	90·3	88·7	90·0	89·9	85·2	84·9	88·1	87·2
		95	60·4	95·1	93·8	94·7	94·8	90·8	90·0	92·9	92·0
		99	73·7	98·8	98·3	98·6	98·6	95·6	95·5	97·1	96·6
	0·9	90	50·9	90·3	88·3	90·0	90·0	85·5	85·3	87·2	87·3
		95	58·8	95·0	93·7	94·7	94·7	90·7	90·0	92·1	92·1
		99	72·0	98·7	98·3	98·6	98·6	95·9	95·2	96·8	96·8
30	0·2	90	59·4	90·1	88·1	88·6	89·2	86·5	86·4	89·1	89·2
		95	67·8	95·1	93·6	94·0	94·3	91·7	91·4	93·8	93·9
		99	80·8	98·9	98·3	98·5	98·7	97·0	96·7	98·0	97·9
	0·5	90	51·7	89·9	88·5	89·2	89·0	87·7	87·5	89·4	88·9
		95	59·8	94·8	93·8	94·2	94·1	92·8	92·5	94·4	93·9
		99	72·9	98·8	98·4	98·6	98·5	97·3	97·2	98·4	98·0
	0·9	90	50·2	89·7	88·0	88·9	89·0	87·4	87·4	88·8	88·9
		95	58·2	94·6	93·6	94·1	94·1	92·5	92·1	93·8	93·8
		99	71·3	98·7	98·4	98·6	98·5	96·8	96·7	98·1	98·1

$n^{-1}W_{e3}$, respectively. The cumulants of $n^{1/2}R_e$ are

$$\kappa_1(n^{1/2}R_e) = n^{-1/2}C_1\left[\frac{\alpha_{11}}{2\alpha_{01}^{3/2}} + \frac{C_2\alpha_2}{\alpha_{01}^{3/2}C_1} - \frac{C_1\alpha_2\alpha_{001}}{6\alpha_{01}^{5/2}}\right] + O(n^{-3/2}) \tag{7}$$

$$\begin{aligned}\kappa_2(n^{1/2}R_e) &= \frac{C_1^2\alpha_2}{\alpha_{01}} + n^{-1}C_1^2\left[\frac{7\alpha_{11}^2}{4\alpha_{01}^2} + \frac{11C_2\alpha_{11}\alpha_2}{\alpha_{01}^2C_1} + \frac{2C_2^2\alpha_2^2}{\alpha_{01}^2C_1^2} + \frac{6C_3\alpha_2^2}{\alpha_{01}^2C_1} + \frac{C_1\alpha_2\alpha_{101}}{\alpha_{01}^2} + \frac{\alpha_{21}}{\alpha_{01}}\right.\\ &+ \frac{2C_2\alpha_3}{\alpha_{01}C_1} + \frac{\alpha_2\alpha_{02}}{\alpha_{01}^2} + \frac{17C_1\alpha_{11}\alpha_2\alpha_{001}}{6\alpha_{01}^3} + \frac{3C_2\alpha_2^2\alpha_{001}}{\alpha_{01}^3C_1} + \frac{8C_2\alpha_2^2\alpha_{001}}{3\alpha_{01}^3} + \frac{C_1\alpha_3\alpha_{001}}{3\alpha_{01}^2} +\\ &- \left.\frac{3\alpha_2^2\alpha_{001}^2}{4\alpha_{01}^4} + \frac{3C_1\alpha_2^2\alpha_{001}^2}{2\alpha_{01}^4} - \frac{C_1^2\alpha_2^2\alpha_{001}^2}{36\alpha_{01}^4} - \frac{\alpha_2^2\alpha_{0001}}{2\alpha_{01}^3} + \frac{C_1\alpha_2^2\alpha_{0001}}{\alpha_{01}^3} - \frac{C_1^2\alpha_2^2\alpha_{0001}}{4\alpha_{01}^3}\right] + O(n^{-2})\end{aligned} \tag{8}$$

$$\kappa_3(n^{1/2}R_e) = n^{-1/2}C_1^2\left[\frac{3C_1\alpha_{11}\alpha_2}{\alpha_{01}^{5/2}} + \frac{6C_2\alpha_2^2}{\alpha_{01}^{5/2}} + \frac{C_1\alpha_3}{\alpha_{01}^{3/2}} + \frac{C_1^2\alpha_2^2\alpha_{001}}{\alpha_{01}^{7/2}}\right] + O(n^{-3/2}) \tag{9}$$

$$\begin{aligned}\kappa_4(n^{1/2}R_e) &= n^{-1}C_1^3\left[\frac{18C_1\alpha_{11}^2\alpha_2}{\alpha_{01}^4} + \frac{84C_2\alpha_{11}\alpha_2^4}{\alpha_{01}^4} + \frac{48C_2^2\alpha_2^3}{\alpha_{01}^4C_1} + \frac{24C_3\alpha_2^3}{\alpha_{01}^4} + \frac{4C_1^2\alpha_2^2\alpha_{101}}{\alpha_{01}^4} +\right.\\ &+ \frac{6C_1\alpha_2\alpha_{21}}{\alpha_{01}^2} + \frac{6C_1\alpha_{11}\alpha_3}{\alpha_{01}^2} + \frac{24C_2\alpha_2\alpha_3}{\alpha_{01}^2} + \frac{C_1\alpha_4}{\alpha_{01}^2} + \frac{3C_1\alpha_2^2\alpha_{02}}{\alpha_{01}^4} + \frac{18C_1^2\alpha_{11}\alpha_2^2\alpha_{001}}{\alpha_{01}^5} +\\ &+ \frac{12C_2\alpha_2^3\alpha_{001}}{\alpha_{01}^5} + \frac{24C_1C_2\alpha_2^3\alpha_{001}}{\alpha_{01}^5} + \frac{4C_1^2\alpha_2\alpha_3\alpha_{001}}{\alpha_{01}^4} - \frac{3C_1\alpha_2^3\alpha_{001}^2}{\alpha_{01}^6} + \frac{6C_1^2\alpha_2^3\alpha_{001}^2}{\alpha_{01}^6} +\\ &+ \left.\frac{C_1^3\alpha_2^3\alpha_{001}^2}{\alpha_{01}^6} - \frac{2C_1\alpha_2^3\alpha_{0001}}{\alpha_{01}^5} + \frac{4C_1^2\alpha_2^3\alpha_{0001}}{\alpha_{01}^5} - \frac{C_1^3\alpha_2^3\alpha_{0001}}{\alpha_{01}^5}\right] + O(n^{-2})\end{aligned} \tag{10}$$

$$\kappa_5(n^{1/2}R_m) = O(n^{-3/2}) \qquad \kappa_j(n^{1/2}R_m) = o(n^{-2}),\ j \geq 6. \tag{11}$$

The signed root R and its cumulants are recovered respectively from R_e and (7)-(11) by setting $C_1 = 1$ and $C_2 = C_3 = 0$.

Appendix 2

The proofs for the Bartlett correctability of W_e and W'_e pivot on the development of formal Edgeworth series for the density of the corresponding signed roots, as outlined in Section 2.2. The construction of the series is straightforward once the second, third, and fourth cumulant of the signed roots exhibit the structure in (3). Therefore, the proofs of Theorem 1 and Theorem 2 are confined to sketch the determination of the constants C_j and $\hat{C}_j, j = 1, 2, 3$, respectively.

Proof of Theorem 1. The constant C_1 in Theorem 1 is obtained by equating to 1 the leading term of $\kappa_2(n^{1/2}R_e)$, whereas C_2 and C_3 by equating to 0 the leading terms of $\kappa_3(n^{1/2}R_e)$ and $\kappa_4(n^{1/2}R_e)$, respectively. The Bartlett factor b_e is obtained by taking termwise expectation in W_e once C_1, C_2, and C_3 are plugged. □

Proof of Theorem 2. The estimate $\hat{C}_1$ admits the expansion $C_1 + r_1 + r_2$, where the reminder terms $r_1 = O_p(n^{-1/2})$ and $r_2 = O_p(n^{-1})$ are obtained by Taylor expanding $\hat{C}_1$ about θ^*, providing

$$
\begin{aligned}
r_1 &= C_1\left[-\frac{A_2}{2\alpha_2}+\frac{A_{01}}{2\alpha_{01}}+\frac{A_1\alpha_{11}}{\alpha_2\alpha_{01}}-\frac{A_1\alpha_{001}}{2\alpha_{01}^2}\right] && (12)\\
r_2 &= C_1\left[\frac{3A_2^2}{8\alpha_2^2}-\frac{A_{01}^2}{8\alpha_{01}^2}-\frac{A_1A_{001}}{2\alpha_{01}^2}+\frac{3A_1^2\alpha_{11}^2}{2\alpha_2^2\alpha_{01}^2}-\frac{A_1A_{01}\alpha_{11}}{2\alpha_2\alpha_{01}^2}-\frac{A_1^2\alpha_{101}}{2\alpha_2\alpha_{01}^2}-\frac{3A_1A_2\alpha_{11}}{2\alpha_2^2\alpha_{01}}+\right. && (13)\\
&+\left.\frac{A_1A_{11}}{\alpha_2\alpha_{01}}-\frac{A_2A_{01}}{4\alpha_2\alpha_{01}}-\frac{A_1^2\alpha_{02}}{2\alpha_2\alpha_{01}^2}+\frac{3A_1A_{01}\alpha_{001}}{4\alpha_{01}^3}+\frac{A_1A_2\alpha_{001}}{4\alpha_2\alpha_{01}^2}-\frac{3A_1^2\alpha_{001}^2}{8\alpha_{01}^4}+\frac{A_1^2\alpha_{0001}}{4\alpha_{01}^3}\right].
\end{aligned}
$$

Once $\hat{C}_1 = C_1 + r_1 + r_2$ is plugged in R_{1e} and R_{2e}, we have that R_{2e} and R_{3e} become $\tilde{R}_{2e} = R_{2e}+r_1A_1\alpha_{01}^{-1/2}$ and $\tilde{R}_{3e} = R_{3e}+r_2A_1\alpha_{01}^{-1/2}+r_1A_1A_{11}\alpha_{01}^{-3/2}/2$. The third cumulant of $n^{1/2}\tilde{R}_e = n^{1/2}[R_{1e} + \tilde{R}_{2e} + \tilde{R}_{3e}]$ is

$$
\kappa_3(n^{1/2}\tilde{R}_e) = n^{-1/2}C_1^3\left[-\frac{6\alpha_{11}\alpha_2}{\alpha_{01}^{5/2}}+\frac{6C_2\alpha_2^2}{\alpha_{01}^{5/2}C_1}-\frac{2\alpha_3}{\alpha_{01}^{3/2}}-\frac{3\alpha_2^2\alpha_{001}}{\alpha_{01}^{7/2}}+\frac{C_1\alpha_2^2\alpha_{001}}{\alpha_{01}^{7/2}}\right]+O(n^{-3/2}),
$$

and by equating the leading term to 0 we obtain the new expression for C_2, which corresponds to that given in Theorem 2 but with sample moments replaced with expected moments and $\hat{C}_1$ with C_1. Similarly to $\hat{C}_1$, we have that $\hat{C}_2$ in Theorem 2 may be expanded as $C_2 + r_3$, where $r_3 = O_p(n^{-1/2})$ is

$$
\begin{aligned}
r_3 = C_1&\left[-\frac{3A_2\alpha_{11}}{2\alpha_2^2}+\frac{A_{11}}{\alpha_2}+\frac{A_1\alpha_{21}}{\alpha_2^2}-\frac{5A_1\alpha_{11}\alpha_3}{3\alpha_2^3}-\frac{A_{01}\alpha_3}{2\alpha_2^2}+\frac{A_{001}}{2\alpha_{01}}-\frac{A_{001}C_1}{6\alpha_{01}}+\right. \qquad (14)\\
&-\frac{3A_1\alpha_{11}^2}{\alpha_2^2\alpha_{01}}-\frac{A_{01}\alpha_{11}}{2\alpha_2\alpha_{01}}+\frac{A_1\alpha_{101}}{\alpha_2\alpha_{01}}+\frac{A_3\alpha_{01}}{3\alpha_2^2}-\frac{5A_2\alpha_3\alpha_{01}}{6\alpha_2^3}+\frac{A_1\alpha_{02}}{\alpha_2\alpha_{01}}+\frac{A_{01}\alpha_{001}}{4\alpha_{01}^2}-\frac{A_1\alpha_{11}\alpha_{001}}{\alpha_2\alpha_{01}^2}\\
&+\left.\frac{A_1C_1\alpha_{11}\alpha_{001}}{3\alpha_2\alpha_{01}^2}-\frac{A_2\alpha_{001}}{4\alpha_2\alpha_{01}}+\frac{A_2C_1\alpha_{001}}{6\alpha_2\alpha_{01}}-\frac{A_1\alpha_3\alpha_{001}}{2\alpha_2^2\alpha_{01}}+\frac{A_1\alpha_{001}^2}{4\alpha_{01}^3}+\frac{A_1\alpha_{0001}}{2\alpha_{01}^2}-\frac{A_1C_1\alpha_{0001}}{6\alpha_{01}^2}\right].
\end{aligned}
$$

Once $\hat{C}_2$ is plugged in $\tilde{R}_{2e}$, we have $\tilde{R}^*_{3e} = \tilde{R}_{3e} + r_3 A_1^2 \alpha_{01}^{-3/2}$. The fourth cumulant of $n^{1/2}\tilde{R}^* = n^{1/2}[R_{1e} + \tilde{R}_{2e} + \tilde{R}^*_{3e}]$ is

$$\begin{aligned}
\kappa_4(n^{1/2}\tilde{R}^*_e) &= n^{-1}C_1^4\left[-\frac{24\alpha_{11}^2\alpha_2}{\alpha_{01}^4} - \frac{60C_2\alpha_{11}\alpha_2^2}{\alpha_{01}^4C_1} + \frac{48C_2^2\alpha_2^3}{\alpha_{01}^4C_1^2} + \frac{24C_3\alpha_2^3}{\alpha_{01}^4C_1} + \frac{12\alpha_2^2\alpha_{101}}{\alpha_{01}^4} + \frac{24\alpha_2\alpha_{21}}{\alpha_{01}^3} + \right.\\
&- \frac{28\alpha_{11}\alpha_3}{\alpha_{01}^3} - \frac{24C_2\alpha_2\alpha_3}{\alpha_{01}^3C_1} - \frac{8\alpha_3^2}{\alpha_2\alpha_{01}^2} + \frac{6\alpha_4}{\alpha_{01}^2} + \frac{12\alpha_2^2\alpha_{02}}{\alpha_{01}^4} + \frac{6\alpha_{11}\alpha_2^2\alpha_{001}}{\alpha_{01}^5} - \frac{10C_1\alpha_{11}\alpha_2^2\alpha_{001}}{\alpha_{01}^5} + \\
&- \frac{36C_2\alpha_2^3\alpha_{001}}{\alpha_{01}^5C_1} + \frac{24C_2\alpha_2^3\alpha_{001}}{\alpha_{01}^5} - \frac{4C_1\alpha_2\alpha_3\alpha_{001}}{\alpha_{01}^4} + \frac{6\alpha_2^3\alpha_{001}^2}{\alpha_{01}^6} - \frac{6C_1\alpha_2^3\alpha_{001}^2}{\alpha_{01}^6} + \frac{C_1^2\alpha_2^3\alpha_{001}^2}{\alpha_{01}^6} + \\
&+ \left.\frac{4\alpha_2^3\alpha_{0001}}{\alpha_{01}^5} - \frac{C_1^2\alpha_2^3\alpha_{0001}}{\alpha_{01}^5}\right] + O(n^{-2}),
\end{aligned}$$

and by equating the leading term to 0 we obtain the new expression for C_3 which corresponds to that given in Theorem 2 but with sample moments replaced with expected moments and $\hat{C}_j$ with C_j, $j = 1, 2$. Note that the leading term of $\kappa_2(n^{1/2}R_e)$ is equal to that of $\kappa_2(n^{1/2}\tilde{R}^*_e)$, and $\kappa_3(n^{1/2}\tilde{R}_e) - \kappa_3(n^{1/2}\tilde{R}^*_e) = O(n^{-2})$. Finally, the Bartlett factor b'_e is obtained by taking termwise expectation of $W'_e = n(\tilde{R}^*)^2 + O_p(n^{-3/2})$ once $\hat{C}_1$, $\hat{C}_2$, and $\hat{C}_3$ are plugged in $n^{1/2}\tilde{R}^*$. □

References

Barndorff-Nielsen, O. and Cox, D. (1994). *Inference and Asymptotics*. Chapman and Hall, London.

Chandler, R. and Bate, S. (2007). Inference for clustered data using the independence loglikelihood. *Biometrika*, **94**, 167–183.

Cox, D. and Reid, N. (2004). A note on pseudolikelihood constructed from marginal densities. *Biometrika*, **91**, 729–737.

DiCiccio, T., Hall, P., and Romano, J. (1991). Empirical likelihood is Bartlett-correctable. *Ann. Statist.*, **19**, 1053–1061.

Jensen, J. (1993). A historical sketch and some new results on the improved log likelihood ratio statistic. *Scand. J. Statist.*, **20**, 1–15.

Kent, J. (1982). Robust properties of likelihood ratio tests. *Biometrika*, **69**, 19–27.

Mykland, P. (1999). Bartlett identities and large deviations in likelihood theory. *Ann. Statist.*, **27**, 1105–1117.

Pace, L., Salvan, A., and Sartori, N. (2011). Adjusting composite likelihood ratio statistics. *Statist. Sinica*, **21**, 129–148.

Rotnitzky, A. and Jewell, N. (1990). Hypothesis testing of regression parameters in semiparametric generalized linear models for cluster correlated data. *Biometrika*, **77**, 485–497.

Varin, C. (2008). On composite marginal likelihoods. *AStA Advances in Statistical Analysis*, **92**, 1–28.

Varin, C. and Vidoni, P. (2005). A note on composite likelihood inference and model selection. *Biometrika*, **92**, 519–528.

Varin, C., Reid, N., and Firth, D. (2011). An overview of composite likelihood methods. *Statist. Sinica*, **21**, 5–42.

Wedderburn, R. (1974). Quasi-likelihood functions, generalized linear models, and the gauss-newton method. *Biometrika*, **61**, 439–447.

White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica*, **50**, 1–25.

Xu, X. and Reid, N. (2011). On the robustness of maximum composite likelihood estimate. *J. Statist. Plann. Inf.*, **141**, 3047–3054.

Acknowledgements

I am fully indebted to A. R. Brazzale, A. Salvan, and N. Sartori for their useful suggestions concerning the contents of the first draft of the manuscript. Since then, the manuscript has been substantially modified leading to the present version. All errors are my own responsibility.

Working Paper Series
Department of Statistical Sciences, University of Padua

You may order paper copies of the working papers by emailing wp@stat.unipd.it
Most of the working papers can also be found at the following url: http://wp.stat.unipd.it

UNIVERSITA DEGLI STUDI DI PADOVA
DIPARTIMENTO DI SCIENZE STATISTICHE

Department of Statistical Sciences
University of Padua
Italy