Statistica Sinica Preprint No: SS-2016-0142	
Title	Composite Likelihood Inference under Boundary Conditions
Manuscript ID	SS-2016-0142
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202016.0142
Complete List of Authors	Jing Huang Yang Ning Yi Cai Kung-Yee Liang and Yong Chen
Corresponding Author	Yong Chen
E-mail	ychen123@upenn.edu
Notice: Accepted version subject to English editing.	

COMPOSITE LIKELIHOOD INFERENCE UNDER BOUNDARY CONDITIONS

Jing Huang[1], Yang Ning[2], Yi Cai[3], Kung-Yee Liang[4]
and Yong Chen[1*]

[1] *The University of Pennsylvania,* [2] *Cornell University,*
[3] *AT&T Services Inc,* [4] *National Health Research Institutes*

Abstract: Often when a data generating process is too complex to fully specify, standard likelihood based inference is not available. Composite likelihood, coined by Bruce Lindsay, can provide inference with partial specification of data generating process. Due to the robustness to model specification and the computational simplicity, the composite likelihood method has been widely used in many applications. The purpose of this paper is to conduct a theoretical investigation of the composite likelihood ratio test (CLRT) when parameters of interest may lie on the boundary of the parameter space. Our main result shows that the limiting distribution of CLRT is equivalent to the likelihood ratio test of a normal mean problem, in which the restricted mean of a multivariate normal distribution is tested based on one observation from a multivariate normal distribution with inverse of Godambe information matrix. Furthermore, we apply our general theoretical result to study a variety of examples. Simulation results confirm that the limiting distribution of CLRT performs well in finite samples.

Key words and phrases: Boundary condition; composite likelihood; hypothesis testing; likelihood ratio test.

1. Introduction

Likelihood based inference is commonly used for modeling complex data. In some applications, fully specifying the data generating process is difficult or not preferred due to the lack of knowledge on the true model or computational challenges. One possible solution is to conduct the inference based on a *partially* specified model. Coined by Lindsay (1988), the composite likelihood approach has drawn a great deal of attention and has been widely used in many areas, such as biomedical research (Heagerty & Lele, 1998; Molenberghs & Verbeke, 2005; Wellner & Zhang, 2000; Henderson & Shimakura, 2003; Guan, 2006; He & Yi,

2011), statistical genetics (McVean *et al.* , 2004; Myers *et al.* , 2005; Larribe & Fearnhead, 2011), geostatistics (Vecchia, 1988; Nott & Rydén, 1999; Stein *et al.* , 2004; Padoan *et al.* , 2010), finance (Bhat *et al.* , 2010; Varin & Vidoni, 2008), social science, and among many others.

Specifically, a composite likelihood is constructed by the product of a set of low-dimensional marginal or conditional densities. This approach is especially useful in modeling correlated data with complex or unknown dependency structure, or reducing a complex likelihood function with high computational cost to a much simpler one. For instance, when a working independence assumption is adopted, the composite likelihood is also called independence likelihood (Chandler & Bate, 2007). Because the data generating process is partially specified, composite likelihood inference is more robust than the standard likelihood based inference. For more discussion on composite likelihood methods, we refer to the review paper by Varin *et al.* (2011) and the references therein.

In standard hypothesis testing problems, regularity conditions require that the parameter values are interior points of the parameter space under the null hypothesis. This assumption guarantees the composite score function being zero at the maximum composite likelihood estimate. Such a first order condition leads to the results that the composite likelihood ratio statistic is asymptotically a mixture of weighted χ^2_1 (Kent, 1982; Molenberghs & Verbeke, 2005). In many important applications, however, the parameters of interest may lie on the boundary of the parameter space. The constraint can arise either from the definition of the parameter space or from previous knowledge on the possible range of the parameter. For example, such a problem is encountered in diagnostic systematic reviews, which aim to evaluate diagnostic accuracy by pooling sensitivity and specificity of a dichotomized diagnostic test from multiple studies (Chen *et al.* , 2014, 2015).

As well acknowledged in the literature, ignoring the boundary constraints often leads to a substantial loss of power (Self & Liang, 1987; Chen & Liang, 2010). However, theoretical results of composite likelihood based inference under boundary conditions have not been established. In this paper, we aim to fill this gap by providing the limiting distributions of the composite likelihood ratio tests under boundary conditions. Furthermore, we apply our general theoretical

result to study the following three examples: (1) stratified case-control study; (2) diagnostic systematic reviews and (3) adverse events detection in medical reports. In our simulation studies, we found that the naive method ignoring the boundary constraints is grossly conservative and leads up to 48% less power compared to the proposed test.

The rest of this paper is organized as follows. In section 2, we review the definition of composite likelihood and introduce applications where boundary problems are involved. In section 3, we first give some regularity conditions and derive the asymptotic distribution of the composite likelihood ratio test followed by the detailed calculation of the limiting distribution when all parameters of interest are on the boundary. We also consider other situations when a subset of parameters of interest are on the boundary. In section 4, we revisit the examples in section 2. In section 5, we provide simulation studies. We finally conclude with a brief discussion in section 6.

2. Notations and Examples

2.1 Composite Likelihood

Let $f(x;\theta)$ be the probability density function of a multidimensional vector random variable X, indexed by a p-dimensional parameter $\theta = (\theta_1, ..., \theta_p)^T$, where θ belongs to the parameter space Ω, a subset of R^p. We assume model identifiability that distinct values of θ correspond to distinct probability distributions. Suppose N independent random variables $X_1, ..., X_N$ are observed from the model $f(x;\theta)$. Let $\{\mathscr{A}_1, ..., \mathscr{A}_K\}$ be a set of marginal or conditional events with associated likelihoods $L_i(\theta;\mathscr{A}_k) \propto \text{pr}(X_i \in \mathscr{A}_k;\theta)$, where $k = 1, 2, ..., K$ and K is the number of events. Following Lindsay (1988), a composite log likelihood can be construct as

$$\ell_c(\theta) = \sum_{i=1}^{N}\sum_{k=1}^{K} \omega_k \log L_i(\theta;\mathscr{A}_k),$$

where ω_k, $k = 1, ..., K$, are nonnegative weights associated with the likelihood $L_i(\theta;\mathscr{A}_k)$. Specially, in the longitudinal data analysis, the composite likelihood can be constructed by pooling marginal densities without considering the correlation between repeated measurements. The use of this likelihood is studied by Chandler & Bate (2007). The maximum of the composite likelihood achieves at the parameter value, $\hat{\theta}_c$, the maximum composite likelihood estimator. De-

note the first two derivatives of $\ell_c(\theta)$ as $U_c(\theta)$ and $h_c(\theta)$ respectively. Since each component of the composite likelihood is a true likelihood, it carries important features of the ordinary likelihood, e.g., Bartlett identities hold for each component

$$E\{S_{ik}(\theta)\} = 0, \quad \text{and} \quad E\{-h_{ik}(\theta)\} = E\{S_{ik}(\theta)\}^2, \quad \text{for} \quad k = 1, 2, \ldots, K,$$

where $S_{ik}(\theta)$ and $h_{ik}(\theta)$ denote the first two derivatives of $\log L_i(\theta; \mathscr{A}_k)$. Since the composite score function $U_c(\theta)$ is a linear combination of component score functions $S_{ik}(\theta)$, $U_c(\theta)$ is an unbiased estimating equation. Assume that we are interested in testing $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$. Standard regularity conditions assume that θ_0 is an interior point of the parameter space. Under some extra regularity conditions, the composite likelihood ratio test CLRT $= 2\{\ell_c(\hat{\theta}_c) - \ell_c(\theta_0)\}$ converges in distribution to a mixture of independent χ^2_1 variables (Varin *et al.* , 2011), where the weights are the eigenvalues of $H(\theta_0)G^{-1}(\theta_0)$ with

$$H(\theta) = -E\Big\{\sum_{k=1}^{K} \omega_k h_{ik}(\theta)\Big\}, \quad J(\theta) = E\Big\{\sum_{k=1}^{K} \omega_k S_{ik}(\theta)\Big\}\Big\{\sum_{k=1}^{K} \omega_k S_{ik}(\theta)\Big\}^T,$$

and $G(\theta_0) = H(\theta_0)J^{-1}(\theta_0)H(\theta_0)$.

However, in many real applications, θ_0 may lie on the boundary of the parameter space, which makes the existing asymptotic results on CLRT invalid. In the following, we present three examples where the boundary problems are encountered.

2.2 Test of Positive Associations in Stratified Case-Control Studies with Sparse Data

Stratified case-control design is widely used in epidemiological and genetic studies. In particular, in the ith stratum, we let $x_{i1}, ..., x_{in_i}$ denote the $p \times 1$ vectors of potential risk factors of n_i cases, and let $x_{in_i+1}, ..., x_{iN_i}$ denote the potential risk factors of m_i controls, where $m_i = N_i - n_i$ and $i = 1, ..., K$. A logistic regression model which accounts for stratum-specific effects is

$$\text{logit pr}(y_{ij} = 1 \text{ in stratum } i | x_{ij}) = \alpha_i + \beta^T x_{ij}, \qquad i = 1, ..., K, \text{ and } j = 1, \ldots, N_i,$$

where $y_{ij} = 1$ if the jth subject in the ith stratum belongs to the case group and $y_{ij} = 0$ otherwise, and the coefficients β quantify the effects of risk factors x_{ij}

on disease status y_{ij}. In the case that n_i and m_i are uniformly bounded while the number of stratum $K \to \infty$, the maximum likelihood estimator is known to be inconsistent. One common technique to solve this problem is to use the composite likelihood method proposed by Liang (1987), which extended the well-known Mantel-Haenszel estimator to logistic regression models with multiple risk factors. Specifically, for the (j, l) case-control pair of subjects in the ith stratum $(j = 1, ..., n_i; l = n_i + 1, ..., N_i)$, the conditional probability that x_{ij} is from the case given the fact that one of x_{ij} and x_{il} is from the case and one of them is from the control can be calculated as

$$\text{pr}(y_{ij} = 1, y_{il} = 0 | y_{ij} + y_{il} = 1, x_{ij}, x_{il}; \alpha_i, \beta) = \frac{e^{\beta^T x_{ij}}}{e^{\beta^T x_{ij}} + e^{\beta^T x_{il}}}.$$

Thus a composite likelihood can be formulated by considering all $n_i m_i$ possible pairs within the ith stratum

$$L_i(\beta) = \prod_{j=1}^{n_i} \prod_{l=n_i+1}^{N_i} \frac{e^{\beta^T x_{ij}}}{e^{\beta^T x_{ij}} + e^{\beta^T x_{il}}}.$$

A composite likelihood combining data from all K strata for β is then constructed by assigning the weight w_i to $L_i(\beta)$,

$$\ell_c(\beta) = \sum_{i=1}^{K} w_i \log \{L_i(\beta)\} = \sum_{i=1}^{K} w_i \log \left\{ \prod_{j=1}^{n_i} \prod_{l=n_i+1}^{N_i} \frac{e^{\beta^T x_{ij}}}{e^{\beta^T x_{ij}} + e^{\beta^T x_{il}}} \right\}.$$

Without loss of generality, we consider the weights $w_i = N_i^{-1}$, in which the maximum composite likelihood estimator reduces to Mantel-Haenszel estimator when only a binary covariate is considered (Liang, 1987). Suppose some of the risk factors are known to be positively associated with the occurrence of disease, then testing the null hypothesis $H_0 : \beta = 0$ is a boundary problem.

2.3 Test of Heterogeneity in Diagnostic Systematic Reviews

Diagnostic systematic review is a vital step in evaluating the accuracy of a diagnostic test. For a dichotomized diagnostic test, it usually involves drawing and comparing the sensitivity (Se) and specificity (Sp) from multiple studies. The procedure of pooling data is not straightforward because of the following two challenges. First, the estimated Se and Sp are typically negatively correlated between studies (Reitsma *et al.* , 2005). Second, there may be substantial

between-study heterogeneity in paired indices (Moses *et al.* , 1993; Irwig *et al.* , 1995; Rutter & Gatsonis, 1995). Such heterogeneity may arise from differences in study population characteristics, variability of assessment and other factors. To account for these challenges, Chen *et al.* (2014, 2015) proposed a composite likelihood based approach that is robust and computationally convenient. The idea is to construct a composite likelihood function by using an independent working assumption between Se and Sp. More specifically, considering a diagnostic review of m studies, denote n_{i11}, n_{i00}, n_{i01}, n_{i10} as the number of true positives, true negatives, false positives, and false negatives, respectively, $i = 1, ..., m$. Let $n_{i1} = n_{i11} + n_{i10}$ and $n_{i0} = n_{i01} + n_{i00}$ be the number of diseased and healthy subjects, respectively, and Se_i and Sp_i be the study-specific Se and Sp, respectively. Assume that

$$n_{i11}|(n_{i1}, Se_i) \sim \text{Binomial}(n_{i1}; Se_i), \qquad n_{i00}|(n_{i0}, Se_i) \sim \text{Binomial}(n_{i0}; Sp_i)$$

$$g(Se_i) = X_i^T\beta_1 + \mu_{i1}, \qquad g(Sp_i) = Z_i^T\beta_2 + \mu_{i2}$$

$$\begin{pmatrix} \mu_{i1} \\ \mu_{i2} \end{pmatrix} \sim \text{BN}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \ \Sigma = \begin{pmatrix} \tau_1^2 & \rho\tau_1\tau_2 \\ \rho\tau_1\tau_2 & \tau_2^2 \end{pmatrix}\right),$$

where $g(\cdot)$ is a known link function such as logit function, X_i and Z_i are vectors of study-level covariates, possibly overlapping, related to Se_i and Sp_i, τ_1^2 and τ_2^2 capture the between-study heterogeneity in Se and Sp, respectively, and ρ describes the correlation between the random effects Se_i and Sp_i in the transformed scale.

Ignoring the correlation between Se_i and Sp_i, a composite likelihood can be constructed by setting $\rho = 0$ in the above model, i.e.,

$$\ell_c(\theta) = \sum_{i=1}^{m} \log \left\{ \int q(n_{i11}|n_{i1}; Se_i)\phi(Se_i; \beta_1, \tau_1^2)dSe_i \right.$$
$$\left. \times \int q(n_{i00}|n_{i0}; Sp_i)\phi(Sp_i; \beta_2, \tau_2^2)dSp_i \right\}$$

where $\theta = (\beta_1^T, \beta_2^T, \tau_1, \tau_2)^T$, $\phi(\cdot; \cdot, \cdot)$ is the probability density function of univariate logit normal distribution and $q(\cdot|\cdot; \cdot)$ is the probability mass function of a binomial distribution. In this model, testing heterogeneity of Se and Sp across all the studies is equivalent to test, $H_0 : \tau_1^2 = \tau_2^2 = 0$. This is a hypothesis testing problem with boundary constraints.

2.4 Signal Detection of Adverse Events (AE) Reporting Rate Across Calendar Years

The Vaccine Adverse Event Reporting System (VAERS) is a national vaccine safety surveillance program that collects information about adverse events (AE) that occur after the administration of vaccines licensed for use in the United States (Chen *et al.* , 1994; Niu *et al.* , 2001; Shimabukuro *et al.* , 2015). The surveillance data are structured in a table format with vaccine-AE combination (i.e., a certain type of AE after a particular vaccination) being the column variable, and reporting year being the row variable (Huang *et al.* , 2011). There are total of I years and J vaccine-AE combinations. The table cell n_{ij} is the number of events reported for the jth vaccine-AE combination during the ith year. The total number of reported cases for ith year is the marginal total of ith row denoted as $n_{i.}$, and the total number of jth vaccine-AE combination for all the years is the marginal total of jth column denoted as $n_{.j}$. Let $n_{..}$ denote the total number of events.

To account for the large amount of zero cells in our data n_{ij}, we consider a zero-inflated Poisson model to deal with the excessive zeros as follow:

$$n_{ij} \mid n_{i.}, p_{ij}, w_{ij} \begin{cases} = 0 \quad \text{with probability} \quad w_{ij}, \\ \sim \text{Poisson}(n_{i.}p_{ij}) \quad \text{with probability} \quad 1 - w_{ij} \end{cases}$$

where w_{ij} is the probability of observing a true zero in the ith year and jth vaccine-AE combination, and p_{ij} is the probability of reported cases in the ith year and jth vaccine-AE combination. In order to account for the heterogeneity of reporting proportion for a fixed jth vaccine-AE combination through year 1990 to 2015 (the most recent reporting year), we further assume that

$$\text{logit}(p_{ij}) \sim \text{N}(\beta_{0j}, \tau_j^2), i = 1, \cdots, I$$

where $\beta_{0j} = \text{logit}(p_{0j})$, p_{0j} is the overall reporting proportion of jth AE-vaccine combination across all the year and τ_j^2 represents the variation of reporting proportion. Similarly, we assume the weight w_{ij} follows

$$\text{logit}(w_{ij}) \sim \text{N}(\alpha_{0j}, \delta_j^2), i = 1, \cdots, I$$

where $\alpha_{0j} = \text{logit}(w_{0j})$, w_{0j} is the overall probability of observing a true zero across all the year and δ_j^2 represents the variation of observing a true zero.

A composite likelihood of the jth vaccine-AE combination can be constructed by multiplying the marginal densities of n_{ij} ignoring their correlations:

$$\ell_c(\theta) = \sum_{i=1}^{I} \log \left\{ \int_0^1 \int_0^1 p(n_{ij}|n_{i\cdot}; p_{ij}, w_{ij}) p(p_{ij}|p_{0j}, \tau_j^2) p(w_{ij}|w_{0j}, \delta_j^2) dp_{ij} dw_{ij} \right\}$$

where $\theta = (w_{0j}, \beta_{0j}, \tau_j^2, \delta_j^2)$. Testing whether the reporting rate is the same across all years is equivalent to testing the hypothesis, $H_0 : \tau_j^2 = \delta_j^2 = 0$, which is a boundary problem.

2.5. Other examples

Other examples of testing problems with boundary constraints in composite likelihood inference include testing spatial correlations in a Gaussian random field model in geostatistics applications (Guan, 2006), and testing serially dependence in time serial data of panel studies (Wellner & Zhang, 2000).

In the next section, we give our main results. The above examples are revisited in Section 4.

3. Main Results

3.1 Regularity Conditions

To derive the asymptotic results, we impose the following regularity conditions.

R1: The first two derivatives of $\ell_c(\theta)$ with respect to θ on the intersection of neighborhoods of the true parameter value, θ_0, and Ω exist and are continuous. If θ_0 is on the boundary of Ω, the derivatives of $\ell_c(\theta)$ are taken from the appropriate side.

R2: The parameter space Ω is compact, the function $E\{\ell_c(\theta)\}$ is continuous and $\theta_0 = \arg\max_{\theta\in\Omega} E\{\ell_c(\theta)\}$ is unique. There exists a function $A(X)$ whose expectation is finite such that $|\sum_{k=1}^{K} \omega_k \log L_i(\theta; \mathscr{A}_k)| \leq A(X_i)$ for any $\theta \in \Omega$.

R3: On the intersection of neighborhoods of θ_0 and Ω, for any $1 \leq j, k \leq p$, it holds that $|\sum_{k=1}^{K} \omega_k \frac{\partial^2 \log L_i(\theta;\mathscr{A}_k)}{\partial\theta_j \partial\theta_k}| \leq C(X_i)$, where $C(X)$ is a function of X whose expectation is finite.

R4: The variability matrix, $J(\theta)$, exists and is positive definite at θ_0.

R5: The sensitivity matrix, $H(\theta)$, exists and is positive definite at θ_0.

When studying the limiting distribution of an estimator for θ_0, we can think it as a local problem because only the values of θ near θ_0 are relevant asymptotically. Thus, we assume that the parameter space near θ_0 can be locally approximated by a cone. This approach was used by Chernoff (1954) and Self & Liang (1987) in the context of likelihood ratio tests.

Definition The set $\Omega \subset R^p$ is approximated at θ_0 by a cone with vertex at θ_0, C_Ω, if

(1) $\inf_{x \in C_\Omega} \| x - y \| = o(\| y - \theta_0 \|)$ for all $y \in \Omega$

and

(2) $\inf_{y \in \Omega} \| x - y \| = o(\| x - \theta_0 \|)$ for all $x \in C_\Omega$

We assume that the parameter space Ω is regular enough to be approximated by a cone with vertex at θ_0, which is mild enough to encompass a wide variety of shapes for Ω. In the sections followed, we first give the asymptotic distribution of the composite likelihood ratio statistics when θ_0 is on the boundary of Ω. Then, we derive the exact form of the asymptotic distribution in several cases.

For notational simplicity, rewrite the hypothesis $H_0 : \theta = \theta_0$ as $H_0 : \theta \in \Omega_0$ where $\Omega_0 = \{\theta : \theta = \theta_0 \in \Omega\}$. The complement of Ω_0 in Ω is denoted by Ω_1. For any subset of R^p, φ, we define $L_\varphi = \sup_{\theta \in \varphi} \ell_c(\theta)$. We also define the maximum composite likelihood estimator in the parameter space φ, $\hat{\theta}_c^{(\varphi)}$, as that value of θ in the closure of φ which maximizes $\ell_c(\theta)$. The composite likelihood ratio test statistic can be written as

$$CLRT = -2(L_{\Omega_0} - L_\Omega).$$

3.2 Limiting Distribution of Composite Likelihood Ratio Statistics

First, we establish the $\sqrt{N}$ consistency of the maximum composite likelihood estimator.

Lemma 1 *If the regularity conditions R1-R5 hold, then as $N \to \infty$, $\hat{\theta}_c$ converges to θ_0 in probability. Moreover, $N^{1/2}(\hat{\theta}_c - \theta_0) = O_p(1)$.*

A proof is given in Section S1 of the online Supplementary Material. Now we derive the asymptotic distribution of the composite likelihood ratio test $CLRT$.

Theorem 1 *If the regularity conditions R1-R5 hold, θ_0 is a limiting point of both Ω_0 and Ω_1, and the sets Ω_0 and Ω_1 are approximated by non-empty cones C_{Ω_0} and C_{Ω_1} respectively, then under the null hypothesis, the asymptotic distribution of the composite likelihood ratio statistic, $CLRT$, is the same as the distribution of the likelihood ratio test of $\theta \in C_{\Omega_0}$ against $\theta \in C_{\Omega_1}$ based on one observation from a population with multivariate normal distribution with mean θ_0 and covariance matrix $H(\theta_0)^{-1}J(\theta_0)H(\theta_0)^{-1}$ while the covariance matrix is misspecified as $H(\theta_0)^{-1}$ in the likelihood ratio test.*

The proof is done by two approximations. First, Ω is approximated by C_Ω locally around θ_0. This is justified by the $\sqrt{N}$ consistency of $\hat{\theta}_c$ and the definition of approximating cones given previously. The second approximation is made through similar argument of Self & Liang (1987), which used a quadratic function to approximate the composite likelihood. More technical details are available in Section S2 of the online Supplementary Material.

In some special cases, the sensitivity matrix may equal to the variability matrix, for example, in the partial likelihood inference for censored data (Cox, 1975). In such cases, the misspecified covariance matrix $H(\theta_0)^{-1}J(\theta_0)H(\theta_0)^{-1}$ equals to $J(\theta_0)^{-1}$, and Theorem 1 reduces to Theorem 3 of Self & Liang (1987). However, in general, the equality is not satisfied (Molenberghs & Verbeke, 2005). The purpose of Theorem 1 is to reduce the general problem of computing the limit distribution of CLRT to a problem of computing the distribution of

$$Q_{C_{\Omega_0}}(Z) - Q_{C_\Omega}(Z), \tag{1}$$

where $Q_\varphi(Z) = \inf_{\theta\in\varphi} \{Z-(\theta-\theta_0)\}^T H(\theta_0) \{Z-(\theta-\theta_0)\}$, $C_\Omega = C_{\Omega_0} \bigcup C_{\Omega_1}$, and $Z \sim MVN(0, G^{-1}(\theta_0))$, where $G(\theta_0) = H(\theta_0)J(\theta_0)^{-1}H(\theta_0)$ is known as the Godambe information matrix.

The limiting distribution of CLRT in equation (1) is still complicated. In the following, we focus on an important special case in which the representation given in equation (1) can be further simplified.

3.3 An Important Special Case

Partition the parameter vector, θ, into two parameters sets $\theta^T = (\gamma^T, \eta^T)$, where γ denotes parameters of interest that will be tested and η denotes the nuisances parameters. Then further partition the parameter vector to four coor-

dinates, here we adopt the notation in Self & Liang (1987): $(p_{11}, p_{12}, p_{21}, p-p_{11}-p_{12}-p_{21})$, where the first p_{11} coordinates of θ represent the parameters of γ with true values on the boundary; the next p_{12} coordinates represent the parameters of γ with true values not on the boundary; the next p_{21} coordinates represent the first p_{21} components of η with true values on the boundary; and finally the remaining $p-p_{11}-p_{12}-p_{21}$ coordinates represent the last $p-p_{11}-p_{12}-p_{21}$ parameters of η with true value not on the boundary. Note that $p_1 = p_{11}+p_{12}$ represents the dimension of γ and $p-p_1$ represents the dimension of η. In the rest of the paper, we denote the boundary value as 0 for ease of presentation. In the following, we consider a special case that two parameters of interest are on the boundary and the rest $(p-2)$-dimensional nuisance parameters are not on the boundary, i.e., $H_0 : \gamma = 0$ with $\gamma = (\gamma_1, \gamma_2)$, $\gamma_1 \geq 0$ and $\gamma_2 \geq 0$. All three examples discussed in Section 2 belong to this setting.

The parameter configuration is give by $(2, 0, 0, p-2)$. Then $C_{\Omega_0} = \{0\}^2 \times R^{p-2}$ and $C_\Omega = [0, +\infty)^2 \times R^{p-2}$. Partitioning $Z^T = (Z_\gamma^T, Z_\eta^T)$ and $G(\theta_0)$ and $H(\theta_0)$ with respect to (γ, η) as

$$G(\theta_0) = \begin{pmatrix} G_{\gamma\gamma} & G_{\gamma\eta} \\ G_{\eta\gamma} & G_{\eta\eta} \end{pmatrix}, \; H(\theta_0) = \begin{pmatrix} H_{\gamma\gamma} & H_{\gamma\eta} \\ H_{\eta\gamma} & H_{\eta\eta} \end{pmatrix},$$
$$G_{\gamma|\eta} = G_{\gamma\gamma} - G_{\gamma\eta} G_{\eta\eta}^{-1} G_{\gamma\eta}^T, \; H_{\gamma|\eta} = H_{\gamma\gamma} - H_{\gamma\eta} H_{\eta\eta}^{-1} H_{\gamma\eta}^T,$$

with some algebra, equation (1) reduces to

$$Z_\gamma^T H_{\gamma|\eta} Z_\gamma - \inf_{\gamma \in [0,+\infty)^2} (Z_\gamma - \gamma)^T H_{\gamma|\eta} (Z_\gamma - \gamma).$$

Let $H_{\gamma|\eta} = P^T P$, where P is a 2×2 nonsingular matrix and denote $\tilde{C}_\gamma = \{\tilde{\gamma} : \tilde{\gamma} = P\gamma \text{ for any } \gamma \in [0, +\infty)^2\}$ and $\tilde{Z}_\gamma = P Z_\gamma$. Then the limiting distribution of CLRT can be rewritten as

$$\| \tilde{Z}_\gamma \|^2 - \inf_{\tilde{\gamma} \in \tilde{C}_\gamma} \| \tilde{Z}_\gamma - \tilde{\gamma} \|^2,$$

where $\| \cdot \|$ is the usual Euclidean metric. Calculation of the second term in the above equation depends on the location of $\tilde{Z}_\gamma$ relative to the boundary of $\tilde{C}_\gamma$. The shaded region in Fig. 1 represents $\tilde{C}_\gamma$, and $(0, 0)$ is the origin. The angle in the shaded area is less than 180°. This is due to the fact that the convexity of

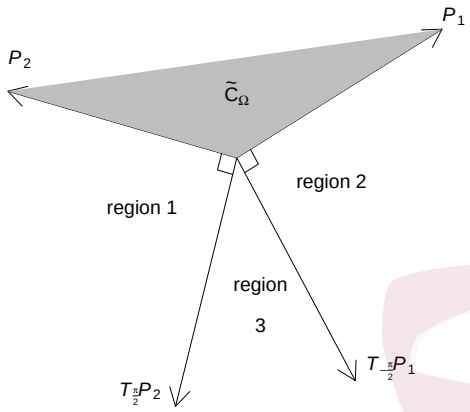

Figure 1: Diagram of the parameter space for the special case.

C_γ is preserved under linear mapping $\gamma \to P\gamma$. Denote the rotation matrix in two dimensional Euclidean space R^2 and the matrix $H_{\gamma|\eta}$ by

$$T_x = \begin{pmatrix} \cos(x) & -\sin(x) \\ \sin(x) & \cos(x) \end{pmatrix}, \quad H_{\gamma|\eta} = \begin{pmatrix} a & b \\ b & d \end{pmatrix},$$

respectively. Further denote the columns of P by P_1 and P_2, and the inner product of vectors a and b by $\langle a, b\rangle = b^T a$. It can be shown that the limiting distribution of CLRT is given by

$$CLRT = \begin{cases} \| \tilde{Z}_\gamma \|^2 \sim U = \lambda_1\chi_1^2 + \lambda_2\chi_1^2 & \text{if } \tilde{Z} \text{ is in the shaded region,} \\ \left\{ \frac{\langle P_2, \tilde{Z}_\gamma\rangle}{\|P_2\|} \right\}^2 \sim d^*\chi_1^2/d & \text{if } \tilde{Z} \text{ is in region 1,} \\ \left\{ \frac{\langle P_1, \tilde{Z}_\gamma\rangle}{\|P_1\|} \right\}^2 \sim a^*\chi_1^2/a & \text{if } \tilde{Z} \text{ is in region 2,} \\ 0 & \text{if } \tilde{Z} \text{ is in region 3,} \end{cases}$$

where λ_1 and λ_2 are eigenvalues of $H_{\gamma|\eta}G_{\gamma|\eta}^{-1}$, a^* and d^* are elements in the matrix $H_{\gamma|\eta}G_{\gamma|\eta}^{-1}H_{\gamma|\eta}$ that

$$H_{\gamma|\eta}G_{\gamma|\eta}^{-1}H_{\gamma|\eta} = \begin{pmatrix} a^* & b^* \\ b^* & d^* \end{pmatrix}.$$

The mixing probabilities for the shaded region, region 1 and region 2, are

$$\pi_s = \cos^{-1}\left[\frac{(G_{\gamma|\eta})_{12}}{\left\{(G_{\gamma|\eta})_{11}(G_{\gamma|\eta})_{22}\right\}^{1/2}}\right]/2\pi,$$

$$\pi_1 = \cos^{-1}\left[\frac{(0,1)G_{\gamma|\eta}H_{\gamma|\eta}^{-1}P^T T_{\pi/2}P_2}{\left\{(0,1)G_{\gamma|\eta}\binom{0}{1}\right\}^{1/2}\left\{P_2^T T_{\pi/2}^T P H_{\gamma|\eta}^{-1}G_{\gamma|\eta}H_{\gamma|\eta}^{-1}P^T T_{\pi/2}P_2\right\}^{1/2}}\right]/2\pi,$$

$$\pi_2 = \cos^{-1}\left[\frac{(1,0)G_{\gamma|\eta}H_{\gamma|\eta}^{-1}P^T T_{-\pi/2}P_1}{\left\{(1,0)G_{\gamma|\eta}\binom{1}{0}\right\}^{1/2}\left\{P_1^T T_{-\pi/2}^T P H_{\gamma|\eta}^{-1}G_{\gamma|\eta}H_{\gamma|\eta}^{-1}P^T T_{-\pi/2}P_1\right\}^{1/2}}\right]/2\pi.$$

Thus the asymptotic distribution of CLRT is a mixture of U, $d^*\chi_1^2/d$, $a^*\chi_1^2/a$, and χ_0^2 with mixing probabilities π_s, π_1, π_2 and $1-\pi_s-\pi_1-\pi_2$, respectively.

4. Examples

4.1 Test of Positive Associations in Stratified Case-Control Studies With Sparse Data

Suppose the covariates x_{ij} are p-dimensional and two of them are known to be positively associated with the occurrence of disease. To test the two positive associations simultaneously, the null hypothesis is $H_0 : \beta_1 = \beta_2 = 0$, and the alternative is H_a : any of $\beta_1, \beta_2 > 0$. In this case, the parameter vector, $\theta = (\beta_1, \beta_2, ...\beta_p)^T$, can be partitioned into two sets $\theta^T = (\gamma^T, \eta^T)$, e.g. $\gamma = (\beta_1, \beta_2)^T$ and $\eta = (\beta_3, ..., \beta_p)^T$. The parameter configuration is given by $(2, 0, 0, p-2)$, $C_{\Omega_0} = \{0\}^2 \times R^{p-2}$ and $C_\Omega = [0, +\infty)^2 \times R^{p-2}$, which is the same as the special case discussed in Section 3.3. Thus the asymptotic distribution of CLRT is a mixture of U, $d^*\chi_1^2/d$, $a^*\chi_1^2/a$, and χ_0^2 with mixing probabilities π_s, π_1, π_2 and $1-\pi_s-\pi_1-\pi_2$, respectively, where the weights and mixing probabilities are calculated correspondingly based on the equations given in Section 3.3.

4.2 Test of heterogeneity in diagnostic systematic reviews

As discussed in Section 2.3, the null hypothesis is $H_0 : \tau_1^2 = \tau_2^2 = 0$, and the parameters involved are $\theta = (\tau_1^2, \tau_2^2, \beta_1, \beta_2)^T$. Partition it into two parameter sets $\theta^T = (\gamma^T, \eta^T)$, where $\gamma = (\tau_1^2, \tau_2^2)^T$ and $\eta = (\beta_1, \beta_2)^T$. The parameter configuration is given by $(2, 0, 0, 2)$, $C_{\Omega_0} = \{0\}^2 \times R^2$ and $C_\Omega = [0, +\infty)^2 \times R^2$, which is the same as the special case in Section 3.3, with $p = 4$. Thus, the limiting distribution of CLRT is obtained by applying the results in Section 3.3.

4.3 Signal detection of AE reporting rate across calendar years

To test whether the reporting rate is the same across all years for jth vaccine-AE combination, the null hypothesis can be written as $H_0 : \tau_j^2 = \delta_j^2 = 0$. Similarly, partition the parameter vector, $\theta = (\tau_j^2, \delta_j^2, \alpha_{0j}, \beta_{0j})^T$, into two parameter sets $\theta^T = (\gamma^T, \eta^T)$, where $\gamma = (\tau_j^2, \delta_j^2)^T$ and $\eta = (\alpha_{0j}, \beta_{0j})^T$. The parameter configuration is given by $(2, 0, 0, 2)$, $C_{\Omega_0} = \{0\}^2 \times R^2$ and $C_\Omega = [0, +\infty)^2 \times R^2$, which is the same as the special case in Section 3.3, with $p = 4$. Thus, the results in Section 3.3 can be applied.

5. Simulation

To explore the finite sample performance of the theoretical findings, we conduct two simulation studies. In the first example, we consider the CLRT in stratified case-control studies as discussed in Section 4.1 with $p = 3$. In the second example, we test the heterogeneity of sensitivities and specificities between multiple studies in diagnostic systematic reviews as discussed in Section 4.2. The details are described as follows.

5.1 Test of Positive Associations in Stratified Case-Control Studies With Sparse Data

In this simulation, we simulate the data with 3 continuous covariates and simultaneously test $\beta_1 = \beta_2 = 0$, where β_1 and β_2 are known to be nonnegative. The nuisance parameter β_3 belongs to $(-\infty, +\infty)$. The three covariates are independently simulated from a standard normal distribution. To mimic the selection procedure of a case-control study, we simulated 1000 subjects for each stratum, and select 5 cases and 5 controls. The number of stratum varies from 25 to 200, and the simulation is repeated 5000 times. We compare the type I error and power of CLRT based on the derived asymptotic distribution with that based on the naive χ_2^2 distribution.

As shown in Table 1, we can see that the method based on the derived asymptotic distribution can control the type I error very well in all scenarios, while the naive method based on the χ_2^2 distribution yields grossly conservative type I error based on 5000 simulations. The power gain is also substantial (up to 48%) by using the derived asymptotic distribution of CLRT. Thus, the naive method is more conservative and less powerful comparing to the method based on the asymptotic distribution of CLRT.

5.2 Test of heterogeneity in diagnostic systematic reviews

As described in Section 2.3 and 4.2, we test the heterogeneity of sensitivities and specificities between multiple studies in diagnostic systematic reviews. We assume the covariates X_i and Z_i are univariate and independently generated from a standard normal distribution. We set $g(\cdot)$ to be the logit function, and set $\beta_1 = g(0.95)$, $\beta_2 = g(0.90)$, and $n_{i1} = n_{i0} = 50$. We simulate μ_{i1}, μ_{i2} from a bivariate normal distribution with $\rho = -0.8$, and $\tau_1^2 = \tau_2^2$ taking increasing values from 0, 0.05, 0.15 to 0.25. Then n_{i11} and n_{i00} are simulated from the binomial distribution as described in Section 2.3. We compare the type I error and power of CLRT based on the derived asymptotic distribution with that based on the naive χ_2^2 distribution under different numbers of studies m.

As shown in Table 2, the method based on the derived asymptotic distribution can control the type I error reasonably at nominal levels of 0.1, 0.05, and 0.01 in all scenarios, whereas the naive method yields grossly conservative type I error in this example based on 5000 simulations. The power gain is achieved up to 15% by using the asymptotic distribution of CLRT.

6. Application to a systematic review of modern imaging technologies for surveillance of melanoma

Melanoma is a type of skin tumor that develops from the pigment-containing cells known as melanocytes. Melanoma is a less common type of skin cancer, but it is much more dangerous when not found early and causes the majority (75%) of deaths related to skin cancer (Lo & Fisher, 2014). Modern imaging technology can be used for the early detection of melanoma metastasis and provides a cost-effective surveillance approach (Jemal *et al.* , 2009). Currently, the most commonly used diagnostic imaging technologies for melanoma include ultrasonography (US), computed tomography (CT), positron emission tomography (PET), and a combination of the latter two (PET-CT). In addition to evaluate the relative performance of these contemporary diagnostic imaging technologies in diagnosis of melanoma for patients at different stages, e.g., regional and distant lesions, it is also important to quantify the heterogeneity of different imaging technologies in terms of their operating characteristics, e.g., variability in sensitivity and specificity across study populations. Xing *et al.* (2011) conducted a diagnostic review based on 98 published studies of 10,528 patients carried out

between January 1st, 1990 and June 30th, 2009. The number of studies for each diagnostic imaging technology and type of cancer (regional and distant metastasis) are shown in Table S1 in the online Supplemental Material.

We applied the proposed composite likelihood model described in Section 2.3 and used the proposed composite likelihood ratio test described in Section 3.3 to test the heterogeneity of both sensitivities and specificities of the imaging technologies for different cancer types across multiple studies for 7 technology-cancer combinations. Our analyses found that, in diagnosis of regional metastatic melanoma, all four imaging technologies have significantly heterogeneous sensitivity or specificity across multiple studies ($p < 0.001$). In diagnosis of distance metastatic melanoma, CT and PET have significant heterogeneity of sensitivity or specificity across multiple studies ($p < 0.05$), while the combination of CT and PET does not ($p > 0.1$). Figure S1 in the online Supplemental Material demonstrates the range of the sensitivities and specificities of the imaging technologies in diagnosis of different cancer types across studies. We observed that the specificities are generally higher than the sensitivities, and the heterogeneity of specificity is smaller than sensitivity across studies for all the diagnosis methods. The heterogeneities of both sensitivity and specificity are higher in diagnosis of regional metastatic melanoma than distant metastatic melanoma. We also compared the results of the proposed methods to the naive method which ignores the boundary constraints. The naive method was more conservative although it still identified the significant heterogeneity of sensitivity or specificity of the four imaging technologies in diagnosis of regional metastatic melanoma, and the PET technology in diagnosis of distant metastatic melanoma, given the magnitude of these heterogeneities are relatively large. However, the naive method failed to identify the significant heterogeneity of sensitivity or specificity of CT technology in diagnosis of distant metastatic melanoma which was identified by the proposed method.

7. Discussion

In this paper, we derive the asymptotic distribution of CLRT when a subset of the testing parameters lie on the boundary of the parameter space, following the work of Self & Liang (1987) and Chen & Liang (2010). The former work studied the asymptotic behavior of the regular likelihood ratio test when param-

eters of interest lie on the boundary, and the later extended the results to deal with situations when a subset of the parameters of interest are tested based on the pseudolikelihood of Gong & Samaniego (1981). Considering the fact that the composite likelihood approach becomes increasingly popular and no such work has been done to study the asymptotic behaviors of CLRT under boundary constraints, our work strives to fill the gap. Notably, the results presented in this paper are very broad, in that any partially specified models can be considered as a special case of composite likelihood.

The composite likelihood ratio based inference under boundary constraints is generally a difficult question. Although the asymptotic results derived in this paper yield well controlled type I error and adequate power, calculation of the test statistic gets more complicated as the number of boundary parameters increase. Another alternative is to use numerical methods. However special attentions are needed in terms of validity, choice of tuning parameters, computational cost, and practical performance. For example, standard nonparametric and parametric bootstrap methods lead to inconsistency estimates when the parameter is on the boundary of the parameter space (Andrews, 2000). The inconsistency is due to the non-smoothness of the empirical distribution from which the bootstrap samples are generated. Andrews (2000) proposed subsampling and m-out-of-n bootstrap methods for obtaining consistent estimators of the limiting distributions of test statistics under boundary constraints. The m-out-of-n bootstrap provides a smoothing operation on the empirical distribution function, by resampling with replacement a smaller sample size m from the original n samples, where m is of a smaller order than n. However, these numerical methods require additional tuning parameters, and the performance varies in practice. For details of m-out-of-n bootstrap, please refer to Politis *et al.* (1999), and references therein.

To examine the practical performance of the m-out-of-n bootstrap method, we implemented the m-out-of-n bootstrap method to estimate the limiting distribution of the composite likelihood ratio statistic and tested positive associations in the example of stratified case-control studies with sparse data as described in Section 5.1. We adopted the method proposed by Bickel & Sakov (2008) to make data-adaptive choice of resample size m, with q = 0.85. The results are

shown in Section S4 of the online Supplemental Material. We observed that the performance of the m-out-of-n bootstrap method was between the proposed method and the naive method. More specifically, the empirical type I error rates of the m-out-of-n bootstrap method were closer to the nominal levels than the naive method, but were still very conservative compared to the proposed method. The power loss of the m-out-n bootstrap method was substantial, compared to the proposed method. In addition, the m-out-of-n bootstrap method required much more computational hours than the naive and the proposed methods. It is of great interest and importance to develop a method that can better balance between computational simplicity and statistical power. We are currently developing a test along the line of work by Susko (2013) using conditional test.

Table 1: Empirical rejection rates (%) in 5000 simulations of CLRT to test for two regression coefficients in stratified case-control study, based on different numbers of stratum K, stratum size N, and effect sizes.

		$\alpha = 0.10$ Rejection (%)		$\alpha = 0.05$ Rejection (%)		$\alpha = 0.01$ Rejection (%)	
$(\beta_1, \beta_2, \beta_3)$	(K,N)	CLRT	Naive	CLRT	Naive	CLRT	Naive
(0,0,0.1)	(25,10)	9.8	0.4	4.7	0.1	0.7	0.0
	(50,10)	10.5	0.3	5.2	0.1	1.0	0.0
	(100,10)	9.5	0.2	4.3	0.0	0.6	0.0
	(200,10)	10.1	0.5	4.9	0.1	1.1	0.0
(0.1,0,0.1)	(25,10)	25.9	2.4	15.7	0.6	4.4	0.0
	(50,10)	35.1	4.0	22.6	1.4	7.4	0.1
	(100,10)	51.4	9.6	37.5	4.1	15.3	0.5
	(200,10)	73.4	25.2	60.9	13.1	35.5	2.2
(0.1,0.1,0.1)	(25,10)	37.9	4.7	25.0	1.6	7.9	0.1
	(50,10)	56.1	11.0	41.4	4.4	17.9	0.5
	(100,10)	78.1	26.6	64.8	13.8	38.3	2.5
	(200,10)	95.1	62.0	90.0	44.7	73.3	16.2
(0.2,0.2,0.2)	(25,10)	76.2	24.5	62.9	12.7	32.7	2.0
	(50,10)	94.1	58.8	88.9	40.9	69.4	14.4
	(100,10)	99.8	93.0	99.3	84.8	96.3	57.3
	(200,10)	100.0	100.0	100.0	99.8	100.0	97.6

Acknowledgment

Research reported in this publication was supported in part by the National Institutes of Health (NIH) under award numbers R01LM012607 and 1R01AI130460.

Table 2: Empirical rejection rates (%) in 5000 simulations for the composite likelihood ratio test for heterogeneity in diagnostic systematic reviews, based on different number of studies m.

		$\alpha = 0.10$ Rejection (%)		$\alpha = 0.05$ Rejection (%)		$\alpha = 0.01$ Rejection (%)	
$\tau_1^2 = \tau_2^2$	m	CLRT	Naive	CLRT	Naive	CLRT	Naive
0	10	0.6	0.0	0.0	0.0	0.0	0.0
	20	2.5	0.3	0.7	0.1	0.0	0.0
	40	4.3	1.3	1.8	0.2	0.1	0.0
	80	7.0	2.3	3.0	0.8	0.3	0.0
	160	9.3	3.2	4.4	1.4	0.6	0.1
0.05	10	33.4	20.3	22.9	12.5	8.8	4.5
	20	33.9	20.8	23.1	13.4	9.1	4.8
	40	47.1	30.2	34.2	20.1	14.6	8.14
	80	57.8	39.7	44.2	29.3	21.7	12.5
	160	74.9	58.6	62.7	47.2	37.5	25.5
0.15	10	52.5	36.2	39.9	26.8	20.9	13.8
	20	68.7	55.6	59.0	46.6	39.2	28.7
	40	71.0	55.8	59.1	46.8	40.7	31.2
	80	97.7	95.7	96.2	93.6	91.6	87.3
	160	100.0	100.0	100.0	99.9	99.8	99.6
0.25	10	70.6	58.6	62.1	50.3	44.1	34.3
	20	92.2	87.2	88.8	82.8	78.2	70.3
	40	97.9	96.6	96.9	95.1	93.6	90.4
	80	100.0	99.8	99.9	99.8	99.7	99.3
	160	100.0	100.0	100.0	100.0	100.0	100.0

Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania, Philadelphia,

E-mail: jing14@upenn.edu

Department of Statistical Science, Cornell University, Ithaca, New York 14853, U.S.A.

E-mail: yn265@cornell.edu

AT&T Chief Data Office, AT&T Services Inc., Plano, Texas, U.S.A.

E-mail: yc010e@att.com

Institute of Population Health Science, National Health Research Institutes

E-mail: kyliang@nhri.org.tw

Department of Biostatistics, Epidemiology and Informatics, University of Penn-

sylvania, Philadelphia, PA 19104, U.S.A.
E-mail: ychen123@upenn.edu

Bibliography

Andrews, Donald WK. 2000. Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space. *Econometrica*, **68**(2), 399–405.

Bhat, Chandra R, Varin, Cristiano, & Ferdous, Nazneen. 2010. A comparison of the maximum simulated likelihood and composite marginal likelihood estimation approaches in the context of the multivariate ordered-response model. *Advances in Econometrics*, **26**, 65.

Bickel, Peter J, & Sakov, Anat. 2008. On the choice of m in the m out of n bootstrap and confidence bounds for extrema. *Statistica Sinica*, 967–985.

Chandler, Richard E, & Bate, Steven. 2007. Inference for clustered data using the independence loglikelihood. *Biometrika*, **94**(1), 167–183.

Chen, Robert T, Rastogi, Suresh C, Mullen, John R, Hayes, Scott W, Cochi, Stephen L, Donlon, Jerome A, & Wassilak, Steven G. 1994. The vaccine adverse event reporting system (VAERS). *Vaccine*, **12**(6), 542–550.

Chen, Yong, & Liang, Kung-Yee. 2010. On the asymptotic behaviour of the pseudolikelihood ratio test statistic with boundary problems. *Biometrika*, **97**(3), 603–620.

Chen, Yong, Liu, Yulun, Ning, Jing, Nie, Lei, Zhu, Hongjian, & Chu, Haitao. 2014. A composite likelihood method for bivariate meta-analysis in diagnostic systematic reviews. *Statistical methods in medical research*, 0962280214562146.

Chen, Yong, Liu, Yulun, Ning, Jing, Cormier, Janice, & Chu, Haitao. 2015. A hybrid model for combining case–control and cohort studies in systematic

reviews of diagnostic tests. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **64**(3), 469–489.

Chernoff, Herman. 1954. On the distribution of the likelihood ratio. *The Annals of Mathematical Statistics*, 573–578.

Cox, David R. 1975. Partial likelihood. *Biometrika*, **62**(2), 269–276.

Gong, Gail, & Samaniego, Francisco J. 1981. Pseudo maximum likelihood estimation: theory and applications. *The Annals of Statistics*, 861–869.

Guan, Yongtao. 2006. A composite likelihood approach in fitting spatial point process models. *Journal of the American Statistical Association*, **101**(476), 1502–1512.

He, Wenqing, & Yi, Grace Y. 2011. A pairwise likelihood method for correlated binary data with/without missing observations under generalized partially linear single-index models. *Statistica Sinica*, **21**(1), 207.

Heagerty, Patrick J, & Lele, Subhash R. 1998. A composite likelihood approach to binary spatial data. *Journal of the American Statistical Association*, **93**(443), 1099–1111.

Henderson, Robin, & Shimakura, Silvia. 2003. A serially correlated gamma frailty model for longitudinal count data. *Biometrika*, **90**(2), 355–366.

Huang, Lan, Zalkikar, Jyoti, & Tiwari, Ram C. 2011. A likelihood ratio test based method for signal detection with application to FDAs drug safety data. *Journal of the American Statistical Association*, **106**(496).

Irwig, Les, Macaskill, Petra, Glasziou, Paul, & Fahey, Michael. 1995. Meta-analytic methods for diagnostic test accuracy. *Journal of clinical epidemiology*, **48**(1), 119–130.

Jemal, Ahmedin, Siegel, Rebecca, Ward, Elizabeth, Hao, Yongping, Xu, Jiaquan, & Thun, Michael J. 2009. Cancer statistics, 2009. *CA: a cancer journal for clinicians*, **59**(4), 225–249.

Kent, John T. 1982. Robust properties of likelihood ratio tests. *Biometrika*, **69**(1), 19–27.

Larribe, F, & Fearnhead, P. 2011. On composite likelihoods in statistical genetics. *Statistica Sinica*, **21**(1), 43.

Liang, Kung-Yee. 1987. Extended Mantel-Haenszel estimating procedure for multivariate logistic regression models. *Biometrics*, 289–299.

Lindsay, Bruce G. 1988. Composite likelihood methods. *Contemporary Mathematics*, **80**(1), 221–39.

Lo, Jennifer A, & Fisher, David E. 2014. The melanoma revolution: from UV carcinogenesis to a new era in therapeutics. *Science*, **346**(6212), 945–949.

McVean, Gilean AT, Myers, Simon R, Hunt, Sarah, Deloukas, Panos, Bentley, David R, & Donnelly, Peter. 2004. The fine-scale structure of recombination rate variation in the human genome. *Science*, **304**(5670), 581–584.

Molenberghs, G., & Verbeke, G. 2005. *Models for discrete longitudinal data*. Springer.

Moses, Lincoln E, Shapiro, David, & Littenberg, Benjamin. 1993. Combining independent studies of a diagnostic test into a summary roc curve: Data-analytic approaches and some additional considerations. *Statistics in medicine*, **12**(14), 1293–1316.

Myers, Simon, Bottolo, Leonardo, Freeman, Colin, McVean, Gil, & Donnelly, Peter. 2005. A fine-scale map of recombination rates and hotspots across the human genome. *Science*, **310**(5746), 321–324.

Niu, Manette T, Erwin, Diane E, & Braun, M Miles. 2001. Data mining in the US Vaccine Adverse Event Reporting System (VAERS): early detection of intussusception and other events after rotavirus vaccination. *Vaccine*, **19**(32), 4627–4634.

Nott, David J, & Rydén, Tobias. 1999. Pairwise likelihood methods for inference in image models. *Biometrika*, **86**(3), 661–676.

Padoan, Simone A, Ribatet, Mathieu, & Sisson, Scott A. 2010. Likelihood-based inference for max-stable processes. *Journal of the American Statistical Association*, **105**(489), 263–277.

Politis, Dimitris N, Romano, Joseph P, & Wolf, Michael. 1999. Subsampling Springer-Verlag New York.

Reitsma, Johannes B, Glas, Afina S, Rutjes, Anne WS, Scholten, Rob JPM, Bossuyt, Patrick M, & Zwinderman, Aeilko H. 2005. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of clinical epidemiology*, **58**(10), 982–990.

Rutter, CM, & Gatsonis, CA. 1995. Regression methods for meta-analysis of diagnostic test data. *Academic radiology*, **2**, S48–56.

Self, Steven G, & Liang, Kung-Yee. 1987. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association*, **82**(398), 605–610.

Shimabukuro, Tom T, Nguyen, Michael, Martin, David, & DeStefano, Frank. 2015. Safety monitoring in the Vaccine Adverse Event Reporting System (VAERS). *Vaccine*, **33**(36), 4398–4405.

Stein, Michael L, Chi, Zhiyi, & Welty, Leah J. 2004. Approximating likelihoods for large spatial data sets. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **66**(2), 275–296.

Susko, Edward. 2013. Likelihood ratio tests with boundary constraints using data-dependent degrees of freedom. *Biometrika*, ast032.

Varin, Cristiano, & Vidoni, Paolo. 2008. Pairwise likelihood inference for general state space models. *Econometric Reviews*, **28**(1-3), 170–185.

Varin, Cristiano, Reid, Nancy, & Firth, David. 2011. An overview of composite likelihood methods. *Statistica Sinica*, **21**(1), 5–42.

Vecchia, Aldo V. 1988. Estimation and model identification for continuous spatial processes. *Journal of the Royal Statistical Society. Series B (Methodological)*, 297–312.

Wellner, J.A., & Zhang, Y. 2000. Two estimators of the mean of a counting process with panel count data. *Annals of Statistics*, 779–814.

Xing, Yan, Bronstein, Yulia, Ross, Merrick I, Askew, Robert L, Lee, Jeffrey E, Gershenwald, Jeffrey E, Royal, Richard, & Cormier, Janice N. 2011. Contemporary diagnostic imaging modalities for the staging and surveillance of melanoma patients: a meta-analysis. *Journal of the National Cancer Institute*, **103**(2), 129–142.