# Composite Likelihood Methods

David Firth

Department of Statistics
University of Warwick

*i-like* Launch Event, Oxford, 2013-01-31

---

## Plan

What are composite likelihoods?

Where are composite likelihoods used?

Some open areas

---

## What are composite likelihoods?

Full likelihood:
$$L(\boldsymbol{\theta}; \mathbf{y})$$

Problem: intractability.

For example, realistic/interesting models often involve unobservables, $\mathbf{u}$ say, that have to be integrated out:

$$L(\boldsymbol{\theta}; \mathbf{y}) = \int (\mathbf{y}|\mathbf{u}; \boldsymbol{\theta}) g(\mathbf{u}) \, \mathrm{d}\mathbf{u}$$

The integral may be very high-dimensional.

Solution: use a more readily computed 'pseudo-likelihood' constructed from low-dimensional (conditional or marginal) views of $\mathbf{y}$.

---

A fairly general setup:

- $\{A_1, \ldots, A_k\}$ a set of marginal or conditional events
- associated ('component') likelihoods $L_k(\boldsymbol{\theta}; \mathbf{y})$.

A *composite likelihood* based on those components is a weighted product

$$L_{\mathsf{C}}(\boldsymbol{\theta}; \mathbf{y}) = \prod_{k=1}^{K} L_k(\boldsymbol{\theta}; \mathbf{y})^{w_k},$$

with weights $w_k \geq 0$ to be chosen.

# Conditional and marginal components

**Conditional** CL: e.g., Besag (1974) on approximate inference for spatial processes,

$$L_{\mathsf{C}}(\boldsymbol{\theta}; \mathbf{y}) = \prod_{r=1}^{m} f(y_r | \{y_s : y_s \text{ is neighbour of } y_r\}; \boldsymbol{\theta}).$$

Other examples in analysis of time series, longitudinal studies, gene expression data, etc; e.g., pooling of full conditional densities,

$$L_{\mathsf{C}}(\boldsymbol{\theta}; \mathbf{y}) = \prod_{r=1}^{m} f(y_r | y_{(-r)}; \boldsymbol{\theta}).$$

**Marginal** CL: Simplest example is the *independence* pseudo-likelihood:

$$L_{\mathsf{ind}}(\boldsymbol{\theta}; \mathbf{y}) = \prod_{r=1}^{m} f(y_r; \boldsymbol{\theta})$$

(or *onewise likelihood*). This neglects dependence structures; still sometimes useful, but can be inefficient.

One step further is a *pairwise* marginal composite likelihood, e.g., based on marginal density of all pairs,

$$L_{\mathsf{pair}}(\boldsymbol{\theta}; \mathbf{y}) = \prod_{r=1}^{m-1} \prod_{s=r+1}^{m} f(y_r, y_s; \boldsymbol{\theta}).$$

This typically *is* informative about dependence parameters.

Possibly combine onewise and pairwise likelihoods for increased efficiency (Cox and Reid, 2004).

# Efficiency?

*Pseudo-likelihood, not likelihood.* Fisher information not the right measure.

Key matrices that generalize the Fisher information are the *sensitivity*

$$H(\boldsymbol{\theta}) = \mathsf{E}\left\{-\nabla^2 \log L_{\mathsf{C}}(\boldsymbol{\theta}; \mathbf{Y})\right\}$$

and *variability*

$$J(\boldsymbol{\theta}) = \mathsf{var}\left\{\nabla \log L_{\mathsf{C}}(\boldsymbol{\theta}; \mathbf{Y})\right\}.$$

These combine to give the Godambe ('sandwich') information matrix,

$$G(\boldsymbol{\theta}) = H(\boldsymbol{\theta})\left[J(\boldsymbol{\theta})^{-1}\right]H(\boldsymbol{\theta}).$$

- An estimate of $G^{-1}$ provides approximate standard errors of maximum composite likelihood estimators.
- Variability matrix $J$ is often difficult to estimate (the most extreme cases being analyses of single time-series, or single spatial arrays, where there is no replication to help with this).
- Efficiency calculations compare $G(\theta)$ with the full-likelihood Fisher information matrix.
- Aim to choose a composite likelihood that makes $G(\theta)$ as 'large' as possible (subject to computational tractability).

i-like.org.uk

# Where are composite likelihoods used?
Lots of application areas already; and still growing rapidly.

Some 'classic' application areas are

- **genetics** (e.g., 2011 review paper by Larribe and Fearnhead)
- **geostatistics** (starting with Hjort and Omre, 1994)
- **spatial extremes** (recent work of A C Davison and others)
- models with **correlated random effects** (spatial models; time series; multivariate/longitudinal data; network dependence models; etc.)
- **financial econometrics**

and various others. (See VRF2011 for some more)

i-like.org.uk

### An indication from 2013 so far: 6 citations of VRF2011

An overview of composite likelihood methods
☐ Search within citing articles

On Graphical Models via Univariate Exponential Family Distributions
E Yang, P Ravikumar, GI Allen, Z Liu - arXiv preprint arXiv:1301.4183, 2013 - arxiv.org
Abstract: Undirected graphical models, or Markov networks, are a popular class of statistical models, used in a wide variety of applications. Popular instances of this class include Gaussian graphical models and Ising models. In many settings, however, it might not be ...
Cite

Marisol Castro, Rajesh Paleti, Chandra R. Bhat
M Castro, R Paleti, CR Bhat - Accident Analysis & Prevention, 2013 - Elsevier
Abstract This paper proposes a flexible econometric structure for injury severity analysis at the level of individual crashes that recognizes the ordinal nature of injury severity categories, allows unobserved heterogeneity in the effects of contributing factors, as well as ...
Cite

Estimation and prediction in spatial models with block composite likelihoods
J Eidsvik, BA Shaby, BJ Reich ... - ... of Computational and ..., 2013 - Taylor & Francis
Abstract A block composite likelihood is developed for estimation and prediction in large spatial datasets. The composite likelihood is constructed from the joint densities of pairs of adjacent spatial blocks. This allows large datasets to be split into many smaller datasets, ...
Cite

[PDF] Estimation in the Presence of Many Nuisance Parameters: composite likelihood and plug-in likelihood
B Wu, Q Yao, S Zhu - 2013 - stats.lse.ac.uk
Abstract We consider the estimation for a small number of parameters of interest for a stationary process in the presence of a large number nuisance parameters. Two estimation methods, namely the maximum composite quasi-likelihood estimation (MCQLE) and the ...
Cite  More

Composite Likelihood Estimation for Latent Variable Models with Ordinal and Continuous, or Ranking Variables
M Katsikatsou - 2013 - uu.diva-portal.org
Dissertation presented at Uppsala University to be publicly examined in Hörsal 2, Ekonomikum, Kyrkogårdsgatan 10, Uppsala, Friday, February 15, 2013 at 10:15 for the degree of Doctor of Philosophy. The examination will be conducted in English.
Cite

Extremes: Spatial Parametric Statistical Modeling
AC Davison - Encyclopedia of Environmetrics - Wiley Online Library
Statistics of spatial extremes has undergone rapid growth in recent years, not least because of the realization that risk analysis for major impacts of climate change depends on modeling rare events with spatial consequences. Obvious examples are prolonged heavy rainfall ...

i-like.org.uk

Cornell University Library

arXiv.org > math > arXiv:1301.4183

**Mathematics > Statistics Theory**

## On Graphical Models via Univariate Exponential Family Distributions

Eunho Yang, Pradeep Ravikumar, Genevera I. Allen, Zhandong Liu

*(Submitted on 17 Jan 2013)*

Undirected graphical models, or Markov networks, are a popular class of statistical models, used in a wide variety of applica... include Gaussian graphical models and Ising models. In many settings, however, it might not be clear which subclass of gra... non-Gaussian and non-categorical data. In this paper, we consider a general sub-class of graphical models where the node... exponential families. This allows us to derive multivariate graphical model distributions from univariate exponential family dis... binomial, and exponential distributions. Our key contributions include a class of M-estimators to fit these graphical model dis... analysis showing that these M-estimators recover the true graphical model structure exactly, with high probability. We provi... networks learned via instances of our class of graphical models derived from Poisson and exponential distributions.

Subjects:  **Statistics Theory (math.ST)**; Machine Learning (stat.ML)
Cite as:  **arXiv:1301.4183 [math.ST]**
        (or **arXiv:1301.4183v1 [math.ST]** for this version)

i-like.org.uk

Accident Analysis & Prevention

Volume 52, 28 March 2013, Pages 188–203

**A spatial generalized ordered response model to examine highway crash injury severity**

Marisol Castro[1], Rajesh Paleti[1], Chandra R. Bhat

The University of Texas at Austin, Department of Civil, Architectural and Environmental Engineering, 301 E. Dean Keeton St. Stop C1761, Austin, TX 78712-1172, United States

http://0-dx.doi.org.pugwash.lib.warwick.ac.uk/10.1016/j.aap.2012.12.009, How to Cite or Link Using DOI

i-like.org.uk
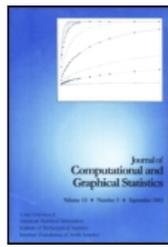
**Journal of Computational and Graphical Statistics**
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/ucgs20

**Estimation and prediction in spatial models with block composite likelihoods**

Jo Eidsvik [a], Benjamin A. Shaby [b], Brian J. Reich [c], Matthew Wheeler [d] & Jarad Niemi [e]

[a] Department of Mathematical Sciences, NTNU, Trondheim, 7491, Norway

[b] Department of Statistics, University of California, Berkeley, Berkeley, California, 94720, U.S.A.

[c] Department of Statistics, North Carolina State University, Raleigh, North Carolina, 27695, U.S.A.

[d] Department of Statistics and Applied Probability, University of California at Santa Barbara, Santa Barbara, California, 93106, U.S.A.

[e] Department of Statistics, Iowa State University, Ames, IA, 50011, U.S.A.
Accepted author version posted online: 16 Jan 2013.

i-like.org.uk

**Estimation in the Presence of Many Nuisance Parameters: composite likelihood and plug-in likelihood**

Billy Wu*     Qiwei Yao*,*     Shiwu Zhu‡

*Department of Statistics, London School of Economics, London, UK

‡School of Economics and Management, Tsinghua University, Beijing, China

Email: q.yao@lse.ac.uk

22 January 2013

i-like.org.uk

*Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences* 86

**Composite Likelihood Estimation for Latent Variable Models with Ordinal and Continuous, or Ranking Variables**

MYRSINI KATSIKATSOU

i-like.org.uk

**Extremes: Spatial Parametric Statistical Modeling**

Anthony C. Davison

Book Title

Published Online: 15 JAN 2013

DOI: 10.1002/9780470057339.vnn015

Copyright © 2002 John Wiley & Sons, Ltd

Encyclopedia of Environmetrics

i-like.org.uk

**Extremes: Spatial Parametric Statistical Modeling**

Anthony C. Davison

Book Title

Published Online: 15 JAN 2013

DOI: 10.1002/9780470057339.vnn015

Copyright © 2002 John Wiley & Sons, Ltd

Encyclopedia of Environmetrics

STATISTICAL SCIENCE

previous :: next

**Statistical Modeling of Spatial Extremes**

**A. C. Davison, S. A. Padoan, and M. Ribatet**

**Source:** Statist. Sci. Volume 27, Number 2 (2012), 161-186.

i-like.org.uk

# Some open areas



Workshops at Warwick CRiSM (2008) and Banff International Research Station (2012). Report on 2012 workshop (and webcasts of some of the talks) available online from

`www.birs.ca`

i-like.org.uk

# Some open areas

There are lots! To mention just a few general ones:

- ▶ choice of components
- ▶ choice of 'weights'
- ▶ robustness? (e.g., Helen Jordan talk at Banff 2012)
- ▶ reliable estimation of variability matrix $J(\theta)$
- ▶ *Bayesian* use of $L_C$? (Pauli+ 2011; Ribatet+ 2012)
- ▶ interplay with other approaches (ABC, simulated likelihoods, …)

etc., etc.

i-like.org.uk