Statistics and Probability Letters 98 (2015) 1–5

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

A note on asymptotic distributions in maximum entropy models for networks

Ting Yan *

Central China Normal University, China

A R T I C L E I N F O

Article history:
Received 26 October 2014
Received in revised form 2 December 2014
Accepted 2 December 2014
Available online 11 December 2014

MSC:
62E20
62F12

Keywords:
Central limit theorem
Diverging number of parameters
Network models

A B S T R A C T

A central limit theorem for a linear combination of all the maximum likelihood estimators with an increasing dimension in maximum entropy models for network data, has been established. Simulation studies illustrate the asymptotic results.

1. Introduction

The degrees of the vertices of the corresponding graphs preliminarily reflect the summarized information of network structures and have been studied by many authors recently. See Blitzstein and Diaconis (2010), Chatterjee et al. (2011), Chung and Lu (2002), Hillar and Wibisono (2013), Janson (2010), Olhede and Wolfe (2012), Perry and Wolfe (2012), Rinaldo et al. (2013) and Zhao et al. (2012). If one only considers the distributions of degrees, a typical realization is putting the degree sequence, i.e., a vector combined by the degrees of all vertices, as the exclusively sufficient statistics for the exponential family distributions. It is called the β-model by Chatterjee et al. (2011) for binary edges (i.e., edges with only two status, "present" or "absent"), an undirected version of the p_1 model for directed graphs by Holland and Leinhardt (1981). Motivated by applications to neuron science, Hillar and Wibisono (2013) have generalized it to weighted (discrete or continuous) edges according to the maximum entropy principle, which will be named "maximum entropy network models" hereafter. These models are useful to reconstruct the network in a situation such as sexually transmitted disease networks, in which only the information of the degree sequence is available due to privacy protection (e.g., Helleringer and Kohler, 2007).

A notable characterization of the maximum entropy network models is that each vertex is assigned an intrinsic parameter, called "potential" by Hillar and Wibisono (2013), which measures the strength of that vertex to form network connections. As the size of networks increases, the number of parameters goes to infinity, making asymptotic analysis challenging (Fienberg, 2012). The asymptotic properties of the maximum likelihood estimators (MLEs) have been little known until recently. As the number of parameters goes to infinity, the MLE has been proved to be uniformly consistent by Chatterjee et al. (2011) in the β-model and by Hillar and Wibisono (2013) in the maximum entropy network models; Yan and Xu (2013) and Yan et al. (2015) derived the asymptotic normality for a fixed number of the MLEs. A natural question

* Correspondence to: Department of Statistics, Central China Normal University, Wuhan, 430079, China.
E-mail address: tingyanty@mail.ccnu.edu.cn.

http://dx.doi.org/10.1016/j.spl.2014.12.003
0167-7152/

appears. Is there similar results for a diverging number of the MLEs? This paper aims to solve this problem. We will establish the central limit theorem for a linear combination of all the MLEs given a sequence of real valued coefficients when the number of parameters goes to infinity. Obviously, the previous results on the fixed number of the estimators is a specific case of the latter. For the remaining part of this paper, we proceed as follows. In Section 2, we have presented the main results. Simulation studies are given in Section 3. We made some discussion in Section 4. All proofs are relegated to online supplementary material (see Appendix A).

2. Main results

Consider an undirected graph $\mathcal{G}$ with no self-loops on n vertices labeled by "1, . . . , n". Let a_{ij} be the weight of edge (i, j) taking values from the set Ω, where Ω could be a finite discrete, infinite discrete or continuous set, and $a_{ij} = a_{ji}$. We set $a_{ii} = 0$ $(i = 1, \ldots, n)$ for convenience. The symmetric matrix $\mathbf{a} = (a_{ij})_{i,j=1}^n$ is called the adjacent matrix of the graph $\mathcal{G}$. Denote $d_i = \sum_{j\neq i} a_{ij}$ by the degree of vertex i, and $\mathbf{d} = (d_1, \ldots, d_n)^T$ by the degree sequence of $\mathcal{G}$. Let $\mathcal{S}$ be a σ-algebra over the set Ω. Assume that there is a canonical σ-finite probability measure ν on $(\Omega, \mathcal{S})$. If Ω is discrete, then ν is the counting measure; if Ω is continuous, then ν is the Borel measure. Let $\nu^{\binom{n}{2}}$ be the product measure on $\Omega^{\binom{n}{2}}$. The maximum entropy network models assume that the probability density function of $\mathbf{a} = (a_{ij})_{i,j=1}^n$ with respective to $\nu^{\binom{n}{2}}$ has the exponential form with the degree sequence as natural sufficient statistics, i.e.,

$$p_\theta(\mathbf{a}) = \exp\big(-\boldsymbol{\theta}^T\mathbf{d} - z(\boldsymbol{\theta})\big), \tag{1}$$

where $z(\boldsymbol{\theta})$ is the normalizing constant and $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)^T$ is the parameter vector. Remark that we use $-\theta$ in the parameterization instead of the classical θ since it will simplify the notations in the later presentation. Model (1) implies that the edges a_{ij}, $1 \le i < j \le n$ are mutually independent.

Let $\widehat{\boldsymbol{\theta}} = (\hat\theta_1, \ldots, \hat\theta_n)^T$ be the MLE of $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)^T$ and $\{c_i\}_{i=1}^\infty$ be a fixed sequence. Denote V_n by the Fisher information matrix of the vector parameter $\boldsymbol{\theta}$. In order to solve the asymptotic distribution of $(\hat\theta_1, \ldots, \hat\theta_r)^T$ with a fixed r, Yan et al. (2015) have obtained an approximate inverse of V_n to derive its approximately explicit expressions as a function of the vertex degrees d_i, $i = 1, \ldots, r$. Here, the approximate inverse is used for representing the linear combination of all the MLEs as a function of the linear combination of all the degrees to derive its asymptotic distribution. Details are given in online supplementary material (see Appendix A). Consider the condition:

$$\sum_{i=1}^{\infty} |c_i| < \infty. \tag{2}$$

The above condition implies $\sum_{i=1}^\infty c_i^2 < \infty$ that is common in the study of asymptotic distributions for a weighted sum of a sequence of independent random variables (e.g. Theorem 4.2 of Billingsley, 1968; Chow and Lai, 1973). The central limit theorems for the linear combination of all the MLEs in the cases of finite discrete, continuous and infinite discrete are given in the following three subsections, respectively.

2.1. Finite discrete weights

In this subsection, we assume that network edges take values from the set $\Omega = \{0, 1, \ldots, q-1\}$ with q a fixed integer. In this case, the edge weights a_{ij} for all $i \neq j$ are independently multinomial random variables with the probability:

$$P(a_{ij} = a) = \frac{e^{a(\theta_i+\theta_j)}}{\sum_{k=0}^{q-1} e^{k(\theta_i+\theta_j)}}, \quad a = 0, 1, \ldots, q-1.$$

This model is a direct generalization of the β-model for dichotomous edges. The normalizing constant is

$$z(\boldsymbol{\theta}) = \sum_{1\le i<j\le n} \log \sum_{a=0}^{q-1} e^{-(\theta_i+\theta_j)a},$$

and the parameter space is $\Theta = R^n$. By direct calculations, V_n has the following representation:

$$v_{ij} = \frac{\sum_{0\le k<l\le q-1} (k-l)^2 e^{(k+l)(\theta_i+\theta_j)}}{\left(\sum_{a=0}^{q-1} e^{a(\theta_i+\theta_j)}\right)^2}, \quad i, j = 1, \ldots, n; i \neq j, \qquad v_{ii} = \sum_{j=1; j\neq i}^{n} v_{ij}. \tag{3}$$

Following Hillar and Wibisono (2013), we assume that $\max_{1\le i\le n} |\theta_i|$ is bounded by a constant. The central limit theorem for the linear combination of all the estimators is stated as follows, whose proof is built on the work of Yan and Xu (2013) and Yan et al. (2015) and relegated to online supplementary material (see Appendix A).