MOMENTS OF PARAMETER ESTIMATES FOR CHUNG-LU RANDOM GRAPH MODELS

Nicholas Arcolano[1], *Karl Ni*[1], *Benjamin A. Miller*[1], *Nadya T. Bliss*[1], *and Patrick J. Wolfe*[2]

[1]MIT Lincoln Laboratory, Lexington, MA 02420, USA
{arcolano, karl.ni, bamiller, nt}@ll.mit.edu
[2]Statistics and Information Sciences Laboratory, Harvard University, Cambridge, MA 02138, USA
patrick@seas.harvard.edu

ABSTRACT

As abstract representations of relational data, graphs and networks find wide use in a variety of fields, particularly when working in non-Euclidean spaces. Yet for graphs to be truly useful in in the context of signal processing, one ultimately must have access to flexible and tractable statistical models. One model currently in use is the Chung-Lu random graph model, in which edge probabilities are expressed in terms of a given expected degree sequence. An advantage of this model is that its parameters can be obtained via a simple, standard estimator. Although this estimator is used frequently, its statistical properties have not been fully studied. In this paper, we develop a central limit theory for a simplified version of the Chung-Lu parameter estimator. We then derive approximations for moments of the general estimator using the delta method, and confirm the effectiveness of these approximations through empirical examples.

Index Terms— graphs and networks, central limit theory, delta method, given expected degree models, parameter estimation

1. INTRODUCTION

A **graph** is defined as a set of nodes (or vertices) and edges, where each edge indicates a relationship between a pair of nodes (or a node and itself, if self-loops are allowed). The simplicity of this definition, however, belies a mathematical object of immense flexibility and expressive power. Today, graphs arise in a wide variety of application domains, and graph analysis has become a pervasive and critical area of research. (For an extensive review of existing graph models, methods, and applications, we direct the reader to [1, 2].)

Let G be an undirected, unweighted graph with n nodes. In order to take full advantage of G as an abstract representation of underlying data, one often will choose to characterize G in terms of a statistical model. One of the simplest approaches is to assume G is a realization of an **Erdős-Rényi** random graph, where the existence of each edge is determined by an independent Bernoulli trial with fixed success probability $p \in (0, 1)$. Equivalently, if we define $\mathbf{A} = \{a_{ij}\}$ to be the adjacency matrix of G, (i.e. $\mathbf{A}$ is a binary symmetric matrix where $a_{ij} = 1$ if and only if there exists an edge between the i-th and j-th nodes), the Erdős-Rényi model assumes $\{a_{ij} : i \geq j\}$ are independent and identically distributed Bernoulli random variables with parameter p.

This work is sponsored by the United States Air Force under contract FA8721-05-C-0002. Opinions, interpretations, recommendations, and conclusions are those of the authors and are not necessarily endorsed by the United States Government.

Unfortunately, though simple and tractable, the Erdős-Rényi model often is not practical for real-world applications. For example, consider the degrees of each node, given by $d_i = \sum_{j=1}^{n} a_{ij}$ for $i = 1, \ldots, n$. Under the Erdős-Rényi model, $d_1, \ldots, d_n$ are identically distributed binomial random variables, with expected value $\mathbb{E}(d_i) = np$ (assuming self-loops are allowed). Consequently, this model is not effective at representing data whose corresponding degree sequence is far from constant, such as any data that exhibits power-law behavior.

One way to address this shortcoming is to define a set of node weights $\{w_1, \ldots, w_n\}$ where $w_i \in (0, 1)$ for $i = 1, \ldots, n$, and assume that edges exist independently between pairs of nodes i and j with probability $p_{ij} = w_i w_j$. This approach—which can be viewed as a generalization of the Erdős-Rényi model—is known as the **Chung-Lu model** [3, 4]. Since its introduction, the Chung-Lu model has proved useful in a number of settings, particularly in the context of modularity theory, where it has been applied to problems such as finding community structure [5], graph partitioning [6], and detection of dense subgraphs in large graphs [7].

As is true for any parametric statistical model, practical application of the Chung-Lu model requires either that we know the model parameters a priori, or (more typically) that we can estimate them given observed data. The standard estimator for the node weights is

$$\hat{w}_i = \frac{d_i}{\left(\sum_{j=1}^{n} d_j\right)^{1/2}}, \qquad (1)$$

where $d_1, \ldots, d_n$ denote the *observed* degrees of each node. Though frequently used in practice [3–7], this estimator has not been well studied, and its theoretical properties have received little attention.

In this paper, we investigate statistics of the estimator in (1), including results for both finite n and asymptotically as $n \to \infty$. We begin by formally defining the Chung-Lu model in Section 2. In Section 3, we develop a central limit theory for a simplified version of the estimator, and derive approximations for moments of the general estimator using an approach known as the **delta method**. Finally, we validate our theoretical results and study the effectiveness of our approximations through a series of empirical examples in Section 4, and close with a summary.

2. THE CHUNG-LU RANDOM GRAPH MODEL

Let G be an undirected, unweighted random graph with n nodes, and let $\mathbf{w} = [w_1 \cdots w_n]^T$ be a vector of weights such that $w_i \in (0, 1)$ for $i = 1, \ldots, n$. To simplify computation we allow G to have self-loops, although our results may be extended to the case where such edges are prohibited.

We say that G follows a Chung-Lu random graph model with n nodes and parameter $\mathbf{w}$, denoted $G(n, \mathbf{w})$, if the presence or absence of each edge is determined by an independent Bernoulli trial, with the probability p_{ij} of an edge between the i-th and j-th nodes given by $p_{ij} = w_i w_j$. Alternatively, if $\mathbf{A} = \{a_{ij}\}$ is the adjacency matrix of G, then an equivalent definition is to say that G follows a Chung-Lu model if $\{a_{ij} : i \geq j\}$ are independent Bernoulli random variables with corresponding parameters $\{p_{ij}\}$.

Let d_i denote the degree of the i-th node, for $i = 1, \ldots, n$. A basic property of the Chung-Lu model is that

$$\mathbb{E}\left(d_i\right) = \sum_{j=1}^{n} \mathbb{E}\left(a_{ij}\right) = \sum_{j=1}^{n} w_i w_j = w_i \left\|\mathbf{w}\right\|_1, \qquad (2)$$

where $\|\mathbf{w}\|_p$ denotes the vector p-norm $\|\mathbf{w}\|_p \equiv \left(\sum_{i=1}^{n} |w_i|^p\right)^{1/p}$. Since the parameter vector $\mathbf{w}$ is proportional to the expected degree sequence, the Chung-Lu model is also referred to as the "given expected degree" model [3].

As discussed in Section 1, in practice one rarely has prior knowledge of $\mathbf{w}$, and thus it must be estimated from data. In the case of the Chung-Lu model, the estimator given in (1) is commonly used to compute $\hat{w}_1 \ldots, \hat{w}_n$ given a sequence of observed degrees $d_1, \ldots, d_n$. Although this approach is used frequently in practice for fitting the Chung-Lu model to empirical data, its characteristics as a parameter estimator have yet to be properly studied. Thus, we proceed by determining expressions for statistics of this estimator and studying some of its asymptotic error properties.

3. STATISTICS OF CHUNG-LU PARAMETER ESTIMATES

For the remainder of the paper, we assume G is an observed random graph generated according to the Chung-Lu model with n nodes and weight vector $\mathbf{w}$. We denote the adjacency matrix of G by the $n \times n$ binary symmetric matrix $\mathbf{A} = \{a_{ij}\}$, and denote the degree sequence of G by $\{d_1, \ldots, d_n\}$ where $d_i = \sum_{j=1}^{n} a_{ij}$.

As a function of the degree sequence of a random graph, the parameter estimator given by (1) is itself a random variable. Accordingly, to determine the effectiveness of $\hat{\mathbf{w}} = [\hat{w}_1 \cdots \hat{w}_n]^T$ as an estimator, we will want to study its distribution for finite n, and determine whether it converges to the true weight vector $\mathbf{w}$ as $n \to \infty$.

3.1. Preliminaries

In addition to the expression for $\mathbb{E}\left(d_i\right)$ given in (2), we will find it necessary to compute second moments of the form $\mathbb{E}\left(d_i\, d_j\right)$ both when $i = j$ and $i \neq j$, as well as the variance $\text{var}\left(d_i\right)$ and the covariance $\text{cov}\left(d_i, d_j\right)$. When $i = j$, we have

$$\begin{aligned}
\mathbb{E}\left(d_i^2\right) &= \mathbb{E}\left[\left(\sum_{j=1}^{n} a_{ij}\right)\left(\sum_{j'=1}^{n} a_{ij'}\right)\right] \\
&= \sum_{j=1}^{n} \mathbb{E}\left(a_{ij}^2\right) + \sum_{j=1}^{n} \sum_{j' \neq j} \mathbb{E}\left(a_{ij} a_{ij'}\right) \\
&= \sum_{j=1}^{n} w_i w_j + \sum_{j=1}^{n} \sum_{j' \neq j} w_i^2 w_j w_{j'} \\
&= w_i \left\|\mathbf{w}\right\|_1 - w_i^2 \sum_{j=1}^{n} w_j^2 + w_i^2 \sum_{j=1}^{n} \sum_{j'=1}^{n} w_j w_{j'} \\
&= w_i \left\|\mathbf{w}\right\|_1 - w_i^2 \left\|\mathbf{w}\right\|_2^2 + w_i^2 \left\|\mathbf{w}\right\|_1^2, \qquad (3)
\end{aligned}$$

and when $i \neq j$, we have

$$\begin{aligned}
\mathbb{E}\left(d_i\, d_j\right) &= \mathbb{E}\left[\left(\sum_{j'=1}^{n} a_{ij'}\right)\left(\sum_{i'=1}^{n} a_{i'j}\right)\right] \\
&= \mathbb{E}\left(a_{ij}^2\right) + \sum_{i'=1}^{n} \sum_{j'=1}^{n} \mathbb{E}\left(a_{ij'} a_{i'j}\right) \mathbb{I}\left(i \neq i', j \neq j'\right) \\
&= w_i w_j + \sum_{i'=1}^{n} \sum_{j'=1}^{n} w_i w_{j'} w_{i'} w_j \,\mathbb{I}\left(i \neq i', j \neq j'\right) \\
&= w_i w_j - w_i^2 w_j^2 + w_i w_j \sum_{i'=1}^{n} \sum_{j'=1}^{n} w_{i'} w_{j'} \\
&= w_i w_j - w_i^2 w_j^2 + w_i w_j \left\|\mathbf{w}\right\|_1^2,
\end{aligned}$$

where $\mathbb{I}\left(\cdot\right)$ denotes the indicator function (i.e. $\mathbb{I}\left(\cdot\right) = 1$ when its argument is true, and $\mathbb{I}\left(\cdot\right) = 0$ otherwise). Finally, we have

$$\text{var}\left(d_i\right) = \mathbb{E}\left(d_i^2\right) - \mathbb{E}^2\left(d_i\right) = w_i \left\|\mathbf{w}\right\|_1 - w_i^2 \left\|\mathbf{w}\right\|_2^2,$$

and

$$\text{cov}\left(d_i, d_j\right) = \mathbb{E}\left(d_i\, d_j\right) - \mathbb{E}\left(d_i\right)\mathbb{E}\left(d_j\right) = w_i w_j - w_i^2 w_j^2.$$

3.2. Error distribution for a simplified estimator

To proceed, let us temporarily assume that $\left\|\mathbf{w}\right\|_1 = \sum_{i=1}^{n} w_i$ is known, as this restricted setting will provide greater insight into the behavior of the general estimator. In this case, we could obtain a simpler estimator by replacing $\sum_{j=1}^{n} d_j$ in (1) with its expected value, yielding

$$\hat{w}_i = \frac{d_i}{\left(\mathbb{E}\left(\sum_{j=1}^{n} d_j\right)\right)^{1/2}} = \frac{d_i}{\left(\sum_{j=1}^{n} \mathbb{E}\left(d_j\right)\right)^{1/2}} = \frac{d_i}{\left\|\mathbf{w}\right\|_1}.$$

Thus, when $\left\|\mathbf{w}\right\|_1$ is known the traditional estimator reduces to a re-scaling of the observed degrees. This estimator is unbiased, as

$$\mathbb{E}\left(\hat{w}_i - w_i\right) = \frac{\mathbb{E}\left(d_i\right)}{\left\|\mathbf{w}\right\|_1} - w_i = 0.$$

Its mean squared error (MSE) is given by

$$\begin{aligned}
\text{MSE}\left(\hat{w}_i\right) = \mathbb{E}\left[\left(\hat{w}_i - w_i\right)^2\right] &= \mathbb{E}\left(\hat{w}_i^2\right) - w_i^2 \\
&= \frac{w_i}{\left\|\mathbf{w}\right\|_1} - w_i^2 \frac{\left\|\mathbf{w}\right\|_2^2}{\left\|\mathbf{w}\right\|_1^2},
\end{aligned}$$

where

$$\mathbb{E}\left(\hat{w}_i^2\right) = \frac{\mathbb{E}\left(d_i^2\right)}{\left\|\mathbf{w}\right\|_1^2} = \frac{w_i}{\left\|\mathbf{w}\right\|_1} - w_i^2 \frac{\left\|\mathbf{w}\right\|_2^2}{\left\|\mathbf{w}\right\|_1^2} + w_i^2,$$

from (3). Consequently, as long as $\left\|\mathbf{w}\right\|_1$ grows without bound as $n \to \infty$ (i.e. as long as $\mathbf{w}$ is not an absolutely summable sequence), we will have $\text{MSE}\left(\hat{w}_i\right) \to 0$ as $n \to \infty$.

We also can prove a central limit theorem (CLT) for this estimator. Recall that for fixed i, $\{a_{i1}, \ldots, a_{in}\}$ is a sequence of independent Bernoulli random variables with corresponding probabilities $\{w_i\, w_1, \ldots, w_i\, w_n\}$. Next, define

$$\begin{aligned}
\mu_j &\equiv \mathbb{E}\left(a_{ij}\right) = w_i w_j, \\
\sigma_j^2 &\equiv \text{var}\left(a_{ij}\right) = w_i w_j - w_i^2 w_j^2,
\end{aligned}$$

and

$$s_n^2 \equiv \sum_{j=1}^{n} \sigma_j^2 = w_i \|\mathbf{w}\|_1 - w_i^2 \|\mathbf{w}\|_2^2 = \text{var}(d_i).$$

It is relatively straightforward to show that Lyapunov's condition [8]

$$\lim_{n\to\infty} \frac{1}{s_n^{2+\delta}} \sum_{j=1}^{n} \mathbb{E}\left(|a_{ij} - \mu_j|^{2+\delta}\right) = 0, \tag{4}$$

holds for $\delta = 1$, in particular by noting that $0 < w_i w_j < 1$ implies

$$\begin{aligned}\mathbb{E}\left(|a_{ij} - \mu_j|^3\right) &= (1 - w_i w_j)^3\, w_i w_j + (w_i w_j)^3\,(1 - w_i w_j)\\ &\le (1 - w_i w_j)\, w_i w_j + (w_i w_j)\,(1 - w_i w_j)\\ &= 2\, w_i w_j - 2\, w_i^2 w_j^2\,,\end{aligned}$$

and thus

$$\frac{1}{s_n^3}\sum_{j=1}^{n} \mathbb{E}\left(|a_{ij} - \mu_j|^3\right) \le \frac{\sum_{j=1}^{n} 2\, w_i w_j - 2\, w_i^2 w_j^2}{s_n^3} = \frac{2}{s_n}\,,$$

which tends to zero as $n \to \infty$ (again, as long as $\|\mathbf{w}\|_1$ grows without bound). Consequently, we can apply the Lyapunov CLT [8], which states that under the condition in (4),

$$\frac{1}{s_n}\sum_{j=1}^{n} (a_{ij} - \mu_j) = \frac{d_i - w_i \|\mathbf{w}\|_1}{s_n} = \frac{\hat{w}_i \|\mathbf{w}\|_1 - w_i \|\mathbf{w}\|_1}{s_n}$$

converges in distribution to a standard normal random variable. In other words, for sufficiently large n the distribution of $\hat{w}_i$ can be approximated arbitrarily well by the normal distribution

$$\mathcal{N}\left(w_i, \frac{w_i}{\|\mathbf{w}\|_1} - w_i^2 \frac{\|\mathbf{w}\|_2^2}{\|\mathbf{w}\|_1^2}\right). \tag{5}$$

3.3. Approximate error statistics for general estimator

We now return to the standard Chung-Lu parameter estimator

$$\hat{w}_i = g(\mathbf{d} \mid i) = \frac{d_i}{\left(\sum_{j=1}^{n} d_j\right)^{1/2}}\,, \tag{6}$$

now denoted as a function $g : \mathbb{R}^n \to \mathbb{R}$ of the observed degree vector $\mathbf{d} = [d_1 \ \cdots \ d_n]^T$ (conditioned on the node index i). Computing the distribution of this estimator is challenging, as g is nonlinear function of a vector whose elements are neither independent nor identically distributed. One potential approach would be to employ the **multivariate delta method** [9], which involves formulating a CLT through a Taylor series expansion of $g(\mathbf{d})$. Unfortunately, this approach requires us to begin with a multivariate CLT for $\mathbf{d}$, which is itself difficult to establish, as the sequence $\{d_1, \ldots, d_n\}$ is both dependent and non-stationary.

Nevertheless, we can still use the Taylor series to obtain approximations of the mean and variance of $g(\mathbf{d})$ under the Chung-Lu model. Furthermore, numerical examples suggest that $\mathbf{d}$ does exhibit CLT-like behavior, and thus normal approximations using these moments will tend to be quite good for large n. (Empirical results will be discussed further in Section 4.)

We begin by expanding g as a multivariate Taylor series, centered at the mean vector $\boldsymbol{\xi} = \mathbb{E}(\mathbf{d})$. We can approximate $g(\mathbf{d})$ by retaining only the first-order terms, yielding

$$g(\mathbf{d}) \approx g(\boldsymbol{\xi}) + \sum_{i=1}^{n} g_i'(\boldsymbol{\xi})\,(d_i - \xi_i)\,, \tag{7}$$

where g_i' denotes the i-th element of the gradient of g,

$$g_i'(\mathbf{d}) \equiv \frac{\partial}{\partial d_i} g(\mathbf{d})\,,$$

for $i = 1, \ldots, n$. Using (7), we can approximate the expected value of $g(\mathbf{d})$ as

$$\mathbb{E}(g(\mathbf{d})) \approx \mathbb{E}\left[g(\boldsymbol{\xi}) + \sum_{i=1}^{n} g_i'(\boldsymbol{\xi})\,(d_i - \xi_i)\right] = g(\boldsymbol{\xi})\,, \tag{8}$$

and the second moment of $g(\mathbf{d})$ as

$$\begin{aligned}\mathbb{E}\left(g^2(\mathbf{d})\right) &\approx \mathbb{E}\left[\left(g(\boldsymbol{\xi}) + \sum_{i=1}^{n} g_i'(\boldsymbol{\xi})\,(d_i - \xi_i)\right)^2\right]\\ &= g^2(\boldsymbol{\xi}) + \sum_{i=1}^{n}\sum_{j=1}^{n} g_i'(\boldsymbol{\xi})\ g_j'(\boldsymbol{\xi})\ \text{cov}(d_i, d_j)\,.\end{aligned}$$

Combining these two expressions, we can approximate the variance of $g(\mathbf{d})$ as

$$\text{var}(g(\mathbf{d})) \approx \sum_{i=1}^{n}\sum_{j=1}^{n} g_i'(\boldsymbol{\xi})\ g_j'(\boldsymbol{\xi})\ \text{cov}(d_i, d_j)\,. \tag{9}$$

To obtain approximate moments of the Chung-Lu estimator, we can apply the approximations in (8)–(9) using the definition of g given in (6). For the expected value of $\hat{w}_i$, this yields

$$\begin{aligned}\mathbb{E}(\hat{w}_i) = \mathbb{E}(g(\mathbf{d} \mid i)) &\approx g(\boldsymbol{\xi} \mid i)\\ &= \frac{\xi_i}{\left(\sum_{j=1}^{n} \xi_j\right)^{1/2}}\\ &= \frac{w_i \|\mathbf{w}\|_1}{\left(\sum_{j=1}^{n} w_j \|\mathbf{w}\|_1\right)^{1/2}}\\ &= w_i\,.\end{aligned}$$

where $\xi_i = \mathbb{E}(d_i) = w_i \|\mathbf{w}\|_1$ from (2). Thus, the standard Chung-Lu estimator is approximately unbiased for finite n.

To compute the approximate variance of the estimator, we first need to calculate partial derivatives g_j' for $j = 1, \ldots, n$. We have

$$g_j'(\mathbf{d} \mid i) = \begin{cases} \frac{1}{\left(\sum_{k=1}^{n} d_k\right)^{1/2}} - \frac{d_i}{2\left(\sum_{k=1}^{n} d_k\right)^{3/2}}, & i = j,\\ -\frac{d_i}{2\left(\sum_{k=1}^{n} d_k\right)^{3/2}}, & \text{otherwise,}\end{cases}$$

and thus

$$g_j'(\boldsymbol{\xi} \mid i) = \begin{cases} \frac{1}{\|\mathbf{w}\|_1} - \frac{w_i}{2\|\mathbf{w}\|_1^2}, & i = j,\\ -\frac{w_i}{2\|\mathbf{w}\|_1^2}, & \text{otherwise.}\end{cases}$$

Next, we substitute these partial derivatives and the degree covariances from Section 3.1 into the approximation

$$\text{var}(\hat{w}_i) \approx \sum_{j=1}^{n}\sum_{k=1}^{n} g_j'(\boldsymbol{\xi} \mid i)\, g_k'(\boldsymbol{\xi} \mid i)\, \text{cov}(d_j, d_k)\,.$$

When evaluating this expression, it will be helpful to consider four distinct cases: (a) $i = j = k$, (b) $i \neq j = k$, (c) $i = j$ or $i = k$ with

$j \neq k$, and (d) $i \neq j, k$ with $j \neq k$. Partitioning the summation according to these cases yields

$$\begin{aligned}\mathrm{var}\left(\hat{w}_i\right) \approx & \left(\tfrac{1}{\|\mathbf{w}\|_1} - \tfrac{w_i}{2\|\mathbf{w}\|_1^2}\right)^2 \left(w_i \|\mathbf{w}\|_1 - w_i^2 \|\mathbf{w}\|_2^2\right) \\ & + \sum_{j \neq i} \left(\tfrac{w_j}{2\|\mathbf{w}\|_1^2}\right)^2 \left(w_j \|\mathbf{w}\|_1 - w_j^2 \|\mathbf{w}\|_2^2\right) \\ & - 2 \sum_{j \neq i} \left(\tfrac{1}{\|\mathbf{w}\|_1} - \tfrac{w_k}{2\|\mathbf{w}\|_1^2}\right) \left(\tfrac{w_i}{2\|\mathbf{w}\|_1^2}\right) \left(w_i w_j - w_i^2 w_j^2\right) \\ & + \sum_{j \neq i} \sum_{k \neq i,j} \left(\tfrac{w_j w_k}{4\|\mathbf{w}\|_1^4}\right) \left(w_j w_k - w_j^2 w_k^2\right).\end{aligned}$$

Simplifying this expression, we obtain

$$\mathrm{var}\left(\hat{w}_i\right) \approx \frac{w_i}{\|\mathbf{w}\|_1} - \frac{w_i^2 \left(\|\mathbf{w}\|_2^2 + 1\right)}{\|\mathbf{w}\|_1^2} + f_1\left(\mathbf{w}\right) - f_2\left(\mathbf{w}\right), \quad (10)$$

where

$$f_1\left(\mathbf{w}\right) = \frac{\|\mathbf{w}\|_3^3 \left(1 + 4w_i^2\right) - 4\|\mathbf{w}\|_2^2 \left(w_i - w_i^3\right) + 4w_i^3 - 4w_i^5}{4\|\mathbf{w}\|_1^3},$$

$$f_2\left(\mathbf{w}\right) = \frac{\|\mathbf{w}\|_2^2 \|\mathbf{w}\|_4^4 - \|\mathbf{w}\|_2^4 + \|\mathbf{w}\|_3^6 + \|\mathbf{w}\|_4^4 - \|\mathbf{w}\|_6^6}{4\|\mathbf{w}\|_1^4}.$$

Since this expression tends to zero with increasing $\|\mathbf{w}\|_1$, we have that the MSE of $\hat{w}_i$ approximately tends to zero as $n \to \infty$.

A simpler approximation can be obtained by noting that as $\|\mathbf{w}\|_1$ grows large, $f_1\left(\mathbf{w}\right)$ and $f_2\left(\mathbf{w}\right)$ quickly become negligible. Ignoring these terms, we obtain

$$\mathrm{var}\left(\hat{w}_i\right) \approx \frac{w_i}{\|\mathbf{w}\|_1} - \frac{w_i^2 \left(\|\mathbf{w}\|_2^2 + 1\right)}{\|\mathbf{w}\|_1^2}, \quad (11)$$

which is quite close to the variance obtained for the simplified estimator given in (5), differing only by $1/\|\mathbf{w}\|_1^2$. As we will see in the next section, empirical results suggest that all three variance expressions—(5), (10), and (11)—can be used as effective approximations of the true estimator variance.

4. EMPIRICAL RESULTS

We conclude by investigating the previously stated moment approximations through an empirical example. Let $n = 1000$, and let $\mathbf{w}$ be a n-length vector whose elements are generated by independent draws from the uniform distribution over $(0, 0.2)$. Thus, we have $\mathbb{E}\left(w_i\right) = 0.1$ and $\mathbb{E}\left(\|\mathbf{w}\|_1\right) = 100$. Given $\mathbf{w}$ fixed, we generated 1000 graphs according to the Chung-Lu model $G\left(n, \mathbf{w}\right)$, and for each graph, we computed the set of Chung-Lu parameter estimates according to (1), given its observed degree sequence.

Let $\hat{w}_1$ denote the estimate of the first node weight, which for this example was $w_1 = 0.1629$. Computed over the 1000 sample estimates, the empirical mean of $\hat{w}_1$ was 0.1617, and the empirical variance was 0.001630. By comparison, the approximate variances as computed by (10) and (11) were 0.001627 and 0.001628. These values are close the variance of the simplified estimator given in (5), which is equal to 0.001631.

In addition to having mean and variance consistent with our approximations, the sample estimates appear to follow CLT-like behavior. Figure 1 shows the empirical distribution of $\hat{w}_1$, obtained by constructing a kernel density estimate using the estimates computed

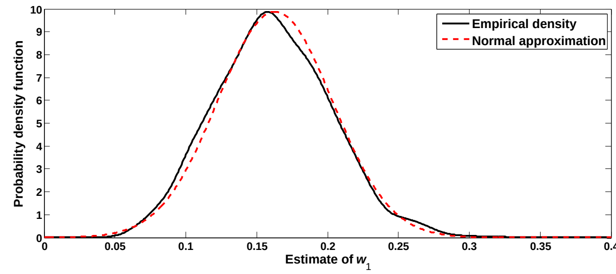

Fig. 1. Kernel density estimate and normal approximation for distribution of $\hat{w}_1$.

for each of the 1000 generated graphs and a Gaussian kernel with a bandwidth of $\sigma = 0.01$. Also plotted is a normal distribution with mean 0.1629 and variance 0.0016. From the figure, we see that the normal distribution appears to provide a good fit to the empirical distribution. This assertion is supported by noting that the skewness and the excess kurtosis of the samples are low (0.2328 and 0.0398, respectively), as is the Kullback-Leibler divergence between the empirical and normal distributions (with a value of 0.0076).

5. SUMMARY

In this paper, we explored statistical properties of a standard estimator used for determining the weight parameter of Chung-Lu random graph models. In addition to developing a central limit theory for a simplified version of the estimator, we derived approximations for moments of the general estimator using the delta method. We also illustrated through an empirical example that these approximations can be effective in practice.

6. REFERENCES

[1] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Rev. Mod. Phys.*, vol. 74, no. 1, pp. 47–97, 2002.

[2] M. E. J. Newman, "The structure and function of complex networks," *SIAM Rev.*, vol. 45, no. 2, pp. 167–256, 2003.

[3] F. Chung and L. Lu, "Connected components in random graphs with given expected degree sequences," *Ann. Comb.*, vol. 6, no. 2, pp. 125–145, 2002.

[4] F. Chung and L. Lu, *Complex graphs and networks*, American Mathematical Society, 2006.

[5] M. E. J. Newman, "Finding community structure in networks using the eigenvectors of matrices," *Phys. Rev. E*, vol. 74, pp. 1–19, 2006.

[6] Y.-T. Chang, D. Pantazis, H. B. Hui, and R. M. Leahy, "Statistically optimal graph partition method based on modularity," in *Proc. IEEE Int. Symp. on Biomed. Imaging*, 2010, pp. 1193–1196.

[7] B. A. Miller, N. T. Bliss, and P. J. Wolfe, "Toward signal processing theory for graphs and non-Euclidean data," in *Proc. ICASSP*, 2010, pp. 5414–5417.

[8] P. Billingsley, *Probability and measure*, John Wiley & Sons, 1995.

[9] G. Casella and R. L. Berger, *Statistical inference*, Duxbury, 2002.