# Using Moments and L-Moments to Characterize Graphical Networks

Fairul Mohd-Zaid* Christine Schubert Kabban†

**Abstract**
Networks that can be modeled graphically may also be summarized through nodal measures, e.g. degree. Nodal measures themselves may be summarized across the network; however, such summaries are often focused on single statistics. Although the mean is the most commonly used statistical summary of a nodal measure, the probability distribution of the nodal measure may be better described using sets of summary measures. The collection of these summary measures may then be used to more fully characterize the network. The purpose of this study was to examine the feasibility of characterizing a network using summary measures of the probability distribution for the nodal network measures. In a large simulation, nodal measures of the degree, betweenness, and closeness for Erdös-Rényi and Watts-Strogatz generated graphs of varying parameter and size were computed. Five summary measures based upon moments and four summary measures based upon L-moments were examined. Through clustering and predictive modeling, the results of the analysis demonstrate that uncorrelated moments and L-moments are dependent on the network type and that the mean alone is not sufficient to characterize the network.

**Key Words:** L-moments, Moments, Random Graphs, Erdos-Renyi, Watts-Strogatz

## 1. Introduction

Many fields of research produce relational data that can be represented in the form of graphs. Areas of research that falls into this category are numerous to include biology, chemistry, engineering, medicine, and sociology, just to name a few. An extensive list of research areas that have seen applications of graph matching were provided by Conte et al. (2004) and Livi and Rizzi (2013).

Most of the graph measures available in the literature, such as degree centrality, betweenness centrality, and closeness centrality, fall short of a full descriptive characterization of the probability distribution of the nodal components even though the graph measures themselves are built from those individual components. Based on a recently conducted literature review, there has not been extensive work examining the statistical characteristics of the distributions resulting from the collection of nodal measures besides merely degree distributions. The need for the probability distribution should be highly considered because, theoretically, it provides a more complete description of the graph measures, and if the graphs can be defined by the probability distribution of its measures, then graph comparisons and matching can be performed by comparing the probability distributions themselves. Differences in distributions might provide a more nuanced explanation about the differences between the networks, or how a network may be changing.

Although some researchers have looked at the probabilistic aspect of a graph by measuring its entropy (Moonesinghe et al. 2007; Mowshowitz and Dehmer 2012), these measures do not capture the uniqueness of a given graph since entropy only measures how diverse and random a particular probability distribution is, thus any comparison of two graphs using entropy measures only compares the randomness of the graphs' structure and does not necessarily imply that they are identical since the measure itself is not unique. One way of summarizing a probability distribution is through its set of moments which

*Air Force Research Lab, 2255 H Street, Bldg 248, Wright-Patterson AFB, OH 45433-7022
†Air Force Institute of Technology, 2950 Hobson Way, Bldg 641, Wright Patterson AFB, OH 45433-7765

describe its central location, scale, symmetry, peakedness, as well as other higher order characteristics.

As will be discussed later, the uniqueness of a probability distribution through its moments and L-moments can be guaranteed under certain conditions. Therefore it is plausible that these characteristics, if applied to network analysis, may provide an approach to detect graph similarity or dissimilarity. However, with the exception of testing the difference on the mean and variance and testing for normality, no formal test of hypothesis have been developed for the other upper moments or L-moments (Pearson et al. 1977; Jarque and Bera 1980; D'Agostino and Belanger 1990; Seier 2002; Harri and Coble 2011; Galvao et al. 2013).

In this paper, the feasibility of using upper moments and L-moments of network measure distributions will be explored. In order to examine this feasibility, a data simulation was performed in order to obtain a sample of moments and L-moments from distributions of network measures for different networks. After a brief description of this simulation, analysis on the feasibility of the use of the moments and L-moments is performed on the collected data. The results from this analysis will help shape the objectives for research on how to structure statistical tests to determine network difference, where a more focused data simulation and a larger array of collected measures from the networks of interest may be required.

## 2. Background

### 2.1 Graph Model

The first commonly used random graph generating algorithm was proposed by Erdös and Rényi (1959) in which the algorithm constructs a graph by connecting any pair of nodes by an edge with probability $p$, and in which each edge is independent from every other edge. This results in a graph of $n$ nodes and $m$ edges having an equal probability of $p^m(1-p)^{\binom{n}{2}-m}$ from all possible undirected simple graphs of $n$ nodes and $m$ edges. One downside to the Erdös-Rényi algorithm is that it is not *scale-free* as shown by Barabási and Albert (1999), a property that many real world networks such as the World Wide Web possess (Albert et al. 1999). A scale-free network is defined as one that has a power law degree distribution between nodes. However, given its history, the Erdös-Rényi algorithm is widely used in the literature as a baseline when making comparisons for network metrics and classifications. Since each node has an equal probability of being connected to other nodes and since the nodes are independent, the Erdös-Rényi random graph has nodal degrees that follow the Binomial distribution (Erdös and Rényi 1959; Wasserman and Faust 1994). In this formation, the probability that the degree $k$ is equal to $c$ for a given node is

$$P(k=c) = \binom{N-1}{c} p^c (1-p)^{N-1-c} \tag{1}$$

and that the distribution converges to the Poisson distribution when N is very large

$$P(k) \sim e^{-K} \frac{K^k}{k!} \tag{2}$$

where $K$ is the mean degree (Albert and Barabási 2002). The random graph also has a group clustering coefficient of $C = \frac{K}{N}$.

A random graph generator model that produces *small-world* properties was introduced by Watts and Strogatz (1998). Small-world networks are networks where the shortest path, $L$, between most pair of nodes in the networks are fairly small and grows proportionately to the logarithm of the network size, $N$, such that $L \propto \log N$. The algorithm for this model functions by first starting with a ring lattice of size $N$. This is then followed by rewiring each edge in the lattice with probability $p$ such that duplicates and self-loops are excluded. Many real world networks such as the neural network of the worm Caenorhabditis elegans, the power grid of the western United States, and the collaboration network of film actors are shown to possess the small-world characteristic of having small average shortest path (Watts and Strogatz 1998). However, the Watts-Strogatz algorithm does not produce a graph that has a scale-free power-law distribution, hence, neither the Barabási-Albert nor the Watts-Strogatz algorithm is fully capable of modeling all real world networks. The degree distribution for the Watts-Strogatz graph was shown by Barrat and Weigt (2000) to be

$$P(k) = \sum_{n=0}^{\min(k-\frac{K}{2},\frac{K}{2})} \binom{\frac{K}{2}}{n} (1-p)^n p^{\frac{K}{2}-n} \frac{(\frac{pK}{2})^{k-\frac{K}{2}-n}}{(k-\frac{K}{2}-n)!} e^{-\frac{pK}{2}} \tag{3}$$

where $k$ is the degree and $K$ is the mean degree such that $k \geq \frac{K}{2}$. The group clustering coefficient for the ring lattice is $C(0) = \frac{3(K-2)}{4(K-1)}$ while the group clustering coefficient for a Watts-Strogatz graph can be approximated with a function of $p$ such that

$$\begin{aligned} C'(p) &\simeq C(0)(1-p)^3 \\ &\simeq \frac{3(K-2)}{4(K-1)}(1-p)^3 \end{aligned} \tag{4}$$

where $p$ is the probability of rewire and $K$ is the mean degree of the graph (Barrat and Weigt 2000).

### 2.2 Graph Measures

Nodal $degree$, $d(n_i)$, is the number of direct links possessed by a node to other nodes and has a range of $[0, N-1]$ for any graph and $[1, N-1]$ for connected graphs. *Betweenness* measures the interactions between two nodes that might depend on other nodes that lie on the path between the two (nodes in the middle). Node betweenness index is the sum of the proportion of the shortest path that goes through node $i$ between all node pairs $j < k$, where $j \neq i, k \neq i$. Node betweenness can be calculated as

$$C_B(n_i) = \sum_{j<k} \frac{g_{jk}(n_i)}{g_{jk}} \tag{5}$$

where $g_{jk}(n_i)$ is the number of shortest paths that contains node $i$ and $g_{jk}$ is the total number of shortest paths between $j$ and $k$. Since it is a measure of proportion, it has the range of $[0, 1]$. *Closeness* is a measure that is related to centrality measures, and it is defined as the inverse of the sum of pairwise distances between the nodes given by

$$C_C(n_i) = \frac{1}{\left(\sum_{j=1}^{N} d(n_i, n_j)\right)} \tag{6}$$

and contains the range of $(0, \frac{1}{N-1}]$.

### 2.3 Moments and L-Moments

In probability and statistics, a moment is a quantitative measure that describes a characteristic of a probability distribution. An extensive set of moments may give a more descriptive summary of a random variable over the traditional approach of reporting the mean and variance since the latter only describes the center mass and scale of the distribution, respectively. It should be noted that while the existence of a moment generating function (MGF) for a random variable implies that there exists an infinite set of moments, the converse is not true. The characterization of a set of moments is not enough to uniquely define a random variable because there may exist another random variable having the same set of moments. However, uniqueness of moments is guaranteed if the random variables have bounded support or if the MGF exists in the neighborhood of zero (Casella and Berger 2002). The $k^{th}$ moment of a continuous probability density function, $f(x)$, is defined as

$$E\left[X^k\right] = \int_{-\infty}^{\infty} x^k f(x)\,dx \tag{7}$$

and for a discrete probability mass function, $P(X = x)$, as

$$E\left[X^k\right] = \sum_{x \in X} x^k P(X = x). \tag{8}$$

The $k^{th}$ central moment is defined as

$$E\left[(X - E\left[X\right])^k\right]. \tag{9}$$

It is often useful to scale the upper moments ($3^{rd}, 4^{th}$, etc.) by a function of the variance so that comparison can be made between different distributions regardless of the variance. A $k^{th}$ standardized moment is defined as

$$\gamma_k = \frac{E\left[(X - E\left[X\right])^k\right]}{\sigma^k}. \tag{10}$$

The $3^{rd}$ and $4^{th}$ standardized moments measure the symmetry and peakedness of the probability density function, respectively. However, not all moments exist for every distribution and no moments exist for some distributions, notably the Cauchy distribution. One set of metrics that solves the issue of nonexistent moments for some distributions is the L-moments. L-moments have a theoretical advantage of being able to characterize a wider range of distributions since the set of L-moments for a random variable exists if and only if the random variable has a finite mean (Hosking 1990). However, this does not solve the problem for distributions where the mean does not exist, instead, other techniques can be used for those distributions such as the trimmed L-moments as defined by Elamir and Seheult (2003).

First proposed by Hosking (1990), L-moments are linear combinations of order statistics that describe the location and shape of the probability distribution analogous to classical moments. The $r^{th}$ L-moment is defined as

$$\lambda_r = \frac{1}{r}\sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} E\left[X_{r-k:r}\right] \tag{11}$$

where $X_{j:n}$ denotes the jth order statistic ($j^{th}$ smallest sample value) in an independent sample of size n. Note that $\lambda_1 = E[X_{1:1}] = E[X] = \mu$. The $r^{th}$ L-moment ratio is defined as

$$\tau_r = \frac{\lambda_r}{\lambda_2}; r = 3, 4, ... \tag{12}$$

and is akin to the standardized conventional moment as defined in equation (10) but has a bound of $(-1, 1)$. The $1^{st}$ and $2^{nd}$ L-moments are referred to as L-mean and L-scale, respectively, whereas the $3^{rd}$ and $4^{th}$ L-moment ratios are referred to as L-skewness and L-kurtosis, respectively. Even though there has not been a term coined for the fifth L-moment ratio, it can be interpreted as a measure of tendency to bimodality and will be referred to as L-hyperkurtosis for the purpose of this research.

Hosking (1990) states that a set of L-moments is unique to a particular distribution as long as the mean of the distribution exists and that the first two L-moments, $\lambda_1$ and $\lambda_2$, as well as the third and fourth L-moment ratios, $\tau_3$ and $\tau_4$, are enough to summarize the main features of a probability distribution. Additionally, the set of L-moments is considered more robust to outliers than conventional moments (Hosking 1990). For example, a distribution with one very outlying point will cause the variance to increase quite notably but does not affect the L-scale to the same extent.

Direct estimators of the first four L-moments were derived by Wang (1996) that circumvent the need for using Probability Weighted Moments (PWMs). These estimators are defined, respectively, as

$$\widehat{\lambda}_1 = \binom{n}{1}^{-1} \sum_{i=1}^{n} x_{(i)} \tag{13}$$

$$\widehat{\lambda}_2 = \frac{1}{2}\binom{n}{2}^{-1} \sum_{i=1}^{n} \left(\binom{i-1}{1} - \binom{n-i}{1}\right) x_{(i)} \tag{14}$$

$$\widehat{\lambda}_3 = \frac{1}{3}\binom{n}{3}^{-1} \sum_{i=1}^{n} \left(\binom{i-1}{2} - 2\binom{i-1}{1}\binom{n-i}{1} + \binom{n-i}{2}\right) x_{(i)} \tag{15}$$

$$\widehat{\lambda}_2 = \frac{1}{4}\binom{n}{4}^{-1} \sum_{i=1}^{n} \left(\binom{i-1}{3} - 3\binom{i-1}{2}\binom{n-i}{1} + 3\binom{i-1}{1}\binom{n-i}{2} - \binom{n-i}{2}\right) x_{(i)}. \tag{16}$$

Additionally, the second L-moment is strictly positive, and the fourth L-moment ratio, L-kurtosis, is shown to have a tighter bound of $\frac{1}{4}(\tau_3^2 - 1) \leq \tau_4 < 1$ (Hosking 1990).

## 3. Analyses

### 3.1 Network Simulations

A dataset comprised of moments and L-moments computed from distributions of network measures of simulated Erdös-Rényi and Watts-Strogatz graphs were generated. The dataset was created to study the feasibility of using moments and L-moments from a distribution of network measures. The simulation was conducted in R using the *igraph* package (Csardi and Nepusz 2006) for network generation and computing the network measures and the *lmom* package (Hosking 2014) for computing the L-moments. The Erdös-Rényi algorithm was selected as a baseline model whereas the Watts-Strogatz algorithm was selected since it closely models many real world networks (Watts and Strogatz 1998).

The Erdös-Rényi algorithm takes in as input the number of nodes, $n$, and the probability, $p$, of connecting any two nodes. The Watts-Strogatz algorithm, on the other hand, takes in the number of nodes, $n$, the dimension of the lattice, $d$, the number of connected nearest-neighbors, $m$, and the probability, $p$, that an additional edge is added to a particular node (i.e. edge rewiring probability). The parameter selection for the simulation is listed in Table 1 where $1000$ independent networks were generated for each combination of parameters.

For each simulated network, nodal measures as well as the moments and L-moments from the distribution of the nodal measures as listed in Table 2 were computed. This produced, for each parameter combination, a sample of $1000$ moments and L-moments for each network measures. Statistical analyses were then conducted using JMP® 10.0.1 on the data as will be shown in the rest of this section.

**Table 1**: Dataset parameters for feasibility study

| | **Erdös-Rényi** | **Watts-Strogatz** |
|---|---|---|
| **Parameters** | $p = \{0.01, 0.03375, 0.0575, 0.08125, 0.105, 0.12875, 0.1525, 0.17625, 0.2\}$<br><br>$p$ is the probability of a new edge | $p = \{0.01, 0.03375, 0.0575, 0.08125, 0.105, 0.12875, 0.1525, 0.17625, 0.2\}$<br>$p$ is the probability of a rewire<br>$d = 1$; the dimension of the starting lattice<br>$m = 4$; the number of closest neighbors for the starting lattice |
| **Size** | $n = 2^k; k = 9, 10, 11$ | $n = 2^k; k = 9, 10, 11, 12$ |

**Table 2**: Computed measures for pilot dataset

| **Network Measures** | **Moments** |
|---|---|
| Degree (Deg), Betweenness (Btw), Closeness (Cls) | $\mu, \sigma^2, \gamma_3, \gamma_4, \gamma_5, \lambda_2, \tau_3, \tau_4, \tau_5$ (mean, var, skew, kurto, hkurto, lscale, lskew, lkurto, lhkurto) |

### 3.2 Empirical Analysis

The purpose of the analyses is to scope the capability of characterizing networks using moments and L-moments from the distribution of network measures. The analyses are not all inclusive in answering the main objective but are merely a collection of various preliminary studies that highlights the feasibility of the research hypothesis.

An initial look at the means of Degree, Betweenness, and Closeness for the Erdös-Rényi as shown in Figure 1 suggests that the graph model is very discriminant based on the first moment (i.e. mean) alone. On the other hand, the same could not be said for the Watts-Strogatz model where the means of the measures do not discriminate the different network very well. However, the distributions of $\lambda_2$ of Degree, $\tau_4$ of Degree, and $\tau_3$ of Betweenness for the Watts-Strogatz graph as shown in Figure 2 suggest that the upper moments and L-moments could be used to discern the difference between the network. Due to this finding, along with the knowledge that the Erdös-Rényi network does not possess real world network characteristics (Barabási and Albert 1999), further analyses of the upper moments and L-moments will not be reported for the Erdös-Rényi graph.

The Watts-Strogatz dataset was then examined to find a possible set of uncorrelated and statistically significant predictors of the network parameter. This is performed by first using Discriminant Analysis to visualize any separation of the network population on a two dimensional plane using the Discriminant scores. Factor Analysis using using Varimax rotation was then applied to find independent measures that account for a substantial amount of the total variance within the collection of moments and L-moments. Using the set of independent measures, logistic regression was then performed on each (size/parameter) combination to find the set of moments and L-moments that were statistically significant in modeling the particular network group.

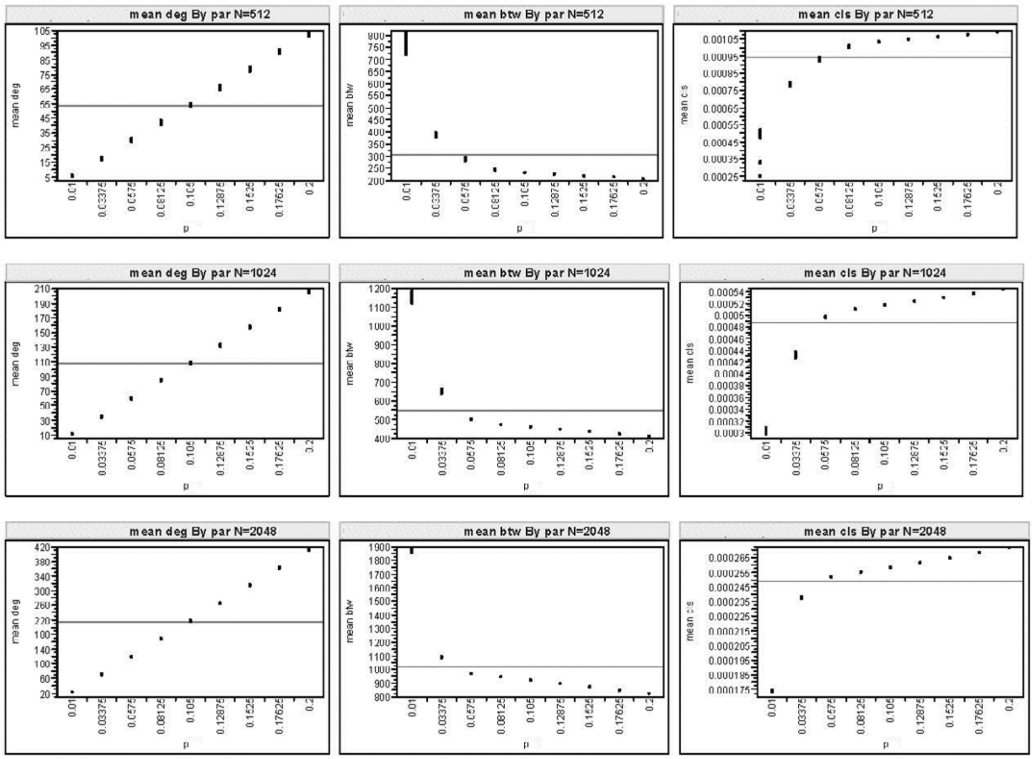



**Figure 1**: Scatterplot of mean Degree, mean Betweenness, and mean Closeness for Erdös-Rényi

### *3.2.1 Discriminant Analysis*

Discriminant Analysis was performed on all moments and L-moments obtained from Degree, Betweenness, and Closeness that are shown to be approximately normally distributed for each network size to see if groupings are present with respect to the parameters. This was performed using the linear method where the covariance for each group was assumed to be the same. Since normality assumptions for the moments and L-moments are required in order to utilize this method, they were first inspected for univariate normality using the Quantile-Quantile plot for each (size/parameter) combination, and it was shown that only $\lambda_{2,Deg}, \tau_{4,Deg}, \tau_{3,Btw}$, and $\tau_{4,Btw}$ are shown to be approximately normal for all of the (size/parameter) combinations. Therefore, only these moments are included for this analysis.

The two dimensional discriminant scores were then plotted as shown in Figure 3 to illustrate the groupings of the scores with respect to the parameters. It is apparent that as the size of the network grows larger, it is easier to discriminate the parameters, where the misclassification rates are $16.67\%, 7.744\%, 2.589\%$, and $0422\%$ for network sizes 512, 1024, 2048, and 4096, respectively. However, discrimination was still easily achieved for smaller networks with parameter $p \leq 0.0575$ where the misclassification is very low although it was harder to completely separate the groupings for the largest network when the parameter $p \geq 0.1525$. As randomness increases (i.e. $p$ increases), the Watts-Strogatz will converge to the Erdös-Rényi graph (Watts and Strogatz 1998). Thus, the property exhibited might be that of the Erdös-Rényi. From these results there appears to be some merit in using the upper L-moments from the distribution of Degree measures in characterizing the network population.

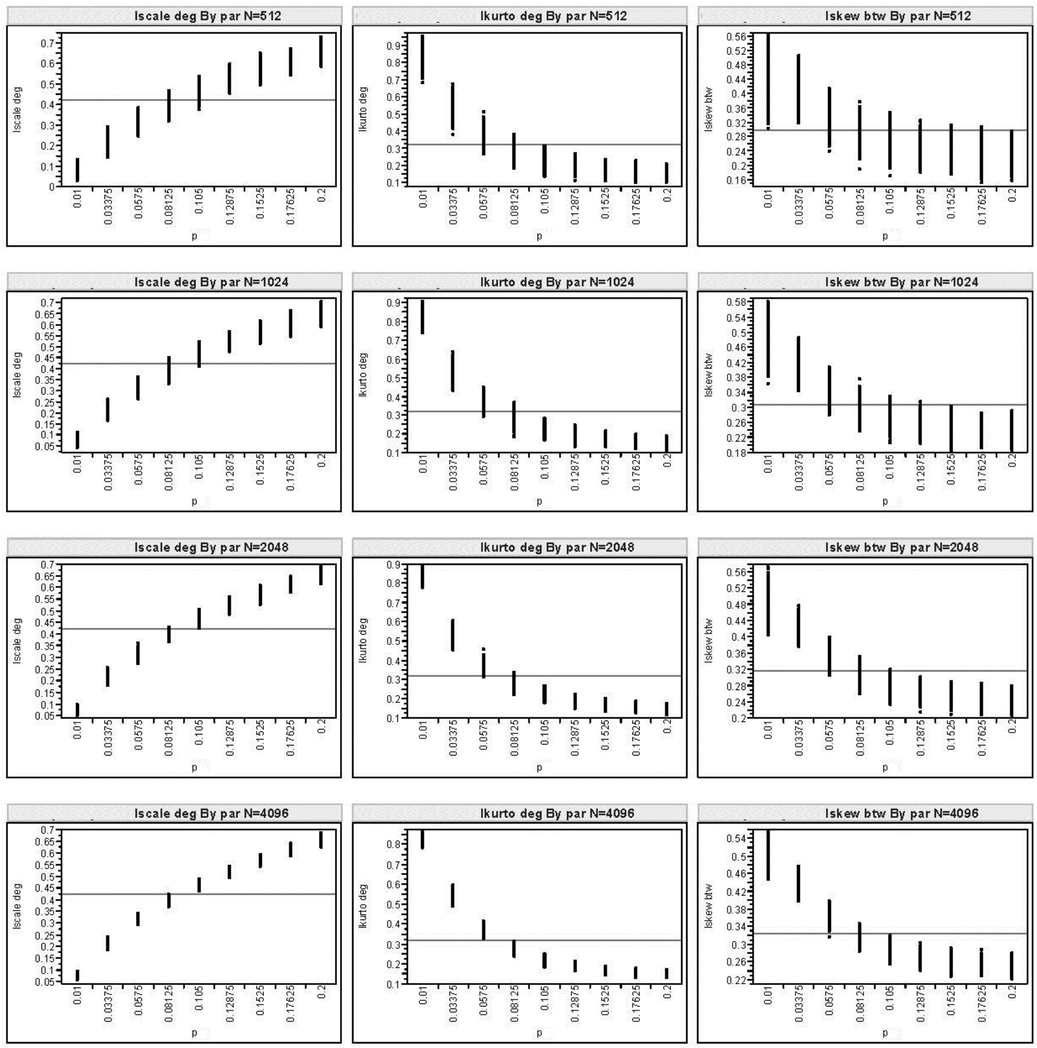
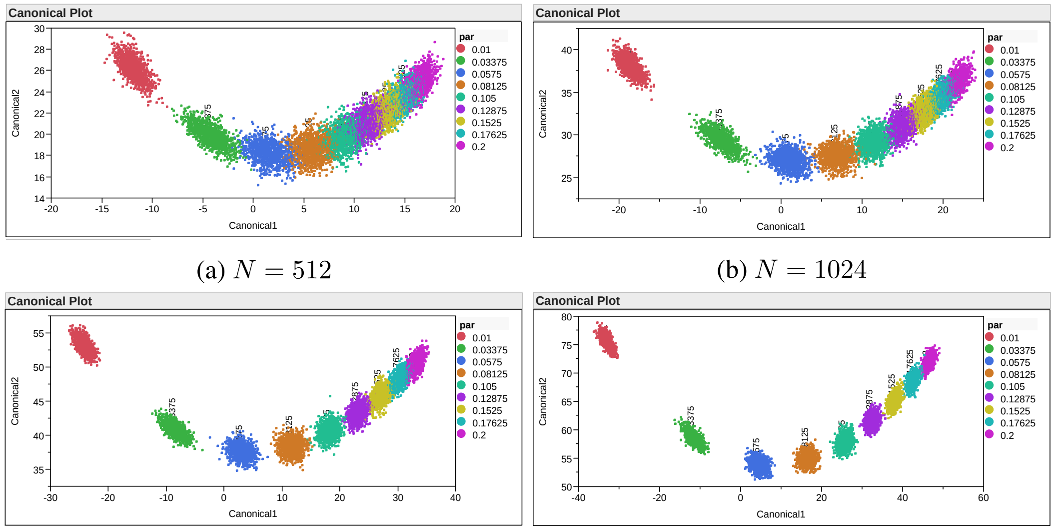



**Figure 2**: Scatterplot of $\lambda_2$ (L-scale) Degree, $\tau_4$ (L-kurtosis) Degree, and $\tau_3$ (L-skew) Betweenness for Watts-Strogatz

### *3.2.2 Factor Analysis*

To obtain the set of independent moments and L-moments for each network population with a specific (size/parameter) combination, Factor Analysis (FA) was conducted on the collected moments and L-moments from the distribution of Degree, Betweenness Centrality, and Closeness Centrality. Principal Components Analysis showed that the first six components captured around $79\% - 89\%$ of the total variance of the collection of moments and L-moments. A Varimax rotation of the measure-space on the first six components was applied to obtain the corresponding Factors. However, each Factor accounted for only $3\% - 28\%$ of the total variance indicating that there was no single Factor that captured a large percentage of the total variance.

Moments or L-moments that correlate the most with each of the first six Factors were selected to be included in the set of independent measures as shown in Table 3. There is no single moment or L-moment that was shown to be highly correlated with the first six Factors for all networks. These correlations ranged between $0.707 - 1$, with lower correlations are more commonly seen in the higher Factors. Additionally, these results showed that $\mu$ does not always account for the largest variance component and that the upper moments and



(a) $N = 512$ (b) $N = 1024$

(c) $N = 2048$ (d) $N = 4096$

**Figure 3**: Canonical plot of $\lambda_{2,Deg}, \tau_{4,Deg}, \tau_{3,Btw}$, and $\tau_{4,Btw}$ for Watts-Strogatz

L-moments may provide additional information.

**Table 3**: Moments with high correlation to the first six rotated factors

| | **Most Highly Correlated Moments** | | | | | |
|---|---|---|---|---|---|---|
| **Size/Parameter** | $1^{st}$ **Factor** | $2^{nd}$ **Factor** | $3^{rd}$ **Factor** | $4^{th}$ **Factor** | $5^{th}$ **Factor** | $6^{th}$ **Factor** |
| 512/0.01 | $\mu_{Cls}$ -0.955 | $\gamma_{3,Btw}$ 0.926 | $\gamma_{3,Deg}$ 0.991 | $\gamma_{3,Cls}$ 0.902 | $\gamma_{4,Cls}$ 0.854 | $\mu_{Deg}$ 1 |
| 512/0.105 | $\gamma_{3,Btw}$ 0.954 | $\gamma_{3,Cls}$ 0.948 | $\mu_{Btw}$ -0.926 | $\tau_{5,Deg}$ 0.968 | $\sigma^2_{Cls}$ 0.925 | $\gamma_{4,Deg}$ 0.810 |
| 512/0.2 | $\gamma_{4,Btw}$ 0.929 | $\gamma_{4,Cls}$ -0.973 | $\gamma_{3,Deg}$ 0.922 | $\lambda_{2,Btw}$ 0.881 | $\mu_{Btw}$ -0.896 | $\tau_{3,Cls}$ 0.733 |
| 1024/0.01 | $\lambda_{2,Deg}$ -0.957 | $\gamma_{3,Btw}$ 0.948 | $\gamma_{3,Deg}$ 0.994 | $\gamma_{5,Cls}$ 0.864 | $\lambda_{2,Cls}$ 0.830 | $\mu_{Deg}$ 1 |
| 1024/0.105 | $\gamma_{3,Btw}$ 0.955 | $\gamma_{3,Cls}$ 0.967 | $\mu_{Btw}$ -0.918 | $\tau_{5,Deg}$ 0.964 | $\lambda_{2,Cls}$ 0.910 | $\gamma_{4,Deg}$ 0.580 |
| 1024/0.2 | $\gamma_{3,Cls}$ 0.984 | $\gamma_{4,Btw}$ 0.927 | $\gamma_{3,Deg}$ 0.916 | $\lambda_{2,Btw}$ 0.908 | $\lambda_{2,Deg}$ 0.952 | $\mu_{Deg}$ 0.557 |
| 2048/0.01 | $\tau_{4,Deg}$ 0.969 | $\gamma_{3,Btw}$ 0.953 | $\gamma_{5,Cls}$ 0.888 | $\gamma_{3,Deg}$ 0.997 | $\lambda_{2,Cls}$ 0.867 | $\mu_{Deg}$ 1 |
| 2048/0.105 | $\gamma_{4,Btw}$ 0.939 | $\lambda_{2,Deg}$ 0.928 | $\gamma_{3,Deg}$ 0.966 | $\sigma^2_{Cls}$ 0.841 | $\gamma_{4,Cls}$ -0.954 | $\tau_{3,Cls}$ 0.832 |
| 2048/0.2 | $\gamma_{3,Cls}$ 0.988 | $\gamma_{4,Btw}$ 0.933 | $\gamma_{3,Deg}$ 0.918 | $\sigma^2_{Btw}$ 0.846 | $\lambda_{2,Deg}$ 0.967 | $\mu_{Deg}$ 0.707 |
| 4096/0.01 | $\lambda_{2,Deg}$ -0.974 | $\gamma_{4,Cls}$ -0.914 | $\tau_{5,Deg}$ 0.979 | $\tau_{3,Btw}$ 0.884 | $\gamma_{5,Btw}$ 0.956 | $\lambda_{2,Cls}$ 0.906 |
| 4096/0.105 | $\gamma_{5,Cls}$ 0.932 | $\gamma_{3,Deg}$ 0.959 | $\sigma^2_{Btw}$ 0.831 | $\gamma_{4,Btw}$ 0.925 | $\lambda_{2,Deg}$ 0.943 | $\tau_{3,Cls}$ 0.823 |
| 4096/0.2 | $\gamma_{3,Cls}$ 0.991 | $\gamma_{4,Btw}$ 0.949 | $\gamma_{3,Deg}$ 0.942 | $\lambda_{2,Btw}$ 0.899 | $\lambda_{2,Deg}$ 0.961 | $\tau_{4,Deg}$ 0.777 |
| % Variance | $17.2 - 28.2$ | $14.6 - 22.5$ | $13.4 - 14.7$ | $10.8 - 14.2$ | $7.4 - 13.1$ | $3.7 - 9.9$ |

NOTE: Deg=Degree; Btw=Betweenness; Cls=Closeness

### 3.2.3 *Logistic Regression*

Using the set of moments and L-moments obtained through FA, logistic regressions were performed on each (size/parameter) combination with its respective independent moments and L-moments. The purpose of this analysis was to find the set of significant and independent moments and L-moments with respect to modeling each network. The logistic model was structured with a two class outcome: the network (size/parameter) combination of interest versus all other combinations. It should be noted that only a subset of the networks were considered for the regressions, specifically $parameter \in \{0.01, 0.105, 0.2\}$, so that the number of (size/parameter) combinations is relatively small and tractable. Significant moments and L-moments based on the Likelihood Ratio $\chi^2$ of the Likelihood Ratio test at $\alpha = 0.05$ is shown in Table 4. Each model was shown to be significant based on the $-2LogLikelihood$ at $\alpha = 0.05$. The Deviance test was unable to reject the null hypothesis that more variables should be added to the model to give a better fit.

These result showed that significant moments and L-moments vary for each network

**Table 4**: Moments with significant effects based on Likelihood Ratio Test through Logistic Regression

| Size/Parameter | Significant Moments | AUC |
|---|---|---|
| 512/0.01 | $\mu_{Cls}$, $\gamma_{3,Btw}$, $\gamma_{3,Deg}$, $\gamma_{4,Cls}$ | 0.935 |
| 512/0.105 | $\gamma_{3,Cls}$, $\mu_{Btw}$, $\gamma_{4,Deg}$ | 0.910 |
| 512/0.2 | $\gamma_{4,Btw}$, $\gamma_{4,Cls}$, $\gamma_{3,Deg}$, $\tau_{3,Cls}$ | 0.833 |
| 1024/0.01 | $\lambda_{2,Deg}$, $\gamma_{3,Btw}$, $\gamma_{3,Deg}$, $\gamma_{5,Cls}$, $\lambda_{2,Cls}$ | 0.967 |
| 1024/0.105 | $\gamma_{3,Btw}$, $\gamma_{3,Cls}$, $\mu_{Btw}$, $\tau_{5,Deg}$, $\lambda_{2,Cls}$ | 0.920 |
| 1024/0.2 | $\gamma_{3,Cls}$, $\lambda_{2,Btw}$, $\lambda_{2,Deg}$ | 0.971 |
| 2048/0.01 | $\tau_{4,Deg}$, $\gamma_{3,Btw}$, $\gamma_{5,Cls}$, $\lambda_{2,Cls}$ | 0.977 |
| 2048/0.105 | $\gamma_{4,Btw}$, $\lambda_{2,Deg}$, $\gamma_{4,Cls}$, $\tau_{3,Cls}$ | 0.552 |
| 2048/0.2 | $\gamma_{3,Cls}$, $\gamma_{4,Btw}$, $\gamma_{3,Deg}$, $\sigma^2_{Btw}$, $\lambda_{2,Deg}$ | 0.961 |
| 4096/0.01 | $\lambda_{2,Deg}$, $\gamma_{4,Cls}$, $\tau_{5,Deg}$, $\tau_{3,Btw}$, $\gamma_{5,Btw}$ | 0.984 |
| 4096/0.105 | $\gamma_{5,Cls}$, $\sigma^2_{Btw}$, $\lambda_{2,Deg}$ | 0.530 |
| 4096/0.2 | $\gamma_{3,Cls}$, $\gamma_{4,Btw}$, $\gamma_{3,Deg}$, $\lambda_{2,Deg}$, $\tau_{4,Deg}$ | 0.964 |

NOTE: Deg=Degree; Btw=Betweenness; Cls=Closeness

and not all independent moments and L-moments are significant for a given network. Although the means are significant in some cases, other upper moments and L-moments provide additional information and also account for greater variability within the network. At least one statistic of Betweenness is included in the set of significant moments for each of the network while no moment or L-moment of Degree is significant for $(2048/0.105)$. Additionally, none of the networks require all six independent moments and L-moments. The Area Under the Curve (AUC) scores in Table 4 suggest that most of the models have very high predictive ability with AUC of $0.9$ or greater. However, there are the exceptions of graphs (2048/0.105) and (4096/0.105) where the AUCs are only slightly above $0.5$.

Although $\gamma_3$ and $\lambda_2$ are shown to be significant more often than the other moments and L-moments as shown in Table 5, results from contingency table analyses by treating $p$ as a category, size as a category, and moments and L-moments as one combined category showed that there are no relationships between the instances of significant moments and L-moments with respect to $p$ or size. And even though there were only two instances of significant $\sigma^2$, the L-moment equivalent, $\lambda_2$, was shown to be significant almost 6 times as often which suggests that the L-moment might be able to capture some information that the conventional moments could not. On the other hand, the number of significant moments and L-moments are distributed evenly between Degree, Betweenness, and Closeness indicating that there was no single graph measure that is more useful than the other two selected measures.

## 4. Result Summary

Based on moments and L-moments from select network measures used on simulated Erdös-Rényi and Watts-Strogatz networks, it was concluded that the Erdös-Rényi network only provided trivial results since the mean of the measures alone are sufficient in characterizing the different network parameters. The Watts-Strogatz network on the other hand, produced some useful patterns that suggest the feasibility of using upper moments and L-moments as a more descriptive measure of the networks.

Discriminant Analysis was used to show that the networks can be separable based on

**Table 5**: Significant moments and L-moments by network measure

| | **Degree** | **Betweenness** | **Closeness** | **Total** |
|---|---|---|---|---|
| $\mu$ | 0 | 2 | 1 | 3 |
| $\sigma^2$ | 0 | 2 | 0 | 2 |
| $\gamma_3$ | 5 | 4 | 5 | 14 |
| $\gamma_4$ | 1 | 4 | 4 | 9 |
| $\gamma_5$ | 1 | 1 | 3 | 5 |
| $\lambda_2$ | 7 | 1 | 3 | 11 |
| $\tau_3$ | 0 | 1 | 2 | 3 |
| $\tau_4$ | 2 | 0 | 0 | 2 |
| $\tau_5$ | 1 | 0 | 0 | 1 |
| **Total** | 17 | 15 | 18 | 50 |

$\mu, \sigma^2, \gamma_3, \gamma_4, \gamma_5, \lambda_2, \tau_3, \tau_4$, and $\tau_5$ of the Degree, Betweenness, and Closeness of the nodes, and separability also becomes prominent as the size of the network increases. Factor Analysis suggests that the set of moments and L-moments that capture the majority of the variance is dependent on the size and parameter of the network. Almost $90\%$ of the variability within the moments and L-moments are captured by the first six factors alone, and the mean does not necessarily correlate with the highest loading factor.

Lastly, Logistic Regression showed that certain moments and L-moments were more prominent than others. Although there was emphasis on the upper classical moments, $\lambda_2$ was shown to be more informative than its classical counterpart, $\sigma^2$, which indicates the usefulness of L-moments. Regardless, these results show that the upper moments are more prominent with respect to the significant effects than the mean and variance alone. These findings support the hypothesis that the networks can be characterized using the upper moments and L-moments of the network measures and will help guide the formation of future research objectives.

## 5. Discussion

Since only two network models were considered for this research, the feasibility of using moments and L-moments of network measures should be studied for other network models such as the Barabási-Albert model (Barabási and Albert 1999) which models the *scale-free* phenomena inherent in some real world networks. Nevertheless, the findings from this study suggest that, for some networks, using the average of well known graph measures may not be enough in characterizing the network.

Considering that most graphical measures comprise nodal measures, then one can create distributions of measures and compute various statistics that may characterize the particular network. Consequently, if the deriviation of the theoretical distribution of a network measure for a given network model is feasible, then it follows that the distribution of the moments and L-moments can also be derived if they exist. From there, a test of hypothesis can be formulated for a collection of moments to compare whether two probability distributions are the same which is a novel idea that has yet to be fully matured. This would have implications for examining similarity and dissimilarity in identifiable network characteristics.

## References

Albert, R., and Barabási, A.-L. (2002), "Statistical mechanics of complex networks," *Rev. Mod. Phys.*, 74, 47–97.

Albert, R., Jeong, H., and Barabási, A.-L. (1999), "Internet: Diameter of the world-wide web," *Nature*, 401(6749), 130–131.

Barabási, A.-L., and Albert, R. (1999), "Emergence of Scaling in Random Networks," *Science*, 286(5439), 509–512.

Barrat, A., and Weigt, M. (2000), "On the properties of small-world network models," *The European Physical Journal B - Condensed Matter and Complex Systems*, 13(3), 547–560.

Casella, G., and Berger, R. L. (2002), *Statistical Inference* Thomson Learning.

Conte, D., Foggia, P., Sansone, C., and Vento, M. (2004), "Thirty Years of Graph Matching in Pattern Recognition," *International Journal of Pattern Recognition and Artificial Intelligence*, 18(03), 265–298.

Csardi, G., and Nepusz, T. (2006), "The igraph software package for complex network research," *InterJournal*, Complex Systems, 1695.
**URL:** *http://igraph.org*

D'Agostino, R. B., and Belanger, A. (1990), "A Suggestion for Using Powerful and Informative Tests of Normality," *The American Statistician*, 44(4), 316–321.
**URL:** *http://www.tandfonline.com/doi/abs/10.1080/00031305.1990.10475751*

Elamir, E. A. H., and Seheult, A. H. (2003), "Trimmed L-moments," *Comput. Stat. Data Anal.*, 43(3), 299–314.

Erdös, P., and Rényi, A. (1959), "On Random Graphs I," *Publicationes Mathematicae Debrecen*, 6, 290–297.

Galvao, A. F., Montes-Rojas, G., Sosa-Escudero, W., and Wang, L. (2013), "Tests for skewness and kurtosis in the one-way error component model," *Journal of Multivariate Analysis*, 122(0), 35 – 52.
**URL:** *http://www.sciencedirect.com/science/article/pii/S0047259X13001346*

Harri, A., and Coble, K. H. (2011), "Normality testing: two new tests using L-moments," *Journal of Applied Statistics*, 38(7), 1369–1379.
**URL:** *http://dx.doi.org/10.1080/02664763.2010.498508*

Hosking, J. R. (1990), "L-moments: analysis and estimation of distributions using linear combinations of order statistics," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 105–124.

Hosking, J. R. M. (2014), *L-moments*. R package, version 2.4.
**URL:** *http://CRAN.R-project.org/package=lmom*

Jarque, C. M., and Bera, A. K. (1980), "Efficient tests for normality, homoscedasticity and serial independence of regression residuals," *Economics Letters*, 6(3), 255 – 259.
**URL:** *http://www.sciencedirect.com/science/article/pii/0165176580900245*

Livi, L., and Rizzi, A. (2013), "The graph matching problem," *Pattern Analysis and Applications*, 16(3), 253–283.

Moonesinghe, H., Valizadegan, H., Fodeh, S., and Tan, P.-N. (2007), A Probabilistic Substructure-Based Approach for Graph Classification,, in *Tools with Artificial Intelligence, 2007. ICTAI 2007. 19th IEEE International Conference on*, Vol. 1, pp. 346–349.

Mowshowitz, A., and Dehmer, M. (2012), "Entropy and the Complexity of Graphs Revisited," *Entropy*, 14(3), 559–570.

Pearson, E. S., D'Agostino, R. B., and Bowman, K. O. (1977), "Tests for departure from normality: Comparison of powers," *Biometrika*, 64(2), 231–246.
**URL:** *http://biomet.oxfordjournals.org/content/64/2/231.abstract*

Seier, E. (2002), "Comparison of tests for univariate normality," *InterStat*, 1, 1–17.

Wang, Q. (1996), "Direct Sample Estimators of L-Moments," *Water Resources Research*, 32(12), 3617–3619.

Wasserman, S., and Faust, K. (1994), *Social Network Analysis: Methods and Applications* Cambridge University Press.

Watts, D. J., and Strogatz, S. H. (1998), "Collective dynamics of 'small-world' networks," *Nature*, 393(6684), 440–442.