GRAPH TWO-SAMPLE TEST VIA EMPIRICAL LIKELIHOOD

A Thesis
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By
Ting Zou

In Partial Fulfillment of the Requirements
for the Degree of
MASTER OF SCIENCE

Major Department:
Statistics

April 2024

Fargo, North Dakota

NORTH DAKOTA STATE UNIVERSITY

Graduate School

Title

GRAPH TWO-SAMPLE TEST VIA EMPIRICAL LIKELIHOOD

By

Ting Zou

The supervisory committee certifies that this thesis complies with North Dakota State University's regulations and meets the accepted standards for the degree of

MASTER OF SCIENCE

SUPERVISORY COMMITTEE:

Prof. Mingao Yuan

Chair

Prof. Zhaohui Liu

Prof. Gang Shen

Approved:

April 12, 2024

Date

Prof. Rhonda Magel

Department Chair

ABSTRACT

In the past two decades, there has been a notable surge in network data. This proliferation has spurred significant advancements in methods for analyzing networks across various disciplines, including computer science, information sciences, biology, bioinformatics, physics, economics, sociology, and health science. Graph two-sample hypothesis testing, aimed at discerning differences between two populations of networks, arises naturally in diverse scenarios. In this paper, we delve into the essential yet intricate task of testing for equivalence between two networks. There are many testing procedures available. For instance, the t-test based on subgraph counts is one of the methods. In this paper, we propose a new test method by using the empirical likelihood. We run extensive simulations to evaluate the performance of the proposed method and apply it a real-world network. Based on the simulation experiments and real data application, the empirical likelihood test consistently outperforms existing subgraph count tests.

ACKNOWLEDGEMENTS

I would like to express my utmost appreciation to my thesis advisor, Dr. Mingao Yuan, for his precious guidance, support and encouragement during my master journey. I am also thankful to the members of my thesis commitee, Dr. Zhaohui Liu and Dr. Gang Shen, for their insightful comments, which enhance the clarity of the thesis and improve the quality of the work.

Special thanks to Dr. Zhaohui Liu for providing an opportunity of research assistantship during the first two years of my master studies, allowing me to finish my coursework without financial concerns.

Gratitude extends to the the professors in the statistics department for their outstanding teaching, enriching our learning experiences with their invaluable insights and guidance.

DEDICATION

This thesis is dedicated to my beloved children and to all those whom I hold dear in my heart.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1. INTRODUCTION

A graph or network, denoted as G, comprises an ordered pair consisting of a set of vertices (nodes) V(G) and a set of edges (links) E(G). Network Science offers a holistic approach to studying complex systems, transcending mere collections of units. Researchers employ it to explore a myriad of systems across diverse domains. For instance, in psychological studies, network nodes represent variables in datasets, with edges indicating pairwise conditional associations between variables(Borsboom et al., 2021); Similarly, in social networks, researchers analyze structures of relationships between individuals or social units, as well as interdependencies in behavior or attitudes(O'Malley and Marsden, 2008); In biological networks, the system-scale behavior of cells emerges from a multitude of molecular interaction webs. These encompass diverse types of interactions, such as protein-protein interaction, metabolic pathways, signaling cascades, and transcriptional regulatory networks. Moreover, Barabási and Oltvai (2004) observed an unexpected universality: the architectural features of molecular interaction networks within a cell are shared to a large degree by other complex systems. This suggests that similar laws govern the development and function of the most complex networks in nature.

The proliferation of network applications has spurred the development of statistically guaranteed methods such as network modeling, community detection, and network dynamics. Consequently, statistical network analysis has emerged as a vibrant area in the statistics literature, driving advancements in applied research (Albert and Barabási, 2002; Abbe and Sandon, 2016; Yuan and Wen, 2023; Maugis et al., 2020).

As recognized, hypothesis testing is a crucial and integral step, as it involves validating assumptions essential for developing statistical parameters. With the widespread adoption of network data, the importance of hypothesis testing of graphs has become pronounced across various scientific studies reliant on network analysis. In real-world scenarios, multiple graphs from diverse populations are common, necessitating the practical examination of differences between them. When presented with two populations of random graphs, determining whether they stem from the same distribution (model) or not becomes imperative. For instance, in a study examining differences in brain connectivity between individuals with Alzheimer's disease and healthy controls, researchers collect fMRI data from both groups(Sanz-Arigita et al., 2010). They create functional brain networks for each participant and then employ a graph two-sample test to compare connectivity patterns. This test helps pinpoint specific brain regions or connections showing significant disparities in connectivity strength or topology between the groups. Such findings offer insights into the neural mechanisms underlying the neurological disorder and may inform the development of diagnostic biomarkers or therapeutic strategies. Consequently, the graph two-sample test has garnered considerable attention in statistical and machine learning communities in recent years.

In this paper, we consider two network populations and concentrate on comparing the difference in population distributions. There are several testing procedures available in literature. For example, in Ghoshdastidar and Luxburg (2018), two novel tests were proposed, leveraging asymptotics for large graphs to address the two-sample testing problem for undirected unweighted graphs defined on a common vertex set. Motivated by the T_{fro} test, Yuan and Wen (2023) introduced a powerful test statistic for weighted graph two-sample hypothesis testing. Under the null hypothesis, the proposed test statistic converges in distribution to the standard normal distribution, and the test's power is theoretically described. Maugis et al. (2020) utilize subgraph counts to assess whether observations in a network sample originate from a specified distribution, model, or share the same model as another network sample.

A subgraph count is the number of copies of a given graph in another graph (Maugis et al., 2020). Subgraph counts method have been proved to be the most powerful tool available to compare networks in comparing large networks, in random graph theory and the study of large graphs (Ali et al., 2014, 2016). Subgraph counts are natural statistics to compare networks because they summarize a network through its fundamental building block and present tractable analytical properties. Rucinski (1988); Bickel et al. (2012); Maugis et al. (2020) proposed a t-test based on subgraph counts.

Originating from Owen (1988), the Empirical Likelihood (EL) method has enjoyed widespread popularity as an inference technique over the past three decades. The Empirical Likelihood method offers a flexible approach, enabling data analysts to apply it without imposing distributional assumptions. Its widespread utilization extends beyond nonparametric models to encompass semiparametric models as well. This versatility capitalizes on the robustness and powerful properties inherent in the likelihood-based approach. Over the past few years, several comprehensive reviews have been conducted on the empirical likelihood method. Nordman and Lahiri (2014) provided a synthesis of advancements in empirical likelihood methodology as applied to time series data. Chen and Van Keilegom (2009) presented a review on the empirical likelihood method for regression models. Additionally, Lazar (2021) conducted a review that delves into the computational aspects of empirical likelihood, alongside its connections to related constructs such as Bayesian empirical likelihood and Bayesian exponentially tilted empirical likelihood.

The Empirical Likelihood (EL) method finds application across diverse research domains, showcasing its versatility and utility. Furthermore, its scope has expanded to encompass two-sample testing scenarios, demonstrating its adaptability and robustness in statistical inference. For example, Wu and Yan (2012) devised a weighted empirical likelihood method, offering a significant computational advantage in inferring the difference of two population means.

In this paper, we apply the Empirical Likelihood method to the graph two-sample test based on subgraph counts. We present findings from extensive simulation study evaluating the finite sample performance of the proposed Empirical Likelihood method, comparing it to the conventional t-test-based approach. Our simulation results demonstrate that the proposed test exhibits higher power under small or moderate sample-size settings. Additionally, we provide a real-world data example to further illustrate the effectiveness of our method.

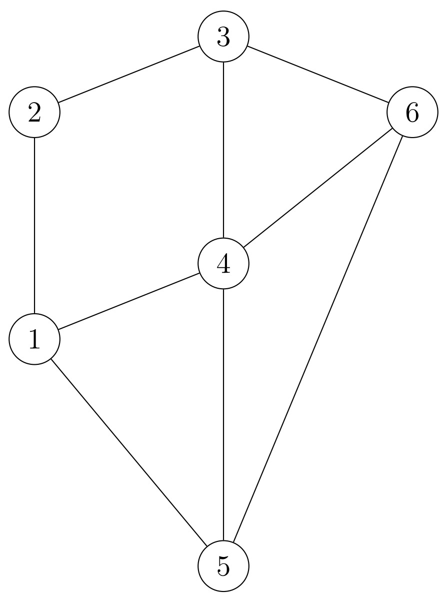
The thesis is organized as follows: Section 2 outlines the standard graph two-sample hypothesis testing problem and presents the empirical likelihood method. Section 3 discusses simulation results and a real data application.

CHAPTER 2. MAIN RESULTS

2.1. Graph two-sample test

Let $V = \{1, 2, \ldots, n\}$ be a vertex (node) set, and $G = (V, E)$ denote an undirected graph on V with edge set E. The adjacency matrix of graph G is a symmetric matrix $A \in 0, 1^{n\times n}$ such that $A_{ij} = 1$ if $(i, j) \in E$ and 0 otherwise. The graph G is binary or unweighted since A_{ij} only records the existence of an edge. If $A_{ij} \sim \text{Bern}(p_{ij})$, where $0 \leq p_{ij} \leq 1$, then the graph G is called an inhomogeneous random graph or inhomogeneous Erdős-Rényi graph.

An example of graph is visualized below. In this graph, there are 6 nodes and 9 edges. The node set is $\{1, 2, 3, 4, 5, 6\}$ and the edge set is $\{(1, 4), (1, 2), (1, 5), (4, 5), (4, 6), (3, 4), (2, 3), (3, 6), (5, 6)\}$.



Random graph models are essential for understanding complex networks like social and biological networks (Drobyshevskiy and Turdakov, 2019). Newman (2010) outlines two primary categories of network structure models: random graph models and mechanistic models, while acknowledging their nuanced boundary. Random graph models typically involve weaker causal associations and generate structures using unconstrained parameters. For example, the Erdős-Rényi random graph model treats each edge as an independent random variable, determining the network's structure. These models allow analytical or numerical computation of network metrics, serving as benchmarks for evaluating empirical data. Comparing observed network metrics with model expectations helps assess the model's adequacy as a null model. Aiello et al. (2000) introduced a random graph model representing sparse random graphs with predefined degree sequences. This model, defined by a few parameters, captures key features of large-scale graphs. It enables derivation of various graph properties, such as the expected distribution of connected component sizes within specific parameter ranges. The model's consistency with real-world massive graphs from telecommunication data is demonstrated, and discussions include topics like the threshold function and random graph evolution.

Now we introduce a popular inhomogeneous random graph (Bollobas et al., 2007). Let $f(x, y)$ be a symmetric function from $[0, 1]^2$ to $[0, 1]$. Let $U_i \sim Unif(0, 1)$, $(1 \leq i \leq n)$. We denote $A \sim Bern(f, n)$ if

$$A_{ij} \sim Bern(f(U_i, U_j)), \quad 1 \leq i < j \leq n,$$

where $A_{ii} = 0$, $A_{ij} = A_{ji}$ and A_{ij} are conditionally independent given U_i, $(1 \leq i \leq n)$.

Given an i.i.d. graph sample $A_1, \ldots, A_{m_1} \sim Bern(f, n_1)$ and an i.i.d. graph sample $B_1, \ldots, B_{m_2} \sim Bern(g, n_2)$, we are interested in the weighted graph two-sample hypothesis testing problem:

$$H_0 : f = g, \qquad H_1 : f \neq g. \tag{2.1}$$

The null hypothesis (H_0) states that the distributions of the two populations are not significantly different. The alternative hypothesis (H_1) suggests that they are significantly different.

Graph two-sample hypotheses test (2.1) has been widely studied. For instance, Tang et al. (2017) addressed testing whether two independent finite-dimensional random dot product graphs ($m_1 = m_2 = 1$) have generating latent positions from the same or related distributions. They proposed a kernel-based test statistic derived from adjacency spectral embedding, demonstrating its consistency across various alternatives. This method enables testing of hypotheses regarding graph distributions, applicable in neuroscience, network analysis, and machine learning. Ghoshdastidar et al. (2017) present a method for two-sample hypothesis testing on undirected graphs, using friendship networks from platforms like Facebook and LinkedIn as examples. Their approach doesn't rely on specific network generation assumptions and introduces a consistent two-sample test based on concentration of network statistics, shown to be minimax optimal for certain statistics. Moreover, in Ghoshdastidar et al. (2020), the authors explored hypothesis testing of graphs in high-dimensional settings, particularly focusing on scenarios with small sample sizes. The study investigates separation rates between two populations of random graph models and highlights conditions under which the testing problem becomes solvable. Additionally, the study introduces near-optimal two-sample tests that adapt to the sparsity level of the graphs. The models in Ghoshdastidar et al. (2017, 2020) assume the edges are independent. Hence the methods developed there may not be applicable to our setting.

Under our setting, one of the most recent test method is the t-test based on subgraph count proposed in Maugis et al. (2020). A subgraph count is the number of copies of a given graph in another graph. Let $X_i(H)$ be the density of a subgraph H in A_i and $Y_i(H)$ be the density of a subgraph H in B_i. For example, let H be a triangle. Then $X_i(H) = \frac{1}{n_1(n_1-1)(n_1-2)} \sum_{j\neq s\neq t} A_{js}A_{st}A_{jt}$, which is the number of triangle divided by the maximum number of triangles in graph A_i. This leads to two samples of subgraph densities $X_i(H)$ $(1 \leq i \leq m)$ and $Y_i(H)$ $(1 \leq i \leq l)$. Maugis et al. (2020) applied the t-test to the two samples test problem (2.1) as follows. Define

$$T = \frac{\bar{X}_{m_1} - \bar{Y}_{m_2}}{S\sqrt{\frac{1}{m_1}+\frac{1}{m_2}}},$$

where

$$S^2 = \frac{(m_1-1)S_1^2+(m_2-1)S_2^2}{m_1+m_2-2},$$

S_1 is the sample variance of $X_i(H)$ $(1 \leq i \leq m_1)$ and S_2 is the sample variance of $Y_i(H)$ $(1 \leq i \leq m_2)$. Reject the null hypothesis if $|T| > t_{(1-\alpha/2),m_1+m_2-2}$. Alternatively, if the p-value is less than a given type I error α, then reject H_0.

2.2. Graph two-sample test via empirical likelihood

In this section, we propose an empirical likelihood approach for the graph two-sample test.

Empirical likelihood has been widely used in classic two-sample inference. In a study by Liu et al. (2021), the empirical likelihood method was utilized to construct confidence regions for the difference in means of two d-dimensional samples. Their findings showcased that the empirical likelihood ratio test follows an asymptotic chi-squared distribution. Moreover, they successfully identified the correct Bartlett correction. Gurevich and Vexler (2011) introduced powerful entropy-based tests for assessing normality, uniformity, and exponentiality, improving upon existing methods by employing a density-based empirical likelihood approach. The method is extended to develop efficient two-sample tests based on sample entropy, which outperform traditional nonparametric tests, particularly in detecting nonconstant shifts. Additionally, Cao and Keilegom (2006) addresses the problem of determining whether two populations follow the same distribution. They utilize kernel estimators of density functions and introduce a test statistic derived from a local empirical likelihood method, establishing its asymptotic distribution. Additionally, the paper proposes a bootstrap technique for test calibration. Furthermore, it discusses the potential extension of the test to multiple samples and multivariate distributions.

Firstly, we introduce the classic empirical likelihood for two-sample test. Let $Y_{11}, ..., Y_{1m_1}$ and $Y_{21}, ..., Y_{2m_2}$ be two independent and identically distributed samples from $Y_1 \sim F$ and $Y_2 \sim G$, respectively, with $\mathrm{E}(Y_1) = \mu_1$, $\mathrm{E}(Y_2) = \mu_2$, $\mathrm{Var}(Y_1) = \sigma_1^2$ and $\mathrm{Var}(Y_2) = \sigma_2^2$. Let $\theta = \mu_1 - \mu_2$ be the parameter of interest. The classic two-sample hypotheses are

$$H_0 : \theta = 0, \quad H_1 : \theta \neq 0.$$

Under H_0, the two means are the same. Under H_1, the means are different.

The empirical likelihood two-sample test is defined as follows. The empirical log-likelihood ratio statistic on the parameter of interest, θ, is defined as

$$r(\theta) = \sum_{j=1}^{m_1} \log\{m_1 \hat{p}_{1j}(\theta)\} + \sum_{j=1}^{m_2} \log\{m_2 \hat{p}_{2j}(\theta)\} \tag{2.2}$$

where $\hat{p}_{1j}(\theta)$ and $\hat{p}_{2j}(\theta)$ maximize (p_1, p_2) subject to the following set of constraints:

$$\sum_{j=1}^{m_1} p_{1j} = 1, \quad \sum_{j=1}^{m_2} p_{2j} = 1, \quad p_{1j} \geq 0, \quad p_{2j} \geq 0, \tag{2.3}$$

$$\sum_{j=1}^{m_1} p_{1j} \; Y_{1j} - \sum_{j=1}^{m_2} p_{2j} \; Y_{2j} = \theta. \tag{2.4}$$

It is shown that $-2r(\theta)$ converges to the chi-square distribution with degree of freedom one under mild conditions. The empirical likelihood test rejects H_0 if

$$-2r(\theta) \geq {\chi^2}_1(\alpha) \tag{2.5}$$

where ${\chi^2}_1(\alpha)$ is the upper $(100\alpha)\%$ quantile from the χ^2_1 distribution.

We apply the empirical likelihood method to the graph two sample test. Let $Y_{1j} = X_i(H)$ and $Y_{2j} = Y_i(H)$. Then we construct the empirical likelihood test statistic as in (2.2). The rejection region is given in (2.5). We will use extensive simulation to evaluate the validity of the proposed test and apply it to a real-world network.

CHAPTER 3. SIMULATION AND REAL DATA APPLICATION

3.1. Simulation and real data application

3.1.1. Simulation

Throughout the simulation, we maintain a nominal type one error $\boldsymbol{\alpha}$ to be 0.05. Empirical size and power metrics are derived from 500 repetitions of the experiment. We consider sample sizes $\boldsymbol{n_1 = n_2 = 20, 30, 50, 70}$ and $\boldsymbol{m_1 = m_2 = 15, 30, 45, 60, 75, 90, 120}$. We compare the graph two-sample empirical likelihood method with the t-test in Maugis et al. (2020), particularly in scenarios of small or moderate sample sizes. Let $\boldsymbol{f(x, y)}$ be one of the following functions:

$$
\begin{aligned}
&\boldsymbol{f(x, y) = xy,}\\
&\boldsymbol{f(x, y) = 1/(1 + xy),}\\
&\boldsymbol{f(x, y) = 1/(1 + x^2y^2),}\\
&\boldsymbol{f(x, y) = 1/(1 + x^2 + y^2),}\\
&\boldsymbol{f(x, y) = e^{xy}/(1 + e^{xy}),}\\
&\boldsymbol{f(x, y) = 1/(1 + sinx * siny),}\\
&\boldsymbol{f(x, y) = 1/(1 + cosx * cosy).}
\end{aligned}
$$

We generate i.i.d. graph sample $\boldsymbol{A_1, \ldots, A_{m_1} \sim Bern(\delta_0 f, n_1)}$ and an i.i.d. graph sample $\boldsymbol{B_1, \ldots, B_{m_2} \sim Bern((\delta_0 + \lambda)f, n_2)}$. Here $\boldsymbol{\delta_0 = 0.2; 0.25}$ and $\boldsymbol{\lambda}$ takes 0, 0.005, 0.010, and 0.015. $\boldsymbol{\lambda = 0}$ corresponds to the type I error, while non-zero values signify the power analysis. The functions described above exhibit symmetry $\boldsymbol{(f(x, y) = f(y, x))}$,with values ranging between $\boldsymbol{0}$ and $\boldsymbol{1}$.

Tables 1 through 7 present a comparative analysis of type I errors and powers between the t-test and empirical likelihood (EL) methods, with $\boldsymbol{\delta_0 = 0.2}$. Correspondingly, Tables 8 through 14 provide a similar comparison for $\boldsymbol{\delta_0 = 0.25}$.

Our observations are as follows: (a) Both the t-test and empirical likelihood method demonstrate commendable performance regarding type I errors, maintaining $\boldsymbol{\alpha = 0.05}$. (b) Increasing $\boldsymbol{\lambda}$

Table 3.1. Simulated type-I error and (size) power with $\boldsymbol{f(x,y) = xy}$, $\boldsymbol{\delta_0 = 0.2}$.

(n_1, n_2)	(m_1, m_2)	$\lambda = 0$		$\lambda = 0.005$		$\lambda = 0.010$		$\lambda = 0.015$	
		t	**EL**	**t**	**EL**	**t**	**EL**	**t**	**EL**
(20, 20)	(15,15)	0.020	0.074	0.048	0.132	0.068	0.136	0.042	0.138
	(30,30)	0.044	0.080	0.050	0.088	0.078	0.132	0.114	0.176
	(45,45)	0.042	0.054	0.052	0.072	0.086	0.116	0.132	0.166
	(60,60)	0.042	0.062	0.064	0.078	0.134	0.170	0.212	0.258
	(75,75)	0.028	0.034	0.074	0.084	0.144	0.170	0.292	0.310
	(90,90)	0.046	0.062	0.098	0.104	0.132	0.154	0.322	0.356
	(120,120)	0.042	0.048	0.080	0.092	0.178	0.200	0.374	0.382
(30, 30)	(15, 15)	0.038	0.068	0.052	0.094	0.110	0.170	0.110	0.210
	(30,30)	0.052	0.072	0.064	0.094	0.118	0.152	0.266	0.334
	(45,45)	0.030	0.048	0.072	0.084	0.206	0.248	0.392	0.450
	(60,60)	0.052	0.056	0.084	0.104	0.276	0.300	0.514	0.532
	(75,75)	0.066	0.074	0.104	0.120	0.344	0.358	0.598	0.608
	(90,90)	0.044	0.052	0.138	0.142	0.338	0.352	0.682	0.686
	(120,120)	0.072	0.072	0.154	0.160	0.464	0.476	0.818	0.822
(50, 50)	(15, 15)	0.034	0.048	0.090	0.132	0.180	0.264	0.340	0.432
	(30,30)	0.044	0.060	0.114	0.140	0.362	0.400	0.654	0.692
	(45,45)	0.062	0.072	0.176	0.182	0.548	0.562	0.880	0.882
	(60,60)	0.046	0.050	0.254	0.278	0.678	0.688	0.942	0.946
	(75,75)	0.054	0.060	0.234	0.242	0.748	0.758	0.980	0.982
	(90,90)	0.040	0.040	0.314	0.318	0.798	0.808	0.988	0.990
	(120,120)	0.040	0.038	0.412	0.422	0.918	0.922	1.000	1.000
(70,70)	(15, 15)	0.036	0.056	0.124	0.172	0.338	0.412	0.650	0.718
	(30,30)	0.056	0.064	0.180	0.202	0.614	0.636	0.946	0.952
	(45,45)	0.046	0.050	0.278	0.298	0.818	0.828	0.990	0.990
	(60,60)	0.046	0.050	0.386	0.394	0.900	0.908	1.000	1.000
	(75,75)	0.038	0.046	0.434	0.436	0.958	0.964	1.000	1.000
	(90,90)	0.044	0.048	0.502	0.508	0.976	0.976	1.000	1.000
	(120,120)	0.066	0.072	0.662	0.674	0.996	0.996	1.000	1.000

from 0.005 to 0.015 leads to proportional enhancements in the powers of both methods. Notably, the empirical likelihood method consistently exhibits superior performance compared to the t-test, consistently achieving higher power. (c) A rise in $\boldsymbol{\delta_0}$ from 0.2 to 0.25 for identical distributions results in a corresponding increase in test power. (d) The powers of both methods experience increments with an increase in the number of nodes ($\boldsymbol{n_1}$ and $\boldsymbol{n_2}$) or sample sizes ($\boldsymbol{m_1}$ and $\boldsymbol{m_2}$). (e) As the numbers of nodes or sample sizes grow sufficiently large, both the empirical likelihood method and t-test approach a power of $\mathbf{1}$. It is therefore reasonable to assert that the empirical likelihood method outperforms the t-test, particularly in scenarios involving small or medium sample sizes.

Table 3.2. Simulated type-I error and (size) power with $\boldsymbol{f(x,y) = 1/(1+xy)}$, $\boldsymbol{\delta_0 = 0.2}$.

$(\boldsymbol{n_1,n_2})$	$(\boldsymbol{m_1,m_2})$	$\boldsymbol{\lambda} = 0$		$\boldsymbol{\lambda} = 0.005$		$\boldsymbol{\lambda} = 0.010$		$\boldsymbol{\lambda} = 0.015$	
		t	**EL**	**t**	**EL**	**t**	**EL**	**t**	**EL**
(20, 20)	(15, 15)	0.054	0.086	0.080	0.122	0.230	0.272	0.480	0.526
	(30,30)	0.030	0.050	0.154	0.182	0.380	0.428	0.730	0.750
	(45,45)	0.042	0.046	0.198	0.214	0.550	0.574	0.880	0.906
	(60,60)	0.058	0.064	0.220	0.242	0.694	0.710	0.958	0.962
	(75,75)	0.038	0.038	0.290	0.302	0.776	0.784	0.988	0.988
	(90,90)	0.058	0.056	0.314	0.322	0.860	0.866	0.996	0.996
	(120,120)	0.062	0.066	0.392	0.396	0.924	0.928	1.000	1.000
(30, 30)	(15, 15)	0.058	0.072	0.170	0.222	0.466	0.532	0.812	0.848
	(30,30)	0.042	0.052	0.290	0.312	0.816	0.840	0.986	0.992
	(45,45)	0.040	0.044	0.404	0.414	0.912	0.924	1.000	1.000
	(60,60)	0.052	0.058	0.478	0.486	0.974	0.976	1.000	1.000
	(75,75)	0.044	0.044	0.614	0.620	0.992	0.992	1.000	1.000
	(90,90)	0.062	0.064	0.688	0.690	1.000	1.000	1.000	1.000
	(120,120)	0.058	0.060	0.818	0.822	1.000	1.000	1.000	1.000
(50, 50)	(15, 15)	0.060	0.078	0.400	0.456	0.912	0.932	1.000	1.000
	(30,30)	0.066	0.070	0.676	0.696	0.994	0.994	1.000	1.000
	(45,45)	0.046	0.048	0.874	0.880	1.000	1.000	1.000	1.000
	(60,60)	0.056	0.060	0.936	0.940	1.000	1.000	1.000	1.000
	(75,75)	0.048	0.050	0.966	0.970	1.000	1.000	1.000	1.000
	(90,90)	0.046	0.046	0.992	0.992	1.000	1.000	1.000	1.000
	(120,120)	0.050	0.050	0.996	0.998	1.000	1.000	1.000	1.000
(70,70)	(15, 15)	0.048	0.058	0.696	0.740	0.996	0.998	1.000	1.000
	(30,30)	0.048	0.056	0.932	0.940	1.000	1.000	1.000	1.000
	(45,45)	0.040	0.048	0.984	0.990	1.000	1.000	1.000	1.000
	(60,60)	0.058	0.062	0.996	0.996	1.000	1.000	1.000	1.000
	(75,75)	0.068	0.066	1.000	1.000	1.000	1.000	1.000	1.000
	(90,90)	0.056	0.056	1.000	1.000	1.000	1.000	1.000	1.000
	(120,120)	0.058	0.060	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.3. Simulated type-I error and (size) power with $\boldsymbol{f(x,y) = 1/(1+x^2y^2)}$, $\boldsymbol{\delta_0 = 0.2}$.

(n_1, n_2)	(m_1, m_2)	$\lambda = 0$		$\lambda = 0.005$		$\lambda = 0.010$		$\lambda = 0.015$	
		t	**EL**	**t**	**EL**	**t**	**EL**	**t**	**EL**
(20, 20)	(15, 15)	0.062	0.080	0.108	0.142	0.246	0.342	0.526	0.608
	(30,30)	0.030	0.040	0.186	0.208	0.490	0.522	0.824	0.848
	(45,45)	0.044	0.050	0.260	0.280	0.670	0.686	0.946	0.954
	(60,60)	0.044	0.044	0.292	0.302	0.768	0.772	0.988	0.990
	(75,75)	0.036	0.038	0.348	0.360	0.860	0.866	0.994	0.998
	(90,90)	0.052	0.054	0.366	0.372	0.930	0.934	0.998	0.998
	(120,120)	0.040	0.042	0.498	0.508	0.972	0.974	1.000	1.000
(30, 30)	(15, 15)	0.066	0.088	0.188	0.232	0.608	0.672	0.880	0.912
	(30,30)	0.046	0.050	0.372	0.390	0.888	0.894	1.000	1.000
	(45,45)	0.040	0.048	0.498	0.506	0.978	0.982	1.000	1.000
	(60,60)	0.052	0.052	0.624	0.632	0.994	0.994	1.000	1.000
	(75,75)	0.046	0.050	0.676	0.684	0.998	0.998	1.000	1.000
	(90,90)	0.038	0.038	0.770	0.776	1.000	1.000	1.000	1.000
	(120,120)	0.044	0.046	0.880	0.882	1.000	1.000	1.000	1.000
(50, 50)	(15, 15)	0.046	0.064	0.508	0.542	0.974	0.980	1.000	1.000
	(30,30)	0.052	0.064	0.762	0.776	1.000	1.000	1.000	1.000
	(45,45)	0.052	0.052	0.926	0.930	1.000	1.000	1.000	1.000
	(60,60)	0.034	0.038	0.964	0.970	1.000	1.000	1.000	1.000
	(75,75)	0.068	0.070	0.996	0.996	1.000	1.000	1.000	1.000
	(90,90)	0.046	0.048	1.000	1.000	1.000	1.000	1.000	1.000
	(120,120)	0.046	0.046	1.000	1.000	1.000	1.000	1.000	1.000
(70,70)	(15, 15)	0.054	0.070	0.756	0.802	1.000	1.000	1.000	1.000
	(30,30)	0.052	0.058	0.966	0.970	1.000	1.000	1.000	1.000
	(45,45)	0.028	0.030	1.000	1.000	1.000	1.000	1.000	1.000
	(60,60)	0.046	0.048	1.000	1.000	1.000	1.000	1.000	1.000
	(75,75)	0.050	0.052	1.000	1.000	1.000	1.000	1.000	1.000
	(90,90)	0.062	0.062	1.000	1.000	1.000	1.000	1.000	1.000
	(120,120)	0.044	0.048	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.4. Simulated type-I error and (size) power with $\boldsymbol{f(x,y) = 1/(1+x^2+y^2)}$, $\boldsymbol{\delta_0 = 0.2}$.

(n_1, n_2)	(m_1, m_2)	$\lambda = 0$		$\lambda = 0.005$		$\lambda = 0.010$		$\lambda = 0.015$	
		t	**EL**	**t**	**EL**	**t**	**EL**	**t**	**EL**
(20, 20)	(15, 15)	0.042	0.072	0.076	0.116	0.148	0.210	0.302	0.376
	(30,30)	0.052	0.066	0.084	0.108	0.270	0.288	0.502	0.564
	(45,45)	0.060	0.062	0.136	0.148	0.390	0.418	0.680	0.698
	(60,60)	0.036	0.042	0.146	0.154	0.464	0.480	0.836	0.840
	(75,75)	0.060	0.064	0.184	0.190	0.544	0.552	0.892	0.896
	(90,90)	0.044	0.050	0.218	0.222	0.628	0.630	0.936	0.942
	(120,120)	0.046	0.048	0.292	0.296	0.770	0.776	0.972	0.974
(30, 30)	(15, 15)	0.062	0.088	0.106	0.138	0.330	0.400	0.626	0.686
	(30,30)	0.038	0.048	0.194	0.222	0.580	0.618	0.926	0.928
	(45,45)	0.042	0.048	0.226	0.250	0.744	0.760	0.988	0.992
	(60,60)	0.052	0.056	0.340	0.356	0.876	0.878	0.996	0.996
	(75,75)	0.060	0.066	0.422	0.434	0.940	0.942	1.000	1.000
	(90,90)	0.048	0.050	0.486	0.492	0.968	0.970	1.000	1.000
	(120,120)	0.052	0.054	0.586	0.590	0.990	0.990	1.000	1.000
(50, 50)	(15, 15)	0.054	0.058	0.236	0.294	0.752	0.788	0.968	0.978
	(30,30)	0.038	0.044	0.452	0.464	0.964	0.976	1.000	1.000
	(45,45)	0.046	0.046	0.660	0.678	0.994	0.996	1.000	1.000
	(60,60)	0.048	0.048	0.756	0.762	1.000	1.000	1.000	1.000
	(75,75)	0.044	0.044	0.850	0.852	1.000	1.000	1.000	1.000
	(90,90)	0.072	0.078	0.916	0.918	1.000	1.000	1.000	1.000
	(120,120)	0.064	0.068	0.950	0.950	1.000	1.000	1.000	1.000
(70,70)	(15, 15)	0.064	0.076	0.408	0.444	0.950	0.960	1.000	1.000
	(30,30)	0.036	0.046	0.748	0.760	0.998	0.998	1.000	1.000
	(45,45)	0.040	0.044	0.886	0.890	1.000	1.000	1.000	1.000
	(60,60)	0.038	0.040	0.954	0.954	1.000	1.000	1.000	1.000
	(75,75)	0.044	0.048	0.980	0.982	1.000	1.000	1.000	1.000
	(90,90)	0.048	0.046	0.998	0.998	1.000	1.000	1.000	1.000
	(120,120)	0.048	0.052	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.5. Simulated type-I error and (size) power with $\boldsymbol{f(x,y) = e^{xy}/(1+e^{xy})}$, $\boldsymbol{\delta_0 = 0.2}$.

(n_1, n_2)	(m_1, m_2)	$\boldsymbol{\lambda} = 0$		$\boldsymbol{\lambda} = 0.005$		$\boldsymbol{\lambda} = 0.010$		$\boldsymbol{\lambda} = 0.015$	
		t	**EL**	**t**	**EL**	**t**	**EL**	**t**	**EL**
(20, 20)	(15, 15)	0.050	0.090	0.054	0.096	0.110	0.166	0.224	0.300
	(30,30)	0.034	0.052	0.090	0.098	0.242	0.260	0.404	0.436
	(45,45)	0.048	0.062	0.110	0.114	0.284	0.312	0.526	0.550
	(60,60)	0.028	0.032	0.146	0.150	0.272	0.298	0.708	0.730
	(75,75)	0.058	0.064	0.168	0.174	0.460	0.470	0.796	0.806
	(90,90)	0.056	0.058	0.178	0.190	0.520	0.532	0.872	0.880
	(120,120)	0.050	0.050	0.252	0.268	0.662	0.668	0.936	0.946
(30, 30)	(15, 15)	0.040	0.062	0.104	0.146	0.254	0.308	0.538	0.608
	(30,30)	0.042	0.042	0.104	0.146	0.486	0.506	0.850	0.862
	(45,45)	0.038	0.040	0.236	0.250	0.710	0.720	0.956	0.958
	(60,60)	0.058	0.062	0.286	0.290	0.792	0.794	0.984	0.986
	(75,75)	0.062	0.062	0.368	0.374	0.874	0.880	0.998	0.998
	(90,90)	0.058	0.060	0.428	0.438	0.940	0.942	0.998	1.000
	(120,120)	0.048	0.052	0.506	0.512	0.984	0.984	1.000	1.000
(50, 50)	(15, 15)	0.056	0.080	0.286	0.338	0.752	0.790	0.966	0.976
	(30,30)	0.046	0.052	0.446	0.474	0.954	0.960	1.000	1.000
	(45,45)	0.048	0.050	0.630	0.652	0.998	0.998	1.000	1.000
	(60,60)	0.032	0.034	0.740	0.754	1.000	1.000	1.000	1.000
	(75,75)	0.050	0.050	0.852	0.858	1.000	1.000	1.000	1.000
	(90,90)	0.056	0.056	0.884	0.892	1.000	1.000	1.000	1.000
	(120,120)	0.058	0.064	0.954	0.956	1.000	1.000	1.000	1.000
(70,70)	(15, 15)	0.052	0.072	0.490	0.538	0.970	0.980	1.000	1.000
	(30,30)	0.056	0.064	0.746	0.762	1.000	1.000	1.000	1.000
	(45,45)	0.044	0.048	0.920	0.926	1.000	1.000	1.000	1.000
	(60,60)	0.046	0.050	0.976	0.976	1.000	1.000	1.000	1.000
	(75,75)	0.054	0.060	0.988	0.988	1.000	1.000	1.000	1.000
	(90,90)	0.040	0.038	1.000	1.000	1.000	1.000	1.000	1.000
	(120,120)	0.060	0.060	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.6. Simulated type-I error and (size) power with $\boldsymbol{f(x,y) = 1/(1+sinx*siny)}$, $\boldsymbol{\delta_0 = 0.2}$.

(n_1, n_2)	(m_1, m_2)	$\lambda = 0$		$\lambda = 0.005$		$\lambda = 0.010$		$\lambda = 0.015$	
		t	**EL**	**t**	**EL**	**t**	**EL**	**t**	**EL**
(20, 20)	(15, 15)	0.050	0.082	0.070	0.108	0.230	0.288	0.444	0.508
	(30,30)	0.044	0.070	0.172	0.198	0.444	0.486	0.764	0.786
	(45,45)	0.050	0.054	0.162	0.186	0.604	0.610	0.908	0.916
	(60,60)	0.050	0.052	0.236	0.248	0.708	0.716	0.972	0.972
	(75,75)	0.058	0.064	0.338	0.342	0.810	0.824	0.994	0.994
	(90,90)	0.040	0.044	0.348	0.354	0.884	0.886	1.000	1.000
	(120,120)	0.052	0.052	0.420	0.426	0.952	0.952	1.000	1.000
(30, 30)	(15, 15)	0.040	0.052	0.164	0.200	0.492	0.546	0.822	0.854
	(30,30)	0.044	0.046	0.304	0.332	0.816	0.826	0.992	0.992
	(45,45)	0.052	0.054	0.410	0.426	0.948	0.952	1.000	1.000
	(60,60)	0.042	0.050	0.558	0.572	0.990	0.990	1.000	1.000
	(75,75)	0.038	0.038	0.662	0.668	0.994	0.994	1.000	1.000
	(90,90)	0.048	0.050	0.690	0.694	0.996	0.996	1.000	1.000
	(120,120)	0.044	0.048	0.850	0.852	1.000	1.000	1.000	1.000
(50, 50)	(15, 15)	0.030	0.058	0.430	0.494	0.942	0.956	1.000	1.000
	(30,30)	0.050	0.060	0.718	0.734	1.000	1.000	1.000	1.000
	(45,45)	0.040	0.044	0.864	0.870	1.000	1.000	1.000	1.000
	(60,60)	0.052	0.054	0.946	0.948	1.000	1.000	1.000	1.000
	(75,75)	0.030	0.036	0.984	0.986	1.000	1.000	1.000	1.000
	(90,90)	0.044	0.046	0.996	0.996	1.000	1.000	1.000	1.000
	(120,120)	0.052	0.054	1.000	1.000	1.000	1.000	1.000	1.000
(70,70)	(15, 15)	0.056	0.080	0.712	0.756	0.998	0.998	1.000	1.000
	(30,30)	0.038	0.042	0.948	0.956	1.000	1.000	1.000	1.000
	(45,45)	0.040	0.042	0.992	0.992	1.000	1.000	1.000	1.000
	(60,60)	0.042	0.044	0.998	0.998	1.000	1.000	1.000	1.000
	(75,75)	0.052	0.058	1.000	1.000	1.000	1.000	1.000	1.000
	(90,90)	0.044	0.054	1.000	1.000	1.000	1.000	1.000	1.000
	(120,120)	0.038	0.038	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.7. Simulated type-I error and (size) power with $\boldsymbol{f(x,y) = 1/(1+cosx*cosy)}$, $\boldsymbol{\delta_0 = 0.2}$.

(n_1, n_2)	(m_1, m_2)	$\lambda = 0$		$\lambda = 0.005$		$\lambda = 0.010$		$\lambda = 0.015$	
		t	**EL**	**t**	**EL**	**t**	**EL**	**t**	**EL**
(20, 20)	(15, 15)	0.052	0.078	0.056	0.096	0.122	0.168	0.250	0.322
	(30,30)	0.048	0.060	0.116	0.126	0.224	0.252	0.418	0.454
	(45,45)	0.060	0.068	0.112	0.118	0.334	0.366	0.668	0.688
	(60,60)	0.044	0.050	0.130	0.144	0.402	0.426	0.746	0.752
	(75,75)	0.062	0.064	0.206	0.216	0.516	0.520	0.856	0.862
	(90,90)	0.062	0.066	0.196	0.200	0.542	0.548	0.892	0.896
	(120,120)	0.056	0.054	0.238	0.244	0.704	0.710	0.956	0.956
(30, 30)	(15, 15)	0.052	0.080	0.106	0.136	0.304	0.376	0.550	0.604
	(30,30)	0.062	0.064	0.222	0.242	0.588	0.606	0.868	0.894
	(45,45)	0.062	0.062	0.254	0.272	0.728	0.746	0.968	0.970
	(60,60)	0.044	0.052	0.304	0.322	0.830	0.840	0.996	0.996
	(75,75)	0.046	0.044	0.362	0.368	0.916	0.918	0.996	0.996
	(90,90)	0.054	0.056	0.456	0.462	0.958	0.962	1.000	1.000
	(120,120)	0.050	0.054	0.552	0.558	0.984	0.986	1.000	1.000
(50, 50)	(15, 15)	0.058	0.076	0.276	0.318	0.750	0.808	0.974	0.988
	(30,30)	0.050	0.062	0.464	0.478	0.976	0.984	1.000	1.000
	(45,45)	0.054	0.062	0.654	0.668	0.988	0.988	1.000	1.000
	(60,60)	0.070	0.078	0.782	0.788	1.000	1.000	1.000	1.000
	(75,75)	0.072	0.072	0.882	0.892	1.000	1.000	1.000	1.000
	(90,90)	0.054	0.052	0.910	0.912	1.000	1.000	1.000	1.000
	(120,120)	0.048	0.048	0.982	0.982	1.000	1.000	1.000	1.000
(70,70)	(15, 15)	0.052	0.072	0.496	0.536	0.968	0.976	1.000	1.000
	(30,30)	0.052	0.062	0.772	0.788	1.000	1.000	1.000	1.000
	(45,45)	0.050	0.052	0.944	0.946	1.000	1.000	1.000	1.000
	(60,60)	0.040	0.042	0.978	0.980	1.000	1.000	1.000	1.000
	(75,75)	0.064	0.066	0.994	0.994	1.000	1.000	1.000	1.000
	(90,90)	0.032	0.034	0.998	0.998	1.000	1.000	1.000	1.000
	(120,120)	0.048	0.052	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.8. Simulated type-I error and (size) power with $\boldsymbol{f(x,y) = xy}$, $\boldsymbol{\delta_0 = 0.25}$.

(n_1, n_2)	(m_1, m_2)	$\lambda = 0$		$\lambda = 0.005$		$\lambda = 0.010$		$\lambda = 0.015$	
		t	**EL**	**t**	**EL**	**t**	**EL**	**t**	**EL**
(20, 20)	(15, 15)	0.032	0.090	0.054	0.116	0.054	0.126	0.072	0.152
	(30,30)	0.054	0.074	0.054	0.082	0.088	0.118	0.112	0.170
	(45,45)	0.048	0.06	0.046	0.066	0.124	0.152	0.182	0.226
	(60,60)	0.046	0.062	0.060	0.082	0.110	0.132	0.190	0.226
	(75,75)	0.048	0.056	0.062	0.072	0.154	0.168	0.264	0.286
	(90,90)	0.052	0.054	0.060	0.080	0.144	0.154	0.264	0.288
	(120,120)	0.056	0.064	0.090	0.098	0.220	0.238	0.376	0.394
(30, 30)	(15, 15)	0.044	0.088	0.072	0.112	0.084	0.142	0.146	0.232
	(30,30)	0.054	0.070	0.060	0.078	0.160	0.208	0.274	0.308
	(45,45)	0.060	0.062	0.088	0.110	0.182	0.204	0.382	0.402
	(60,60)	0.062	0.070	0.098	0.108	0.236	0.252	0.478	0.502
	(75,75)	0.042	0.044	0.100	0.110	0.308	0.318	0.598	0.606
	(90,90)	0.048	0.058	0.106	0.118	0.366	0.376	0.668	0.678
	(120,120)	0.038	0.042	0.146	0.148	0.464	0.474	0.766	0.768
(50, 50)	(15, 15)	0.052	0.080	0.082	0.122	0.168	0.230	0.324	0.402
	(30,30)	0.066	0.076	0.126	0.150	0.310	0.354	0.606	0.644
	(45,45)	0.044	0.048	0.148	0.150	0.454	0.464	0.796	0.812
	(60,60)	0.060	0.066	0.204	0.218	0.596	0.602	0.886	0.894
	(75,75)	0.040	0.046	0.214	0.224	0.682	0.692	0.952	0.954
	(90,90)	0.046	0.044	0.246	0.256	0.746	0.752	0.974	0.978
	(120,120)	0.058	0.060	0.362	0.372	0.872	0.874	0.998	0.998
(70,70)	(15, 15)	0.062	0.082	0.118	0.144	0.306	0.356	0.580	0.638
	(30,30)	0.026	0.034	0.162	0.190	0.558	0.584	0.878	0.892
	(45,45)	0.036	0.044	0.234	0.256	0.708	0.726	0.974	0.976
	(60,60)	0.052	0.054	0.342	0.352	0.818	0.826	0.992	0.994
	(75,75)	0.044	0.046	0.376	0.390	0.906	0.908	0.998	0.998
	(90,90)	0.044	0.042	0.430	0.438	0.962	0.964	1.000	1.000
	(120,120)	0.044	0.046	0.518	0.524	0.992	0.992	1.000	1.000

Table 3.9. Simulated type-I error and (size) power with $\boldsymbol{f(x,y) = 1/(1+xy)}$, $\boldsymbol{\delta_0 = 0.25}$.

(n_1, n_2)	(m_1, m_2)	$\lambda = 0$		$\lambda = 0.005$		$\lambda = 0.010$		$\lambda = 0.015$	
		t	**EL**	**t**	**EL**	**t**	**EL**	**t**	**EL**
(20, 20)	(15, 15)	0.044	0.064	0.100	0.132	0.220	0.282	0.416	0.494
	(30,30)	0.056	0.060	0.150	0.166	0.374	0.400	0.708	0.728
	(45,45)	0.050	0.060	0.162	0.174	0.534	0.552	0.862	0.874
	(60,60)	0.042	0.054	0.250	0.262	0.688	0.700	0.964	0.968
	(75,75)	0.034	0.040	0.304	0.308	0.804	0.810	0.986	0.988
	(90,90)	0.058	0.060	0.338	0.346	0.820	0.822	0.996	0.996
	(120,120)	0.048	0.050	0.414	0.424	0.930	0.932	1.000	1.000
(30, 30)	(15, 15)	0.038	0.058	0.156	0.188	0.434	0.502	0.800	0.848
	(30,30)	0.050	0.068	0.256	0.284	0.762	0.784	0.974	0.976
	(45,45)	0.050	0.046	0.382	0.392	0.912	0.918	0.996	0.996
	(60,60)	0.056	0.056	0.468	0.484	0.966	0.966	1.000	1.000
	(75,75)	0.046	0.052	0.588	0.594	0.990	0.992	1.000	1.000
	(90,90)	0.042	0.046	0.674	0.676	0.996	0.996	1.000	1.000
	(120,120)	0.050	0.050	0.772	0.776	0.998	0.998	1.000	1.000
(50, 50)	(15, 15)	0.072	0.084	0.360	0.412	0.872	0.898	0.998	0.998
	(30,30)	0.050	0.058	0.630	0.656	0.992	0.994	1.000	1.000
	(45,45)	0.032	0.032	0.806	0.816	1.000	1.000	1.000	1.000
	(60,60)	0.030	0.034	0.898	0.908	1.000	1.000	1.000	1.000
	(75,75)	0.054	0.054	0.960	0.962	1.000	1.000	1.000	1.000
	(90,90)	0.054	0.056	0.986	0.986	1.000	1.000	1.000	1.000
	(120,120)	0.056	0.056	0.996	0.998	1.000	1.000	1.000	1.000
(70,70)	(15, 15)	0.048	0.066	0.572	0.634	0.996	0.996	1.000	1.000
	(30,30)	0.064	0.064	0.866	0.872	1.000	1.000	1.000	1.000
	(45,45)	0.042	0.042	0.968	0.972	1.000	1.000	1.000	1.000
	(60,60)	0.048	0.052	0.990	0.990	1.000	1.000	1.000	1.000
	(75,75)	0.054	0.056	1.000	1.000	1.000	1.000	1.000	1.000
	(90,90)	0.046	0.048	1.000	1.000	1.000	1.000	1.000	1.000
	(120,120)	0.046	0.048	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.10. Simulated type-I error and (size) power with $\boldsymbol{f(x,y) = 1/(1+x^2y^2)}$, $\boldsymbol{\delta_0 = 0.25}$.

(n_1, n_2)	(m_1, m_2)	$\lambda = 0$		$\lambda = 0.005$		$\lambda = 0.010$		$\lambda = 0.015$	
		t	**EL**	**t**	**EL**	**t**	**EL**	**t**	**EL**
(20, 20)	(15, 15)	0.040	0.062	0.100	0.122	0.268	0.322	0.494	0.564
	(30,30)	0.062	0.076	0.164	0.192	0.478	0.504	0.804	0.830
	(45,45)	0.054	0.068	0.202	0.220	0.650	0.670	0.950	0.954
	(60,60)	0.064	0.068	0.308	0.314	0.742	0.754	0.974	0.980
	(75,75)	0.064	0.064	0.380	0.394	0.856	0.860	0.990	0.990
	(90,90)	0.056	0.060	0.384	0.394	0.918	0.922	0.998	0.998
	(120,120)	0.044	0.052	0.476	0.486	0.976	0.978	1.000	1.000
(30, 30)	(15, 15)	0.058	0.066	0.164	0.208	0.536	0.600	0.888	0.910
	(30,30)	0.036	0.048	0.320	0.340	0.850	0.864	0.996	0.996
	(45,45)	0.058	0.068	0.470	0.478	0.950	0.954	1.000	1.000
	(60,60)	0.046	0.056	0.560	0.574	0.994	0.994	1.000	1.000
	(75,75)	0.042	0.044	0.682	0.688	0.998	0.998	1.000	1.000
	(90,90)	0.060	0.060	0.760	0.764	1.000	1.000	1.000	1.000
	(120,120)	0.044	0.044	0.858	0.860	1.000	1.000	1.000	1.000
(50, 50)	(15, 15)	0.044	0.052	0.426	0.482	0.946	0.952	1.000	1.000
	(30,30)	0.050	0.056	0.728	0.750	1.000	1.000	1.000	1.000
	(45,45)	0.060	0.060	0.900	0.910	1.000	1.000	1.000	1.000
	(60,60)	0.058	0.062	0.952	0.954	1.000	1.000	1.000	1.000
	(75,75)	0.040	0.042	0.982	0.986	1.000	1.000	1.000	1.000
	(90,90)	0.070	0.070	0.996	0.996	1.000	1.000	1.000	1.000
	(120,120)	0.036	0.038	0.998	0.998	1.000	1.000	1.000	1.000
(70,70)	(15, 15)	0.050	0.064	0.674	0.726	1.000	1.000	1.000	1.000
	(30,30)	0.064	0.074	0.950	0.954	1.000	1.000	1.000	1.000
	(45,45)	0.058	0.070	0.992	0.992	1.000	1.000	1.000	1.000
	(60,60)	0.062	0.064	1.000	1.000	1.000	1.000	1.000	1.000
	(75,75)	0.050	0.052	1.000	1.000	1.000	1.000	1.000	1.000
	(90,90)	0.048	0.050	1.000	1.000	1.000	1.000	1.000	1.000
	(120,120)	0.044	0.044	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.11. Simulated type-I error and (size) power with $\boldsymbol{f(x,y) = 1/(1+x^2+y^2)}$, $\boldsymbol{\delta_0 = 0.25}$.

(n_1, n_2)	(m_1, m_2)	$\boldsymbol{\lambda} = 0$		$\boldsymbol{\lambda} = 0.005$		$\boldsymbol{\lambda} = 0.010$		$\boldsymbol{\lambda} = 0.015$	
		t	**EL**	**t**	**EL**	**t**	**EL**	**t**	**EL**
(20, 20)	(15, 15)	0.048	0.070	0.060	0.088	0.124	0.170	0.230	0.316
	(30,30)	0.054	0.068	0.092	0.112	0.258	0.296	0.490	0.538
	(45,45)	0.056	0.066	0.136	0.146	0.348	0.380	0.646	0.656
	(60,60)	0.048	0.048	0.142	0.154	0.482	0.498	0.802	0.812
	(75,75)	0.050	0.062	0.176	0.180	0.570	0.590	0.908	0.910
	(90,90)	0.062	0.066	0.222	0.226	0.650	0.660	0.934	0.936
	(120,120)	0.050	0.048	0.250	0.256	0.742	0.746	0.976	0.978
(30, 30)	(15, 15)	0.034	0.052	0.134	0.170	0.334	0.400	0.608	0.670
	(30,30)	0.046	0.056	0.190	0.220	0.574	0.584	0.868	0.884
	(45,45)	0.044	0.048	0.276	0.296	0.712	0.728	0.952	0.954
	(60,60)	0.048	0.054	0.312	0.334	0.826	0.828	0.996	0.996
	(75,75)	0.060	0.064	0.366	0.374	0.872	0.880	0.996	0.996
	(90,90)	0.040	0.040	0.418	0.426	0.960	0.960	1.000	1.000
	(120,120)	0.058	0.058	0.532	0.544	0.990	0.990	1.000	1.000
(50, 50)	(15, 15)	0.058	0.080	0.248	0.292	0.654	0.706	0.948	0.960
	(30,30)	0.036	0.040	0.402	0.422	0.910	0.916	1.000	1.000
	(45,45)	0.042	0.050	0.550	0.558	0.990	0.990	1.000	1.000
	(60,60)	0.062	0.064	0.666	0.676	1.000	1.000	1.000	1.000
	(75,75)	0.050	0.050	0.796	0.804	1.000	1.000	1.000	1.000
	(90,90)	0.042	0.046	0.854	0.856	1.000	1.000	1.000	1.000
	(120,120)	0.046	0.048	0.930	0.934	1.000	1.000	1.000	1.000
(70,70)	(15, 15)	0.054	0.076	0.364	0.400	0.874	0.904	0.996	0.998
	(30,30)	0.064	0.078	0.632	0.642	0.998	0.998	1.000	1.000
	(45,45)	0.052	0.058	0.808	0.810	1.000	1.000	1.000	1.000
	(60,60)	0.056	0.058	0.900	0.906	1.000	1.000	1.000	1.000
	(75,75)	0.054	0.056	0.952	0.954	1.000	1.000	1.000	1.000
	(90,90)	0.036	0.038	0.986	0.986	1.000	1.000	1.000	1.000
	(120,120)	0.066	0.070	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.12. Simulated type-I error and (size) power with $\boldsymbol{f(x,y) = e^{xy}/(1+e^{xy})}$, $\boldsymbol{\delta_0 = 0.25}$.

(n_1, n_2)	(m_1, m_2)	$\lambda = 0$		$\lambda = 0.005$		$\lambda = 0.010$		$\lambda = 0.015$	
		t	**EL**	**t**	**EL**	**t**	**EL**	**t**	**EL**
(20, 20)	(15, 15)	0.052	0.080	0.048	0.080	0.114	0.184	0.238	0.304
	(30,30)	0.054	0.064	0.086	0.108	0.214	0.240	0.408	0.460
	(45,45)	0.046	0.056	0.122	0.132	0.298	0.312	0.598	0.620
	(60,60)	0.054	0.058	0.136	0.144	0.410	0.418	0.726	0.738
	(75,75)	0.070	0.068	0.192	0.194	0.472	0.484	0.834	0.836
	(90,90)	0.042	0.046	0.178	0.180	0.556	0.574	0.890	0.898
	(120,120)	0.040	0.042	0.216	0.224	0.654	0.666	0.254	0.260
(30, 30)	(15, 15)	0.042	0.070	0.108	0.138	0.278	0.340	0.522	0.590
	(30,30)	0.070	0.084	0.164	0.182	0.520	0.526	0.830	0.846
	(45,45)	0.048	0.048	0.226	0.242	0.680	0.686	0.958	0.962
	(60,60)	0.034	0.042	0.320	0.332	0.796	0.804	0.988	0.990
	(75,75)	0.040	0.044	0.296	0.308	0.886	0.890	0.998	0.998
	(90,90)	0.054	0.056	0.388	0.394	0.934	0.936	1.000	1.000
	(120,120)	0.036	0.038	0.510	0.514	0.962	0.966	1.000	1.000
(50, 50)	(15, 15)	0.038	0.060	0.236	0.296	0.702	0.740	0.966	0.982
	(30,30)	0.050	0.062	0.408	0.432	0.940	0.944	1.000	1.000
	(45,45)	0.038	0.044	0.408	0.432	0.962	0.966	1.000	1.000
	(60,60)	0.058	0.060	0.716	0.726	1.000	1.000	1.000	1.000
	(75,75)	0.060	0.062	0.824	0.836	1.000	1.000	1.000	1.000
	(90,90)	0.034	0.034	0.874	0.880	1.000	1.000	1.000	1.000
	(120,120)	0.054	0.054	0.944	0.944	1.000	1.000	1.000	1.000
(70,70)	(15, 15)	0.058	0.084	0.428	0.476	0.938	0.946	1.000	1.000
	(30,30)	0.036	0.042	0.762	0.770	1.000	1.000	1.000	1.000
	(45,45)	0.056	0.060	0.882	0.884	1.000	1.000	1.000	1.000
	(60,60)	0.044	0.048	0.966	0.972	1.000	1.000	1.000	1.000
	(75,75)	0.050	0.052	0.986	0.986	1.000	1.000	1.000	1.000
	(90,90)	0.052	0.054	0.998	1.000	1.000	1.000	1.000	1.000
	(120,120)	0.048	0.050	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.13. Simulated type-I error and (size) power with $\boldsymbol{f(x,y) = 1/(1 + sinx * siny)}$, $\boldsymbol{\delta_0 = 0.25}$.

(n_1, n_2)	(m_1, m_2)	$\lambda = 0$		$\lambda = 0.005$		$\lambda = 0.010$		$\lambda = 0.015$	
		t	**EL**	**t**	**EL**	**t**	**EL**	**t**	**EL**
(20, 20)	(15, 15)	0.046	0.066	0.094	0.116	0.212	0.272	0.412	0.478
	(30,30)	0.042	0.044	0.144	0.162	0.362	0.398	0.734	0.758
	(45,45)	0.054	0.058	0.162	0.178	0.588	0.608	0.882	0.890
	(60,60)	0.048	0.056	0.242	0.248	0.688	0.694	0.966	0.970
	(75,75)	0.052	0.054	0.262	0.268	0.788	0.792	0.988	0.990
	(90,90)	0.048	0.046	0.330	0.334	0.882	0.886	0.998	0.998
	(120,120)	0.058	0.060	0.456	0.458	0.948	0.950	1.000	1.000
(30, 30)	(15, 15)	0.050	0.066	0.152	0.208	0.460	0.522	0.822	0.856
	(30,30)	0.054	0.062	0.234	0.270	0.796	0.812	0.982	0.982
	(45,45)	0.048	0.064	0.404	0.418	0.936	0.938	0.998	0.998
	(60,60)	0.048	0.046	0.498	0.512	0.976	0.976	1.000	1.000
	(75,75)	0.030	0.032	0.580	0.588	0.994	0.994	1.000	1.000
	(90,90)	0.046	0.048	0.668	0.676	0.998	0.998	1.000	1.000
	(120,120)	0.058	0.058	0.776	0.778	1.000	1.000	1.000	1.000
(50, 50)	(15, 15)	0.038	0.066	0.322	0.392	0.912	0.936	0.998	0.998
	(30,30)	0.056	0.066	0.654	0.666	0.996	0.998	1.000	1.000
	(45,45)	0.032	0.040	0.810	0.828	1.000	1.000	1.000	1.000
	(60,60)	0.054	0.054	0.912	0.914	1.000	1.000	1.000	1.000
	(75,75)	0.050	0.048	0.976	0.980	1.000	1.000	1.000	1.000
	(90,90)	0.038	0.038	0.988	0.990	1.000	1.000	1.000	1.000
	(120,120)	0.062	0.062	0.998	0.998	1.000	1.000	1.000	1.000
(70,70)	(15, 15)	0.042	0.064	0.576	0.618	0.994	0.994	1.000	1.000
	(30,30)	0.034	0.036	0.904	0.906	1.000	1.000	1.000	1.000
	(45,45)	0.040	0.042	0.984	0.984	1.000	1.000	1.000	1.000
	(60,60)	0.054	0.054	0.994	0.994	1.000	1.000	1.000	1.000
	(75,75)	0.042	0.042	1.000	1.000	1.000	1.000	1.000	1.000
	(90,90)	0.040	0.040	1.000	1.000	1.000	1.000	1.000	1.000
	(120,120)	0.052	0.054	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.14. Simulated type-I error and (size) power with $\boldsymbol{f(x,y) = 1/(1 + cosx * cosy)}$, $\boldsymbol{\delta_0 = 0.25}$.

(n_1, n_2)	(m_1, m_2)	$\lambda = 0$		$\lambda = 0.005$		$\lambda = 0.010$		$\lambda = 0.015$	
		t	**EL**	**t**	**EL**	**t**	**EL**	**t**	**EL**
(20, 20)	(15, 15)	0.042	0.068	0.108	0.144	0.196	0.280	0.480	0.554
	(30,30)	0.058	0.072	0.158	0.178	0.420	0.464	0.758	0.778
	(45,45)	0.046	0.054	0.196	0.210	0.586	0.620	0.886	0.900
	(60,60)	0.058	0.060	0.230	0.232	0.696	0.698	0.962	0.962
	(75,75)	0.056	0.064	0.278	0.288	0.820	0.824	0.980	0.982
	(90,90)	0.046	0.050	0.322	0.330	0.854	0.850	1.000	1.000
	(120,120)	0.048	0.050	0.406	0.410	0.960	0.960	1.000	1.000
(30, 30)	(15, 15)	0.054	0.070	0.160	0.204	0.474	0.538	0.806	0.852
	(30,30)	0.056	0.062	0.274	0.290	0.770	0.796	0.988	0.990
	(45,45)	0.042	0.048	0.388	0.396	0.922	0.964	0.998	0.998
	(60,60)	0.042	0.046	0.470	0.478	0.968	0.970	1.000	1.000
	(75,75)	0.044	0.044	0.592	0.594	0.992	0.992	1.000	1.000
	(90,90)	0.056	0.062	0.662	0.664	0.998	0.998	1.000	1.000
	(120,120)	0.046	0.046	0.800	0.806	1.000	1.000	1.000	1.000
(50, 50)	(15, 15)	0.050	0.066	0.386	0.436	0.886	0.920	1.000	1.000
	(30,30)	0.042	0.046	0.634	0.662	0.998	0.998	1.000	1.000
	(45,45)	0.054	0.062	0.812	0.826	1.000	1.000	1.000	1.000
	(60,60)	0.036	0.040	0.912	0.914	1.000	1.000	1.000	1.000
	(75,75)	0.042	0.044	0.984	0.984	1.000	1.000	1.000	1.000
	(90,90)	0.054	0.052	0.992	0.992	1.000	1.000	1.000	1.000
	(120,120)	0.056	0.058	0.998	0.998	1.000	1.000	1.000	1.000
(70,70)	(15, 15)	0.058	0.074	0.650	0.706	0.994	0.996	1.000	1.000
	(30,30)	0.048	0.060	0.882	0.894	1.000	1.000	1.000	1.000
	(45,45)	0.044	0.048	0.968	0.968	1.000	1.000	1.000	1.000
	(60,60)	0.036	0.038	0.994	0.994	1.000	1.000	1.000	1.000
	(75,75)	0.054	0.056	1.000	1.000	1.000	1.000	1.000	1.000
	(90,90)	0.038	0.038	1.000	1.000	1.000	1.000	1.000	1.000
	(120,120)	0.052	0.056	1.000	1.000	1.000	1.000	1.000	1.000

3.1.2. Real data application

We will now employ the proposed Empirical Likelihood method to analyze real-life data accessible for download from a publicly available database(`http://fcon_1000.projects.nitrc.org/indi/retro/cobre.html`). This dataset consists of raw anatomical and functional scans obtained from 146 subjects, including 72 patients diagnosed with schizophrenia and 74 healthy controls. After undergoing preprocessing steps conducted by Relin et al. (2013), the dataset was refined to include only 124 subjects (70 patients with schizophrenia and 54 healthy controls) for subsequent analysis.

Each subject's dataset initially contains 34,453 elements. Following a meticulous one-to-one mapping procedure, the data for each subject is transformed into a square symmetric matrix, featuring 263 rows and 263 columns. Consequently, 263 brain regions of interest are identified as nodes. The connectivity between these nodes is quantified through edge weights, representing the Fisher-transformed correlation between the functional MRI time series of the nodes, post-rank transformation (Relion and Kessler, 2019). The healthy group (54 networks) and unhealthy group (70 networks) have different sample sizes, but our proposed Empirical Likelihood Method is still able to deal with such scenario. We construct the EL test statistic with the number of quadrilateral to compute the p-value at the specified type-I error level $\boldsymbol{\alpha}$=0.05. The detailed results are presented in the Table 3.15. As observed from the table, the computed p-value using the Empirical Likelihood

Table 3.15. Empirical Likelihood Test and T-test Results at $\boldsymbol{\alpha}$=0.05

Method	**Test Statistic**	**P-Value**	**Significance**
Empirical Likelihood Test	4.271717	0.02978398	Significant
t-test	1.975263	0.05049616	Not Significant

Test is below the pre-specified significance level ($\boldsymbol{\alpha}$=0.05), In contrast, the p-value obtained from the t-test slightly exceeds the significant level, potentially leading to a different conclusion regarding the significance of differences between the healthy and unhealthy populations at the specified level of level $\boldsymbol{\alpha = 0.05}$. Notably, the p-value derived from the t-test is approximately twice as large as the p-value obtained from the proposed Empirical Likelihood Test. This indicates that our proposed EL test method exhibits greater power in detecting population disparities.

BIBLIOGRAPHY

Abbe, E. and Sandon, C. (2016). Achieving the KS threshold in the general stochastic block model with linearized acyclic belief propagation. In *Advances in Neural Information Processing Systems*, volume 29.

Aiello, W., Chung, F., and Lu, L. (2000). A random graph model for massive graphs. In *In Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing*, pages 171–180.

Albert, R. and Barabási, A. L. (2002). Statistical mechanics of complex networks. *Rev. Modern Phys*, pages 47–97.

Ali, T., Rito, T., Reinert, G., Sun, F., and Deane, C. M. (2014). Alignment-free protein interaction network comparison. *Bioinformatics*, 30:i430–i437.

Ali, W., Wegner, A. E., Gaunt, R. E., Deane, C. M., and Reinert, G. (2016). Comparison of large networks with sub-sampling strategies. *Scientific Reports*, 6:28955.

Barabási, A. L. and Oltvai, Z. (2004). Network biology: understanding the cell's functional organization. *Nat Rev Genet*, 5:101–113.

Bickel, P. J., Chen, A., and Levina, E. (2012). The method of moments and degree distributions for network models. *Ann. Statist*, 39:2280–2301.

Bollobas, B., Janson, S., and Riordan, O. (2007). The phase transition in inhomogeneousrandom graphs. *Random Structures & Algorithms*, 31:3–122.

Borsboom, D., Deserno, M. K., Rhemtulla, M., et al. (2021). Network analysis of multivariate data in psychological science. *Nat Rev Methods Primers*, 1:58.

Cao, R. and Keilegom, I. (2006). Empirical likelihood tests for two-sample problems via nonparametric density estimation. *Canadian Journal of Statistics*, 34:61–77.

Chen, S. X. and Van Keilegom, I. (2009). A review on empirical likelihood methods for regression. *TEST*, 18:415–447.

Drobyshevskiy, M. and Turdakov, D. (2019). Random graph modeling: A survey of the concepts. *Association for Computing Machinery*, 52(6).

Ghoshdastidar, D., Gutzeit, M., Carpentier, A., and von Luxburg, U. (2017). Two-sample tests for large random graphs using network statistics. *Proceedings of Machine Learning Research*, 65:1–24.

Ghoshdastidar, D., Gutzeit, M., Carpentier, A., and von Luxburg, U. (2020). Two-sample hypothesis testing for inhomogeneous random graphs. *The Annals of Statistics*, 48(4):2208–2229.

Ghoshdastidar, D. and Luxburg, U. V. (2018). Practical methods for graph two-sample testing. In *Neural Information Processing Systems*, pages 3019–3028.

Gurevich, G. and Vexler, A. (2011). A two-sample empirical likelihood ratio test based on samples entropy. *Statistics and Computing*, 21:657–670.

Lazar, N. A. (2021). A review of empirical likelihood. *Annual Review of Statistics and its Application*, 8:329–344.

Liu, Y., Zou, C., and Zhang, R. (2021). Empirical likelihood for the two-sample mean problem. *Statistics & Probability Letters*, 78(5):548–556.

Maugis, P.-A. G., Olhede, S. C., Priebe, C. E., and Wolfe, P. J. (2020). Testing for equivalence of network distribution using subgraph counts. *Journal of Computational and Graphical Statistics*, 29(3):455–465.

Newman, M. (2010). *Networks: An Introduction*. Oxford University Press, Oxford.

Nordman, D. and Lahiri, S. (2014). A review of empirical likelihood methods for time series. *Journal of Statistical Planning and Inference*, 155:1–18.

O'Malley, A. J. and Marsden, P. V. (2008). The analysis of social networks. *Health Serv Outcomes Res Methodol*, 8(4):222–269.

Relin, A., Jess, D., Kessler, D., Levina, E., and Taylor, S. F. (2013). Network classification with applications to brain connectomics. *Ann. Appl. Stat*, 13:1648–1677.

Relion, J. A. and Kessler, D. (2019). Graphclass: Network classification. *R package version 1.1.*

Rucinski, A. (1988). When are small subgraphs of a random graph normally distributed? *Probability Theory Related Fields*, 78(1):1–10.

Sanz-Arigita, E. J., Schoonheim, M. M., Damoiseaux, J. S., Rombouts, S. A., Maris, E., Barkhof, F., Scheltens, P., and Stam, C. J. (2010). Loss of 'small-world' networks in alzheimer's disease: graph analysis of fmri resting-state functional connectivity. *PLoS One*, 5(11):e13788.

Tang, M., Athreya, A., Sussman, D. L., Lyzinski, V., and Priebe, C. E. (2017). A nonparametric two-sample hypothesis testing problem for random graphs. *Bernoulli*, 23(3):1599–1630.

Wu, C. and Yan, Y. (2012). Empirical likelihood inference for two-sample problems. *Statistics and Its Interface*, 5.

Yuan, M. and Wen, Q. (2023). A practical two-sample test for weighted random graphs. *Journal of Applied Statistics*, 50(3):495–511.