8-2018

Power Comparisons of the Rician and Gaussian Random Fields Tests for Detecting Signal from Functional Magnetic Resonance Images

Hasni Idayu Binti Saidi

UNIVERSITY OF NORTHERN COLORADO

Greeley, Colorado

The Graduate School

POWER COMPARISONS OF THE RICIAN AND GAUSSIAN RANDOM FIELDS TESTS FOR DETECTING SIGNAL FROM FUNCTIONAL MAGNETIC RESONANCE IMAGES

A Dissertation Submitted in Partial Fulfillment
of the Requirement for the Degree of
Doctor of Philosophy

Hasni Idayu Binti Saidi

College of Education and Behavioral Sciences
Department of Applied Statistics and Research Methods

August 2018

This dissertation by: Hasni Idayu Binti Saidi

Entitled: *Power Comparisons of the Rician and Gaussian Random Fields Tests for Detecting Signal from Functional Magnetic Resonance Images*

has been approved as meeting the requirement for the Degree of Doctor of Philosophy in College of Education and Behavioral Sciences in Department of Applied Statistics and Research Methods

Accepted by the Doctoral Committee

Khalil Shafie Holighi, Ph.D., Research Advisor

Trent Lalonde, Ph.D., Committee Member

Jay Schaffer, Ph.D., Committee Member

Heng-Yu Ku, Ph.D., Faculty Representative

Date of Dissertation Defense

Accepted by the Graduate School

Linda L. Black, Ed.D.
Associate Provost and Dean
Graduate School and International Admissions

ABSTRACT

Saidi, Hasni Idayu Binti. *Power Comparisons of the Rician and Gaussian Random Fields Tests for Detecting Signal from Functional Magnetic Resonance Images.* Published Doctor of Philosophy dissertation, University of Northern Colorado, 2018.

The functional magnetic resonance imaging (fMRI) data are known to be complex valued. The real and imaginary components are assumed to be independently and normally distributed. After image reconstructions, these components are separated into two components, namely magnitude and phase. Usually, only the magnitude component is used in the analysis and it is assumed to be normally, or Gaussian, distributed. The statistical analysis of fMRI data using random field theory also assumed that the data are Gaussian distributed. However, the magnitude component is actually Rician distributed and no work has been found on the Rician random field. In this dissertation, Rician random field was defined, in general, and simulated in a two-dimensional image. A new test statistic to detect a signal from the functional magnetic resonance image, $R_{\max}$, which follows the Rician random field, was introduced. The power of $R_{\max}$ was calculated using Monte Carlo simulation, and compared to the Gaussian test statistic, $Z_{\max}$. The effects of factors known to influence the power of $R_{\max}$, namely amplitude, scale and location of the signal, were also studied. The amplitude was shown to be the most influencing factor on the power of $R_{\max}$, followed by the scale of the signal.

The location of the signal did not seem to affect the power of the $R_{\max}$. However, the power of $R_{\max}$ did not outperform the power of $Z_{\max}$. Future studies are required to provide more information on the properties and behaviors of $R_{\max}$.

ACKNOWLEDGEMENTS

In the name of Allah, the Most Merciful and the Most Loving. I am thankful for His Mercy that has helped me get through my dissertation and Ph.D. study. Without His guidance, this study would not have been possible. I would like to thank my parents, Abah and Mak, for being my supporters in every way possible, and for always believing in me even when I did not believe in myself. Words cannot express how lucky and thankful I am to have such supportive parents, who are always by my side, through thick and thin, from the beginning till the end. May Allah continue to shower His blessings on both of them. Thanks also goes to my sisters, Ijan and Kakak, and my brother, Abang Sham, for their words of encouragement and reminders that have helped me get back on track every time I lost my way.

I would like to express my heartfelt thank you to my advisor, Dr. Shafie, for being a father, a teacher, a guide, and a friend to me throughout my studies, especially during the hardest part for me, the dissertation writing phase. This dissertation would not be possible without his patience, guidance, confidence and kindness. He is the best teacher that I could possibly have to guide me through this dissertation. I am thankful also to my committee members, Dr. Lalonde, Dr. Jay, and Dr. Ku, for their valuable comments, suggestions, and input which have helped refine this dissertation.

I am very thankful to my friends who have been with me throughout this journey, especially Dr. Gadir and her husband Dr. Mohammad, Zabedah, Sam and her family, Dr. Amani, Samy, Maaly, Dr. Lu (Annie), Dr. Niloofar, and Dr. Yang (Katie). I would like to give special thanks to Emily, my counselor, and the group at the Counseling Center for helping me get through the final rough years at UNC. I would also like to express special thanks to Keyleigh, the department's secretary, for her endless help throughout my program. Without her, this journey would not have been smooth. Also, I would like to thank Dr. Hutchinson, Dr. Lahman, Dr. Raheem, and Dr. Pearson, whom I have had the chance to learn from. Many thanks also to my Malaysian-Colorado friends/family and friends at the Department of Applied Statistics and Research Methods for their support and encouragement that have helped me all this time.

Last but not the least, I would like to thank my sponsor, the Malaysian Ministry of Higher Education, and my employer, Universiti Putra Malaysia, for giving me the chance and financial support to study in the United States. I would like to give special thanks to the Department of Biomedical Science, especially the current Head of Department, Prof. Nazrul Hakim, former Head of Department, Dr. Latifah and Dr. Sabrina, colleagues and staff of the department for their endless support since I started this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Functional magnetic resonance imaging (fMRI) is a neuroimaging technique to study neural activity in the brain using an indirect measure called blood oxygenation level-dependent (BOLD). The fMRI can be used to identify location of activation due to certain activities, to determine functional networks that correspond to brain function, and to predict a person's disease or psychological state (Ashby, 2011; Lindquist, 2008). The fMRI uses the magnetic resonance imaging (MRI) scanner to scan and produce images of the brain at different time points while performing the tasks. The series of functional images produced contain the functional data (BOLD signals) and have high temporal resolution but low spatial resolution (Ashby, 2011).

An fMRI experiment can include single or multiple participants, and each participant is typically scanned across several sessions. Each session can consists of several runs, and each run consists of a series of whole brain images at the designated time points. Each three-dimensional whole brain image consists of units called voxels, which is equivalent to pixels if the image is a two-dimensional image. Each voxel consists of a time series of functional data at the specific scanned time (Lindquist, 2008).

The fMRI data are also noisy from various nuisance factors, which makes it harder to detect the signal (Lindquist, 2008). As such, the fMRI data have to undergo preprocessing steps to clean and prepare the data for statistical analysis. One of the steps is spatial smoothing. Spatial smoothing increases the signal-to-noise ratio (SNR). However, the fMRI image becomes blurry and the neighboring voxels become dependent on each other. The huge and autocorrelated nature of fMRI data poses challenges in the statistical analysis of fMRI data.

Functional Magnetic Resonance Imaging Statistical Analysis

The approach in modeling huge fMRI data is to model the BOLD signal at each voxel. This approach is called massive univariate approach. The idea is to fit a general linear model (GLM) to the BOLD signal at every voxel. As such, each voxel will have a statistic, for example a z-score or a t-statistic. These statistics are then constructed into an image called a statistical parametric map (SPM). The SPM displays the value of the statistic at each voxel using a range of colors (see Figure 1).

Since there are thousands of statistics, a problem known as multiple comparisons arises. Therefore, a strategy to avoid inflation of the Type I error rate while determining significance needs to be employed. Random field theory (RFT) is one of the methods that can be used to control the Type I error rate. One of the advantages of using random field theory is that it takes care of the spatial dependence of the voxels due to spatial smoothing.

The raw fMRI data are complex valued due to the nature of collecting the magnetic resonance signal (Lindquist, 2008). The real and imaginary components

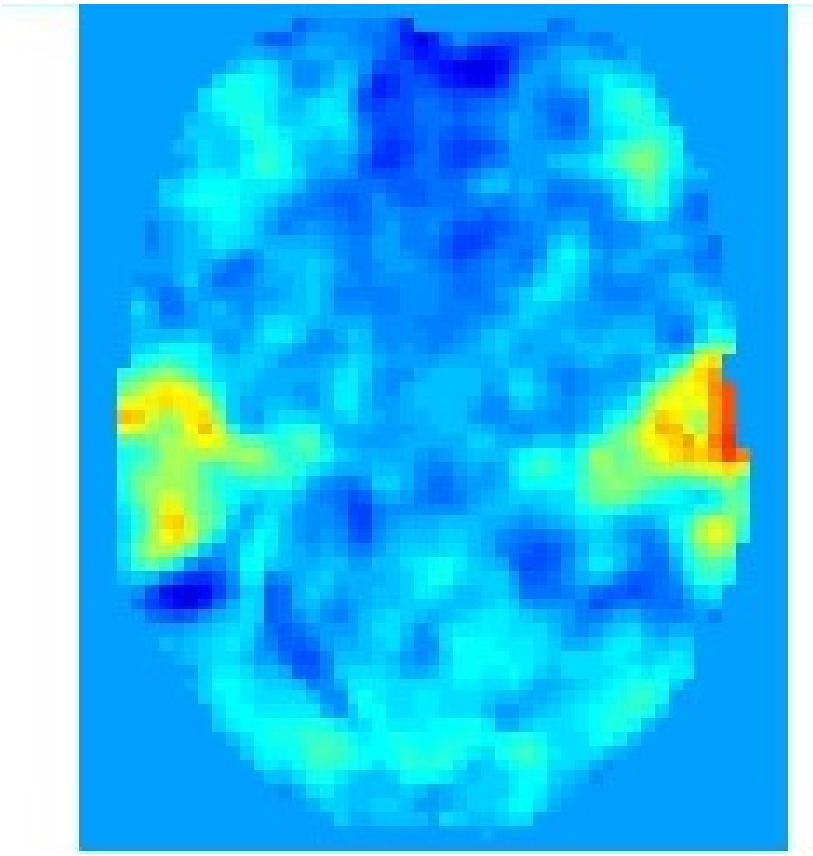

Figure 1. A Statistical Parametric Map (SPM) for a visual recognition experiment. The data used to generate the image were retrieved from https://openfmri.org/dataset/ds000105/.

are assumed to be independently and normally distributed. After image reconstruction, these components are separated into magnitude and phase components. The magnitude component of the signal is used in the analysis while the phase component is discarded. The magnitude component is not normally distributed. It follows the Rician distribution (Gudbjartsson & Patz, 1995). When the signal-to-noise ratio is high, the Rician distribution is approximately normally distributed. Statistical analysis of fMRI data usually assumes that the data will follow the normal distribution.

Problem Statement

The magnitude component of the fMRI data is the data that are usually used for statistical analysis. The data are assumed to be normally distributed. However, this assumption may not hold due to the commonly low signal-to-noise ratio in the data. Furthermore, the data are actually complex valued and follow Rician distribution (Gudbjartsson & Patz, 1995). As such, some researchers have proposed methods to analyze the complex-valued fMRI data. Rowe and Logan (2004), Rowe (2005a), Rowe (2005b), and Adrian, Maitra, and Rowe (2013) were among the studies that focused on the analysis of complex-valued fMRI data. In these studies, the researchers used another method to handle multiple comparisons problem. The method is known as false discovery rate (FDR).

Controlling the Type I error rate using random field theory such as Gaussian random field and χ^2 random field was established by the late Keith Worsley and his colleagues in various papers (Worsley, 1994, 2001; Worsley, Evans, Marrett, & Neelin, 1992; Worsley, Taylor, Tomaiuolo, & Lerch, 2004). However, no work has been found on Rician random field. Usually, Gaussian random field is assumed to be the distribution of the fMRI data. Since the actual data is Rician distributed, there is a need to define Rician random field.

Purpose of the Study

To the author's knowledge, there are no published studies on power comparisons of Gaussian and Rician random fields tests to detect signals from fMRI images. As such, the current study proposed a definition of a Rician random field and simulated data that follows Rician random field. Then, power comparisons

between the Gaussian and the Rician random fields tests under various conditions, namely location, amplitude and scale of the signals, were conducted in the current study.

Research Questions

Q1 How is Rician random field defined in general, and simulated in two-dimensional image?

Q2 How can the power of the Rician test statistic, $R_{\max}$, be calculated using Monte Carlo simulation?

Q3 How does the amplitude of the signal affect the power of the Rician test statistic?

Q4 How does the scale of the signal affect the power of the Rician test statistic?

Q5 How does the location of the signal affect the power of the Rician test statistic?

Q6 How is the power of the Rician test statistic as compared to the Gaussian test statistic under the conditions specified in previous research questions?

Delimitation of the Study

The current study was limited to certain conditions. First, the simulated error fields were assumed to be Rician random field, which was defined in the current study. Second, only one signal was assumed to exist in the data. Third, the simulation only used two-dimensional images. Finally, the critical value and the power for the Rician random field test statistic were determined from the Monte Carlo simulation method. As such, the results from the current study needed to be used with caution and might not be generalized to all conditions in fMRI studies.

Definition of Terms

Blood oxygenation level-dependent (BOLD) contrast. A measure of the ratio of oxygenated hemoglobin to deoxygenated hemoglobin.

Euler characteristic (EC). A topological measure of the statistical parametric map after thresholding.

Excursion set. A set of points where a random field exceeds a fixed threshold value.

Finite dimensional distributions. A collection of distribution functions for a random field.

Full width at half maximum (FWHM). A measure of the width of the smoothing kernel when the maximum height of the kernel is equal to half.

Functional magnetic resonance imaging (fMRI). A neuroimaging technique used to study brain functions.

Gaussian kernel. A smoothing kernel used to smooth data.

Gaussian random field. A type of random field where all of the finite dimensional distributions are multivariate joint Gaussian (normal) distributions.

Pixel. A measure of unit in a two-dimensional image.

Random fields. A collection of random variables defined over a subset of N-dimensional Euclidean space.

Rician random field. A type of random field where all of the finite dimensional distributions are multivariate joint Rician distribution.

Spatial smoothing. A procedure that replaces the BOLD signal in a voxel with the average of BOLD signals from neighboring voxels.

Statistical parametric map (SPM). A color-coded brain map that represents the numerical value of the test statistics at each voxel.

Voxel. A measure of unit in a three-dimensional image.

Dissertation Structure

This dissertation is organized in chapters. Chapter I is an introduction to the functional magnetic resonance imaging and its statistical analysis. It also includes problem statement, purpose of the study, research questions, delimitation of the study, and the definition of terms. Chapter II is the review of the literature, where the author studied the fMRI and its statistical data analysis, the random field theory including Gaussian, Chi-square, and Rician random fields, the modeling of the fMRI using random fields, the modeling of the complex valued fMRI data, and the power calculation and comparison in fMRI. Chapter III describes the methodology used in the study, which includes data generation, the test statistics, the power calculation and the schemes of the parameters. Chapter IV presents the results of the study. Chapter V includes brief discussions on the results of the study, conclusions and suggestions for future research. Finally, the list of references and the R code for simulation are documented.

CHAPTER II

REVIEW OF LITERATURE

This chapter begins with an introduction to functional magnetic resonance imaging (fMRI). It is followed by brief information on fMRI preprocessing, and then on the fMRI statistical analysis. The next section discusses the random field theory, including Gaussian random fields and Rician random fields. The following section explores on the topic of modeling fMRI using random field theory, and modeling fMRI as complex valued data. Afterwards, a section on power calculation and comparison in fMRI is presented. Finally the chapter is concluded with a summary of the literature review.

Functional Magnetic Resonance Imaging

Functional magnetic resonance imaging (fMRI) is a neuroimaging technique used to study neural activity in the brain. The most common use of fMRI is for localization, which is to identify location of activation due to certain tasks or activities. FMRI is also used to determine functional networks that correspond to brain function using connectivity analysis, and to predict a person's disease or psychological state (Ashby, 2011; Lindquist, 2008). The use of fMRI for localization of activation is the focus of this dissertation.

FMRI uses the blood oxygenation level-dependent (BOLD) signal proposed by Ogawa, Lee, Kay, and Tank (1990) as an indirect measure of neural activity.

BOLD signal is a measure of blood flow changes in the brain. FMRI is a noninvasive procedure and does not involve harmful materials. Other neuroimaging techniques namely computed tomography (CT) uses x-rays while positron emission tomography (PET) uses radioactive drugs to scan the brain, which pose some health risks to the participants (Ashby, 2011). As such, fMRI is becoming a popular technique in the neuroimaging field. Additionally, fMRI provides images with high temporal and spatial resolutions as compared to other techniques (Ashby, 2011; Bullmore & Suckling, 2001; Matthews & Jezzard, 2004).

In a typical fMRI study, each participant will alternately rest and perform tasks, such as tapping fingers and viewing pictures, while lying down inside the magnetic resonance imaging (MRI) scanner for a duration of time. The MRI scanner will scan each participant's brain and produce a structural brain image and a series of functional brain images. The functional images are images that contain the BOLD signal information (Ashby, 2011). Figure 2 shows an example of functional images from a visual recognition experiment. It can be seen that functional images have low spatial resolution but high temporal resolution.

The functional brain images for one whole brain are usually collected in slices (Lindquist, 2008). Each slice is divided into cubes called voxels (volume elements) for a three-dimensional image, or pixels (picture elements) for a two-dimensional image. Each voxel (or pixel) contains a BOLD signal at a specific repetition time (TR). Repetition time is the time between two consecutive whole brain scans (Ashby, 2011). As such, one whole brain image contains thousands of voxels. For a single run of an experiment, there will be a time series of BOLD signals at each of

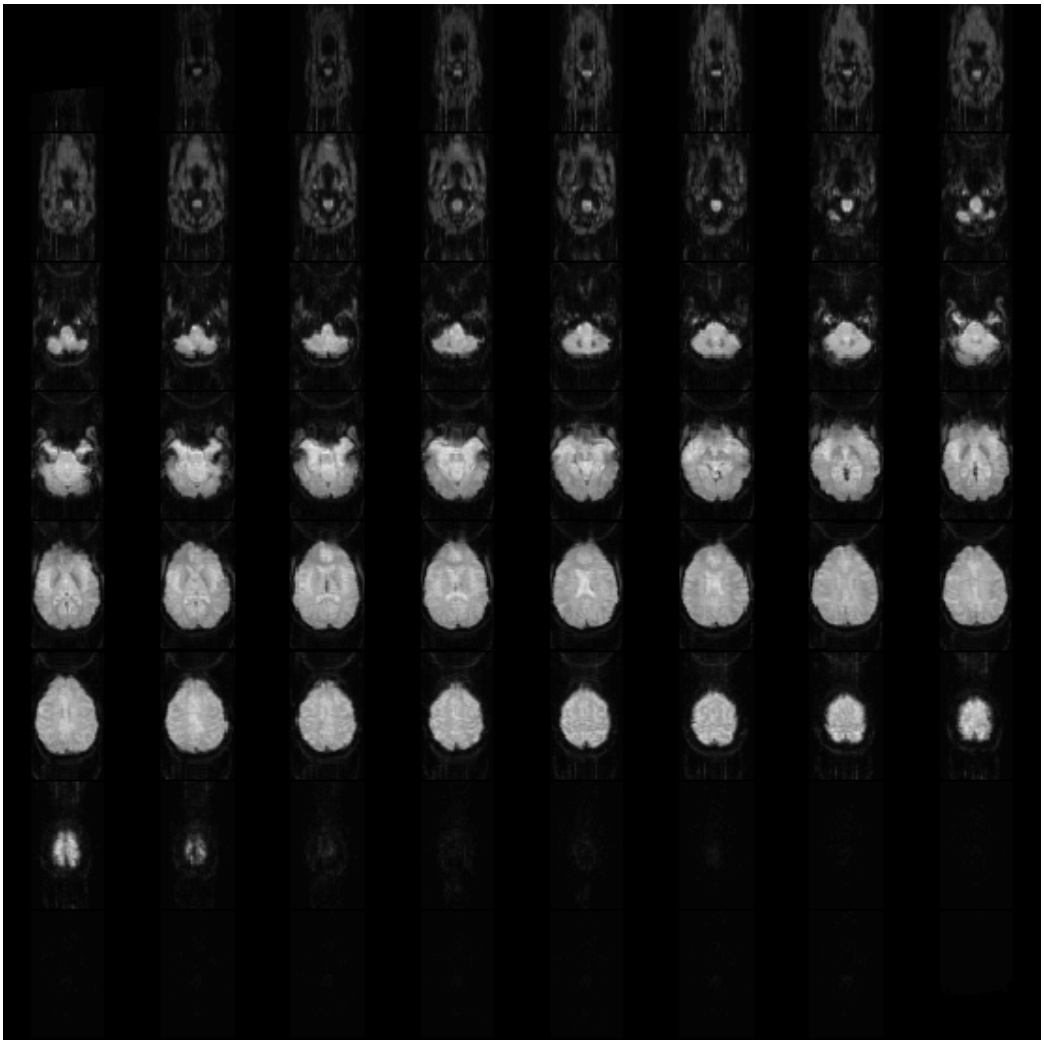

Figure 2. A series of functional brain images from one subject over one run for a visual recognition experiment. The data used to generate the image were retrieved from https://openfmri.org/dataset/ds000105/.

the voxels for each of the subjects. These huge amounts of autocorrelated data impose difficulties in data analysis.

Functional Magnetic Resonance Imaging Preprocessing

Preprocessing involves some procedures conducted prior to statistical analysis to remove artifacts and reduce noise that are inherited in the fMRI data during data collection. It is an important step to prepare the data set for statistical

analysis. Common preprocessing procedures are slice timing correction, head motion correction, coregistration, spatial normalization, and spatial smoothing (Ashby, 2011).

Slice timing correction is a procedure to correct slight time differences due to collecting brain images in slices. This procedure is crucial in fMRI study that uses event-related design, where the rests and the tasks are alternated quickly. In an fMRI study that uses block design, where the rests and tasks are alternated in a block duration, the slice timing correction is not as crucial since the participants are doing the same thing in each block for a certain amount of time (Ashby, 2011).

Head motion correction is conducted to correct the head movement effect on the fMRI data. As an fMRI experiment could be running for a certain amount of time, the participants are bound to move their heads while lying down in the MRI scanner. A slight movement can greatly affect the location of the BOLD signals. Thus, this is a critical step in preprocessing (Poldrack, Mumford, & Nichols, 2011).

Coregistration means alignment of the functional and structural brain images. It is conducted so that we can localize the BOLD signal to its anatomical location. Functional brain images is known to have poor spatial resolution (Ashby, 2011). As such, even if we found significant signals in the voxel, we could not be certain of its anatomical location. Therefore, coregistration needs to be conducted on these images.

The shape and size of human brains vary across individuals. In order to make comparisons between individuals, the brain images should be standardized. Spatial normalization is a preprocessing procedure to 'standardize' every brain

image to templates. The most common templates are known as the MNI templates, which were developed at the Montreal Neurological Institute (Poldrack et al., 2011).

Smoothing means averaging BOLD signal in a voxel using signals from neighboring voxels (Ashby, 2011). It is done by convolving the BOLD signals with a Gaussian kernel. This consequently blurs the image and improves the signal-to-noise ratio (SNR). To determine the neighboring voxels, a full width at half maximum (FWHM) is used. It is a measure of the width of the Gaussian kernel at half maximum (Ashby, 2011). The SNR is maximized when the size and shape of the kernel is the same as the BOLD signal. This is known as the matched filter theorem of signal processing.

Functional Magnetic Resonance Imaging Statistical Data Analysis

Statistical analysis of fMRI data is conducted after preprocessing to determine significance of signals in the brain. The most common method is using a general linear model (GLM) to test the hypothesis of no activation in the whole brain against the hypothesis of activation in any location of the brain. The fMRI data can be modeled as

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

where $\mathbf{Y}$ is an $(IJ \times K)$ matrix of BOLD signals, $\mathbf{X}$ is the design matrix, $\boldsymbol{\beta}$ is the vector of parameters and $\boldsymbol{\epsilon}$ is a matrix of errors. The errors are assumed to be distributed as Gaussian with mean zero and variance-covariance matrix $\boldsymbol{\Sigma}$. Each element in $\mathbf{Y}$, say $Y_{ij,k}$, represents the BOLD signal for subject i at time j and voxel k, where $i = 1, 2, ..., I$, $j = 1, 2, ..., J$ and $k = 1, 2, ..., K$ (Ashby, 2011).

The fMRI data are massive and autocorrelated across times and voxels. However, it is a common practice in fMRI to analyze the data separately at each voxel by assuming that the voxels are independent. This is known as the massive univariate approach. It is also common for the analysis to be conducted in two levels, where the first level deals with each subject and the second level deals with groups of subjects (Lindquist, 2008), as shown in Figure 3. Nevertheless, multivariate approaches have also been proposed, such as multivariate analysis of covariance (MANCOVA), principal components analysis (PCA) and independent component analysis (ICA) (Ashby, 2011; Friston, Frith, Frackowiak, & Turner, 1995; McKeown & Sejnowski, 1998).

Using the massive univariate approach, we calculate a test statistic at each voxel, U_k. All of the calculated test statistics can be presented visually through the statistical parametric map (SPM). The map is color-coded according to the numerical value of the test statistics (see Figure 1 in the previous chapter). To determine statistical significance, we look for voxels that have large test statistics and try to find a height threshold, u, that will control a family of Type I error rates known as family-wise error rate (FWER) (Brett, Penny, & Kiebel, 2004). This is known as the multiple comparisons problem. The two known methods for controlling the FWER in fMRI are Bonferroni correction and random field theory (RFT). Another method to deal with the multiple comparisons problem is by controlling the false positives rate. This method is known as false discovery rate (FDR).

Bonferroni correction uses probability rules to calculate the P-value. For example, say that we have 10,000 test statistics to compare and we want to have a

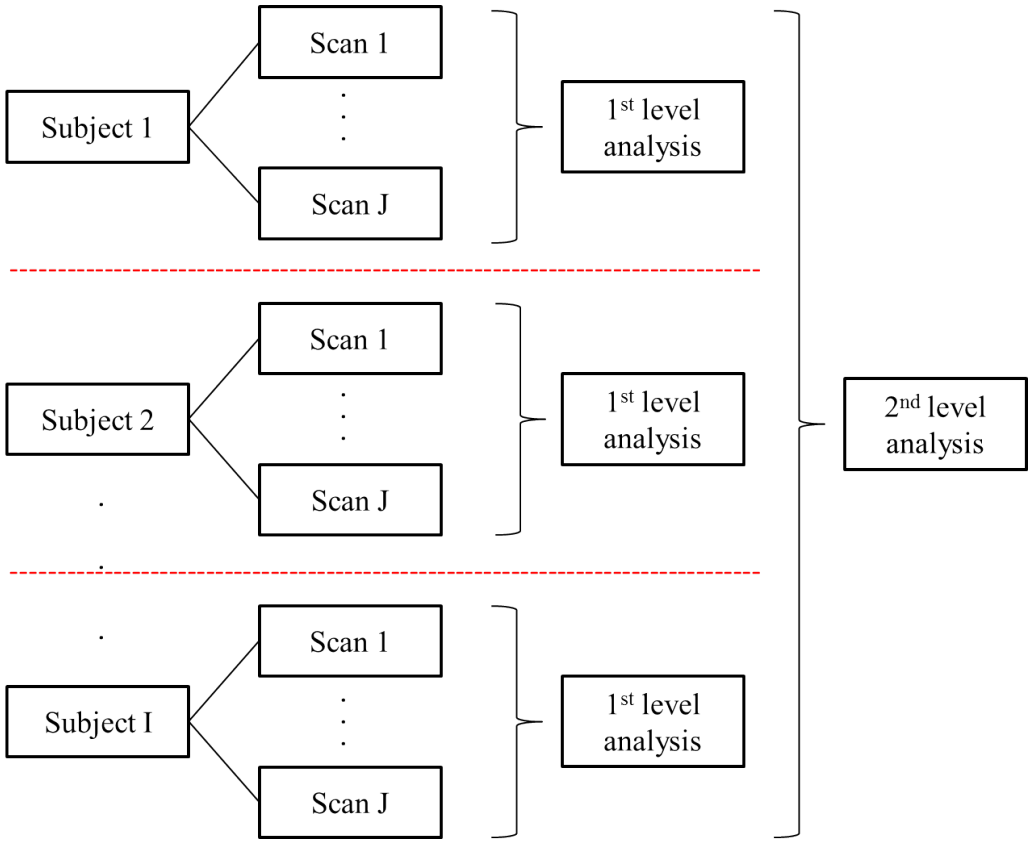

Figure 3. A common statistical analysis approach for fMRI data. The analysis is conducted in two levels; within each subject and between groups of subjects.

FWER of 5%. Therefore at every voxel, the test statistic is significant if its P-value is less than $0.05/10000 = 0.000005$. This is usually called corrected P-value. However, Bonferroni correction assumes that each of the test statistics is independent of each other. This is not the case in fMRI since spatial correlation is embedded in the data due to spatial smoothing during preprocessing. As such, Bonferroni correction is a very conservative method of multiple comparisons for fMRI studies.

Random field theory (RFT) is an alternative to Bonferroni correction. It uses the smoothness and the number of resolution elements (resels) in the statistical parametric map to approximately determine the P-value, which happens to be the

expected value of Euler characteristic (EC) (Worsley, Evans, Marrett, & Neelin, 1992). We will discuss this in detail in the next section.

Another multiple comparisons procedure that is popular in fMRI is called false discovery rate (FDR). FDR controls the expected value of the number of falsely rejected tests (false positives) over the total number of rejected tests (Benjamini & Hochberg, 1995). The use of FDR in fMRI studies was introduced by Genovese, Lazar, and Nichols (2002). They defined FDR as the "proportion of declared-active voxels which are false positives" (Genovese, Lazar, & Nichols, 2002, p.7). The FDR is zero if we fail to declare any activation in any voxels.

Random Field Theory

A random field or a stochastic process is a collection of random variables, U, defined over a parameter space, T, where T is a subset of N-dimensional Euclidean space, $\Re^N$ (Adler, 1981). Any point in T is denoted by $\mathbf{t}$. As such, an N-dimensional random field can be denoted as $U(\mathbf{t}), \mathbf{t} \in T$, where T is a subset of $\Re^N$, $N \geq 1$. When $N = 1$, the random field is defined over the real line, $\Re$. For example, a random field defined over time is called stochastic process. When $N = 2$, the random field is defined over a surface and is known as random surface. When $N = 3$, the random field is defined over a volume and commonly known as spatial process (Worsley, 2006).

Random field theory was brought to popularity by Adler (1981) through his book, which is a continuity from his doctoral dissertation. However, random field theory was first seen being applied in ocean waves study by Longuet-Higgins (1952). Since then, random field theory has been applied in various fields including

neuroimaging (Friston, Jezzard, & Turner, 1994; Worsley, 2001, 2007), earth sciences (Christakos, 1992), and astrophysics (Bertschinger, 2001).

A random field U has a few unique characteristics. As mentioned, random field is a collection of random variables. Each of the random variables corresponds to a distribution function (Adler, 1981). The collection of the distribution functions is called the finite dimensional distributions for the random field U. The finite dimensional distributions determine the distributional properties of the random field U and are summarized usually by the mean function

$$\mu(\mathbf{t}) = E[U(\mathbf{t})],$$

and the covariance function

$$Q(\mathbf{s}, \mathbf{t}) = Cov[U(\mathbf{s}), U(\mathbf{t})] = E([U(\mathbf{s}) - \mu(\mathbf{s})]'[U(\mathbf{t}) - \mu(\mathbf{t})]).$$

Besides that, another characteristic of random field is homogeneity, or commonly known in stochastic process as stationarity. A random field is strongly homogeneous, or strongly stationary, when the joint distributions do not change with any shift in points $\mathbf{t}$ (Adler, 1981). A random field is weak or second-order stationary when the random field has a constant mean function and the covariance function only depends on the difference between two points $(\mathbf{s} - \mathbf{t})$. If the random field is stationary and the covariance function depends only on the Euclidean distance between two points $\|\mathbf{s} - \mathbf{t}\|$, the random field is isotropic.

There are several specific types of random fields that have been recognized by researchers. Only those which are relevant to this dissertation are discussed here. Random fields are recognized by their finite dimensional distributions. In general, a random field that is commonly used to deal with fMRI data is either Gaussian random field or Gaussian-related random field.

Gaussian Random Fields

A specific type of random field is Gaussian random field. By definition, an N-dimensional random field is called a Gaussian field when the finite dimensional distributions are all multivariate joint Gaussian distributions. Gaussian field is completely determined by its mean function and its covariance function. This means that once the mean and the covariance functions are specified, all the finite dimensional distributions of a real-valued Gaussian field is determined (Adler, 1981). For example, let

$$Y(\mathbf{t}) = \mu(\mathbf{t}) + e(\mathbf{t}),$$

where $e(\mathbf{t})$ is a Gaussian field with mean zero and covariance function

$$Cov[e(\mathbf{s}), e(\mathbf{t})] = \exp\left(-\frac{\|\mathbf{s}-\mathbf{t}\|^2}{2\sigma^2}\right).$$

Since $e(\mathbf{t})$ has constant mean and its covariance function depends only on the Euclidean distance between two points, $\|\mathbf{s}-\mathbf{t}\|$, $e(\mathbf{t})$ is a stationary and isotropic Gaussian field.

χ^2 Random Fields

χ^2 random field, or simply χ^2 field, is an example of Gaussian-related random fields. Let $X_1(\mathbf{t}), X_2(\mathbf{t}), ..., X_n(\mathbf{t}), \mathbf{t} \in \Re^N$, each be a stationary and independent Gaussian random field with mean zero and covariance function, $R(\mathbf{t})$, with variance $\sigma^2 = 1$. Now, let

$$Y(\mathbf{t}) = [X_1(\mathbf{t})]^2 + [X_2(\mathbf{t})]^2 + ... + [X_n(\mathbf{t})]^2.$$

For each fixed $\mathbf{t} \in \Re^N$, $Y(\mathbf{t})$ has a χ^2 distribution with n degrees of freedom.

Also, for each $\mathbf{t} \in \Re^N$, the mean function is

$$\mu(\mathbf{t}) = n,$$

and the covariance function is

$$Q^*(\mathbf{s}, \mathbf{t}) = Cov[Y(\mathbf{s}), Y(\mathbf{t})] = 2nR^2(\mathbf{s}, \mathbf{t}),$$

where $Q(\mathbf{s}, \mathbf{t})$ is the common covariance function of the X_i (Adler, 1981).

Rician Random Fields

The Rice distribution was derived by Rice (1944, 1945, 1948) to describe a sine wave plus random noise. Sometimes, Rice distribution is also known as Rician or Ricean distribution. Let X and Y be two independent random variables, each follows a normal distribution with respective mean, μ_X and μ_Y, and same variance,

σ^2. Now let R be defined as

$$R = \sqrt{X^2 + Y^2} \geq 0.$$

Thus, R follows a Rician distribution and the probability density function (PDF) is given by

$$f_R(r) = \frac{re^{-(r^2+\mu^2)/2\sigma^2}}{2\pi\sigma^2} I_0\left(\frac{r\mu}{\sigma^2}\right), \qquad r \geq 0,$$

where

$$I_0(x) = \frac{1}{\pi}\int_0^\pi e^{x\cos\phi} d\phi, \qquad -\infty < x < \infty$$

is the modified Bessel function of the first kind and zeroth order. Also the parameter μ, where

$$\mu = \sqrt{\mu_X^2 + \mu_Y^2},$$

is the distance between the center of the bivariate normal distribution (μ_X, μ_Y) and the origin $(0, 0)$. For the case where both X and Y have mean zero, R follows a Rayleigh distribution (Kobayashi, Mark, & Turin, 2011).

We can extend this definition to random field. Let $X(\mathbf{t})$ and $Y(\mathbf{t})$ be two independent Gaussian fields, with respective mean functions, $\mu_X(\mathbf{t})$ and $\mu_Y(\mathbf{t})$, and same covariance functions, $Q(\mathbf{s}, \mathbf{t})$. Thus for each $\mathbf{t} \in \Re^N$,

$$R(\mathbf{t}) = \sqrt{X^2(\mathbf{t}) + Y^2(\mathbf{t})} \geq 0 \tag{1}$$

is a Rician field. When both $X(\mathbf{t})$ and $Y(\mathbf{t})$ have mean zero, $R(\mathbf{t})$ is a Rayleigh field.

Other Random Fields

Other Gaussian-related fields that have been identified are t and F fields (Worsley, 1994). A random field which the marginal distribution for fixed $\mathbf{t} \in \Re^N$ follows a t-distribution with degrees of freedom m is called t field and can be denoted by $T_m(\mathbf{t})$. A random field which the marginal distribution for fixed $\mathbf{t} \in \Re^N$ follows an F-distribution with degrees of freedom m and n, $m + n > N$, is defined as F field, $F(\mathbf{t})$.

Modeling Functional Magnetic Resonance Imaging using Random Fields

General linear model is the most commonly used method in fMRI modeling to determine significant local activation in the brain. Most of the time, we do not know the location of the task-related activation. Therefore, we are searching for task-related activation in the whole brain. The statistical model at each voxel $\mathbf{t}$ in a D-dimensional brain region C can be written as

$$\mathbf{Y}(\mathbf{t}) = \mathbf{X}\boldsymbol{\beta}(\mathbf{t}) + \boldsymbol{\epsilon}(\mathbf{t}),$$

where $\mathbf{Y}(\mathbf{t})$ is a vector of BOLD signals, $\mathbf{X}$ is the design matrix, $\boldsymbol{\beta}(\mathbf{t})$ is the vector of regression coefficients, and $\boldsymbol{\epsilon}(\mathbf{t})$ is a vector of errors. The columns of $\mathbf{X}$ are the predictor variables while the $\boldsymbol{\beta}(\mathbf{t})$ represents the average changes in $\mathbf{Y}(\mathbf{t})$ with a unit change in $\mathbf{X}$. Also, we assume that the errors, $\boldsymbol{\epsilon}(\mathbf{t})$, are independent and identically distributed isotropic Gaussian field with zero mean and variance σ^2 (Cao & Worsley, 2001).

At each voxel $\mathbf{t}$, we calculate a statistic, say $U(\mathbf{t})$. Since we calculate a statistic at each voxel, we have thousands of statistics in an SPM. Thus, the question arises on which statistics should be chosen as the test statistic to assess the significant effects of the predictors. Worsley et al. (1992) proposed the maximum of the statistics across all $\mathbf{t}$ voxels in the search region, U_{max}, as the test statistic so that we can threshold the statistical parametric map at height threshold u. The distribution of this random variable depends on several criteria, namely the number of responses, the knowledge on standard deviation, and the number of contrasts being tested (Cao & Worsley, 2001).

Now that we have established the test statistic and its distribution, we need to find the height threshold u such that the P-value takes care of the multiple comparisons problem. The P-value can be approximated by the expected value of Euler characteristic of the excursion set of the random field U above u, $E[\chi(A_u)]$ (Adler, 2000). Excursion set, A_u, is a set of voxels $\mathbf{t}$ where the random field U is above the threshold u (Cao & Worsley, 2001). Euler characteristic (EC) is a topological measure of the statistical parametric map after thresholding (Brett et al., 2004). The P-value for a smooth statistical parametric map is given by

$$P(U_{\max} \geq u) \approx E[\chi(A_u)] = \sum_{d=0}^{D} \mathrm{Resel}_d(C)\mathrm{EC}_d(u),$$

where D is the dimension of the search region C, $\mathrm{Resel}_d(C)$ is the number of d-dimensional resels in the search region C, and $\mathrm{EC}_d(u)$ is the d-dimensional EC

density of the statistical parametric map (Worsley, 2004; Worsley, Taylor, Tomaiuolo, & Lerch, 2004).

Resel measures the resolution elements of a smoothed statistical parametric map (Brett et al., 2004). In a three-dimensional SPM, resel is "a block of voxels with dimensions equal to the effective FWHM" of the SPM (Worsley et al., 1992, p. 901). In a two-dimensional SPM, resel is a block of pixels. For example, the number of resel in a three-dimensional search region C when $d = 3$ can be calculated as

$$\text{Resel}_3 = \frac{\text{V}}{\text{FWHM}_\text{x} \times \text{FWHM}_\text{y} \times \text{FWHM}_\text{z}},$$

where V is the volume of the search region C, and $\text{FWHM}_\text{x}, \text{FWHM}_\text{y}$, and FWHM_z are the effective full width at half maximum of the Gaussian kernel in direction x, y, and z, respectively. The Gaussian kernel is the kernel that is used to smooth the statistical parametric map (Worsley, 2004).

The FWMH has relationship with the variance matrix $\mathbf{\Lambda}$ (Worsley et al., 1992). Under the assumption that the SPM has no activation, the smoothed SPM can be generated by convolving a white noise Gaussian random field with a Gaussian kernel in the form of $k(\mathbf{x}) \propto \exp[-\mathbf{x}^T\mathbf{\Sigma}^{-1}\mathbf{x}/2]$. This gives

$$\mathbf{\Lambda} = \mathbf{\Sigma}^{-1}/2 = \begin{pmatrix} 1/\text{FWHM}_x^2 & 0 & 0 \\ 0 & 1/\text{FWHM}_y^2 & \\ 0 & 0 & 1/\text{FWHM}_z^2 \end{pmatrix} (4 \ln\ 2).$$

The EC density depends on the type of random field for the SPM and the threshold u. As an example, the EC density for a Gaussian random field when $d = 3$

given by Worsley et al. (1996) is

$$\mathrm{EC}_3(x) = \frac{(4\log_e 2)^{3/2}}{(2\pi)^2} e^{-x^2/2}(x^2-1).$$

A list of $\mathrm{Resel}_d(C)$ and the $\mathrm{EC}_d(x)$ for $d = 0, 1, 2, 3$ for Gaussian, χ^2 and F fields can be found in Worsley et al. (1996).

Modeling Complex Valued Functional Magnetic Resonance Imaging Data

FMRI raw data are complex valued due to data acquisition using the MRI scanner. This means that fMRI data consist of real and imaginary components. Moreover, the raw data are collected in a frequency domain known as k-space. As such, inverse Fourier transformation (IFT) has to be applied to the raw data in k-space to reconstruct an fMRI image in image space, where the data are analyzed. The IFT is a linear transformation. As such, the reconstructed fMRI image will also be complex valued. These processes are repeated until a sequence of brain images are collected (Lindquist, 2008).

Rowe and Logan (2004) were among the first to propose a complex model for fMRI data analysis. They believed that since the reconstructed fMRI images contain complex valued data, both the real and imaginary information should be modeled. They defined the complex valued image measured over time j at a particular voxel as

$$\rho_{mj} = [\mathbf{x}_\mathbf{j}'\boldsymbol{\beta}\cos\theta_j + \eta_{Re,j}] + i[\mathbf{x}_\mathbf{j}'\boldsymbol{\beta}\sin\theta_j + \eta_{Im,j}], \tag{2}$$

where $i^2 = -1$, $\mathbf{x_j'}\boldsymbol{\beta} = \beta_0 + \beta_1 x_{1j} + ... + \beta_q x_{qj}$, $\mathbf{x_j'}$ is the jth row, $j = 1, 2, ..., J$, of a $J \times (q+1)$ design matrix $\mathbf{X}$, θ_j represents the phase imperfection at time j, $(\eta_{Re,j}, \eta_{Im,j})' \sim N(\mathbf{0}, \boldsymbol{\Sigma})$ and $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}_2$ (Adrian, Maitra, & Rowe, 2013; Rowe & Logan, 2004).

The model in Equation 2 can also be written in a vector form as

$$\begin{pmatrix} y_{Re,j} \\ \\ y_{Im,j} \end{pmatrix} = \begin{pmatrix} \mathbf{x_j'}\boldsymbol{\beta}\cos\theta_j \\ \\ \mathbf{x_j'}\boldsymbol{\beta}\sin\theta_j \end{pmatrix} + \begin{pmatrix} \eta_{Re,j} \\ \\ \eta_{Im,j} \end{pmatrix}, \qquad j = 1, ..., J,$$

and further into

$$\mathbf{y} = \begin{pmatrix} \mathbf{X} & \mathbf{0} \\ \\ \mathbf{0} & \mathbf{X} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}\cos\theta_j \\ \\ \boldsymbol{\beta}\sin\theta_j \end{pmatrix} + \boldsymbol{\eta}, \tag{3}$$

where the observed vector of data $\mathbf{y} = (\mathbf{y_{Re}'}, \mathbf{y_{Im}'})$ is the vector of observed real values stacked on the vector of observed imaginary values, and the vector of errors $\boldsymbol{\eta} = (\boldsymbol{\eta_{Re}'}, \boldsymbol{\eta_{Im}'}) \sim N(\mathbf{0}, \boldsymbol{\Sigma} \otimes \boldsymbol{\Phi})$ is the vector of observed real errors stacked on the vector of imaginary errors. It is assumed that $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I_2}$ and $\boldsymbol{\Phi} = \mathbf{I_n}$ (Rowe & Logan, 2004). The $\otimes$ is called the Kronecker product operation, which means every element of $\boldsymbol{\Sigma}$ is multiplied to the entire $\boldsymbol{\Phi}$ (Rowe, 2005a).

Most of the fMRI data analyses only model the magnitude of the complex valued image, $|\rho_{mj}|$, defined as

$$r_j = |\rho_{mj}| = \sqrt{(\mathbf{x_j'}\boldsymbol{\beta}\cos\theta_j + \eta_{Re,j})^2 + (\mathbf{x_j'}\boldsymbol{\beta}\sin\theta_j + \eta_{Im,j})^2},$$

while the phase component, defined as

$$\phi_j = \tan^{-1}\left[\frac{y_{Im,j}}{y_{Re,j}}\right] = \tan^{-1}\left[\frac{\mathbf{x_j'}\boldsymbol{\beta}\sin\theta_j + \eta_{Im,j}}{\mathbf{x_j'}\boldsymbol{\beta}\cos\theta_j + \eta_{Re,j}}\right],$$

is discarded because the phase component has not been considered to contain relevant signal information (Lindquist, 2008; Rowe & Logan, 2004). The magnitude component, r_j, is known to follow a Rician distribution, $r_j \sim Rician(\mathbf{x_j'}\boldsymbol{\beta}, \sigma)$, where $\mathbf{x_j'}\boldsymbol{\beta}$ represents the signal and σ represents the noise (Adrian, Maitra, & Rowe, 2013). The PDF for the Rician distribution is given by

$$f(r_j|\mathbf{x_j}, \boldsymbol{\beta}, \sigma^2) = \frac{r_j}{\sigma^2}\exp\left(-\frac{1}{2\sigma^2}\left[r_j^2 + (\mathbf{x_j'}\boldsymbol{\beta})^2\right]\right) \times \int_{-\pi}^{\pi}\frac{1}{2\pi}\exp\left(\frac{r_j\mathbf{x_j'}\boldsymbol{\beta}}{\sigma^2}\cos(\phi_j - \theta_j)\right)d\phi_j \tag{4}$$

for $r_j \geq 0, \mathbf{x_j'}\boldsymbol{\beta} \geq 0, \sigma^2 > 0$ (Adrian, Maitra, & Rowe, 2013; Rowe, 2005b). The integral factor is the modified zeroth order Bessel function of the first kind and is usually denoted as $I_O(r_j\mathbf{x_j'}\boldsymbol{\beta}/\sigma^2)$ (Gudbjartsson & Patz, 1995). The log-likelihood function of the Rician distribution is given as

$$LL_R(\boldsymbol{\beta}, \sigma^2|\mathbf{r}) = \sum_{j=1}^{J}\left[\log(r_j/\sigma^2) - \frac{r_j^2 + (\mathbf{x_j'}\boldsymbol{\beta})^2}{2\sigma^2} + \log I_0\left(\frac{r_j\mathbf{x_j'}\boldsymbol{\beta}}{\sigma^2}\right)\right], \tag{5}$$

where $\mathbf{r} = [r_1, r_2, ..., r_J]'$ and J is the number of scans (Adrian, Maitra, & Rowe, 2013; Rowe, 2005b).

The signal-to-noise ratio (SNR) within the voxel, $\mathbf{x_j'}\boldsymbol{\beta}/\sigma$, determines the shape of the Rician distribution. If there is no signal or activation (SNR=0), then

the magnitude component r_j follows a Rayleigh distribution (Lindquist, 2008; Rowe, 2005b). When the SNR is large enough, the magnitude component r_j is approximately Gaussian distributed, $r_j \sim N(\mathbf{x_j'}\boldsymbol{\beta}, \sigma^2)$, with PDF

$$f(r_j|\mathbf{x_j}, \boldsymbol{\beta}, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}\left[r_j - (\mathbf{x_j'}\boldsymbol{\beta})\right]^2\right) \tag{6}$$

(Gudbjartsson & Patz, 1995; Lindquist, 2008; Rowe, 2005b).

Since the magnitude component r_j is approximately Gaussian only at high SNRs, Rowe (2005b) argued that it is not appropriate to assume a Gaussian distribution for low SNRs, which is the case in fMRI data. Rowe (2005b) mentioned that it was difficult to maximize the log-likelihood function of Rician distribution in Equation 5. Thus, he proposed a model that approximates the magnitude component using a truncated Taylor series expansion (Rowe, 2005b),

$$f(r_j|\mathbf{x_j}, \boldsymbol{\beta}, \sigma^2) = \sqrt{\frac{r_j}{\mathbf{x_j'}\boldsymbol{\beta}}} \frac{1}{\sqrt{2\pi\sigma^2}} \times \exp\left(-\frac{1}{2\sigma^2}[r_j - (\mathbf{x_j'}\boldsymbol{\beta})]^2\right), \tag{7}$$

which he found by replacing the cosine term in Equation 4 with the first two terms of its Taylor series expansion, $\cos(\phi_j - \theta) = 1 - (\phi_j - \theta)^2/2$. Nevertheless, Rowe's proposed Taylor model in Equation 7 has been argued to not give optimal test by Adrian and his colleagues (Adrian, Maitra, & Rowe, 2013).

Adrian, Maitra, and Rowe (2013) further proposed to use the exact Rician distribution by maximizing the log-likelihood of Rician distribution in Equation 5 using Expectation Maximization (EM) and Newton-Raphson (NR) iterations, with

the Gaussian-model estimates as starting values. For test of no activation, $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{0}$, against activated, $H_1 : \mathbf{C}\boldsymbol{\beta} \neq \mathbf{0}$, they calculated the restricted and unrestricted maximum likelihood estimators (MLEs) for the Rician log-likelihood in Equation 5.

Under unrestricted maximization, their EM method gave the $(k+1)$th step estimates of $\hat{\boldsymbol{\beta}}_R^{(k)}$ and $\hat{\sigma}_R^{2(k)}$ respectively as

$$\hat{\boldsymbol{\beta}}_R^{(k+1)} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\hat{u}^{(k)},$$

and

$$\hat{\sigma}_R^{2(k+1)} = [\mathbf{r}'\mathbf{r} - (\mathbf{X}'\hat{\mathbf{u}}^{(k)})'(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\hat{\mathbf{u}}^{(k)})]/(2n),$$

where $\hat{\mathbf{u}}^{(k)}$ is a vector of length J with jth entry $\hat{u}_j^{(k)} = r_j \mathrm{A}(\mathbf{x}_\mathbf{j}'\hat{\boldsymbol{\beta}}_R^{(k)} r_j/\hat{\sigma}_R^{2(k)})$, $j = 1, 2, ..., J$ and $\mathrm{A}(\cdot) = \mathrm{I}_1(\cdot)/\mathrm{I}_0(\cdot)$. Under restricted maximization, their method gave

$$\tilde{\boldsymbol{\beta}}_R^{(k+1)} = \Psi(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\tilde{u}^{(k)},$$

and

$$\tilde{\sigma}_R^{2(k+1)} = [\mathbf{r}'\mathbf{r} - (\mathbf{X}'\tilde{\mathbf{u}}^{(k)})'\Psi(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\tilde{\mathbf{u}}^{(k)})]/(2n),$$

where $\Psi = \mathbf{I}_q - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}'[\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}']^{-1}\mathbf{C}$ and $\tilde{\mathbf{u}}^{(k)}$ is a vector of length J with jth entry $\tilde{u}_j^{(k)} = r_j \mathrm{A}(\mathbf{x}_\mathbf{j}'\tilde{\boldsymbol{\beta}}_R^{(k)} r_j/\tilde{\sigma}_R^{2(k)})$, $j = 1, 2, ..., J$ (Adrian, Maitra, & Rowe, 2013). Hence, the likelihood ratio test (LRT) statistic for the Rician model in Equation 4 is given by

$$\Lambda_R = 2[LL_R(\hat{\boldsymbol{\beta}}_R, \hat{\sigma}_R^2) - LL_R(\tilde{\boldsymbol{\beta}}_R, \tilde{\sigma}_R^2)]. \tag{8}$$

Power Calculation and Comparison in Functional Magnetic Resonance Imaging

Power is defined as the probability of rejecting null hypothesis when the alternative hypothesis is true. In fMRI context, "power is the probability of detecting a true task-related activation" (Ashby, 2011, p. 178). Power is also known as sensitivity, which is the probability of accepting the alternative hypothesis correctly for a given specificity. Specificity can be defined as the probability of rejecting the null hypothesis correctly (Friston, Holmes, Poline, Price, & Frith, 1996).

Usually the receiver operator characteristic (ROC) curve is used to display power, either against sample size (Desmond & Glover, 2002; Hayasaka, Peiffer, Hugenschmidt, & Laurienti, 2007) or against $(1 - \text{specificity})$ (Friston et al., 1996). Power of a model has been shown to be influenced by several factors, namely the width of the kernel (FWHM) and the smoothness, the standard deviation and the amplitude of the signal (Friston et al., 1996), and the signal-to-noise ratio (SNR) (Rowe, 2005b).

Studies on power analysis for fMRI have been conducted through various approaches. Some focused on the Gaussian random field theory models (Friston et al., 1996; Hayasaka et al., 2007), while some focused on other approaches such as non-central distributions at each voxel (Van Horn, Ellmore, Esposito, & Berman, 1998; Zarahn & Slifstein, 2001) and simulation and resampling (Desmond & Glover, 2002; Murphy & Garavan, 2004). The power analysis studies have also been

extended to complex models (Adrian, Maitra, & Rowe, 2013; Rowe, 2005a, 2005b; Rowe & Logan, 2004).

Rowe and Logan (2004) compared the power of their proposed complex valued model (Equation 3) against the power of the Gaussian approximated magnitude model (Equation 6). They showed that their complex model had constant power regardless of the SNR, while the power for the normal approximate magnitude model reduced as SNR decreased. They compared the power curves using three thresholding procedures, which are unadjusted threshold (Type I error rate), false discovery rate (FDR), and Bonferroni procedure.

In a subsequent paper, Rowe (2005b) investigated the estimators for the magnitude and the complex-valued models. This time around, he introduced the Taylor approximated magnitude model (Equation 7) and compared it to the Gaussian approximated magnitude model (Equation 6) and the complex valued model (Equation 3). His results showed that the Gaussian approximated magnitude model (Equation 6) failed to achieve unbiasedness when SNR is lower than 10, while the estimators for the Taylor approximated magnitude model (Equation 7) were unbiased until about SNR of 7.5.

Nevertheless, Adrian et al. (2013) were able to show that the Rowe's Taylor magnitude model (Equation 7) was not usable due to the incompatibility between its false detection rate and its desired significance level. Instead of approximating the Rician magnitude model, Adrian et al. (2013) used the Expectation Maximization (EM) and Newton-Raphson (NR) iterations to calculate the MLEs for the Rician magnitude model (Equation 4) and used the LRT statistic (Equation

8), as shown in previous section. Then they compared the area under the receiver operating characteristic curve (AUC) of their Rician magnitude model (Equation 4) to the Gaussian approximated magnitude model (Equation 6). They showed that the Rician magnitude model (Equation 4) had better performance, in terms of AUC, than the Gaussian magnitude model (Equation 6) only when the SNR is about 0.6. This SNR was too low and not common in fMRI studies. Thus, they argued that the Gaussian approximation of the noise was appropriate in fMRI studies.

Rowe (2005b) and Adrian et al. (2013) compared the power of Rician magnitude model to the power of Gaussian magnitude model and used false discovery rate (FDR) to correct for multiple comparisons. Currently, there is no published study that compares the power of the Rician magnitude model to the Gaussian magnitude model using random field theory. This is the motivation for this dissertation. In this dissertation, a new test statistic using Rician random field was proposed. Then, using simulation data, the power of the Rician test in detecting a signal from functional magnetic resonance images was compared to the power of Gaussian random fields test.

Summary

BOLD fMRI is becoming a popular neuroimaging method. The fMRI data are known to be huge, complex valued and autocorrelated. These impose a difficulty in analyzing the data. Nevertheless, statistical methods have been proposed to cater to these demands. The most common statistical analysis method is using general linear model at each voxel with application of random field theory. It is assumed that the magnitude component of the fMRI data is Gaussian distributed, while the

phase component is discarded. However, the magnitude component is actually following Rician distribution. As such, Rowe (2005b) suggested a Taylor-approximated Rician model and showed that it was a better model than the Gaussian-approximated Rician model. However, a later study by Adrian et al. (2013) proved that the Taylor-approximated model was only better than the Gaussian-approximated model under unusually low value of SNRs. Both studies used the FDR method to handle the multiple comparisons problem. To the author's knowledge, a similar approach has not been studied in random field theory setting. As such, the current study proposed a new test statistic using Rician random field, and, using simulation, compared its power in detecting a signal from functional magnetic resonance images to the power of the Gaussian field test.

CHAPTER III

METHODOLOGY

The objectives of the current study were to propose a new test statistic using Rician random field, $R_{\max}$, in detecting a signal from functional magnetic resonance images, and to compare its power to the power of the Gaussian random field test, $Z_{\max}$, using simulation. Factors that might affect the performance of the Rician test statistic, $R_{\max}$, namely the location, amplitude, and the scale of the signal, were also investigated.

This chapter outlines the methodology that was used to answer the following research questions:

Q1 How is Rician random field defined in general, and simulated in two-dimensional image?

Q2 How can the power of the Rician test statistic, $R_{\max}$, be calculated using Monte Carlo simulation?

Q3 How does the amplitude of the signal affect the power of the Rician test statistic?

Q4 How does the scale of the signal affect the power of the Rician test statistic?

Q5 How does the location of the signal affect the power of the Rician test statistic?

Q6 How is the power of the Rician test statistic as compared to the Gaussian test statistic under the conditions specified in previous research questions?

This chapter starts with the explanation on data generation. Then the Rician test statistic, $R_{\max}$, and the Gaussian test statistic, $Z_{\max}$, is determined. For a specified α, the critical value to test the significance of the Rician test statistic is obtained using simulation, while the critical value to test the significance of the Gaussian test statistic is calculated using approximation as described by Adler (1981). Later, the power for each of the test statistics is calculated and compared.

Data Generation

In the current study, the smoothed Rician random field with signal was generated as

$$R(\mathbf{t}) = \sqrt{[\mu(\mathbf{t}) + \epsilon_{Re}(\mathbf{t})]^2 + [\mu(\mathbf{t}) + \epsilon_{Im}(\mathbf{t})]^2}. \tag{9}$$

As can be seen in Equation 9, Rician random field at location $\mathbf{t}$ is comprised of a smoothed signal, $\mu(\mathbf{t})$, a smoothed Gaussian white noise for the real component, $\epsilon_{Re}(\mathbf{t})$, and a smoothed Gaussian white noise for the imaginary component, $\epsilon_{Im}(\mathbf{t})$. Lu (2015) wrote a detailed explanation on how to generate a smoothed Gaussian white noise and two smoothed signals in a two-dimensional image. The current study used the same method but two smoothed Gaussian white noise and one smoothed signal were generated instead.

The simulation started with generating two two-dimensional independent Gaussian white noise images, each with mean zero and variance one, using `rnorm()` function in R Version 3.4.2.. The two Gaussian noise images represented the real

and imaginary component of an fMRI data. Each noise image was Fourier transformed using `fft()` function. Then, the Gaussian kernel, defined as

$$f\left(\frac{\mathbf{t}-\mathbf{t_0}}{\sigma}\right) = \exp\left(-\frac{\|\mathbf{t}-\mathbf{t_0}\|^2}{2\sigma^2}\right), \tag{10}$$

was used to smooth the noise. The noise was also centered. An example of a two-dimensional Gaussian white noise image is shown in Figure 4.

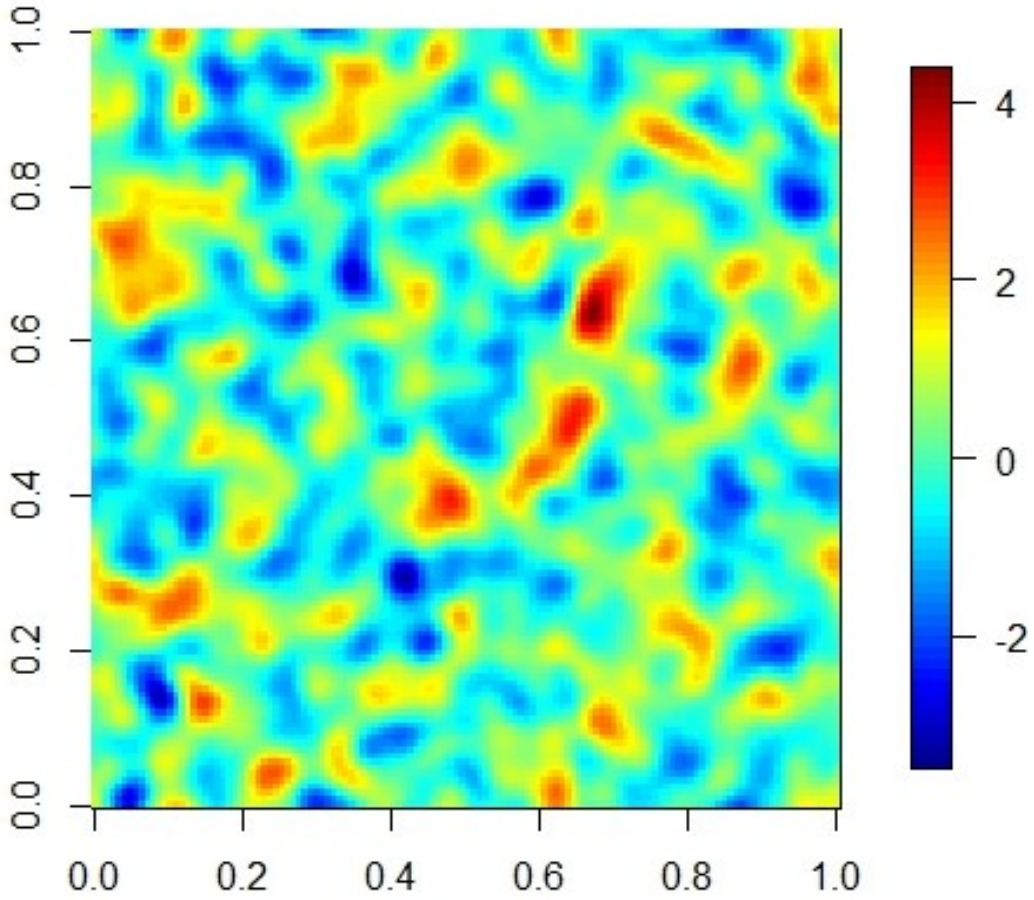

Figure 4. Two-dimensional 128 × 128 smoothed Gaussian white noise image.

The next part was to generate a smoothed signal with specified amplitude (ξ), scale (σ), and location ($\mathbf{t}$). Lu (2015, p. 39) defined the mean of a smoothed

signal for a Gaussian scale space random field as

$$\mu(\mathbf{t},\sigma) = \frac{2\xi\sigma\sigma_0}{\sigma^2+\sigma_0^2}\exp\left(\frac{-\|\mathbf{t}-\mathbf{t_0}\|^2}{2(\sigma^2+\sigma_0^2)}\right). \tag{11}$$

Simplifying Equation 11 for a fixed scale, where $\sigma_0 = \sigma$, the mean of a smoothed signal for a Gaussian random field, $\mu(\mathbf{t})$, was calculated as

$$\begin{aligned}\frac{2\xi\sigma\sigma_0}{\sigma^2+\sigma_0^2}\exp\left(\frac{-\|\mathbf{t}-\mathbf{t_0}\|^2}{2(\sigma^2+\sigma_0^2)}\right) &= \frac{2\xi\sigma^2}{2\sigma^2}\exp\left(\frac{-\|\mathbf{t}-\mathbf{t_0}\|^2}{2(2\sigma^2)}\right)\\ &= \xi\exp\left(\frac{-\|\mathbf{t}-\mathbf{t_0}\|^2}{4\sigma^2}\right).\end{aligned} \tag{12}$$

Based on previous studies by Rowe and Logan (2004) and Rowe (2005a), the signals from the real and imaginary component were assumed to be the same in the current study. An example of the two-dimensional signal is shown in Figure 5.

Finally, the smoothed signal was added to each of the smoothed noise. Then following Equation 9, a smoothed Rician random field with signal was generated. An example of a two-dimensional 128 $\times$ 128 smoothed Rician random field with signal is shown in Figure 6. The maximum of the Rician random field image, as proposed by Worsley et al. (1992), was used as the test statistic to test

$$H_0 : \xi = 0$$

$$H_1 : \xi > 0.$$

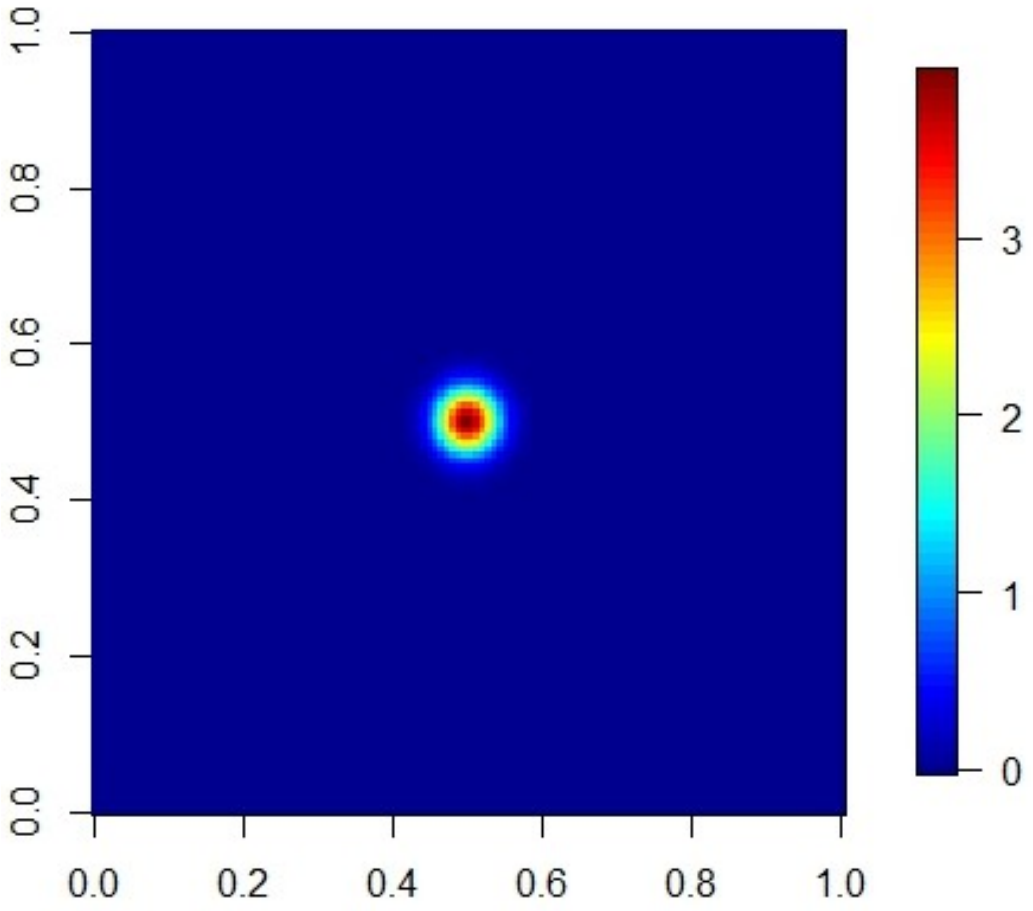

Figure 5. Two-dimensional smoothed signal image with amplitude $\xi = 4$ and scale $\sigma = .02$ at location $\mathbf{t} = (0, 0)$.

Rician Test Statistic, $R_{\max}$

Rician test statistic was defined as

$$R_{\max} = \max_{\mathbf{t}} R(\mathbf{t}).$$

To determine the significance of $R_{\max}$, a critical value needs to be determined from the empirical distribution of $R_{\max}$ under the null hypothesis. To do so, a smoothed

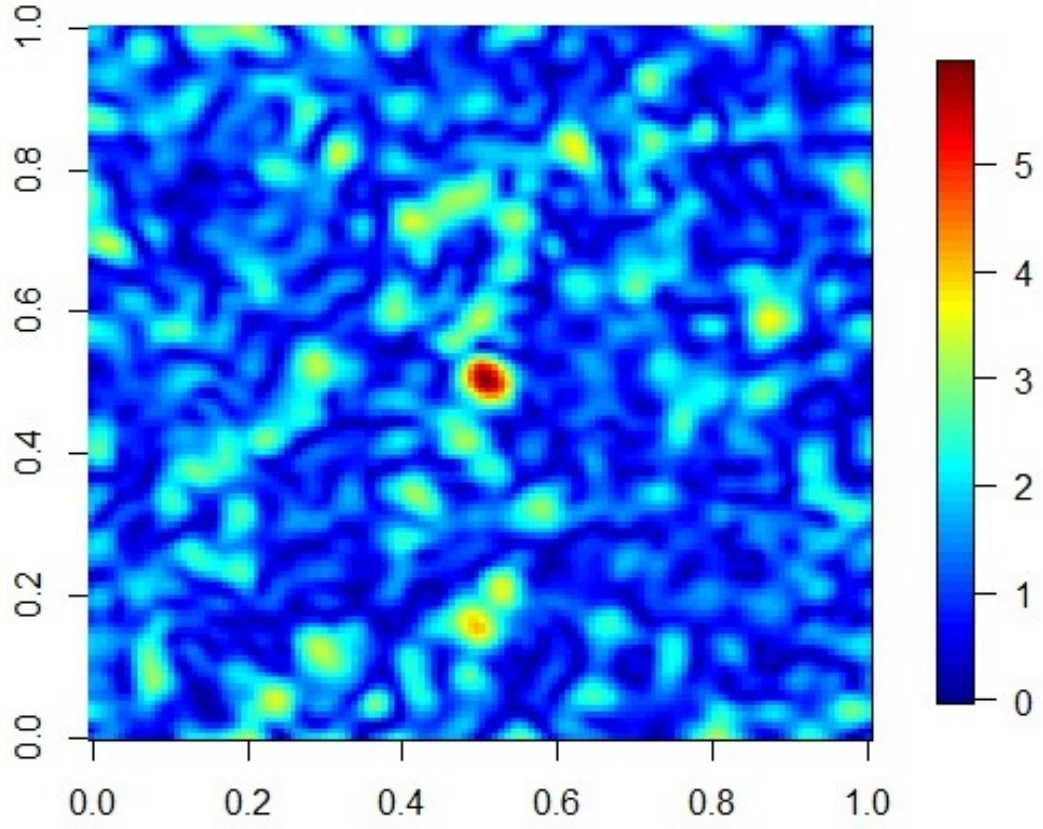

Figure 6. Two-dimensional 128 × 128 smoothed Rician random field with amplitude $\xi = 4$ and scale $\sigma = .02$ at location $\mathbf{t} = (0, 0)$.

Rician random field with no signal was generated as

$$R_0(\mathbf{t}) = \sqrt{[\epsilon_{Re}(\mathbf{t})]^2 + [\epsilon_{Im}(\mathbf{t})]^2}. \tag{13}$$

The generation of the smoothed Rician random field with no signal was replicated for 5,000 times to create the empirical distribution of test statistics under null hypothesis, say $R_{\max_0}$. Then, the critical value was determined as the $(1 - \alpha)$th percentile of the empirical distribution of $R_{\max_0}$. The same approached was used by Lu (2015).

Gaussian Test Statistic, $Z_{\max}$

In this study, the Gaussian test statistic was defined as

$$Z_{\max} = \max_{\mathbf{t}} R(\mathbf{t}).$$

To determine the significance of $Z_{\max}$, the critical value was obtained by equating the two-dimensional result for the P-value of the maximum of a Gaussian random field defined by Worsley et al. (1992, p. 906) to the alpha (α) value,

$$P(Z_{max} > z) = R(4\ln 2)(2\pi)^{-3/2} z e^{-\frac{1}{2}z^2} = \alpha, \tag{14}$$

where R is the number of resels in the search area,

$$R = \frac{\text{SearchArea}}{\text{FWHM}_m * \text{FWHM}_n}. \tag{15}$$

The sigma of the kernel at direction m, σ_m, has a relationship with the FWHM at direction m (Worsley, 1995), where

$$\text{FWHM}_m = \sigma_m \sqrt{8\ln 2}. \tag{16}$$

Using Equation 16 into Equation 15, the z value in Equation 14 was determined.

Power of the Random Field Test Statistic

Empirical power was defined as the number of times the null hypothesis is rejected when the alternative is true divided by the number of replications. Power

for the Rician test statistic and the Gaussian test statistic was calculated for every combination of amplitude (ξ), scale (σ), and location ($\mathbf{t}$) of the signal, as shown in Table 1. Also, the current study considered three values of alpha (α), 0.01, 0.05 and 0.10, for image resolution 128×128. Hence, the total combinations became 126. Each combination was replicated for 5,000 times. The 5,000 replications was chosen based on study by Lu (2015). The power for Rician and Gaussian test statistics was then compared and presented for each of the combinations.

The amplitude (ξ), scale (σ), and location ($\mathbf{t}$) of the signal were used in calculating the mean of the smoothed signal, as was seen in Equation 12. Thus, these factors influenced the size of the mean signal. The amplitude (ξ) and scale (σ) values were chosen based on the previous study by Lu (2015). In the current study, the three values of amplitude (ξ) were considered as small amplitude ($\xi = 0.5$), medium amplitude ($\xi = 2$) and large amplitude ($\xi = 4$). The location ($\mathbf{t}$) of the signal were chosen arbitrarily such that the center of the image and the points around the center of the image were included but the points did not touch the border of the image. These locations were chosen to resemble the locations of the signal used by Rowe and Logan (2004) and Rowe (2005a).

Table 1

Schemes of the Parameters

Scheme No.	Amplitude (ξ)	Signal Location	Scale (σ)
1	0.5	(0, 0)	0.02
2			0.04
3		(0.3, 0)	0.02
4			0.04
5		(-0.3, 0)	0.02
6			0.04
7		(0.2, 0.3)	0.02
8			0.04
9		(-0.2, 0.3)	0.02
10			0.04
11		(0.2, -0.3)	0.02
12			0.04
13		(-0.2,-0.3)	0.02
14			0.04
15	2	(0, 0)	0.02
16			0.04
17		(0.3, 0)	0.02
18			0.04
19		(-0.3, 0)	0.02
20			0.04
21		(0.2, 0.3)	0.02
22			0.04
23		(-0.2, 0.3)	0.02
24			0.04
25		(0.2, -0.3)	0.02
26			0.04
27		(-0.2,-0.3)	0.02
28			0.04
29	4	(0, 0)	0.02
30			0.04
31		(0.3, 0)	0.02
32			0.04
33		(-0.3, 0)	0.02
34			0.04
35		(0.2, 0.3)	0.02
36			0.04
37		(-0.2, 0.3)	0.02
38			0.04
39		(0.2, -0.3)	0.02
40			0.04
41		(-0.2,-0.3)	0.02
42			0.04

CHAPTER IV

RESULTS

This chapter presents the result of the simulation study as proposed in Chapter III. The main objective of the study was to compare the power of Rician test statistic, $R_{\max}$, to the power of Gaussian test statistic, $Z_{\max}$. The empirical distribution of $R_{\max}$ under each condition in Table 1 of Chapter III, for each of the three α values, was examined. The power of $R_{\max}$ under each condition was calculated and the effects of amplitude, scale and location of the signal on the power of $R_{\max}$ were investigated. The power of $R_{\max}$ was then compared to the power of $Z_{\max}$.

The chapter begins with the discussion on empirical distribution of $R_{\max}$. The following section is focused on the power of $R_{\max}$, which includes the effects of amplitude, scale and location of the signal on the power of $R_{\max}$. The next section is focused on the power comparisons between $R_{\max}$ and $Z_{\max}$.

The first research question is answered as follows:

Q1 How is Rician random field defined in general, and simulated in two-dimensional image?

Rician random field is defined, in general, in Equation (1). The simulation procedure of two-dimensional Rician random field is explained in Chapter III under section Data Generation. The simulation code is given in Appendix A.

Empirical Distribution of the $R_{\max}$

Under $H_0 : \xi = 0$

Under the null hypothesis, the Rician random field with no signal (Equation 13) was generated and replicated for 5000 times to create the empirical distribution of test statistics under the null hypothesis, $R_{\max_0}$. This was done for every condition in Table 1 of Chapter III. The 5000 values of $R_{\max_0}$ were sorted into an empirical distribution, where the $(1-\alpha)$th percentile of the empirical distribution was taken as the critical value, $R^{\alpha}_{\max_0}$, to determine significance of the test statistic $R_{\max}$. Figure 7 shows the empirical distributions of $R_{\max_0}$ for different amplitudes (ξ) and signal locations ($\mathbf{t}$) when the scale (σ) is 0.02 at $\alpha = 0.01$. The empirical distributions were slightly skewed to the right regardless of amplitudes and locations. The critical values, $R^{0.01}_{\max_0}$, were above 5.04 for all locations.

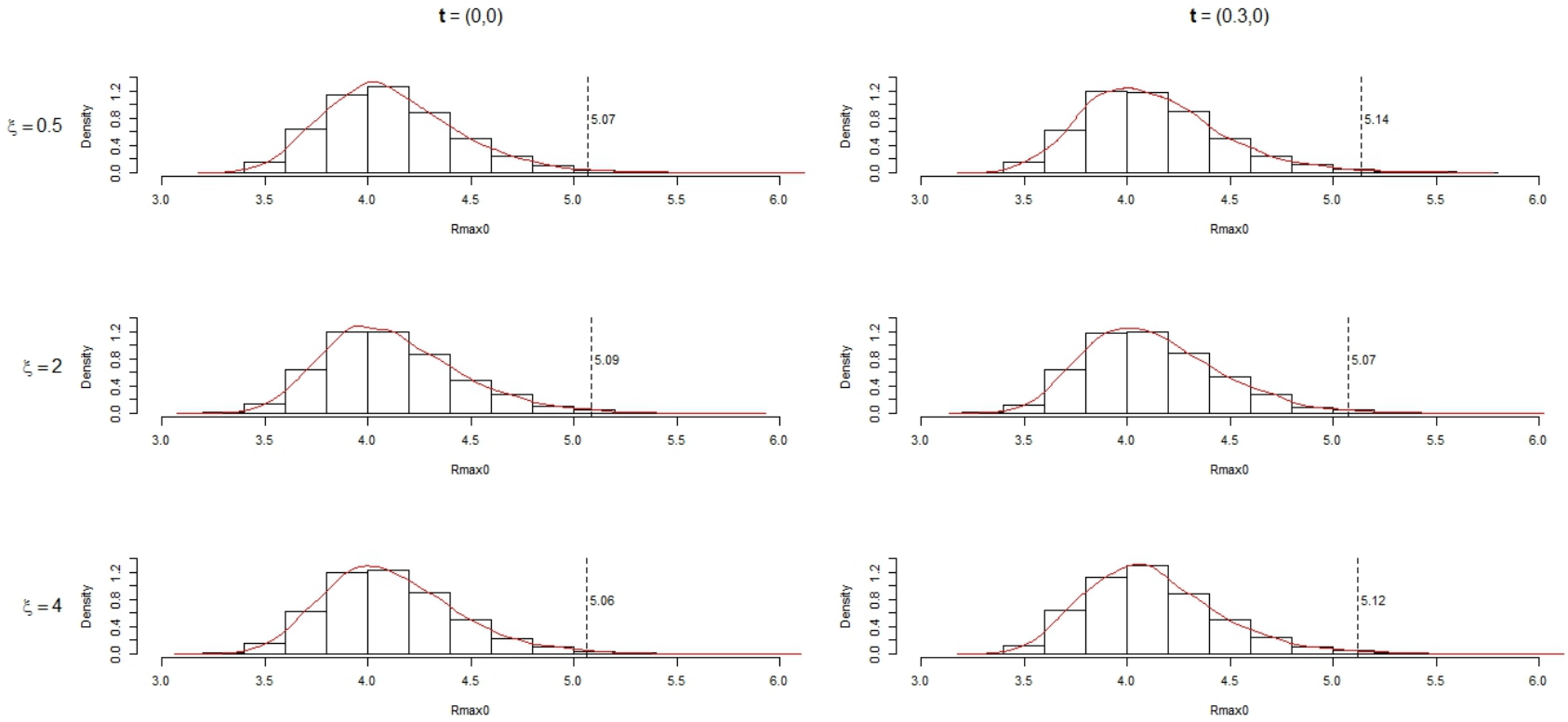

Figure 7. The empirical distributions of $R_{\max_0}$ for different ξ and $\mathbf{t}$ when $\sigma = 0.02$ at $\alpha = 0.01$.

The empirical distributions of $R_{\max_0}$ for different ξ and $\mathbf{t}$ when $\sigma = 0.04$ at $\alpha = 0.01$ are shown in Figure 8. The empirical distributions were slightly skewed to the right regardless of ξ and $\mathbf{t}$. However the $R_{\max_0}$ values were slightly lower than the $R_{\max_0}$ when $\sigma = 0.02$ (Figure 7), which resulted in the slight shift of the distributions to the left and lower critical values. The critical values, $R^{0.05}_{\max_0}$, were above 4.71 for all locations.

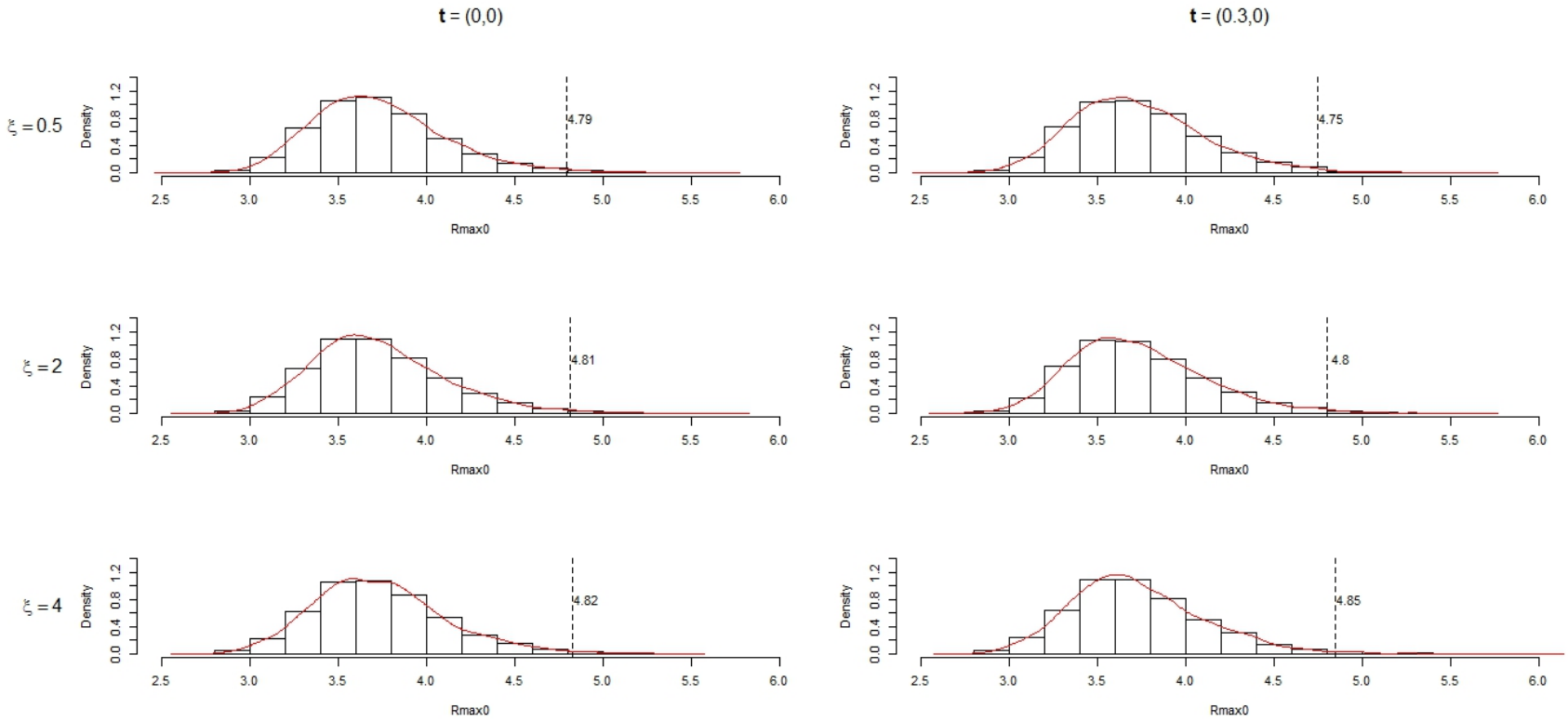

Figure 8. The empirical distributions of $R_{\max_0}$ for different ξ and $\mathbf{t}$ when $\sigma = 0.04$ at $\alpha = 0.01$.

The empirical distributions of $R_{\max_0}$ for different ξ and $\mathbf{t}$ values when $\sigma = 0.02$ at $\alpha = 0.05$ are shown in Figure 9, while the empirical distributions of $R_{\max_0}$ when $\sigma = 0.04$ are shown in Figure 10. All the distributions were skewed to the right across amplitudes and locations. The $R_{\max_0}$ when $\sigma = 0.02$ were bigger than the $R_{\max_0}$ when $\sigma = 0.04$. The critical values when $\sigma = 0.02$ were larger than 4.71, whereas the critical values when $\sigma = 0.04$ were larger than 4.40.

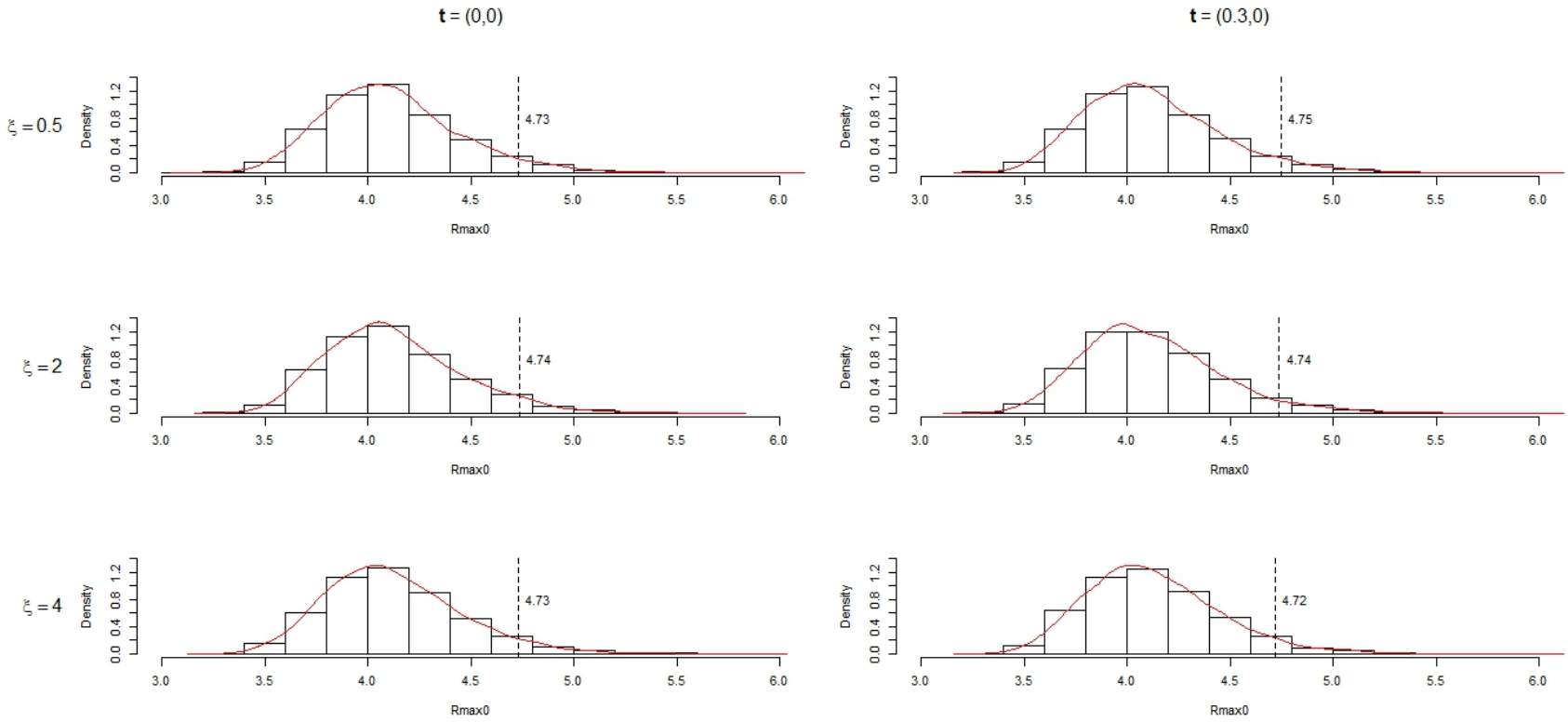

Figure 9. The empirical distributions of $R_{\max_0}$ for different ξ and $\mathbf{t}$ when $\sigma = 0.02$ at $\alpha = 0.05$.

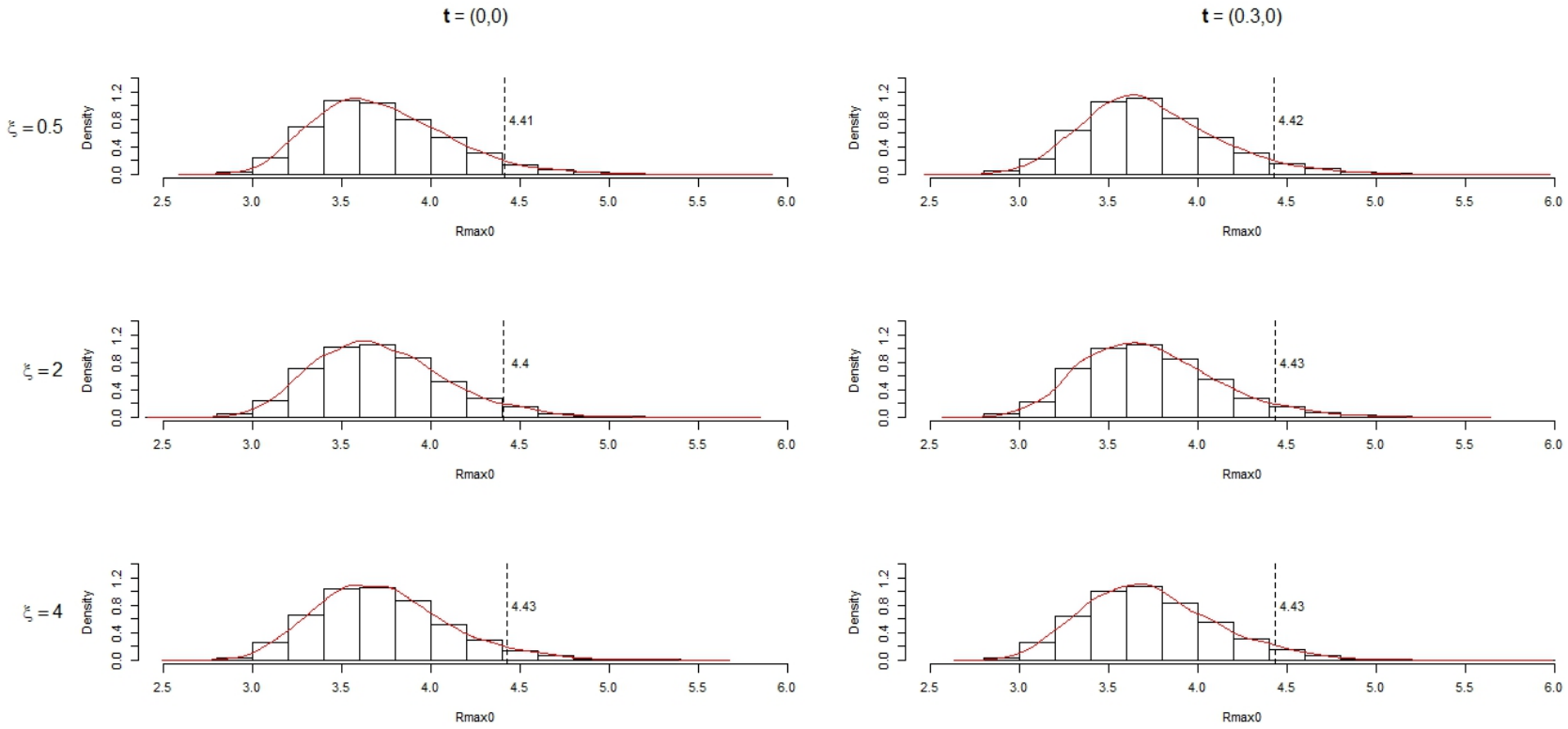

Figure 10. The empirical distributions of $R_{\max_0}$ for different ξ and $\mathbf{t}$ when $\sigma = 0.04$ at $\alpha = 0.05$.

The empirical distributions of $R_{\max_0}$ for different ξ and $\mathbf{t}$ values when $\sigma = 0.02$ at $\alpha = 0.10$ are shown in Figure 11, while the empirical distributions of $R_{\max_0}$ when $\sigma = 0.04$ are shown in Figure 12. The distributions were skewed to the right across amplitudes and locations. The $R_{\max_0}$ when $\sigma = 0.02$ were larger than the $R_{\max_0}$ when $\sigma = 0.04$. The critical values when $\sigma = 0.02$ were larger than 4.54, while the critical values when $\sigma = 0.04$ were larger than 4.20.

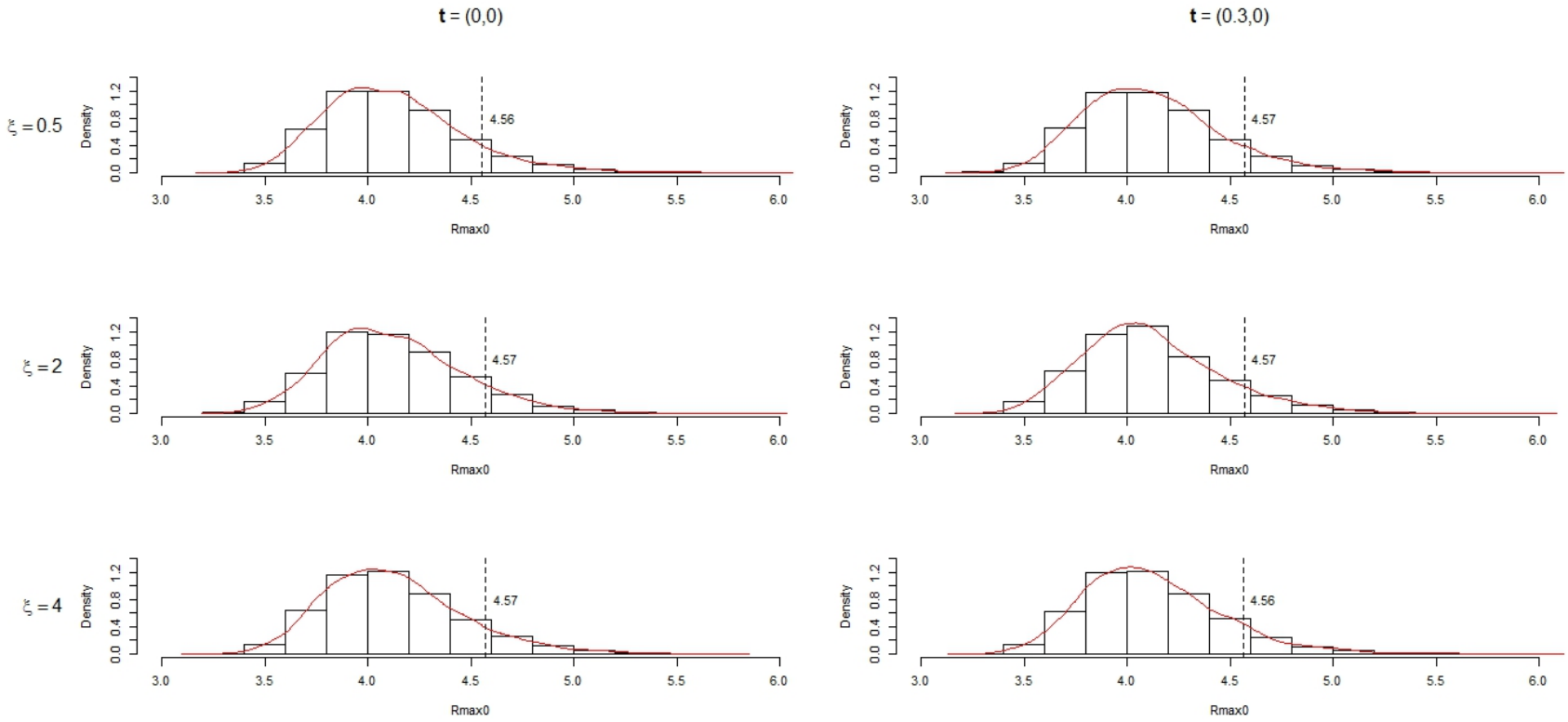

Figure 11. The empirical distributions of $R_{\max_0}$ for different ξ and $\mathbf{t}$ when $\sigma = 0.02$ at $\alpha = 0.10$.

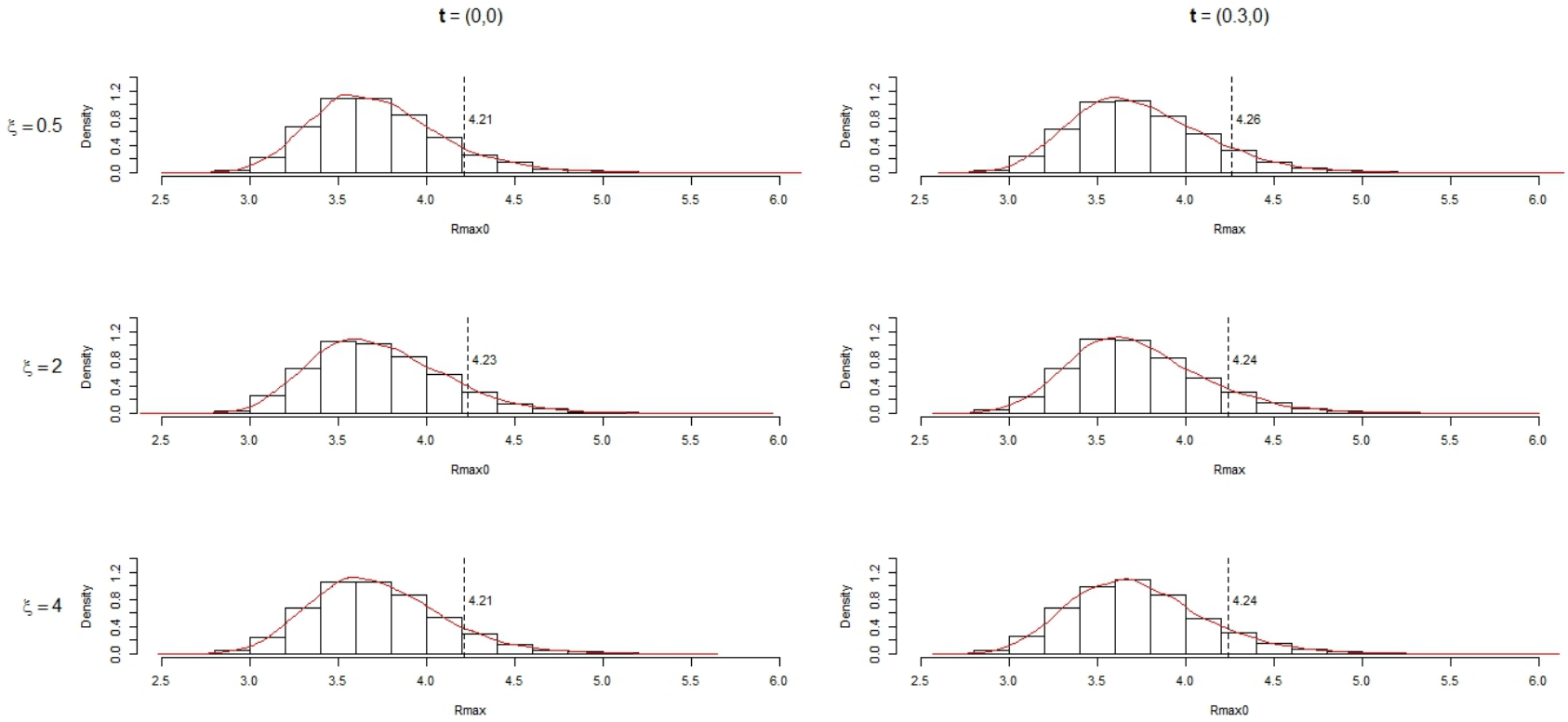

Figure 12. The empirical distributions of $R_{\max_0}$ for different ξ and $\mathbf{t}$ when $\sigma = 0.04$ at $\alpha = 0.10$.

Under $H_1 : \xi > 0$

Figure 13 shows the histograms and density plots of $R_{\max}$ for different ξ and $\mathbf{t}$ when $\sigma = 0.02$ at $\alpha = 0.01$. The empirical distributions of $R_{\max}$ across the three amplitudes at $\mathbf{t} = (0, 0)$ and $\mathbf{t} = (0.3, 0)$ looked very much alike. The empirical distributions were slightly skewed to the right when the amplitude was small ($\xi = 0.5$) and medium ($\xi = 2$). The empirical distributions became symmetrical and wider when the amplitude was large ($\xi = 4$). The critical values to determine significance of $R_{\max}$ for these particular conditions ranged between 5.07 and 5.14.

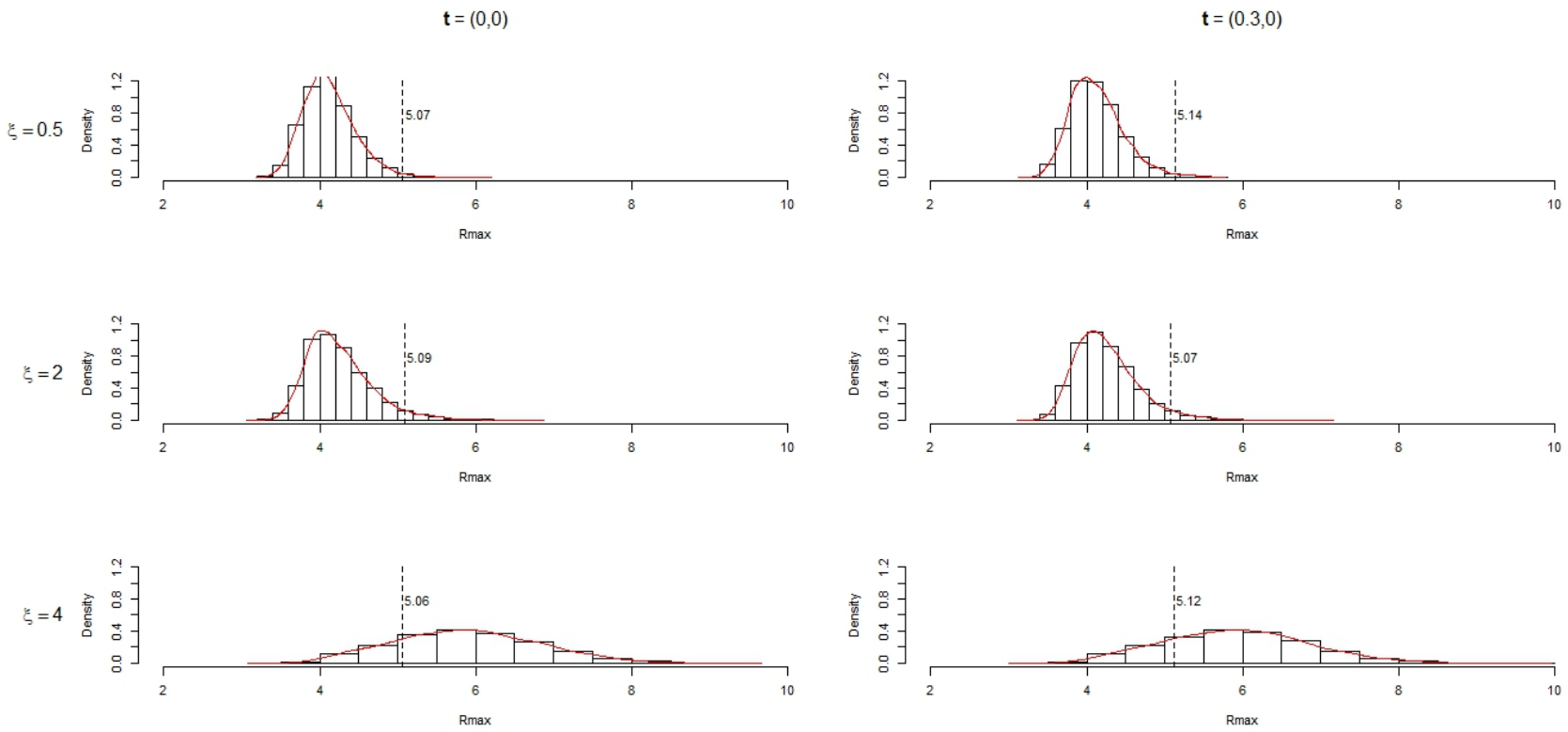

Figure 13. The empirical distributions of $R_{\max}$ for different ξ and $\mathbf{t}$ when $\sigma = 0.02$ at $\alpha = 0.01$.

When $\sigma = 0.04$ and $\alpha = 0.01$, the histograms and density plots of $R_{\max}$ for the three amplitudes (ξ) at two signal locations, $\mathbf{t} = (0, 0)$ and $\mathbf{t} = (0.3, 0)$, are shown in Figure 14. For both locations, the distributions were slightly skewed to the right when $\xi = 0.5$ and when $\xi = 2$, but were symmetrical and wider when $\xi = 4$. The empirical distributions of $R_{\max}$ when $\sigma = 0.04$ (Figure 14) were very similar to the distributions of $R_{\max}$ when $\sigma = 0.02$ (Figure 13). Nevertheless, the critical values had wider range between 4.75 and 5.15.

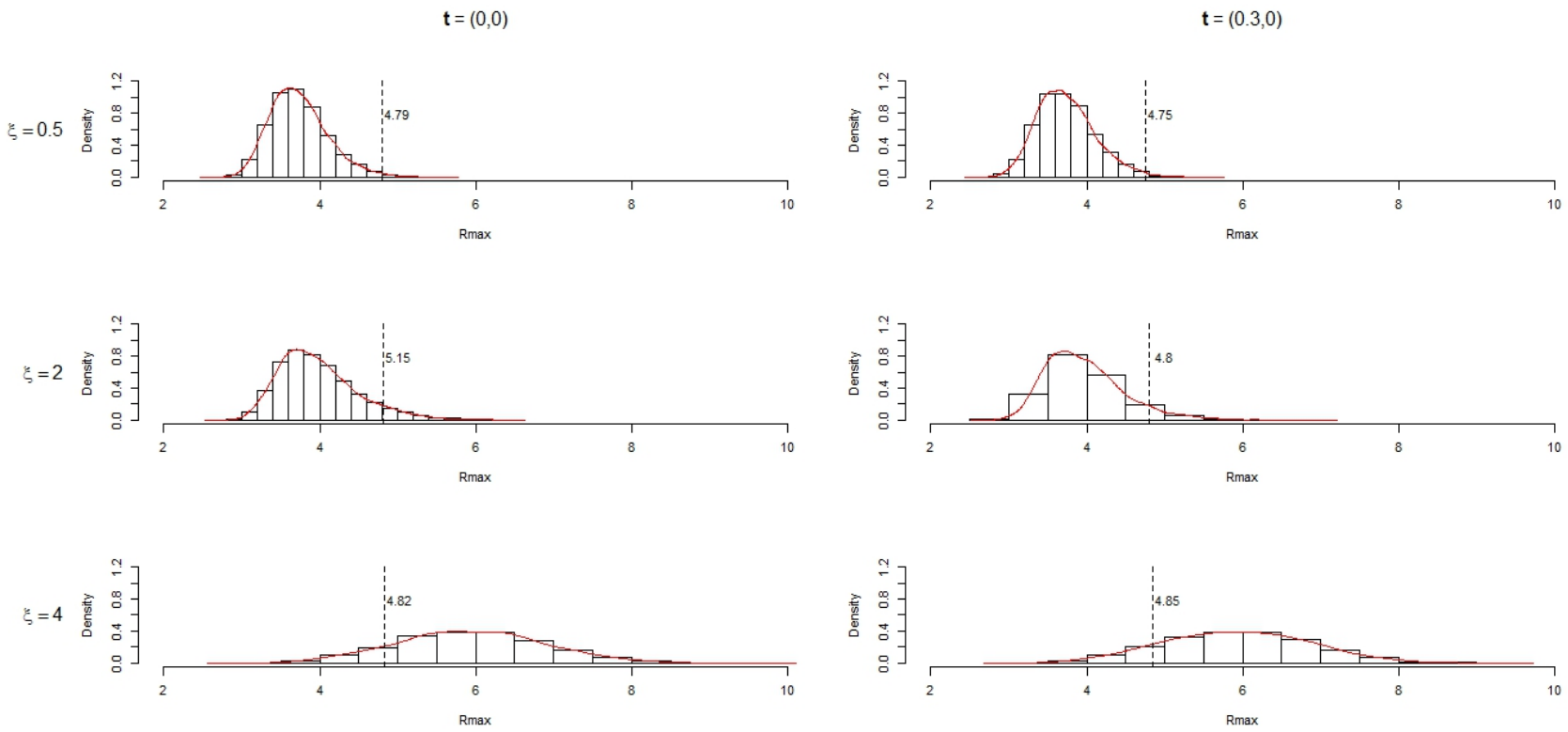

Figure 14. The empirical distributions of $R_{\max}$ for different ξ and $\mathbf{t}$ when $\sigma = 0.04$ at $\alpha = 0.01$.

At $\alpha = 0.05$, the histograms and density plots of $R_{\max}$ for the three amplitudes (ξ) at two signal locations, $\mathbf{t} = (0, 0)$ and $\mathbf{t} = (0.3, 0)$, when $\sigma = 0.02$ are shown in Figure 15. For the same parameters but when $\sigma = 0.04$, the histograms and density plots of $R_{\max}$ are shown in Figure 16. Both figures display similar patterns in the distribution of $R_{\max}$. The distributions were slightly skewed to the right when $\xi = 0.5$ and when $\xi = 2$, but were symmetrical and wider when $\xi = 4$. The critical values for $R_{\max}$ when $\sigma = 0.02$ were around 4.70. These were slightly higher than the critical values for $R_{\max}$ when $\sigma = 0.04$, which were around 4.40.

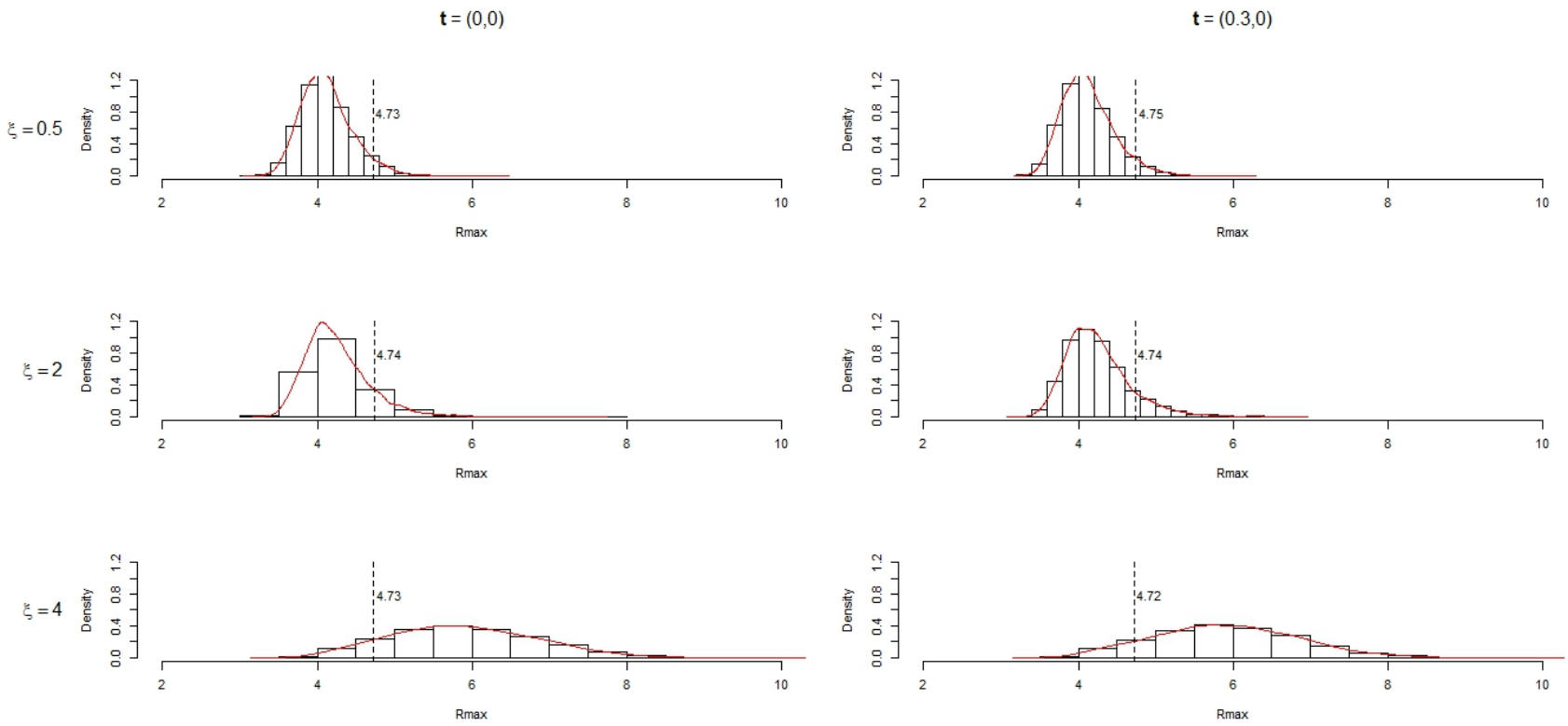

Figure 15. The empirical distributions of $R_{\max}$ for different ξ and $\mathbf{t}$ when $\sigma = 0.02$ at $\alpha = 0.05$.

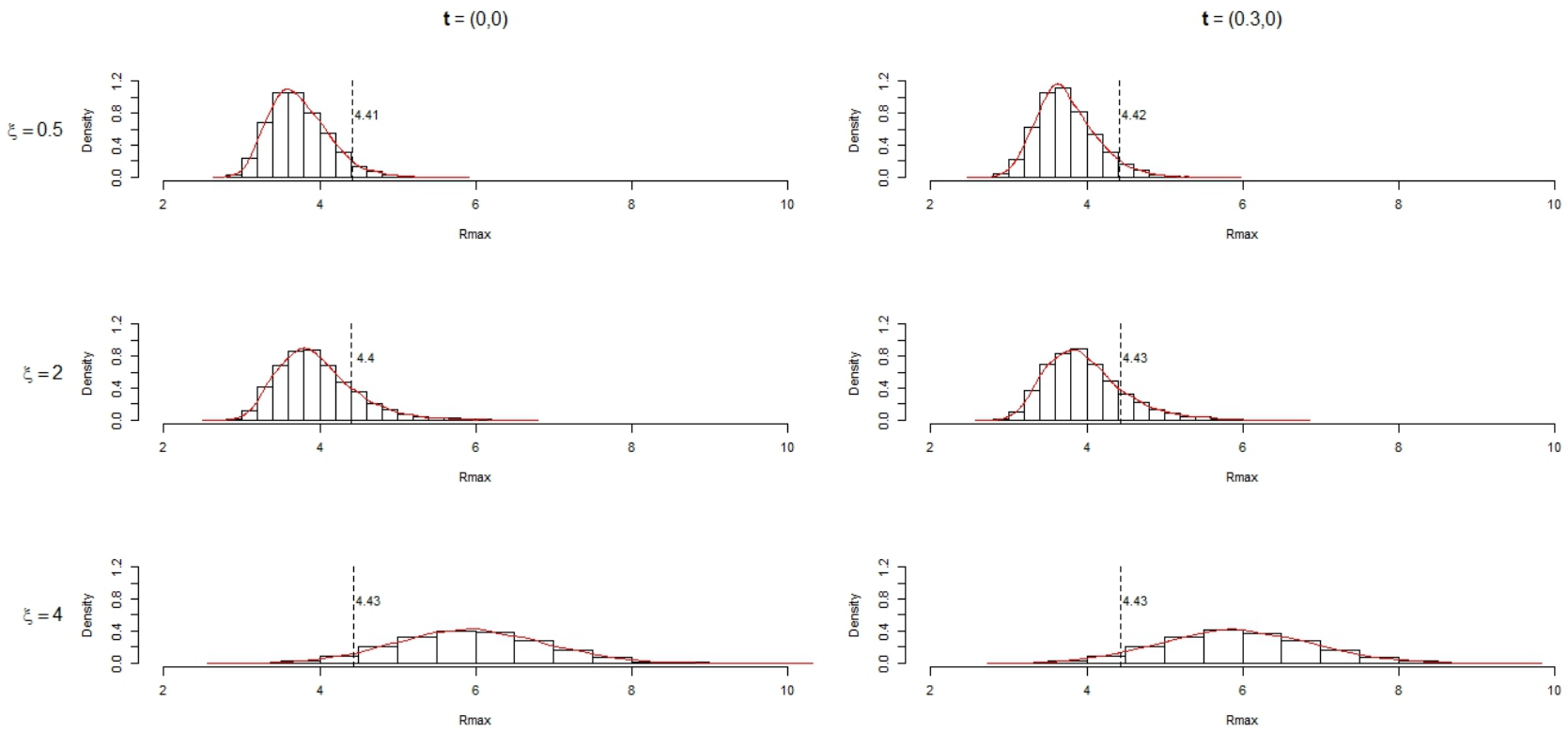

Figure 16. The empirical distributions of $R_{\max}$ for different ξ and $\mathbf{t}$ when $\sigma = 0.04$ at $\alpha = 0.05$.

At $\alpha = 0.10$, Figure 17 shows the histogram and density plots of $R_{\max}$ for the three amplitudes (ξ) at two signal locations, $\mathbf{t} = (0, 0)$ and $\mathbf{t} = (0.3, 0)$, when $\sigma = 0.02$, while Figure 18 shows the histogram and density plots of $R_{\max}$ for the same parameters when $\sigma = 0.04$. In both figures, the distributions were slightly skewed to the right when $\xi = 0.5$ and when $\xi = 2$, but were symmetrical and wider when $\xi = 4$. The critical values for $R_{\max}$ when $\sigma = 0.02$ were either 4.56 or 4.57, whereas the critical values for $R_{\max}$ when $\sigma = 0.04$ ranged between 4.21 to 4.26.

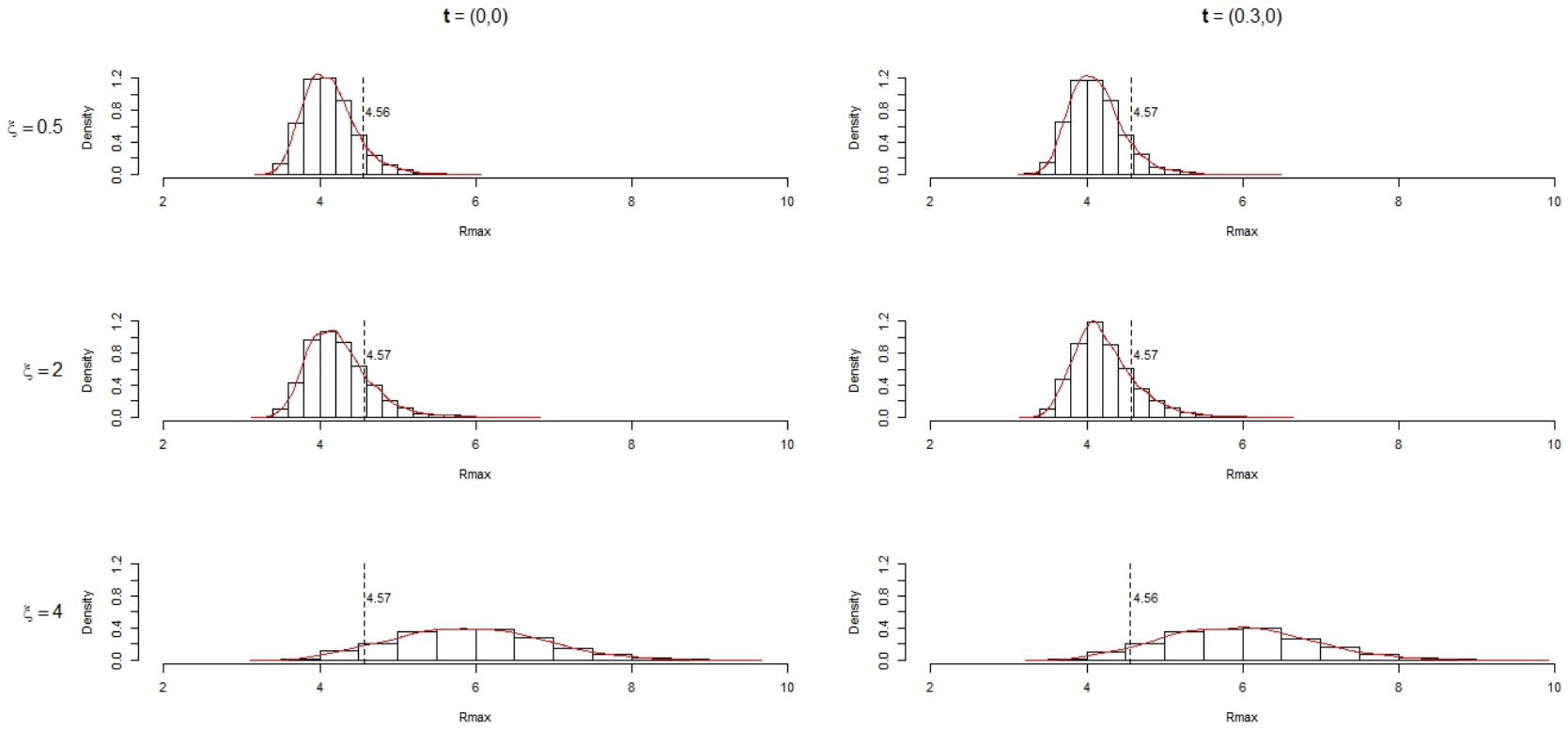

Figure 17. The empirical distributions of $R_{\max}$ for different ξ and $\mathbf{t}$ when $\sigma = 0.02$ at $\alpha = 0.10$.

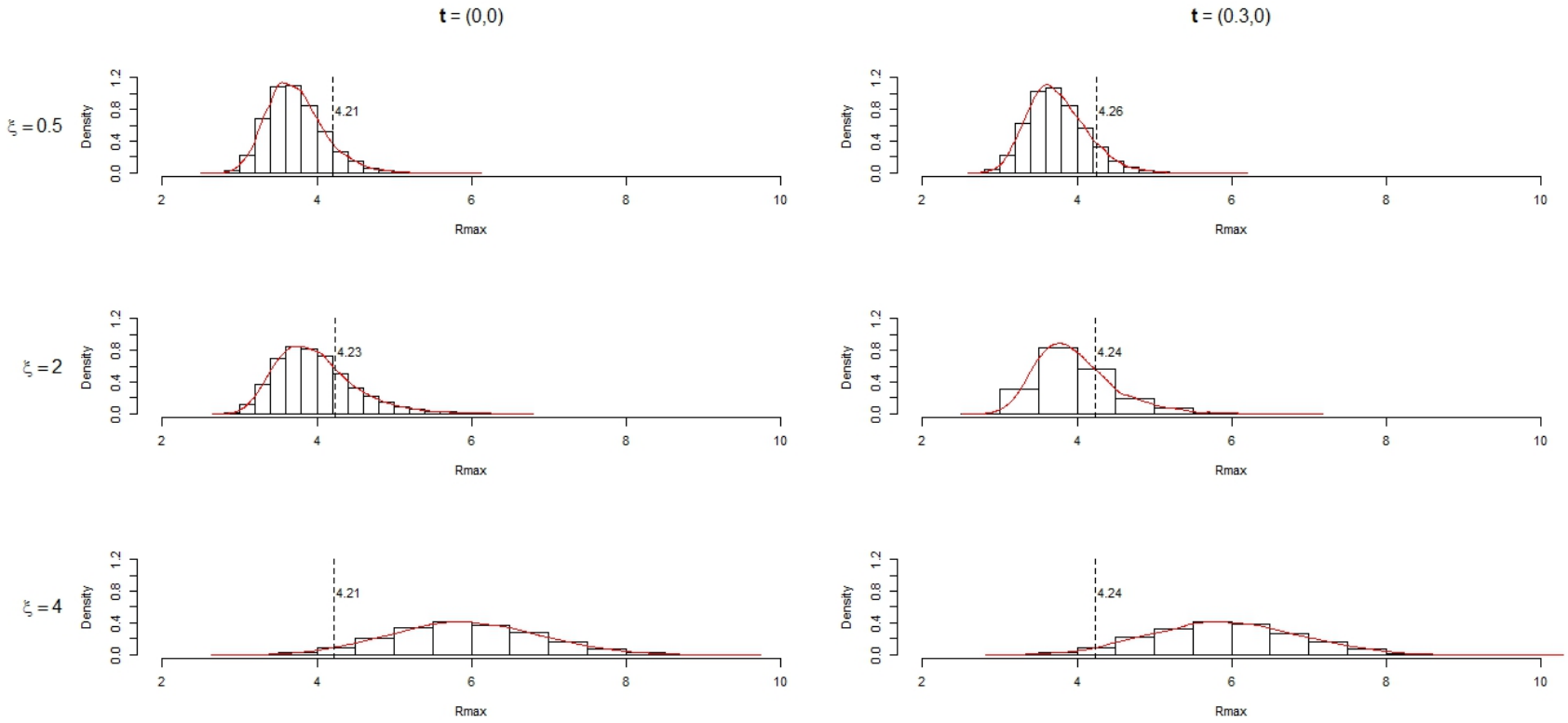

Figure 18. The empirical distributions of $R_{\max}$ for different ξ and $\mathbf{t}$ when $\sigma = 0.04$ at $\alpha = 0.10$.

For all three α values chosen in this study, the empirical distributions of $R_{\max}$ at similar amplitude showed similar behavior across both locations and scale of the signal. This suggests that amplitude plays a big role in controlling the distribution of $R_{\max}$. The distribution of $R_{\max}$ becomes less skewed, symmetrical and has a bell shaped curve as the amplitude becomes larger.

Empirical Power of the Rician Test Statistic, $R_{\max}$

Q2 How can the power of the Rician test statistic, $R_{\max}$, be calculated using Monte Carlo simulation?

The empirical power of the Rician test statistic was calculated as the number of times the test statistic, $R_{\max}$, is greater than the critical value at a specified significance level α, $R^{\alpha}_{\max_0}$, divided by the number of replications. Tables 2, 3, and 4 show the empirical power of $R_{\max}$ and $Z_{\max}$ under various σ, ξ, and $\mathbf{t}$ at α of 0.01, 0.05 and 0.10, respectively. From these tables, it can be seen that the amplitude, the scale and the location of the signal have their own effects on the power of $R_{\max}$. In this dissertation, the power of $R_{\max}$ at $\alpha = 0.01$ is labeled $R_{\max}99$, the power of $R_{\max}$ at $\alpha = 0.05$ is labeled $R_{\max}95$, and the power of $R_{\max}$ at $\alpha = 0.10$ is labeled $R_{\max}90$. Also, the power of $Z_{\max}$ at $\alpha = 0.01$ is labeled $Z_{\max}99$, the power of $Z_{\max}$ at $\alpha = 0.05$ is labeled $Z_{\max}95$, and the power of $Z_{\max}$ at $\alpha = 0.10$ is labeled $Z_{\max}90$.

Table 2

Power Table at $\alpha = 0.01$

σ	ξ	**t**	$R_{\max}99$	$Z_{\max}99$
0.02	0.5	(0, 0)	0.0102	0.0934
		(0.3, 0)	0.0104	0.0914
		(-0.3, 0)	0.0100	0.0896
		(0.2, 0.3)	0.0100	0.0980
		(-0.2, 0.3)	0.0102	0.0926
		(0.2, -0.3)	0.0102	0.0894
		(-0.2, -0.3)	0.0104	0.0952
	2	(0, 0)	0.0452	0.1818
		(0.3, 0)	0.0404	0.1742
		(-0.3, 0)	0.0352	0.1788
		(0.2, 0.3)	0.0404	0.1786
		(-0.2, 0.3)	0.0370	0.1726
		(0.2, -0.3)	0.0344	0.1760
		(-0.2, -0.3)	0.0402	0.1782
	4	(0, 0)	0.8106	0.9200
		(0.3, 0)	0.7978	0.9230
		(-0.3, 0)	0.7908	0.9130
		(0.2, 0.3)	0.7926	0.9210
		(-0.2, 0.3)	0.7822	0.9206
		(0.2, -0.3)	0.7956	0.9212
		(-0.2, -0.3)	0.7970	0.9212
0.04	0.5	(0, 0)	0.0100	0.0950
		(0.3, 0)	0.0102	0.0958
		(-0.3, 0)	0.0104	0.0974
		(0.2, 0.3)	0.0106	0.0992
		(-0.2, 0.3)	0.0096	0.0984
		(0.2, -0.3)	0.0098	0.1044
		(-0.2, -0.3)	0.0100	0.0920
	2	(0, 0)	0.0712	0.2630
		(0.3, 0)	0.0770	0.2564
		(-0.3, 0)	0.0596	0.2434
		(0.2, 0.3)	0.0684	0.2482
		(-0.2, 0.3)	0.0754	0.2508
		(0.2, -0.3)	0.0712	0.2560
		(-0.2, -0.3)	0.0744	0.2472
	4	(0, 0)	0.8738	0.9618
		(0.3, 0)	0.8794	0.9656
		(-0.3, 0)	0.8754	0.9624
		(0.2, 0.3)	0.8712	0.9578
		(-0.2, 0.3)	0.8792	0.9636
		(0.2, -0.3)	0.8750	0.9620
		(-0.2, -0.3)	0.8818	0.9642

Table 3

Power Table at $\alpha = 0.05$

σ	ξ	**t**	$R_{\max}95$	$Z_{\max}95$
0.02	0.5	(0, 0)	0.0496	0.3672
		(0.3, 0)	0.0504	0.3648
		(-0.3, 0)	0.0504	0.3698
		(0.2, 0.3)	0.0502	0.3536
		(-0.2, 0.3)	0.0498	0.3548
		(0.2, -0.3)	0.0504	0.3614
		(-0.2, -0.3)	0.0506	0.3754
	2	(0, 0)	0.1186	0.4796
		(0.3, 0)	0.1224	0.4742
		(-0.3, 0)	0.1140	0.4902
		(0.2, 0.3)	0.1208	0.4816
		(-0.2, 0.3)	0.1168	0.4700
		(0.2, -0.3)	0.1236	0.4784
		(-0.2, -0.3)	0.1144	0.4786
	4	(0, 0)	0.8942	0.9764
		(0.3, 0)	0.8892	0.9740
		(-0.3, 0)	0.8902	0.9748
		(0.2, 0.3)	0.8948	0.9736
		(-0.2, 0.3)	0.8914	0.9760
		(0.2, -0.3)	0.8898	0.9766
		(-0.2, -0.3)	0.9014	0.9756
0.04	0.5	(0, 0)	0.0516	0.3670
		(0.3, 0)	0.0514	0.3678
		(-0.3, 0)	0.0516	0.3564
		(0.2, 0.3)	0.0508	0.3678
		(-0.2, 0.3)	0.0522	0.3652
		(0.2, -0.3)	0.0524	0.3682
		(-0.2, -0.3)	0.0520	0.3708
	2	(0, 0)	0.1664	0.5544
		(0.3, 0)	0.1662	0.5604
		(-0.3, 0)	0.1650	0.5448
		(0.2, 0.3)	0.1752	0.5718
		(-0.2, 0.3)	0.1666	0.5664
		(0.2, -0.3)	0.1710	0.5608
		(-0.2, -0.3)	0.1838	0.5568
	4	(0, 0)	0.9462	0.9894
		(0.3, 0)	0.9436	0.9900
		(-0.3, 0)	0.9434	0.9924
		(0.2, 0.3)	0.9494	0.9924
		(-0.2, 0.3)	0.9436	0.9924
		(0.2, -0.3)	0.9446	0.9922
		(-0.2, -0.3)	0.9384	0.9916

Table 4

Power Table at $\alpha = 0.10$

σ	ξ	$\mathbf{t}$	$R_{\max}90$	$Z_{\max}90$
0.02	0.5	(0, 0)	0.0998	0.5918
		(0.3, 0)	0.1002	0.5860
		(-0.3, 0)	0.1012	0.5908
		(0.2, 0.3)	0.1014	0.5904
		(-0.2, 0.3)	0.1008	0.5792
		(0.2, -0.3)	0.1016	0.5754
		(-0.2, -0.3)	0.1004	0.5890
	2	(0, 0)	0.1838	0.6794
		(0.3, 0)	0.1872	0.6770
		(-0.3, 0)	0.1908	0.6804
		(0.2, 0.3)	0.1808	0.6924
		(-0.2, 0.3)	0.1914	0.6916
		(0.2, -0.3)	0.1890	0.6946
		(-0.2, -0.3)	0.1846	0.6756
	4	(0, 0)	0.9200	0.9866
		(0.3, 0)	0.9290	0.9880
		(-0.3, 0)	0.9252	0.9926
		(0.2, 0.3)	0.9206	0.9892
		(-0.2, 0.3)	0.9158	0.9904
		(0.2, -0.3)	0.9288	0.9906
		(-0.2, -0.3)	0.9312	0.9880
0.04	0.5	(0, 0)	0.1020	0.5800
		(0.3, 0)	0.1020	0.5752
		(-0.3, 0)	0.1034	0.5782
		(0.2, 0.3)	0.1020	0.5734
		(-0.2, 0.3)	0.1014	0.5846
		(0.2, -0.3)	0.1028	0.5748
		(-0.2, -0.3)	0.1022	0.5858
	2	(0, 0)	0.2548	0.7260
		(0.3, 0)	0.2648	0.7396
		(-0.3, 0)	0.2562	0.7358
		(0.2, 0.3)	0.2604	0.7424
		(-0.2, 0.3)	0.2578	0.7370
		(0.2, -0.3)	0.2536	0.7384
		(-0.2, -0.3)	0.2624	0.7398
	4	(0, 0)	0.9642	0.9970
		(0.3, 0)	0.9682	0.9966
		(-0.3, 0)	0.9666	0.9978
		(0.2, 0.3)	0.9616	0.9978
		(-0.2, 0.3)	0.9640	0.9964
		(0.2, -0.3)	0.9612	0.9964
		(-0.2, -0.3)	0.9624	0.9972

Figure 19 illustrates $R_{\max}99$, $R_{\max}95$ and $R_{\max}90$ at every location ($\mathbf{t}$) against amplitude (ξ), for both scale of the signal, $\sigma = 0.02$ and $\sigma = 0.04$. This figure is used to show the effect of amplitude, scale and location of the signal on the power of $R_{\max}$.

The Effect of Amplitude on the Power of $R_{\max}$

Q3 How does the amplitude of the signal affect the power of the Rician test statistic?

The figure shows that the power of $R_{\max}$ increases as the amplitude is increased. $R_{\max}99$ was around 0.01 when the amplitude was 0.5. $R_{\max}99$ increased to between 0.03 to 0.07 when $\xi = 2$. When $\xi = 4$, $R_{\max}99$ ranged between 0.78 to 0.88. $R_{\max}95$ was around 0.5 when $\xi = 0.5$. When $\xi = 2$, $R_{\max}95$ ranged between 0.10 and 0.20. $R_{\max}95$ ranged between 0.88 to 0.95 when $\xi = 4$. $R_{\max}90$ was around 0.10 when $\xi = 0.5$. $R_{\max}90$ increased to between 0.18 to 0.26 when $\xi = 2$. $R_{\max}90$ increased again to between 0.91 to 0.97 when $\xi = 4$.

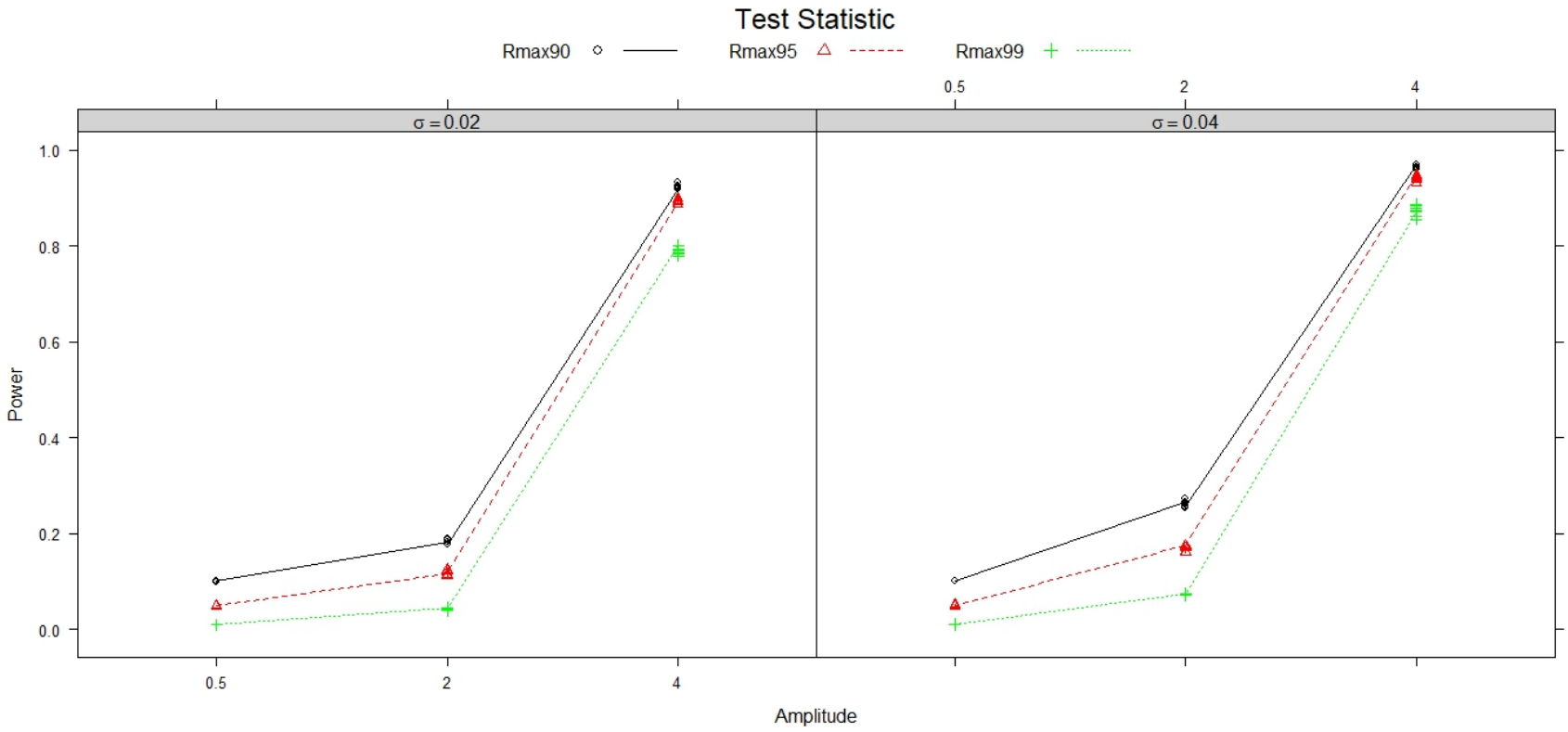

Figure 19. Empirical power of $R_{\max}$ at every location vs. amplitude for both scale of the signal (σ).

The Effect of Scale on the Power of $R_{\max}$

Q4 How does the scale of the signal affect the power of the Rician test statistic?

Figure 19 also shows that the power of $R_{\max}$ barely increases when the scale of the signal (σ) is increased from 0.02 to 0.04. When $\sigma = 0.02$, $R_{\max}99$ ranged from 0.01 to 0.80. When $\sigma = 0.04$, $R_{\max}99$ ranged from 0.01 to 0.89. At $\alpha = 0.05$, $R_{\max}95$ ranged from 0.05 to 0.89 when $\sigma = 0.02$, while $R_{\max}95$ ranged from 0.05 to 0.95 when $\sigma = 0.04$. $R_{\max}90$ ranged from 0.10 to 0.92 when $\sigma = 0.02$, whereas $R_{\max}90$ ranged from 0.10 to 0.97 when $\sigma = 0.04$.

The Effect of Location on the Power of $R_{\max}$

Q5 How does the location of the signal affect the power of the Rician test statistic?

The power of $R_{\max}$ at every location ($\mathbf{t}$) against amplitude is plotted in Figure 19. The $R_{\max}99$, $R_{\max}95$, and $R_{\max}90$ did not differ much from one location to another when the amplitude (ξ) and scale (σ) were held constant. As an example, Table 5 shows $R_{\max}99$, $R_{\max}95$ and $R_{\max}90$ at all six locations when $\xi = 0.5$, for both $\sigma = 0.02$ and $\sigma = 0.04$. When $\sigma = 0.02$, $R_{\max}99$ at $\mathbf{t} = (0, 0)$ was 0.0102, whereas at $\mathbf{t} = (-0.3, 0)$ $R_{\max}99$ was 0.0100. $R_{\max}99$ at $\mathbf{t} = (0.2, 0.3)$ was 0.0100 and at $\mathbf{t} = (-0.2, -0.3)$ $R_{\max}99$ was 0.0104. Even when $\sigma = 0.04$, $R_{\max}99$ did not change much either. When $\sigma = 0.04$, $R_{\max}99$ at $\mathbf{t} = (0, 0)$ was 0.0102 while $R_{\max}99$ at $\mathbf{t} = (0.2, -0.3)$ was 0.0096.

Table 5

Power Table of $R_{\max}$ when $\xi = 0.5$

σ	ξ	**t**	$R_{\max}99$	$R_{\max}95$	$R_{\max}90$
0.02	0.5	(0, 0)	0.0102	0.0496	0.0998
		(0.3, 0)	0.0104	0.0504	0.1002
		(-0.3, 0)	0.0100	0.0504	0.1012
		(0.2, 0.3)	0.0100	0.0502	0.1014
		(-0.2, 0.3)	0.0102	0.0498	0.1008
		(0.2, -0.3)	0.0102	0.0504	0.1016
		(-0.2, -0.3)	0.0104	0.0506	0.1004
0.04	0.5	(0, 0)	0.0100	0.0516	0.1020
		(0.3, 0)	0.0102	0.0514	0.1020
		(-0.3, 0)	0.0104	0.0516	0.1034
		(0.2, 0.3)	0.0106	0.0508	0.1020
		(-0.2, 0.3)	0.0096	0.0522	0.1014
		(0.2, -0.3)	0.0098	0.0524	0.1028
		(-0.2, -0.3)	0.0100	0.0520	0.1022

Power Comparisons between the $R_{\max}$ and the $Z_{\max}$

Q6 How is the power of the Rician test statistic as compared to the Gaussian test statistic under the conditions specified in previous research questions?

Tables 2, 3, and 4 which are mentioned earlier in this chapter show the empirical power of $R_{\max}$ and $Z_{\max}$ under various ξ, σ, and **t** at α of 0.01 ($R_{\max}99, Z_{\max}99$), 0.05 ($R_{\max}95, Z_{\max}95$), and 0.10 ($R_{\max}90, Z_{\max}90$), respectively. From these tables, it can be seen that the empirical power of $R_{\max}$ is lower than the empirical power of $Z_{\max}$ at every condition. To illustrate this, the plot of power of $R_{\max}$ and power of $Z_{\max}$ for each three alpha values (i.e. $R_{\max}99$, $Z_{\max}99$, $R_{\max}95$, $Z_{\max}95$, $R_{\max}90$, and $Z_{\max}90$) at $\mathbf{t} = (0, 0)$ is shown in Figure 20. Similar plots but at another three signal locations, namely at $\mathbf{t} = (0.3, 0)$ (Figure 21), at $\mathbf{t} = (-0.2, 0.3)$ (Figure 22), and at $\mathbf{t} = (-0.2, -0.3)$ (Figure 23), are also presented.

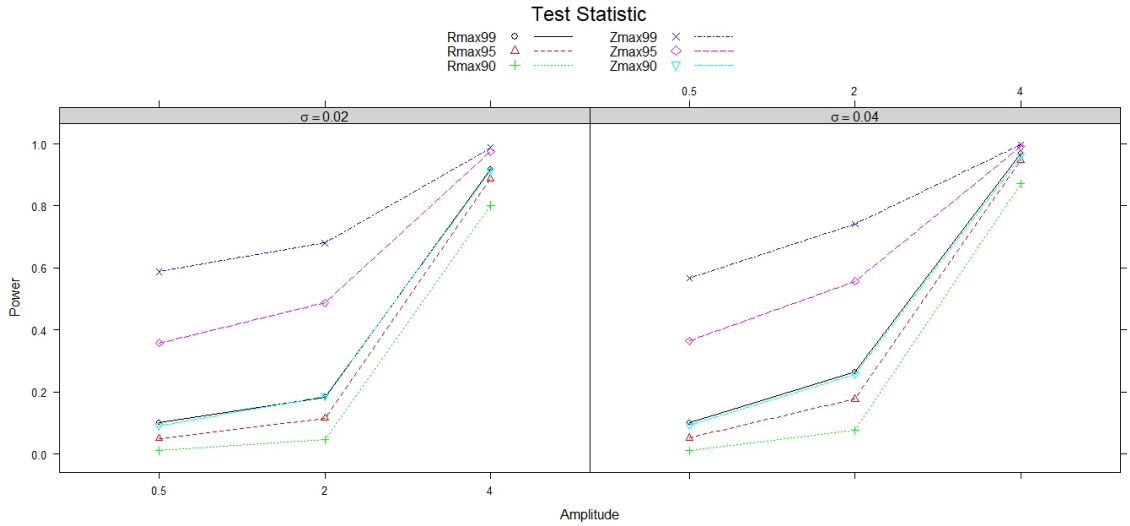

Figure 20. Empirical power vs. amplitude at $\mathbf{t} = (0, 0)$.

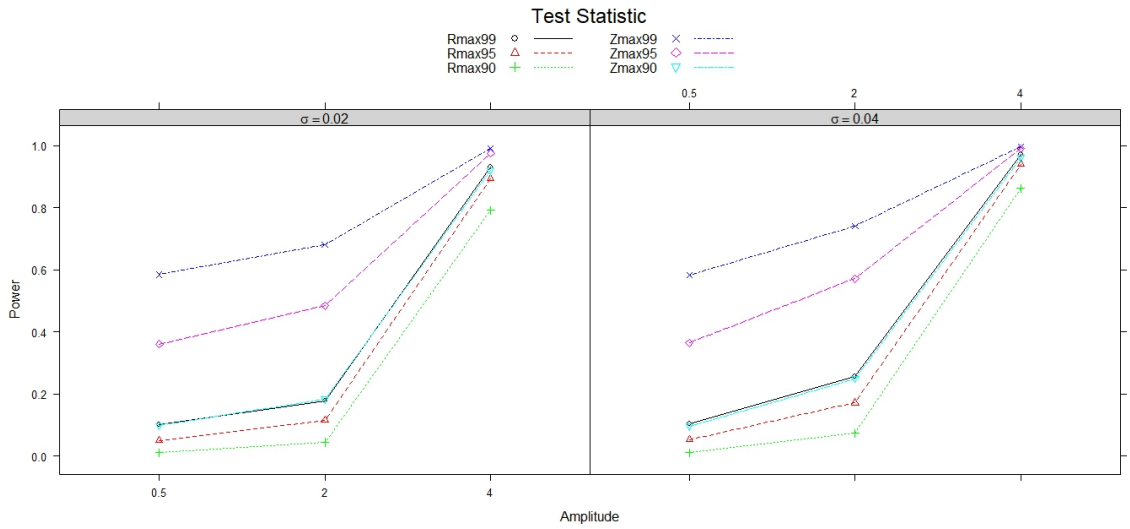

Figure 21. Empirical power vs. amplitude at $\mathbf{t} = (0.3, 0)$.

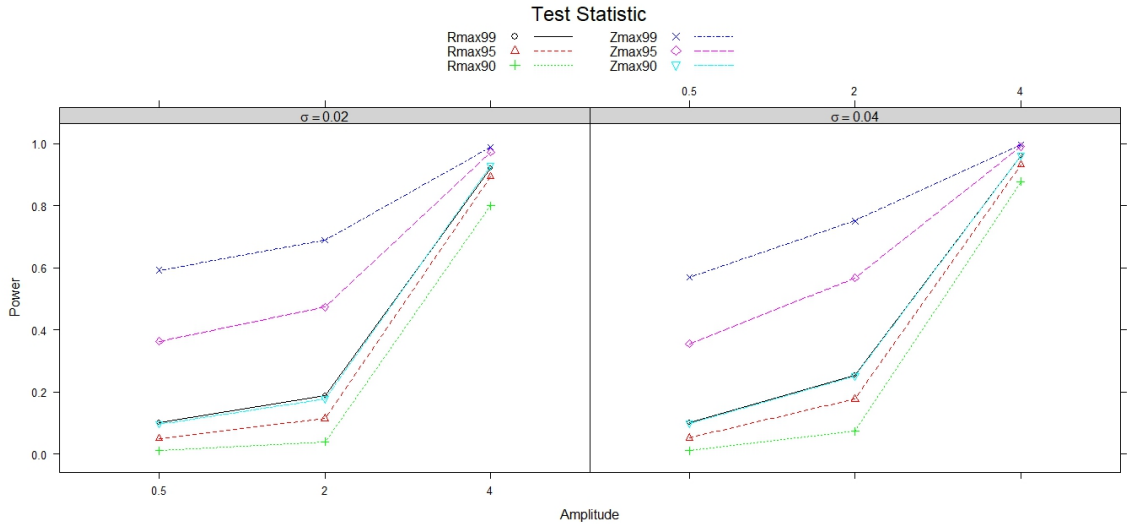

Figure 22. Empirical power vs. amplitude at $\mathbf{t} = (-0.2, 0.3)$.

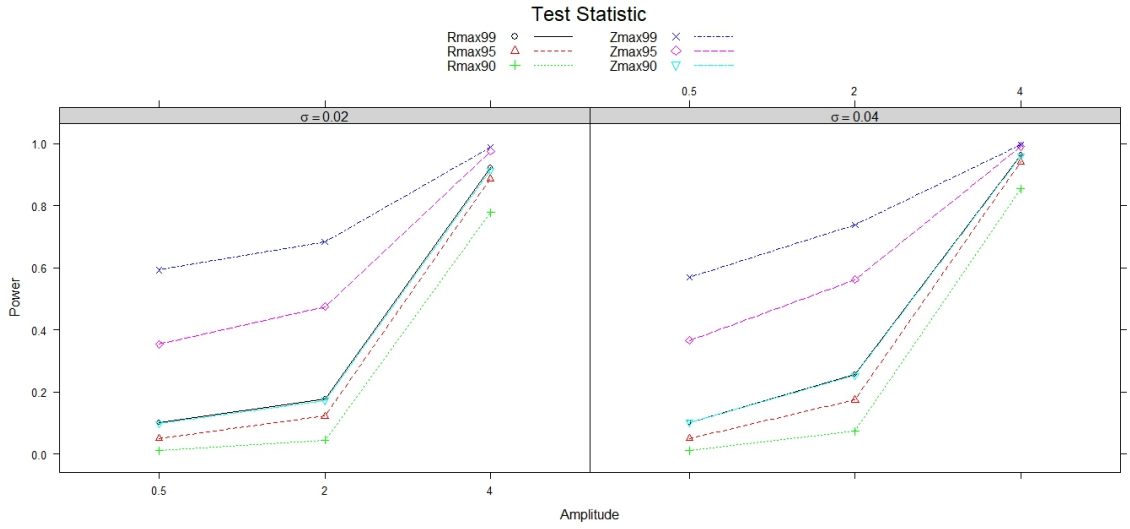

Figure 23. Empirical power vs. amplitude at $\mathbf{t} = (-0.2, -0.3)$.

From these plots, it can be seen that the power of $R_{\max}$ is always lower than the power of $Z_{\max}$. In fact, when the amplitude was small ($\xi = 0.5$) or medium ($\xi = 2$), the power of $R_{\max}$ was extremely low. The power of $R_{\max}$ and $Z_{\max}$ were closer to each other only when the amplitude was high ($\xi = 4$). When the scale of the signal (σ) changed from 0.02 to 0.04, the power of both $R_{\max}$ and $Z_{\max}$ increased a little. Also, the power of $R_{\max}$ and $Z_{\max}$ at different signal location ($\mathbf{t}$) were very close to each other. These suggest that the power of $R_{\max}$ and the power of $Z_{\max}$ do not change much from one signal location to the other.

Critical Values of $R_{\max}$ and $Z_{\max}$

Figure 24 shows the critical values for $R_{\max}$ and $Z_{\max}$ at every location ($\mathbf{t}$) against amplitude (ξ) for both scale of the signal, $\sigma = 0.02$ and $\sigma = 0.04$. It can be seen in the plot that the critical values for each of the $R_{\max}$ did not change much when ξ is increased, or when $\mathbf{t}$ is changed. However, the critical values for $R_{\max}$ were higher when $\sigma = 0.02$ than when $\sigma = 0.04$. The critical value for $Z_{\max}$ only depends on the scale of the signal (σ) and the significance level (α), as shown in Equation (14), Equation (15) and Equation (16) of Chapter III. As such, there is only one critical value for the same σ and α. The exact critical values for $R_{\max}$ at $\alpha = 0.01, 0.05$ and 0.10 are listed in Table 6, whereas the exact critical values for $Z_{\max}$ at $\alpha = 0.01, 0.05$ and 0.10 are listed in Table 7.

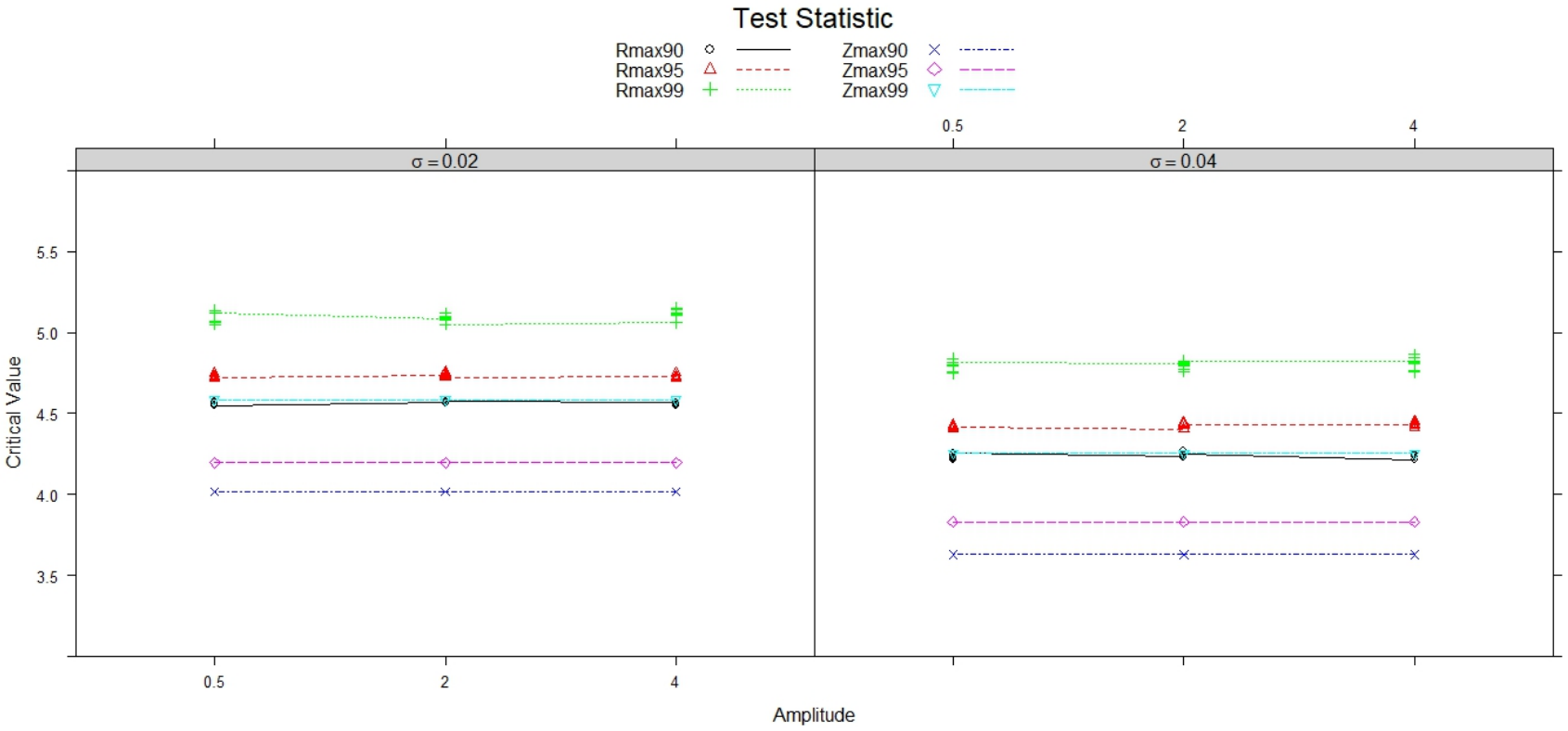

Figure 24. Critical values for $R_{\max}$ and $Z_{\max}$ at every location vs. amplitude for both scale of the signal (σ).

Table 6

Critical Values for $R_{\max}$

			α		
σ	ξ	**t**	0.01	0.05	0.10
0.02	0.5	(0, 0)	5.0665	4.7280	4.5553
		(0.3, 0)	5.1379	4.7462	4.5693
		(-0.3, 0)	5.0639	4.7150	4.5564
		(0.2, 0.3)	5.0498	4.7208	4.5452
		(-0.2, 0.3)	5.1166	4.7464	4.5722
		(0.2, -0.3)	5.0613	4.7430	4.5731
		(-0.2, -0.3)	5.1209	4.7241	4.5461
	2	(0, 0)	5.0854	4.7357	4.5680
		(0.3, 0)	5.0733	4.7373	4.5704
		(-0.3, 0)	5.0963	4.7281	4.5739
		(0.2, 0.3)	5.0837	4.7320	4.5597
		(-0.2, 0.3)	5.1199	4.7541	4.5640
		(0.2, -0.3)	5.0889	4.7463	4.5702
		(-0.2, -0.3)	5.0479	4.7209	4.5784
	4	(0, 0)	5.0610	4.7304	4.5715
		(0.3, 0)	5.1211	4.7168	4.5622
		(-0.3, 0)	5.1105	4.7204	4.5658
		(0.2, 0.3)	5.1408	4.7506	4.5459
		(-0.2, 0.3)	5.1097	4.7223	4.5666
		(0.2, -0.3)	5.1076	4.7260	4.5556
		(-0.2, -0.3)	5.1527	4.7260	4.5543
0.04	0.5	(0, 0)	4.7942	4.4107	4.2098
		(0.3, 0)	4.7488	4.4231	4.2556
		(-0.3, 0)	4.8160	4.4161	4.2459
		(0.2, 0.3)	4.8367	4.4162	4.2222
		(-0.2, 0.3)	4.7602	4.4064	4.2361
		(0.2, -0.3)	4.7988	4.4271	4.2299
		(-0.2, -0.3)	4.8177	4.4184	4.2566
	2	(0, 0)	4.8105	4.4028	4.2341
		(0.3, 0)	4.8016	4.4296	4.2358
		(-0.3, 0)	4.7575	4.4405	4.2232
		(0.2, 0.3)	4.8174	4.4427	4.2379
		(-0.2, 0.3)	4.7964	4.4422	4.2695
		(0.2, -0.3)	4.7718	4.4269	4.2333
		(-0.2, -0.3)	4.8236	4.4316	4.2499
	4	(0, 0)	4.8243	4.4266	4.2138
		(0.3, 0)	4.8460	4.4323	4.2410
		(-0.3, 0)	4.7668	4.4145	4.2500
		(0.2, 0.3)	4.8118	4.4369	4.2347
		(-0.2, 0.3)	4.8064	4.4458	4.2325
		(0.2, -0.3)	4.7537	4.4524	4.2525
		(-0.2, -0.3)	4.8640	4.4352	4.2409

Table 7

Critical Values for $Z_{\max}$

			α		
σ	ξ	$\mathbf{t}$	0.01	0.05	0.10
0.02	0.5	(0, 0)			
		(0.3, 0)			
		(-0.3, 0)			
		(0.2, 0.3)			
		(-0.2, 0.3)			
		(0.2, -0.3)			
		(-0.2, -0.3)			
	2	(0, 0)			
		(0.3, 0)			
		(-0.3, 0)			
		(0.2, 0.3)	4.5829	4.1962	4.0168
		(-0.2, 0.3)			
		(0.2, -0.3)			
		(-0.2, -0.3)			
	4	(0, 0)			
		(0.3, 0)			
		(-0.3, 0)			
		(0.2, 0.3)			
		(-0.2, 0.3)			
		(0.2, -0.3)			
		(-0.2, -0.3)			
0.04	0.5	(0, 0)			
		(0.3, 0)			
		(-0.3, 0)			
		(0.2, 0.3)			
		(-0.2, 0.3)			
		(0.2, -0.3)			
		(-0.2, -0.3)			
	2	(0, 0)			
		(0.3, 0)			
		(-0.3, 0)			
		(0.2, 0.3)	4.2522	3.8277	3.6274
		(-0.2, 0.3)			
		(0.2, -0.3)			
		(-0.2, -0.3)			
	4	(0, 0)			
		(0.3, 0)			
		(-0.3, 0)			
		(0.2, 0.3)			
		(-0.2, 0.3)			
		(0.2, -0.3)			
		(-0.2, -0.3)			

CHAPTER V

CONCLUSIONS AND FUTURE RESEARCH

Conclusions

In this dissertation, Rician random field was defined in general, and simulated in a two-dimensional image. The empirical power of the Rician test statistic, $R_{\max}$, were also calculated using Monte Carlo simulation study of 5000 replicates under different values of amplitude (ξ), scale (σ), and location ($\mathbf{t}$) of the signal, at three different α values. The effect of amplitude, scale, and location of the signal on the power of $R_{\max}$ were investigated. The empirical power and the critical values of $R_{\max}$ were also compared to the empirical power and the critical values of Gaussian test statistic, $Z_{\max}$, respectively.

The simulation study showed that the power of $Z_{\max}$ outperformed the power of $R_{\max}$ at every condition studied. This suggests that the use of $Z_{\max}$ is sufficient and should be preferred over $R_{\max}$. This result is unexpected given that the data were Rician distributed. Moreover, this finding is contrary to the results on the LRT statistics by Rowe (2005b) and Adrian et al. (2013), where the power of Rician LRT outperformed the power of Gaussian LRT when SNR was small.

However Adrian et al. (2013) also found that for the power of Rician LRT to outperform the power of Gaussian LRT, the SNR needed to be smaller than the

ones usually found in fMRI data. Hence, they preferred the Gaussian LRT for its simplicity rather than the Rician LRT. This seems to be true for $R_{\max}$ too, as shown in the current study. Therefore, the use of $Z_{\max}$ should be preferred than $R_{\max}$, given similar conditions as in the current study.

As shown in the result chapter, amplitude has the biggest effect on the power of $R_{\max}$, followed by the scale of the signal. High power is observed when both the amplitude and the scale of the signal are high. Unfortunately, change in the location of the signal, while holding the amplitude and scale constant, does not have much effect on the power of $R_{\max}$. This could be because there was only one signal generated at one location in each replication. Thus, the power of the test statistic to detect the signal is the same regardless of the location.

The empirical distributions of $R_{\max}$ show that $R_{\max}$ is slightly right skewed distributed. The empirical distribution of $R_{\max}$ becomes more symmetrical as the amplitude increases. This is similar to the findings on Rician distribution of noisy MRI data by Gudbjartsson and Patz (1995), where it was right skewed and as the SNR is increased the distribution approximated the Gaussian distribution. This convinces the author that the simulated data were following Rician distribution.

The critical value at a specific α, $R^{\alpha}_{\max}$, to determine significance of $R_{\max}$ was obtained from the 5000 replicates of the Monte Carlo simulation study at each combination of amplitude, scale, and location of the signal. As such, the critical value changes accordingly. Nevertheless, the critical values at the same α and σ were very close to each other. Since the critical value for $Z_{\max}$ only depends on α and σ, it might be the same case for $R_{\max}$.

In conclusion, the current study has defined a new random field named Rician random field. From there, a new test statistic named Rician test statistic, $R_{\max}$, was defined and introduced as the test statistic to detect signal in fMRI images. The empirical distribution of the $R_{\max}$ under several conditions were examined. This study showed that $R_{\max}$ is not as powerful in detecting a signal in an fMRI image as compared to the Gaussian test statistic, $Z_{\max}$. As such, it is suggested to use $Z_{\max}$ in detecting a signal in an fMRI image. Nevertheless, an insight on the effect of amplitude, scale and location of the signal on $R_{\max}$ was investigated and the power of $R_{\max}$ was compared to the power of $Z_{\max}$. Also, the critical value to determine significance of $R_{\max}$ was examined and compared to the critical value for $Z_{\max}$. The current study could be regarded as the first step into a more in depth study of $R_{\max}$ as a test statistic to detect signal in fMRI images.

Future Research and Limitations

In this dissertation, the Rician random field was defined, in general, and simulated in a two-dimensional image. The Rician test statistic, $R_{\max}$, was also defined and introduced. This was the first attempt to use $R_{\max}$ as a test statistic to detect signal in an image. Therefore, this study only focused on the power of $R_{\max}$ under several levels of amplitude, scale and location of the signal. This might be the reason for the extremely low power of the $R_{\max}$ as compared to the Gaussian test statistic, $Z_{\max}$. It is of future interest to investigate other levels of amplitude, scale and location of the signal that might be able to show the true power of $R_{\max}$. Furthermore, studies on the critical value and the probability of Type I error are also of future interest, especially since the power of $R_{\max}$ has been shown to be low.

The current study was mainly focused on simulation rather than the theoretical aspects of the Rician random field and the Rician test statistic. This is due to the fact that there are limited literatures on Rician random field and Rician test statistic. As such, a simulation study was considered for the current dissertation. Future studies that focus on the theoretical parts are highly suggested. For example, a study on whether the $R_{\max}$ is a likelihood ratio test statistic will be much appreciated to further explore and investigate the properties of the Rician test statistic.

Moreover, the Rician test statistic in this dissertation was used only in a fixed kernel width scenario. It is of interest to know if the Rician test statistic can be used in scale space setting, or even rotation space, as was done for the Gaussian test statistic (Shafie, Sigal, Siegmund, & Worsley, 2003; Siegmund & Worsley, 1995). In this dissertation, the study on the power of $R_{\max}$ was conducted with the assumption that only one signal occurred in the image. Further works in cases where multiple signals occurred in the image can be conducted, such as the work done by Lu (2015) for Gaussian test statistic.

REFERENCES

Adler, R. J. (1981). *The geometry of random fields*. New York, NY: John Wiley & Sons, Ltd.

Adler, R. J. (2000). On excursion sets, tube formulas and maxima of random fields. *The Annals of Applied Probability*, *10*(1), 1–74.

Adrian, D. W., Maitra, R., & Rowe, D. B. (2013). Ricean over Gaussian modelling in magnitude fMRI analysis - added complexity with negligible practical benefits. *Stat*, *2*(1), 303–316.

Ashby, F. G. (2011). *Statistical analysis of fMRI data*. Cambridge, MA: MIT Press.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 289–300.

Bertschinger, E. (2001). Multiscale Gaussian random fields and their application to cosmological simulations. *The Astrophysical Journal Supplement Series*, *137*, 1–20.

Brett, M., Penny, W., & Kiebel, S. (2004). Introduction to random field theory. In R. S. Frackowiak et al. (Eds.), *Human brain function* (2nd ed., pp. 867–879). Burlington, MA: Academic Press, Inc.

Bullmore, E. T., & Suckling, J. (2001). Functional magnetic resonance imaging. *International Review of Psychiatry*, *13*(1), 24–33.

Cao, J., & Worsley, K. J. (2001). Applications of random fields in human brain mapping. In M. Moore (Ed.), *Spatial statistics: Methodological aspects and applications* (pp. 169–182). New York, NY: Springer New York.

Christakos, G. (1992). *Random field models in earth sciences*. San Diego, CA: Academic Press, Inc.

Desmond, J. E., & Glover, G. H. (2002). Estimating sample size in functional MRI (fMRI) neuroimaging studies: Statistical power analyses. *Journal of Neuroscience Methods*, *118*(2), 115–128.

Friston, K. J., Frith, C. D., Frackowiak, R. S., & Turner, R. (1995). Characterizing dynamic brain responses with fMRI: A multivariate approach. *NeuroImage*, *2*(2), 166–172.

Friston, K. J., Holmes, A., Poline, J.-B., Price, C. J., & Frith, C. D. (1996). Detecting activations in PET and fMRI: Levels of inference and power. *NeuroImage*, *4*(3), 223–235.

Friston, K. J., Jezzard, P., & Turner, R. (1994). Analysis of functional MRI time-series. *Human Brain Mapping*, *1*(2), 153–171.

Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, *15*(4), 870–878.

Gudbjartsson, H., & Patz, S. (1995). The Rician distribution of noisy MRI data. *Magnetic Resonance in Medicine*, *34*(6), 910–914.

Hayasaka, S., Peiffer, A. M., Hugenschmidt, C. E., & Laurienti, P. J. (2007). Power and sample size calculation for neuroimaging studies by non-central random

field theory. *NeuroImage*, *37*(3), 721–730.

Kobayashi, H., Mark, B. L., & Turin, W. (2011). *Probability, random processes, and statistical analysis: Applications to communications, signal processing, queueing theory and mathematical finance*. New York, NY: Cambridge University Press.

Lindquist, M. A. (2008). The statistical analysis of fMRI data. *Statistical Science*, *23*(4), 439–464.

Longuet-Higgins, M. S. (1952). On the statistical distributions of sea waves. *Journal of Marine Research*, *11*(3), 245–265.

Lu, P. (2015). *Testing for two signals with unknown locations in functional magnetic resonance images using Gaussian random field: A Monte Carlo simulation study* (Doctoral dissertation, University of Northern Colorado). Retrieved from https://search-proquest-com.unco.idm.oclc.org/docview/1733968327?accountid=12832

Matthews, P. M., & Jezzard, P. (2004). Functional magnetic resonance imaging. *Journal of Neurology, Neurosurgery & Psychiatry*, *75*(1), 6–12. Retrieved from http://jnnp.bmj.com/content/75/1/6.abstract

McKeown, M. J., & Sejnowski, T. J. (1998). Independent component analysis of fMRI data: Examining the assumptions. *Human Brain Mapping*, *6*, 368–372.

Murphy, K., & Garavan, H. (2004). An empirical investigation into the number of subjects required for an event-related fMRI study. *NeuroImage*, *22*(2), 879–885.

Ogawa, S., Lee, T.-M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic

resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences*, *87*(24), 9868–9872.

Poldrack, R. A., Mumford, J. A., & Nichols, T. E. (2011). *Handbook of functional MRI data analysis*. New York, NY: Cambridge University Press.

Rice, S. O. (1944). Mathematical analysis of random noise. *Bell System Technical Journal*, *23*(3), 282–332.

Rice, S. O. (1945). Mathematical analysis of random noise. *Bell System Technical Journal*, *24*(1), 46–156.

Rice, S. O. (1948). Statistical properties of a sine wave plus random noise. *Bell System Technical Journal*, *27*(1), 109–157.

Rowe, D. B. (2005a). Modeling both the magnitude and phase of complex-valued fMRI data. *NeuroImage*, *25*(4), 1310–1324.

Rowe, D. B. (2005b). Parameter estimation in the magnitude-only and complex-valued fMRI data models. *NeuroImage*, *25*(4), 1124–1132.

Rowe, D. B., & Logan, B. R. (2004). A complex way to compute fMRI activation. *NeuroImage*, *23*(3), 1078–1092.

Shafie, K., Sigal, B., Siegmund, D., & Worsley, K. (2003). Rotation space random fields with an application to fmri data. *The Annals of Statistics*, *31*(6), 1732–1771.

Siegmund, D. O., & Worsley, K. J. (1995). Testing for a signal with unknown location and scale in a stationary gaussian random field. *The Annals of Statistics*, *23*(2), 608–639.

Van Horn, J. D., Ellmore, T. M., Esposito, G., & Berman, K. F. (1998). Mapping

voxel-based statistical power on parametric images. *NeuroImage*, 7(2), 97–107.

Worsley, K. J. (1994). Local maxima and the expected euler characteristic of excursion sets of χ^2, F and t fields. *Advances in Applied Probability*, 13–42.

Worsley, K. J. (1995). Estimating the number of peaks in a random field using the Hadwiger characteristic of excursion sets, with applications to medical images. *The Annals of Statistics*, *23*(2), 640–669.

Worsley, K. J. (2001). Testing for signals with unknown location and scale in a χ^2 random field, with an application to fMRI. *Advances in Applied Probability*, 773–793.

Worsley, K. J. (2004). Developments in random field theory. In R. S. Frackowiak et al. (Eds.), *Human brain function* (2nd ed., pp. 881–886). Burlington, MA: Academic Press, Inc.

Worsley, K. J. (2006). Random field, Gaussian. In *Encyclopedia of Environmetrics.* Retrieved from http://dx.doi.org/10.1002/9780470057339.var008

Worsley, K. J. (2007). Random field theory. In K. Friston, J. Ashburner, S. Kiebel, T. Nichols, & W. Penny (Eds.), *Statistical parametric mapping* (pp. 232–236). London: Academic Press, Inc.

Worsley, K. J., Evans, A. C., Marrett, S., & Neelin, P. (1992). A three-dimensional statistical analysis for CBF activation studies in human brain. *Journal of Cerebral Blood Flow & Metabolism*, *12*(6), 900–918.

Worsley, K. J., Marrett, S., Neelin, P., Vandal, A. C., Friston, K. J., & Evans, A. C. (1996). A unified statistical approach for determining significant signals in

images of cerebral activation. *Human Brain Mapping*, *4*(1), 58–73.

Worsley, K. J., Taylor, J. E., Tomaiuolo, F., & Lerch, J. (2004). Unified univariate and multivariate random field theory. *NeuroImage*, *23*, S189–S195.

Zarahn, E., & Slifstein, M. (2001). A reference effect approach for power analysis in fMRI. *NeuroImage*, *14*(3), 768–779.

APPENDIX A

R CODE FOR SIMULATION

```
#########################################################
#install.packages('fields')
#install.packages('rootSolve')
library(fields)
library(rootSolve)
library(gsubfn)
set.seed(98765)

data.generate <- function(nsim,N,sigma,t1,t2,amp){
 data <- matrix(0,nrow=nsim,ncol=2)
 for (i in 1:nsim){

 #### Gaussian noises; real and imaginary part
 noise1 <- matrix(rnorm(N*N),N,N)
 fnoise1 <- fft(noise1)
 noise2 <- matrix(rnorm(N*N),N,N)
 fnoise2 <- fft(noise2)

 #### Generate Gaussian smoothing kernel: filter
 x <- ((row(noise1)-1)/(N-1)) - 0.5
 y <- ((col(noise1)-1)/(N-1)) - 0.5
 Q <- exp(-0.5*(x^2+y^2)/sigma^2)
 filter <- Q/sqrt(sum(Q^2))
 ffilter <- Mod(fft(filter))

 #### Smoothing Gaussian noise
 sz1 <- Re(fft(fnoise1*ffilter,inverse=T))/N/N
 sz2 <- Re(fft(fnoise2*ffilter,inverse=T))/N/N

 #### Smoothed signal
 mu <- amp*exp(-1/4*((x-t1)^2+(y-t2)^2)/(sigma^2))

 #### Rician+signal
 rice <- sqrt((sz1+mu)^2+(sz2+mu)^2)
 Rmax <- max(rice)

 #### Rician CV under null hypothesis: no signal
 rice0 <- sqrt(sz1^2+sz2^2)
 Rmax0 <- max(rice0)

 data[i,1] <- Rmax
 data[i,2] <- Rmax0
}
return(data)
}
```

```
nsim <- 5000
N <- 128
alpha <- rep(c(0.01,0.05,0.10),each=42)
amp <- rep(c(0.5,2,4),each=14)
sigma <- rep(c(0.02,0.04),each=7,times=3)
location <- matrix(c(0,0,0.3,0,-0.3,0,0.2,0.3,-0.2,0.3,
        0.2,-0.3,-0.2,-0.3),nrow=7,ncol=2,byrow=T)
colnames(location) <- c("t1", "t2")
condition = cbind(data.frame(alpha,amp,sigma),location)
maxR <- array(0,c(nrow(condition),nsim,2))
strt <- Sys.time()
output <- matrix(0,nrow=nrow(condition),ncol=9)
colnames(output) <- c("alpha","amplitude","scale","t1",
 "t2","CVRician","PowerR","CVGaussian","PowerZ")

for (j in 1:nrow(condition)){
 alpha <- condition[j,1]
 amp <- condition[j,2]
 sigma <- condition[j,3]
 t1 <- condition[j,4]
 t2 <- condition[j,5]

 data <- data.generate(nsim,N,sigma,t1,t2,amp)
 colnames(data) <- c("Rmax", "Rmax0")
 maxR[j,,] <- data

 #### CV at alpha: Simulation
 CVr <- quantile(sort(data[,2]),probs=1-alpha)

 #### Gaussian CV under null hypothesis: no signal
 FWHM <- sigma*sqrt(8*log(2))
 resel <- 1/(FWHM*FWHM)  #Search area=1
 ttW2 <- seq(from=-6,to=6,by=0.001)
 ffW2 <- function(t)
   resel*4*log(2)*(2*pi)^(-3/2)*t*exp(-0.5*t^2)-alpha
 CVzF <- max(uniroot.all(ffW2, c(0,10)))

 #### Power
 PowerR <- sum(data[,1]>CVr)/nsim
 PowerZF <- sum(data[,1]>CVzF)/nsim

 output[j,1] <- alpha
 output[j,2] <- amp
 output[j,3] <- sigma
```

```
 output[j,4] <- t1
 output[j,5] <- t2
 output[j,6] <- CVr
 output[j,7] <- PowerR
 output[j,8] <- CVzF
 output[j,9] <- PowerZF
}
print(Sys.time()-strt)
```