2016

Score Test Derivations and Implementations for Bivariate Probability Mass and Density Functions with an Application to Copula Functions

Roy Bower
University of South Carolina

Score Test Derivations and Implementations for Bivariate Probability Mass and Density Functions with an Application to Copula Functions

by

Roy Bower

Bachelor of Science
John Carroll University 2009

Master of Science
John Carroll University 2011

Master of Science
Villanova University 2013

Master of Science
University of South Carolina 2015

Submitted in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy in

Biostatistics

The Norman J. Arnold School of Public Health

University of South Carolina

2016

Accepted by:

James Hardin, Major Professor

James Hussey, Committee Member

Jiajia Zhang, Committee Member

Joseph Quattro, Committee Member

Cheryl L. Addy, Vice Provost and Dean of the Graduate School

Acknowledgments

Over the past twelve years of college, I have dedicated my life to education in all aspects. I have worked tirelessly to excel in whatever I do and to achieve above and beyond what I thought was possible. I've pushed myself to the brink; even to the point where I wanted to throw in the towel a few times. However, I've learned to persevere and to never give up, no matter what. I'd like to think I could have achieved all this on my own, but I know better. There are so many people to thank, and there is so much to be thankful for, that it is truly hard to find the appropriate time and words to express my gratitude. First and foremost, I want to thank my parents. Without them, I wouldn't be where I am today. They have always been there to support whatever decision I have made in life, school based or otherwise. I imagine it hasn't been easy for them watching their son spend a majority of his life in college. Even more difficult to imagine is having to support this son through it all. I imagine that they will be thrilled that I'm about to get a real job. Furthermore, it takes a special set of parents to support me the way they have; without question and with total devotion. I thank them and love them with all my heart. A major extension of my parents is my grandparents. Grandma and grandpa Parks and grandma and grandpa Bower are big fans of mine. They have supported me and cheered for me through it all. Grandpa Bower has two master's degrees in math, but I think he will finally admit that I've surpassed him! One reason I've rushed to finish this degree is for him. His time is running out and the last thing he wants before he passes is to see me receive my PhD. He jokingly likes to admit "hurry up!" Well pap, you don't have to wait much longer.

I've said this in the past, and I'll say it again: graduate students are products of their professors, and I've been extremely lucky to have some of the best professors possible. I'd like to thank Dr. Hardin for all his hard work, guidance, and advice. Without him, I would have been lost. I admit that I'm not a natural born genius or some sort of statistical prodigy. I have to work very hard at everything I do. Nothing comes easy. It was very refreshing to have an advisor who understood this and who was willing to work with me. I'd also like to thank my committee members: Dr. Hussey, Dr. Zhang, and Dr. Quattro. They have been there when I've needed them, and have offered me invaluable insight that has helped to shape my dissertation and gradate experience. I'd also like to send a special shout out to my mentor of many years, Dr. Short. If there is one professor who has truly shaped who I am and the direction my life is taking, it would be Dr. Short. To this day, Dr. Short and I still remain in contact and he still gives me advice. He has been a great resource and mentor, but an even better friend. Soon, he will be a colleague. I look forward to possibly working with Dr. Short on a project or two. Regardless, I consider Dr. Short to be a lifelong friend.

I consider myself to be the luckiest graduate student at USC. I've honestly made some lifelong friends. I'd like to thank my best friend Will Cipolli for his statistical insights, but more importantly for his emotional support. If you haven't thought about quitting at least once and you haven't had a friend there to put it all in perspective and reign you back in, then you probably aren't a PhD student. I'd also like to thank my current roommate Michael Roney for putting up with all my statistical weirdness. My best friend from John Carroll University, Anthony Vecci, also deserves some credit. This mathematical and statistical journey of mine began at John Carroll University twelve years ago, and since that time Anthony and I have remained great friends. His occasional text of support about my studies means the world to me. For some reason unknown to me, he wants to come watch my defense.

I'm appreciative, but I'm not sure he will be after an hour of material that is sure to put him to sleep! I'd also like to send a special shout out to Kitty, the owner of Cool Beans. If you want to know where I spent all my time writing this dissertation, just ask Kitty!

In conclusion, I am sincerely humbled and thankful to have such special professors, friends, and family. This twelve-year journey is finally coming to an end, but an even longer journey is about to begin. Thanks to the support of everyone aforementioned, I am prepared to excel and live a long, happy, and prosperous life. I've learned so much more than just how to derive a score test or code up a simulation; I've learned who I am and what I'm meant to do for the rest of my life. This wouldn't have been possible without the support of everyone in my life. For this, I am forever grateful, and I look forward to providing the same support to the next generation.

Abstract

This dissertation is comprised and grounded in statistical theory with an application to solving real world problems. In particular, the development and implementation of multiple score tests under a variety of scenarios are derived, applied, and interpreted. In chapter 2, I propose a score test for independence of the marginals based on Lakshminarayana's bivariate Poisson distribution. Each marginal distribution of the bivariate model is a univariate Poisson distribution, and the parameters of the bivariate distribution can be estimated using maximum likelihood methods. The simulation study shows that the score test maintains size close to the nominal level. To assess the efficiency of the derived score test, the estimated significance levels and powers of the likelihood ratio and Wald tests are compared. A relevant data set is used to illustrate the application of the bivariate Poisson model and the proposed score test for independence. In chapter 3, two score tests are proposed: one for testing independence based on Sankaran and Nair's bivariate Pareto distribution and one for testing whether Sankaran and Nair's parameterization reduces to the more popular bivariate Pareto distribution introduced by Lindley–Singpurwalla. The marginal distributions of both bivariate parameterizations are univariate Pareto II distributions, and the parameters of the bivariate distribution are estimated using numerical methods. The simulation studies show that both score tests maintain a significance level close to the nominal size. To check the efficiency of the derived score tests, the estimated significance levels and powers of the likelihood ratio and Wald tests are also compared. One real world data set is used to illustrate the application of both score tests. In chapter 4, an increasingly popular approach to model the dependence between random variables via the use of copula functions is explored. A score test for testing independence of response variables is proposed for the specific case where the marginal distributions are known to be Poisson. The simulation study shows the test keeps the significance level close to the nominal one. Similarly, the estimated significance levels and powers of the likelihood ratio and Wald tests are also compared to show our test is numerically stable. A real world data set is used to demonstrate the application of the test.

Table of Contents

List of Tables

Chapter 1

Introduction

Methods to derive hypothesis tests based on finite sample criteria are available, but yield optimal tests (i.e., uniformly most powerful) for only a small collection of problems. It is of greater interest to formulate large sample approaches to hypothesis tests (Radhakrishna Rao, 1948) such as the approach advocated here.

Suppose $Y_1, Y_2, \ldots, Y_n$ comprise a collection of independent and identically distributed random variables from $f_Y(y|\theta)$, where $\theta \in \Theta \subseteq \Re^1$, i.e. θ is a scalar parameter. The likelihood function is $L(\theta|\mathbf{Y}) = f_{\mathbf{Y}}(\mathbf{y}|\theta) = f_{Y_1}(y_1|\theta) f_{Y_2}(y_2|\theta) \ldots f_{Y_n}(y_n|\theta) = \prod_{i=1}^{n} f_{Y_i}(y_i|\theta)$, where the joint distribution given by the product of the univariate distributions is a consequence of the independence of the sample observations. The log-likelihood function can then be written as $\mathscr{L} \equiv \ln[L(\theta|\mathbf{Y})] = \sum_{i=1}^{n} \ln[f_{Y_i}(y_i|\theta)]$.

The likelihood is a function of θ with the data considered fixed. The maximum of the likelihood is the same as the maximum for the log of the likelihood and is usually easier to work. The derivative of the log of the likelihood is known as the score function, and when the score function is set equal to zero, it is known as the estimating equation. The estimating equation is so named because when solved, it yields the maximum likelihood estimate denoted $\widehat{\theta}$. This estimate can be viewed as the value of θ that maximizes the probability of the observed data, $\mathbf{y}$. Similarly, $\widetilde{\theta}$ will be used to denote the maximum likelihood estimate of θ under the null hypothesis. In most cases, the estimating equation must be solved numerically to find $\widehat{\theta}$. That is, an analytic solution to the estimating equation is not usually available.

The score function, when viewed as random with independent and identically distributed random variables, is $S(\theta|\mathbf{Y}) = \frac{\partial \mathscr{L}}{\partial \theta} = \frac{\partial}{\partial \theta}\sum_{i=1}^{n} \ln[f_{Y_i}(y_i|\theta)]$. By the Cramer–Rao Inequality Theorem it can be shown, as in Casella and Berger (2002), that

1. $E_\theta[S(\theta|\mathbf{Y})] = 0$
2. $Var_\theta[S(\theta|\mathbf{Y})] = E\left\{\frac{\partial}{\partial \theta}\ln[f_{\mathbf{Y}}(\mathbf{y}|\theta)]\right\}^2 = nE\left\{\frac{\partial}{\partial \theta}\ln[f_Y(y|\theta)]\right\}^2 \equiv nI_1(\theta)$

By the Central Limit Theorem, $\frac{S(\theta|\mathbf{Y})}{\sqrt{I_n(\theta)}} \xrightarrow{d} N(0,1)$ as $n \to \infty$. That is, the distribution of the score function divided by the square root of the Fisher information based on all n observations converges to a $N(0,1)$ distribution as $n \to \infty$. This forms the basis for the score test.

Consider testing $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$. For this simple hypothesis, the parameter space is $\Theta = \{\theta_0\}$, a singleton. When H_0 is assumed to be true, the score statistic is $Z_n^s = \frac{S(\theta_0|\mathbf{Y})}{\sqrt{I_n(\theta_0)}} \xrightarrow{d} N(0,1)$ as $n \to \infty$, and the rejection region is $R = \{|(Z_n^s)| \geq Z_{1-\frac{\alpha}{2}}\}$. An alternative formula is given as $\xi_S = S(\Theta_0; y)^T J^{-1}(\Theta_0) S(\Theta_0; y)$, where $S(\Theta_0; y)$ is the score vector and $J^{-1}(\Theta_0)$ is the inverse of the expected Fisher information matrix, both evaluated under the null parameter space $\theta \in \Theta_0 \subseteq \Re^1$. It can be shown that $\xi_S \xrightarrow{d} \chi_1^2$ as $n \to \infty$ and the rejection region is $R = \{\xi_S > \chi^2_{1,1-\alpha}\}$.

Suppose a vector of parameters, say $\boldsymbol{\theta}$, is of interest. Then instead of the first deriviatve we have a vector of first partial derivatives. This is sometimes referred to as the gradient vector. For d parameters in the model, define this vector as

$$\nabla \mathscr{L}(\boldsymbol{\theta}) \equiv \begin{bmatrix} \dfrac{\partial \mathscr{L}(\boldsymbol{\theta})}{\partial \theta_1} \\ \dfrac{\partial \mathscr{L}(\boldsymbol{\theta})}{\partial \theta_2} \\ \vdots \\ \dfrac{\partial \mathscr{L}(\boldsymbol{\theta})}{\partial \theta_d} \end{bmatrix}$$

where $\mathscr{L}(\boldsymbol{\theta})$ is the log-likelihood function. Similarly, instead of one second derivative we have a matrix of second partial deriviatives

$$\nabla^2 \mathscr{L}(\boldsymbol{\theta}) \equiv \begin{bmatrix} \frac{\partial^2 \mathscr{L}(\boldsymbol{\theta})}{\partial \theta_1 \partial \theta_1} & \frac{\partial^2 \mathscr{L}(\boldsymbol{\theta})}{\partial \theta_1 \partial \theta_2} & \cdots & \cdots & \frac{\partial^2 \mathscr{L}(\boldsymbol{\theta})}{\partial \theta_1 \partial \theta_d} \\ \frac{\partial^2 \mathscr{L}(\boldsymbol{\theta})}{\partial \theta_2 \partial \theta_1} & \frac{\partial^2 \mathscr{L}(\boldsymbol{\theta})}{\partial \theta_2 \partial \theta_2} & \cdots & \cdots & \frac{\partial^2 \mathscr{L}(\boldsymbol{\theta})}{\partial \theta_2 \partial \theta_d} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ \frac{\partial^2 \mathscr{L}(\boldsymbol{\theta})}{\partial \theta_d \partial \theta_1} & \frac{\partial^2 \mathscr{L}(\boldsymbol{\theta})}{\partial \theta_d \partial \theta_2} & \cdots & \cdots & \frac{\partial^2 \mathscr{L}(\boldsymbol{\theta})}{\partial \theta_d \partial \theta_d} \end{bmatrix}.$$

As with the one-parameter case, we can write $Var\left\{\nabla \mathscr{L}(\boldsymbol{\theta})\right\} = -E\left\{\nabla^2 \mathscr{L}(\boldsymbol{\theta})\right\} = -E\left\{\frac{\partial^2 \mathscr{L}(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j}\right\}$, where $i, j = 1, 2, \cdots, d$. This is the expected Fisher information and is typically denoted as $I_n(\boldsymbol{\theta})$. Thus, in the multivariate setting, the score statistic can be defined as $\xi_S = \left[\nabla \mathscr{L}(\boldsymbol{\theta_0})\right]^T I_n^{-1}(\boldsymbol{\theta_0}) \left[\nabla \mathscr{L}(\boldsymbol{\theta_0})\right]$, evaluated under the null parameter space $\boldsymbol{\theta_0} \in \Re^d$. As seen before, $\xi_S \xrightarrow{d} \chi^2_1$ as $n \to \infty$ and the rejection region is $R = \{\xi_S > \chi^2_{1,1-\alpha}\}$. It should be noted that the degrees of freedom is one as I will only be assessing one parameter under the null hypothesis for each score test. Furthermore, for this dissertation, this score statistic will be used to assess independence of response variables (without covariates) under a variety of different scenarios and will also be applied to real world data.

Chapter 2

The Score Test for Independence of Two Marginal Poisson Variables

We are interested in developing a score test based on Lakshminarayana's bivariate Poisson probability mass function (Lakshminarayana et al., 1999). This bivariate distribution, for which the marginals are Poisson, is created as a product of Poisson marginals with a multiplicative dependency parameter λ. Furthermore, the parameterization given by Lakshminarayana (Lakshminarayana et al., 1999) is more flexible compared to the parameterizations of other bivariate Poisson distributions in that the correlation between the two variables can be positive, zero, or negative, depending on the value of λ. In healthcare, for example, bivariate Poisson data arise when examining the number emergency visits and the number of inpatient visits of individual patients. In marketing, the bivariate Poisson distribution can be useful to model the number of purchases of substitute products. In the former example, the correlation between visit types is positive. In the latter case, the correlation between numbers of purchases of substitute products is expected to be negative (Yahav and Shmueli, 2007).

This chapter is organized as follows: in Section 2.1 I develop the score test for whether two independent Poisson models should be modeled as correlated Poisson counts using Lakshminarayana's bivariate Poisson probability mass function. In Section 2.2 a simulation study is presented to show the estimated significance level and power of the score test when compared to the likelihood ratio and Wald tests. In

Section 2.3 a real world example will be provided to show the application and implementation of the test. I will conclude in Section 2.4.

2.1 The Score Test

Let $f(y_1, y_2)$ be the joint probability mass function of (y_1, y_2) with marginal distributions $f(y_1)$ and $f(y_2)$ respectively. In particular, let $f(y_1)$ be a Poisson probability mass function with mean and variance λ_1, and let $f(y_2)$ be a Poisson probability mass function with mean and variance λ_2. Thus, $0 \leq y_1 < \infty; \lambda_1 > 0$ and $0 \leq y_2 < \infty; \lambda_2 > 0$. The bivariate Poisson probability mass function as given by Lakshminarayana et al. (1999) is

$$f(y_1, y_2) = \frac{\lambda_1^{y_1} \lambda_2^{y_2} e^{-\lambda_1 - \lambda_2} \left[1 + \lambda \left(e^{-y_1} - e^{-(1-e^{-1})\lambda_1}\right)\left(e^{-y_2} - e^{-(1-e^{-1})\lambda_2}\right)\right]}{y_1! y_2!}, \quad (2.1)$$

where λ is the multiplicative dependency parameter. The quantity $e^{-(1-e^{-1})\lambda_t}(t = 1, 2)$ is the expectation $E(e^{-Y_t})$ under the Poisson marginal distribution. This term guarantees that the distribution defined in Equation 2.1 has Poisson marginals for values of λ for which the quantity in brackets is non-negative. The covariance between Y_1 and Y_2 is $\lambda\lambda_1\lambda_2(1 - e^{-1})^2 e^{-(1-e^{-1})(\lambda_1+\lambda_2)}$. Thus, the correlation coefficient is $\rho = \lambda\sqrt{\lambda_1\lambda_2}(1 - e^{-1})^2 e^{-(1-e^{-1})(\lambda_1+\lambda_2)}$, which can take on both positive and negative values depending on λ (Famoye, 2010). Also, as described in Lakshminarayana et al. (1999), λ should lie in the range $|\lambda| \leq \frac{1}{(1-A)(1-B)}$ and ρ should lie in the range $|\rho| \leq \frac{\sqrt{\lambda_1\lambda_2}AB(1-e^{-1})^2}{(1-A)(1-B)}$, where $A = e^{-\lambda_1(1-e^{-1})}$ and $B = e^{-\lambda_2(1-e^{-1})}$.

We are interested in testing $H_0 : \lambda = 0$ versus $H_1 : \lambda \neq 0$. When the processes being studied are correlated and follow a bivariate Poisson distribution, e.g. bivariate growth models, it is of interest to fit such a model. However, when the processes aren't correlated (i.e. independent) then reducing a bivariate Poisson distribution to the product of independent marginal Poisson distributions is appropriate. If $\lambda = 0$, then

$f(y_1, y_2) = \frac{\lambda_1^{y_1}\lambda_2^{y_2}e^{-\lambda_1-\lambda_2}}{y_1!} = \frac{e^{-\lambda_1}\lambda_1^{y_1}}{y_1!} \cdot \frac{e^{-\lambda_2}\lambda_2^{y_2}}{y_2!} = f(y_1)f(y_2)$, and so the joint distribution of (y_1, y_2) reduces to the product of the marginals, and Y_1 and Y_2 are independently Poisson distributed.

Consider n observations $[(y_{i1}, y_{i2}); i = 1, 2, \ldots, n]$ comprising a random sample from the bivariate Poisson distribution

$$f(y_{i1}, y_{i2}) = \frac{\lambda_1^{y_{i1}}\lambda_2^{y_{i2}}e^{-\lambda_1-\lambda_2}[1 + \lambda(e^{-y_{i1}} - e^{-(1-e^{-1})\lambda_1})(e^{-y_{i2}} - e^{-(1-e^{-1})\lambda_2})]}{y_{i1}!y_{i2}!}.$$

The log-likelihood is given by

$$\begin{aligned}\mathscr{L} &= \sum_{i=1}^{n} \ln[f(y_{i1}, y_{i2})] \\ &= \sum_{i=1}^{n}\{y_{i1}\ln(\lambda_1) + y_{i2}\ln(\lambda_2) - \lambda_1 - \lambda_2 \\ &\quad +\ln[1 + \lambda(e^{-y_{i1}} - e^{-(1-e^{-1})\lambda_1})(e^{-y_{i2}} - e^{-(1-e^{-1})\lambda_2})] - \ln(y_{i1}!y_{i2}!)\}.\end{aligned}$$

The first partial derivatives of $\mathscr{L}$ with respect to λ_1, λ_2, and λ are given by

1. $\frac{\partial \mathscr{L}}{\partial \lambda_1} = \frac{\lambda_1(A+B)+C}{D}$

   a) $A = -2\lambda \exp\left(\left(\frac{1}{e} - 1\right)\lambda_1 + \left(\frac{1}{e} - 1\right)\lambda_2 + y_1 + y_2 + 1\right) - e\lambda$
   $+ \lambda e^{-\frac{(e-1)(\lambda_1+\lambda_2)}{e}+y_1+y_2} - \lambda e^{\left(\frac{1}{e}-1\right)\lambda_1+y_1}$

   b) $B = 2\lambda e^{\left(\frac{1}{e}-1\right)\lambda_1+y_1+1} + \lambda e^{\left(\frac{1}{e}-1\right)\lambda_2+y_2+1} - e^{y_1+y_2+1}$

   c) $C = ey_1\left(\lambda + \lambda e^{-\frac{(e-1)(\lambda_1+\lambda_2)}{e}+y_1+y_2} - \lambda e^{\left(\frac{1}{e}-1\right)\lambda_1+y_1} - \lambda e^{\left(\frac{1}{e}-1\right)\lambda_2+y_2} + e^{y_1+y_2}\right)$

   d) $D = e\lambda_1\left(\lambda + \lambda e^{-\frac{(e-1)(\lambda_1+\lambda_2)}{e}+y_1+y_2} - \lambda e^{\left(\frac{1}{e}-1\right)\lambda_1+y_1} - \lambda e^{\left(\frac{1}{e}-1\right)\lambda_2+y_2} + e^{y_1+y_2}\right)$

2. $\frac{\partial \mathscr{L}}{\partial \lambda_2} = \frac{\lambda_2(E+F)+G}{H}$

   a) $E = -2\lambda \exp\left(\left(\frac{1}{e} - 1\right)\lambda_1 + \left(\frac{1}{e} - 1\right)\lambda_2 + y_1 + y_2 + 1\right) - e\lambda$
   $+ \lambda e^{-\frac{(e-1)(\lambda_1+\lambda_2)}{e}+y_1+y_2} + \lambda e^{\left(\frac{1}{e}-1\right)\lambda_1+y_1+1} - \lambda e^{\left(\frac{1}{e}-1\right)\lambda_2+y_2}$

   b) $F = 2\lambda e^{\left(\frac{1}{e}-1\right)\lambda_2+y_2+1} - e^{y_1+y_2+1}$

c) $G = ey_2 \left(\lambda + \lambda e^{-\frac{(e-1)(\lambda_1+\lambda_2)}{e}+y_1+y_2} - \lambda e^{\left(\frac{1}{e}-1\right)\lambda_1+y_1} - \lambda e^{\left(\frac{1}{e}-1\right)\lambda_2+y_2} + e^{y_1+y_2} \right)$

d) $H = e\lambda_2 \left(\lambda + \lambda e^{-\frac{(e-1)(\lambda_1+\lambda_2)}{e}+y_1+y_2} - \lambda e^{\left(\frac{1}{e}-1\right)\lambda_1+y_1} - \lambda e^{\left(\frac{1}{e}-1\right)\lambda_2+y_2} + e^{y_1+y_2} \right)$

3. $\frac{\partial \mathscr{L}}{\partial \lambda} = \frac{\left(e^{-y_1}-e^{\left(\frac{1}{e}-1\right)\lambda_1}\right)\left(e^{-y_2}-e^{\left(\frac{1}{e}-1\right)\lambda_2}\right)}{\lambda\left(e^{-y_1}-e^{\left(\frac{1}{e}-1\right)\lambda_1}\right)\left(e^{-y_2}-e^{\left(\frac{1}{e}-1\right)\lambda_2}\right)+1}$

The first partial derivatives of $\mathscr{L}$ with respect to λ_1, λ_2, and λ, evaluated under the restriction that $\lambda = 0$, are given by

$$\left.\frac{\partial \mathscr{L}}{\partial \lambda_1}\right|_{\lambda=0} = \sum_{i=1}^{n} \left[\frac{y_{i1}}{\lambda_1} - 1 \right] \tag{2.2}$$

$$\left.\frac{\partial \mathscr{L}}{\partial \lambda_2}\right|_{\lambda=0} = \sum_{i=1}^{n} \left[\frac{y_{i2}}{\lambda_2} - 1 \right] \tag{2.3}$$

$$\left.\frac{\partial \mathscr{L}}{\partial \lambda}\right|_{\lambda=0} = \sum_{i=1}^{n} \left(e^{-y_{i1}} - e^{\left(e^{-1}-1\right)\lambda_1} \right) \left(e^{-y_{i2}} - e^{\left(e^{-1}-1\right)\lambda_2} \right)$$

Setting the score equations 2.2 and 2.3 to zero and solving for the unknown parameters yields the restricted maximum likelihood estimates of λ_1 and λ_2. These are given as $\widetilde{\lambda}_1 = \overline{y}_1$ and $\widetilde{\lambda}_2 = \overline{y}_2$, respectively. Furthermore, the null parameter space and gradient vector evaluated under such a space are $\boldsymbol{\theta_0} = \left[\widetilde{\lambda}_1 \ \widetilde{\lambda}_2 \ 0\right]^T = \left[\overline{y}_1 \ \overline{y}_2 \ 0\right]^T$ and $\nabla \mathscr{L}\left(\boldsymbol{\theta_0}\right) = \left[0 \ 0 \ \left.\frac{\partial \mathscr{L}}{\partial \lambda}\right|_{\lambda=0}\right]^T$, respectively. To calculate the expected Fisher information matrix, the second partial derivatives are needed along with their respective expectations evaluated under the restriction that $\lambda = 0$.

The second partial derivatives are given by:

1. $\frac{\partial^2 \mathscr{L}}{\partial \lambda_1 \partial \lambda_1^T} = -\frac{I+J}{K}$

a) $I = e^2 y_1 I^*$

b) $I^* = \left(\lambda + \lambda e^{-\frac{(e-1)(\lambda_1+\lambda_2)}{e}+y_1+y_2} - \lambda e^{\left(\frac{1}{e}-1\right)\lambda_1+y_1} - \lambda e^{\left(\frac{1}{e}-1\right)\lambda_2+y_2} + e^{y_1+y_2} \right)^2$

c) $J = J_1 J_2$

d) $J_1 = (e-1)^2\lambda_1^2\lambda e^{\left(\frac{1}{e}-1\right)\lambda_1 - 2\lambda_2 + y_1}$

e) $J_2 = \left(e^{\lambda_2} - e^{\frac{\lambda_2}{e}+y_2}\right)\left(e^{\lambda_2}\lambda + \lambda\left(-e^{\frac{\lambda_2}{e}+y_2}\right) + e^{\lambda_2+y_1+y_2}\right)$

f) $K = e^2\lambda_1^2 K^*$

g) $K^* = \left(\lambda + \lambda e^{-\frac{(e-1)(\lambda_1+\lambda_2)}{e}+y_1+y_2} - \lambda e^{\left(\frac{1}{e}-1\right)\lambda_1+y_1} - \lambda e^{\left(\frac{1}{e}-1\right)\lambda_2+y_2} + e^{y_1+y_2}\right)^2$

2. $\frac{\partial^2 \mathscr{L}}{\partial\lambda_2\partial\lambda_2^T} = -\frac{L+M}{N}$

a) $L = e^2 y_2 L^*$

b) $L^* = \left(\lambda + \lambda e^{-\frac{(e-1)(\lambda_1+\lambda_2)}{e}+y_1+y_2} - \lambda e^{\left(\frac{1}{e}-1\right)\lambda_1+y_1} - \lambda e^{\left(\frac{1}{e}-1\right)\lambda_2+y_2} + e^{y_1+y_2}\right)^2$

c) $M = M_1 M_2$

d) $M_1 = (e-1)^2\lambda_2^2\lambda e^{-2\lambda_1+\frac{\lambda_2}{e}-\lambda_2+y_2}\left(e^{\lambda_1} - e^{\frac{\lambda_1}{e}+y_1}\right)$

e) $M_2 = \left(e^{\lambda_1}\lambda + \lambda\left(-e^{\frac{\lambda_1}{e}+y_1}\right) + e^{\lambda_1+y_1+y_2}\right)$

f) $N = e^2\lambda_2^2 N^*$

g) $N^* = \left(\lambda + \lambda e^{-\frac{(e-1)(\lambda_1+\lambda_2)}{e}+y_1+y_2} - \lambda e^{\left(\frac{1}{e}-1\right)\lambda_1+y_1} - \lambda e^{\left(\frac{1}{e}-1\right)\lambda_2+y_2} + e^{y_1+y_2}\right)^2$

3. $\frac{\partial^2 \mathscr{L}}{\partial\lambda\partial\lambda^T} = -\frac{\left(e^{-y_1}-e^{\left(\frac{1}{e}-1\right)\lambda_1}\right)^2\left(e^{-y_2}-e^{\left(\frac{1}{e}-1\right)\lambda_2}\right)^2}{\left(\lambda\left(e^{-y_1}-e^{\left(\frac{1}{e}-1\right)\lambda_1}\right)\left(e^{-y_2}-e^{\left(\frac{1}{e}-1\right)\lambda_2}\right)+1\right)^2}$

4. $\frac{\partial^2 \mathscr{L}}{\partial\lambda_1\partial\lambda^T} = -\frac{(e-1)e^{\frac{\lambda_1}{e}-\lambda_1+2y_1+y_2-1}\left(e^{\left(\frac{1}{e}-1\right)\lambda_2+y_2}-1\right)}{\left(\lambda+\lambda e^{-\frac{(e-1)(\lambda_1+\lambda_2)}{e}+y_1+y_2}-\lambda e^{\left(\frac{1}{e}-1\right)\lambda_1+y_1}-\lambda e^{\left(\frac{1}{e}-1\right)\lambda_2+y_2}+e^{y_1+y_2}\right)^2}$

5. $\frac{\partial^2 \mathscr{L}}{\partial\lambda_2\partial\lambda^T} = -\frac{(e-1)e^{\frac{\lambda_2}{e}-\lambda_2+y_1+2y_2-1}\left(e^{\left(\frac{1}{e}-1\right)\lambda_1+y_1}-1\right)}{\left(\lambda+\lambda e^{-\frac{(e-1)(\lambda_1+\lambda_2)}{e}+y_1+y_2}-\lambda e^{\left(\frac{1}{e}-1\right)\lambda_1+y_1}-\lambda e^{\left(\frac{1}{e}-1\right)\lambda_2+y_2}+e^{y_1+y_2}\right)^2}$

6. $\frac{\partial^2 \mathscr{L}}{\partial\lambda_1\partial\lambda_2^T} = \frac{(e-1)^2\lambda\exp\left(\left(\frac{1}{e}-1\right)\lambda_1+\left(\frac{1}{e}-1\right)\lambda_2+2y_1+2y_2-2\right)}{\left(\lambda+\lambda e^{-\frac{(e-1)(\lambda_1+\lambda_2)}{e}+y_1+y_2}-\lambda e^{\left(\frac{1}{e}-1\right)\lambda_1+y_1}-\lambda e^{\left(\frac{1}{e}-1\right)\lambda_2+y_2}+e^{y_1+y_2}\right)^2}$

The second partial derivaitves under $\lambda = 0$ are given by:

1. $\frac{\partial^2 \mathscr{L}}{\partial\lambda_1\partial\lambda_1^T} = -\frac{y_1}{\lambda_1^2}$

2. $\frac{\partial^2 \mathscr{L}}{\partial\lambda_2 \partial\lambda_2^T} = -\frac{y_2}{\lambda_2^2}$

3. $\frac{\partial^2 \mathscr{L}}{\partial\lambda \partial\lambda^T} = -\left(e^{-y_1} - e^{\left(\frac{1}{e}-1\right)\lambda_1}\right)^2 \left(e^{-y_2} - e^{\left(\frac{1}{e}-1\right)\lambda_2}\right)^2$

4. $\frac{\partial^2 \mathscr{L}}{\partial\lambda_1 \partial\lambda^T} = -(e-1)e^{\frac{\lambda_1}{e}-\lambda_1-y_2-1}\left(e^{\left(\frac{1}{e}-1\right)\lambda_2+y_2} - 1\right)$

5. $\frac{\partial^2 \mathscr{L}}{\partial\lambda_2 \partial\lambda^T} = -(e-1)e^{\frac{\lambda_2}{e}-\lambda_2-y_1-1}\left(e^{\left(\frac{1}{e}-1\right)\lambda_1+y_1} - 1\right)$

6. $\frac{\partial^2 L}{\partial\lambda_1 \partial\lambda_2^T} = 0$

The expectation of the second partial derivatives evaluated under the restriction that $\lambda = 0$ are given by:

$$-E\left[\frac{\partial^2 \mathscr{L}}{\partial\lambda_1 \partial\lambda_1^T}\right]\Bigg|_{\lambda=0} = E\left[\sum_{i=1}^{n} \frac{y_{i1}}{\lambda_1^2}\right] \tag{2.4}$$

$$-E\left[\frac{\partial^2 \mathscr{L}}{\partial\lambda_2 \partial\lambda_2^T}\right]\Bigg|_{\lambda=0} = E\left[\sum_{i=1}^{n} \frac{y_{i2}}{\lambda_2^2}\right] \tag{2.5}$$

$$-E\left[\frac{\partial^2 \mathscr{L}}{\partial\lambda \partial\lambda^T}\right]\Bigg|_{\lambda=0} = \sum_{i=1}^{n}\left[e^{-\frac{2(e-1)(\lambda_1+\lambda_2)}{e}}\left(e^{\frac{(e-1)^2\lambda_1\lambda_1}{e^2}} - 1\right)\left(e^{\frac{(e-1)^2\lambda_2}{e^2}} - 1\right)\right] \tag{2.6}$$

$$-E\left[\frac{\partial^2 \mathscr{L}}{\partial\lambda_1 \partial\lambda^T}\right]\Bigg|_{\lambda=0} = -E\left[\frac{\partial^2 \mathscr{L}}{\partial\lambda \partial\lambda_1^T}\right]\Bigg|_{\lambda=0} = 0$$

$$-E\left[\frac{\partial^2 \mathscr{L}}{\partial\lambda_2 \partial\lambda^T}\right]\Bigg|_{\lambda=0} = -E\left[\frac{\partial^2 \mathscr{L}}{\partial\lambda \partial\lambda_2^T}\right]\Bigg|_{\lambda=0} = 0$$

$$-E\left[\frac{\partial^2 \mathscr{L}}{\partial\lambda_1 \partial\lambda_2^T}\right]\Bigg|_{\lambda=0} = -E\left[\frac{\partial^2 \mathscr{L}}{\partial\lambda_2 \partial\lambda_1^T}\right]\Bigg|_{\lambda=0} = 0$$

Under $\boldsymbol{\theta_0}$, the inverse of the expected Fisher information matrix, $I_n^{-1}(\boldsymbol{\theta_0})$, is expressed by the inverse of the diagonal matrix with entries given by Equations 2.4, 2.5, and 2.6. The score statistic for testing the null is then calculated as

$$\xi_S = \left[\nabla \mathscr{L}(\boldsymbol{\theta_0})\right]^T I_n^{-1}(\boldsymbol{\theta_0}) \left[\nabla \mathscr{L}(\boldsymbol{\theta_0})\right]$$
$$= \frac{\left\{\sum_{i=1}^n \left(e^{-y_{i1}} - e^{\left(e^{-1}-1\right)\widetilde{\lambda}_1}\right)\left(e^{-y_{i2}} - e^{\left(e^{-1}-1\right)\widetilde{\lambda}_2}\right)\right\}^2}{\sum_{i=1}^n \left[e^{-\frac{2(e-1)\left(\widetilde{\lambda}_1+\widetilde{\lambda}_2\right)}{e}}\left(e^{\frac{(e-1)^2\widetilde{\lambda}_1}{e^2}} - 1\right)\left(e^{\frac{(e-1)^2\widetilde{\lambda}_2}{e^2}} - 1\right)\right]}.$$

2.2 Simulation Results

To check the efficiency of the score test, a simulation study compared the likelihood ratio, Wald, and the proposed score test. In every iteration, the score, likelihood ratio, and Wald tests were calculated to test $\lambda = 0$. These procedures were repeated 1000 times independently for each value of λ_1 and λ_2, and the significance level (when $\lambda = 0$) and power (when $\lambda > 0$) were estimated for sample sizes $n = 10, 15, 25, 50, 100$ and 200. For each pair of means, I used 5 different correlations, ρ, between 0 and an upper limit defined as $\min(0.50, \text{maximum possible}\,\rho)$. Tables 2.1, 2.2, 2.3, and 2.4 summarize the results for the nominal significance level 0.05.

For small sample sizes ($n = 10, 15$ and 25), the score statistic performs better than both the likelihood ratio and Wald tests in maintaining the nominal significance level. In fact, the likelihood ratio and Wald tests appear to be extremely sensitive to rejecting the null hypothesis and thus both reject too often. For large sample sizes ($n = 50, 100$ and 200) and for a majority of cases, the score test maintains the nominal 0.05 significance level. The likelihood ratio and Wald tests also maintain the 0.05 error rate in most cases. For each sample size and each pair of λ_1 and λ_2, the power of the proposed test increased as ρ increased. For large samples, the power of the score test is close to that of the likelihood ratio and Wald tests.

2.3 Illustration

Indices based on larval fish abundances are often used to track changes in adult spawning stock biomass, and to define spawning times and areas (Smith and Moser, 2003). Plankton samples are typically collected using fine mesh nets towed from research vessels, and larval fish are removed then stored after sample preservation. Samples are frequently collected using the paired bongo net, which consists of two usually round net frames joined at a central point, and towed either obliquely or vertically through the water column (Habtes et al., 2014). Mesopelagic fish families such as the Myctophidae are some of the most specious and abundant in the worlds oceans (Smith and Moser, 2003). Previous analyses of larval fish assemblages from surveys in the northern Gulf of Mexico showed the Myctophidae to be the most abundant family in the dataset, in aggregate accounting for 14.6% of all larvae collected (Muhling et al., 2012). Myctophid larvae were also present in most samples collected, except those from shallower waters on the inner continental shelf.

In this example, I compared the abundances of myctophid larvae from samples taken from the spatially-independent sides of a paired bongo net. The use of this very abundant family reduced the risk of zero-inflation, which occurs commonly with the larvae of more rarely encountered taxa. The Myctophidae counts from the left and right sides of each bongo net are a result of 30+ years of sampling in the Gulf of Mexico, therefore both the time and the location of each paired sample (right versus left 'bongo') is identical. Furthermore, Y_1 is the count of myctophid larvae in the left side of the bongo net and Y_2 is the count of myctophid larvae sampled in the right. A total of 261 paired samples are used in this example (Lyczkowski-Shultz et al., 2013). From the aforementioned data, the p-value corresponding to our score test is $p < 0.001$. Thus, if the significance level 0.05 is used, our test strongly rejects the independence of Y_1 and Y_2. This result is not surprising. Even though each bongo net is being fished through a different area in the Gulf of Mexico, we would expect the number of Myctophidae in the left and right side of each net to be highly correlated.

2.4 Conclusions

I derived the score test for testing independence in Lakshminarayana's bivariate Poisson distribution. The performance of the proposed score test was examined and compared to that of the likelihood ratio and Wald tests under a variety of sample sizes and λ_1 and λ_2 values. The score test performs better in maintaining the nominal significance level as compared to the likelihood ratio and Wald tests in smaller samples ($n = 10, 15, 25, 50$). Larger values of correlation result in higher power for all sample sizes and all combinations of λ_1 and λ_2.

Table 2.1: Estimated significance level and power of tests for testing independence in Lakshminarayana's bivariate Poisson distribution at the nominal size $\alpha = 0.05$ for $n = 10$, $n = 15$ and $n = 25$, and varying values of λ_1 and λ_2

			$n = 10$			$n = 15$			$n = 25$		
λ_1	λ_2	ρ	LR	ξ_S	Wald	LR	ξ_S	Wald	LR	ξ_S	Wald
0.50	0.50	0.00	0.270	0.059	0.112	0.138	0.046	0.070	0.077	0.055	0.083
		0.12	0.257	0.062	0.115	0.148	0.067	0.089	0.097	0.085	0.133
		0.25	0.364	0.132	0.223	0.261	0.193	0.235	0.304	0.279	0.339
		0.38	0.472	0.260	0.365	0.462	0.395	0.429	0.625	0.607	0.654
		0.50	0.599	0.395	0.492	0.645	0.589	0.615	0.820	0.804	0.810
0.50	1.00	0.00	0.223	0.053	0.103	0.120	0.051	0.074	0.069	0.043	0.079
		0.12	0.274	0.077	0.118	0.201	0.096	0.095	0.181	0.132	0.186
		0.25	0.374	0.125	0.193	0.338	0.186	0.167	0.350	0.286	0.349
		0.38	0.540	0.218	0.282	0.533	0.349	0.293	0.678	0.585	0.551
		0.50	0.727	0.365	0.386	0.768	0.537	0.413	0.907	0.835	0.670
0.50	1.50	0.00	0.238	0.052	0.120	0.154	0.048	0.075	0.093	0.049	0.076
		0.12	0.291	0.078	0.122	0.208	0.074	0.071	0.173	0.098	0.144
		0.25	0.413	0.109	0.167	0.434	0.172	0.149	0.473	0.299	0.343
		0.38	0.597	0.205	0.227	0.691	0.322	0.218	0.807	0.569	0.470
		0.50	0.710	0.287	0.315	0.862	0.513	0.347	0.975	0.796	0.605
0.50	2.00	0.00	0.253	0.059	0.143	0.171	0.056	0.083	0.095	0.045	0.065
		0.11	0.298	0.049	0.126	0.241	0.054	0.087	0.218	0.087	0.132
		0.21	0.390	0.087	0.125	0.401	0.125	0.121	0.426	0.196	0.276
		0.32	0.482	0.153	0.190	0.606	0.237	0.192	0.793	0.412	0.514
		0.42	0.577	0.188	0.212	0.700	0.330	0.263	0.874	0.590	0.568
1.00	1.00	0.00	0.221	0.052	0.151	0.104	0.050	0.093	0.057	0.040	0.078
		0.12	0.250	0.074	0.149	0.121	0.077	0.128	0.142	0.108	0.178
		0.25	0.340	0.149	0.223	0.290	0.196	0.295	0.371	0.328	0.434
		0.38	0.498	0.271	0.366	0.584	0.449	0.532	0.723	0.664	0.748
		0.50	0.744	0.470	0.580	0.838	0.666	0.740	0.959	0.890	0.922

Table 2.2: Estimated significance level and power of tests for testing independence in Lakshminarayana's bivariate Poisson distribution at the nominal size $\alpha = 0.05$ for $n = 50$, $n = 100$ and $n = 200$, and varying values of λ_1 and λ_2

			$n = 50$			$n = 100$			$n = 200$		
λ_1	λ_2	ρ	LR	ξ_S	Wald	LR	ξ_S	Wald	LR	ξ_S	Wald
0.50	0.50	0.00	0.043	0.034	0.063	0.061	0.058	0.070	0.056	0.055	0.059
		0.12	0.126	0.125	0.134	0.199	0.202	0.212	0.396	0.396	0.398
		0.25	0.458	0.449	0.482	0.774	0.773	0.776	0.955	0.956	0.955
		0.38	0.866	0.861	0.867	0.987	0.988	0.987	1.000	1.000	1.000
		0.50	0.981	0.980	0.981	1.000	1.000	1.000	1.000	1.000	1.000
0.50	1.00	0.00	0.042	0.041	0.055	0.052	0.048	0.059	0.046	0.047	0.052
		0.12	0.202	0.183	0.254	0.329	0.301	0.364	0.552	0.549	0.576
		0.25	0.569	0.548	0.626	0.854	0.840	0.871	0.992	0.989	0.994
		0.38	0.908	0.880	0.912	0.996	0.993	0.996	1.000	1.000	1.000
		0.50	0.997	0.990	0.950	1.000	1.000	1.000	1.000	1.000	1.000
0.50	1.50	0.00	0.066	0.051	0.082	0.046	0.043	0.055	0.051	0.052	0.054
		0.12	0.175	0.139	0.220	0.293	0.269	0.348	0.505	0.477	0.533
		0.25	0.692	0.574	0.744	0.938	0.909	0.953	0.999	0.998	0.999
		0.38	0.974	0.935	0.935	1.000	0.999	0.999	1.000	1.000	1.000
		0.50	1.000	0.994	0.944	1.000	1.000	0.998	1.000	1.000	1.000
0.50	2.00	0.00	0.054	0.047	0.067	0.066	0.059	0.077	0.043	0.045	0.047
		0.11	0.216	0.132	0.281	0.331	0.303	0.412	0.569	0.547	0.641
		0.21	0.636	0.477	0.672	0.883	0.799	0.912	0.991	0.975	0.992
		0.32	0.955	0.797	0.928	1.000	0.989	1.000	1.000	1.000	1.000
		0.42	0.993	0.926	0.970	1.000	0.999	1.000	1.000	1.000	1.000
1.00	1.00	0.00	0.052	0.051	0.070	0.065	0.057	0.072	0.066	0.068	0.069
		0.12	0.209	0.202	0.235	0.325	0.310	0.344	0.581	0.573	0.592
		0.25	0.613	0.584	0.645	0.874	0.870	0.889	0.997	0.996	0.997
		0.38	0.964	0.949	0.967	1.000	0.998	1.000	1.000	1.000	1.000
		0.50	0.999	0.996	0.981	1.000	1.000	0.998	1.000	1.000	1.000

Table 2.3: Estimated significance level and power of tests for testing independence in Lakshminarayana's bivariate Poisson distribution at the nominal size $\alpha = 0.05$ for $n = 10$, $n = 15$ and $n = 25$, and varying values of λ_1 and λ_2

			$n = 10$			$n = 15$			$n = 25$		
λ_1	λ_2	ρ	LR	ξ_S	Wald	LR	ξ_S	Wald	LR	ξ_S	Wald
1.00	1.50	0.00	0.229	0.057	0.138	0.129	0.047	0.099	0.106	0.054	0.115
		0.09	0.252	0.058	0.167	0.152	0.059	0.110	0.133	0.093	0.159
		0.18	0.302	0.116	0.201	0.215	0.121	0.177	0.242	0.209	0.281
		0.26	0.417	0.182	0.277	0.378	0.249	0.340	0.434	0.353	0.482
		0.35	0.512	0.212	0.358	0.561	0.394	0.501	0.730	0.609	0.746
1.00	2.00	0.00	0.268	0.037	0.183	0.148	0.039	0.112	0.127	0.061	0.122
		0.06	0.240	0.059	0.144	0.153	0.066	0.118	0.124	0.073	0.129
		0.13	0.269	0.072	0.170	0.182	0.087	0.142	0.179	0.119	0.190
		0.19	0.328	0.111	0.205	0.253	0.134	0.203	0.317	0.223	0.306
		0.25	0.408	0.167	0.260	0.406	0.265	0.326	0.487	0.380	0.498
1.50	1.50	0.00	0.275	0.045	0.182	0.162	0.054	0.121	0.121	0.057	0.105
		0.06	0.289	0.073	0.198	0.168	0.070	0.142	0.099	0.056	0.107
		0.12	0.295	0.077	0.210	0.192	0.105	0.157	0.158	0.134	0.183
		0.18	0.380	0.119	0.283	0.253	0.135	0.221	0.288	0.222	0.323
		0.24	0.390	0.166	0.287	0.345	0.249	0.301	0.412	0.358	0.462
1.50	2.00	0.00	0.263	0.034	0.182	0.167	0.050	0.104	0.115	0.054	0.098
		0.04	0.290	0.055	0.213	0.158	0.055	0.118	0.119	0.065	0.119
		0.09	0.292	0.091	0.222	0.194	0.088	0.155	0.132	0.096	0.138
		0.13	0.327	0.090	0.241	0.240	0.103	0.186	0.210	0.156	0.237
		0.17	0.389	0.141	0.290	0.315	0.186	0.275	0.315	0.268	0.333
2.00	2.00	0.00	0.310	0.056	0.219	0.192	0.042	0.116	0.132	0.045	0.104
		0.03	0.354	0.054	0.264	0.202	0.067	0.148	0.128	0.070	0.106
		0.06	0.323	0.084	0.254	0.200	0.088	0.139	0.167	0.105	0.154
		0.09	0.339	0.086	0.260	0.225	0.103	0.184	0.140	0.104	0.135
		0.12	0.426	0.136	0.337	0.263	0.117	0.207	0.198	0.150	0.210

Table 2.4: Estimated significance level and power of tests for testing independence in Lakshminarayana's bivariate Poisson distribution at the nominal size $\alpha = 0.05$ for $n = 50$, $n = 100$ and $n = 200$, and varying values of λ_1 and λ_2

			$n = 50$			$n = 100$			$n = 200$		
λ_1	λ_2	ρ	LR	ξ_S	Wald	LR	ξ_S	Wald	LR	ξ_S	Wald
1.00	1.50	0.00	0.064	0.047	0.084	0.051	0.050	0.067	0.047	0.043	0.052
		0.09	0.142	0.133	0.177	0.218	0.220	0.236	0.399	0.395	0.410
		0.18	0.381	0.351	0.432	0.639	0.631	0.665	0.925	0.917	0.927
		0.26	0.693	0.646	0.732	0.941	0.926	0.953	1.000	1.000	1.000
		0.35	0.953	0.913	0.959	0.998	0.998	0.999	1.000	1.000	1.000
1.00	2.00	0.00	0.083	0.050	0.104	0.053	0.049	0.076	0.050	0.045	0.057
		0.06	0.106	0.096	0.131	0.146	0.133	0.172	0.230	0.225	0.243
		0.13	0.235	0.207	0.277	0.345	0.333	0.376	0.597	0.596	0.621
		0.19	0.482	0.412	0.506	0.771	0.728	0.800	0.960	0.951	0.963
		0.25	0.758	0.685	0.779	0.955	0.929	0.955	1.000	0.998	1.000
1.50	1.50	0.00	0.063	0.041	0.080	0.056	0.048	0.074	0.072	0.068	0.074
		0.06	0.093	0.079	0.112	0.131	0.133	0.140	0.238	0.233	0.238
		0.12	0.196	0.199	0.224	0.322	0.324	0.336	0.581	0.588	0.580
		0.18	0.438	0.412	0.494	0.741	0.716	0.748	0.962	0.945	0.961
		0.24	0.690	0.657	0.717	0.919	0.905	0.923	1.000	0.998	1.000
1.50	2.00	0.00	0.098	0.049	0.103	0.052	0.036	0.062	0.051	0.048	0.072
		0.04	0.095	0.078	0.117	0.119	0.123	0.130	0.144	0.164	0.143
		0.09	0.160	0.142	0.177	0.178	0.186	0.190	0.279	0.286	0.280
		0.13	0.268	0.256	0.296	0.437	0.444	0.446	0.732	0.715	0.729
		0.17	0.480	0.453	0.519	0.748	0.718	0.759	0.967	0.957	0.965
2.00	2.00	0.00	0.124	0.057	0.112	0.080	0.050	0.087	0.066	0.055	0.095
		0.03	0.097	0.060	0.101	0.070	0.059	0.077	0.086	0.089	0.086
		0.06	0.148	0.133	0.162	0.217	0.230	0.222	0.351	0.378	0.331
		0.09	0.153	0.165	0.169	0.261	0.275	0.264	0.457	0.474	0.443
		0.12	0.270	0.277	0.305	0.459	0.480	0.461	0.720	0.724	0.716

Chapter 3

Multiple Score Tests for a Bivariate Pareto Density Function

3.1 Introduction

The Score Test (Radhakrishna Rao, 1948) is a common large sample approache to hypothesis testing. This test can be used to evaluate a model parameter in the univariate case as well as to evaluate model parameters in a multivariate setting. It is well known that if we consider testing $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$, then the parameter space is a singleton, and when H_0 is assumed to be true, the score statistic is $Z_n^s \equiv \frac{S(\theta_0|\mathbf{Y})}{\sqrt{I_n(\theta_0)}} \xrightarrow{d} N(0,1)$ as $n \to \infty$, with corresponding rejection region $R \equiv \{|(Z_n^s)| \geq Z_{1-\frac{\alpha}{2}}\}$. An alternative formula is given as $\xi_S = S(\Theta_0; y)^T J^{-1}(\Theta_0) S(\Theta_0; y)$, where $S(\Theta_0; y)$ is the score vector and $J^{-1}(\Theta_0)$ is the inverse of the expected Fisher information matrix, both evaluated under the null parameter space $\theta \in \Theta_0 \subseteq \Re^1$. It can be shown that $\xi_S \xrightarrow{d} \chi_1^2$ as $n \to \infty$ and the rejection region is $R = \{\xi_S > \chi^2_{1,1-\alpha}\}$. The multivariate equivalent is $\xi_S = [\nabla \mathscr{L}(\boldsymbol{\theta_0})]^T I_n^{-1}(\boldsymbol{\theta_0}) [\nabla \mathscr{L}(\boldsymbol{\theta_0})]$, evaluated under the null parameter space $\boldsymbol{\theta_0} \in \Re^k$.

I am interested in developing two score tests based on Sankaran and Nair's bivariate Pareto probability density function (Sankaran and Nair, 1993) and the bivariate Pareto model introduced by Lindley–Singpurwalla (Lindley and Singpurwalla, 1986). In Section 3.2 a brief background and applicability of both parameterizations is given. In Section 3.3 two score tests are developed: one for testing whether two independent Pareto models should be modeled as correlated Pareto observations, and one for testing whether the more parsimonious bivariate Pareto model introduced by Lindley–Singpurwalla (Lindley and Singpurwalla, 1986) can be used to analyze the data. In Section 3.4 parameter estimation issues and data generation are discussed. In Section 3.5 simulation studies are presented to show the estimated significance level and power of the score tests, likelihood ratio, and Wald tests. In Section 3.6 a real world example will be provided to show the application and implementation of both tests. I will conclude in Section 3.7.

3.2 Background

Let $f(y_1, y_2)$ be the joint probability density function of (y_1, y_2) with marginal distributions $f(y_1)$ and $f(y_2)$, respectively. In particular, let $f(y_1)$ and $f(y_2)$ be Pareto II probability density functions. That is, for $y_i > 0, i = 1, 2,\ f_{Y_i}(y_i) = \frac{\alpha_i \theta}{(1+\alpha_i y_i)^{\theta+1}}$; α_i is the scale parameter and θ is the shape parameter. For $\theta > 1$ (the mean exists only if $\theta > 1$), $E(Y_1) = \frac{1}{\alpha_1(\theta-1)}$ and $E(Y_2) = \frac{1}{\alpha_2(\theta-1)}$. Similarly, for $\theta > 2$ (the variance exists only if $\theta > 2$), $Var(Y_1) = \frac{\theta}{(\theta-1)^2(\theta-2)\alpha_1^2}$ and $Var(Y_2) = \frac{\theta}{(\theta-1)^2(\theta-2)\alpha_2^2}$.

Lindley and Singpurwalla (1986) introduced a bivariate Pareto density, which has the following parameterization for $y_1 > 0, y_2 > 0, \alpha_1 > 0, \alpha_2 > 0$, and $\theta > 0$;

$$f(y_1, y_2) = \frac{\theta \alpha_1 \alpha_2 (\theta + 1)}{(1 + \alpha_1 y_1 + \alpha_2 y_2)^{\theta+2}}. \tag{3.1}$$

Lindley and Singpurwalla (1986) reviewed and discussed several useful results as well as important theoretical properties and justifications of the distribution given in Equation 3.1. Sankaran and Nair (1993) proposed a bivariate Pareto distribution which has the following parameterization for $y_1 > 0, y_2 > 0, \alpha_1 > 0, \alpha_2 > 0, \theta > 0$ and $0 \le \alpha_0 \le (\theta + 1)\alpha_1\alpha_2$,

$$f(y_1, y_2) = \frac{\theta \left[\theta(\alpha_1 + \alpha_0 y_2)(\alpha_2 + \alpha_0 y_1) + \alpha_1\alpha_2 - \alpha_0\right]}{(1 + \alpha_1 y_1 + \alpha_2 y_2 + \alpha_0 y_1 y_2)^{\theta+2}}. \tag{3.2}$$

To have a well-defined bivariate Pareto distribution, $\theta > 2$ so the second moments exist. It is clear that Lindley and Singpurwalla's distribution in Equation 3.1 can be obtained as a special case of Sankaran and Nair's distribution in Equation 3.2 by letting $\alpha_0 = 0$. Furthermore, I would like to test whether the submodel from Equation 3.1 can be used to analyze the data. A score test for $H_0 : \alpha_0 = 0$ versus $H_1 : \alpha_0 > 0$ is developed in Section 3.3. However, as described in Dykstra and El Barmi (1997), the usual limiting distribution of the likelihood ratio test statistic (χ_1^2) is not sufficient in the case where α_0 lies on the boundary of the hypothesized region. Under this scenario, the limiting distribution is known as "chi-bar-square", denoted as $\overline{\chi}_1^2$. Furthermore, $\xi_S \xrightarrow{d} \overline{\chi}_1^2$ as $n \to \infty$. Since the sizes in Tables 3.10, 3.11, 3.12, 3.13, 3.14, 3.15, 3.16, and 3.17 were calculated using the cutoff of $\chi^2_{1,0.95}$, these values will be compared to 0.025 instead of 0.05 for the likelihood ratio test results. The results for the proposed score test will be compared to the usual 0.05 level.

The marginals of Sankaran and Nair's distribution are also Pareto II marginals. Since this bivariate distribution has an extra parameter, α_0, it is more flexible than Lindley and Singpurwalla's distribution, and thus can be used more effectively in modeling bivariate survival or reliability data (Sankaran and Kundu, 2014). Lindley and Singpurwalla (1986) also discuss the relevance of the bivariate Pareto distribution, especially in regards to assessing the reliability of series and parallel systems based on their component reliabilities. Moreover, under certain cases, the joint life lengths of the components being tested have Pareto II marginals (Lindley and Singpurwalla, 1986).

From Equation 3.2, it easily follows that if $\alpha_0 = \alpha_1\alpha_2$, then $f(y_1, y_2) = f(y_1)f(y_2)$ and so Y_1 and Y_2 are independent. A score test for $H_0 : \alpha_0 = \alpha_1\alpha_2$ versus $H_1 : \alpha_0 \neq \alpha_1\alpha_2$ is developed in Section 3.3. Deriving a score test for independence is of practical importance since assuming responses are not correlated when in reality this may not be true could result in misleading conclusions. Testing for independence is one special feature of Sankaran and Nair's distribution that cannot be achieved using Lindley and Singpurwalla's. Also, when $\alpha_0 = \alpha_1\alpha_2$, it is clear that the correlation coefficient, ρ, between Y_1 and Y_2 is 0. For other values of α_0, ρ cannot be expressed in an explicit form. Sankaran and Kundu (2014) provide a more detailed explanation of the non-trivial relationship between the correlation coefficient of Y_1 and Y_2 for different values of θ and α_0 for when $\alpha_1 = \alpha_2 = 1$. Lastly, Lindley and Singpurwalla's distribution allows only positive correlation between Y_1 and Y_2. On the other hand, the more flexible model by Sankaran and Nair allows the correlation between Y_1 and Y_2 to be both positive and negative (Sankaran and Kundu, 2014). For more details about the correlation coefficient and its relationship to the model parameters, namely θ, see Balakrishnan and Lai (2009).

3.3 Score Tests

Testing $\alpha_0 = \alpha_1\alpha_2$

Consider n observations $[(y_{i1}, y_{i2}); i = 1, 2, \ldots, n]$ comprising a random sample from the bivariate Pareto distribution

$$f(y_{i1}, y_{i2}) = \frac{\theta\left[\theta(\alpha_1 + \alpha_0 y_{i2})(\alpha_2 + \alpha_0 y_{i1}) + \alpha_1\alpha_2 - \alpha_0\right]}{(1 + \alpha_1 yi1 + \alpha_2 y_{i2} + \alpha_0 y_{i1} y_{i2})^{\theta+2}}.$$

The log-likelihood is given by

$$\begin{aligned}\mathscr{L} \;&=\; \sum_{i=1}^{n}\Bigg\{\ln\theta + \ln\left[\theta\left(\alpha_1 + \alpha_0 y_{i2}\right)\left(\alpha_2 + \alpha_0 y_{i1}\right) + \alpha_1\alpha_2 - \alpha_0\right] \\ &\quad -(\theta+2)\ln\left(1 + \alpha_1 y_{i1} + \alpha_2 y_{i2} + \alpha_0 y_{i1} y_{i2}\right)\Bigg\}. \qquad (3.3)\end{aligned}$$

I am interested in deriving a score test for testing independence between Y_1 and Y_2 and thus testing $H_0 : \alpha_0 = \alpha_1\alpha_2$ versus $H_0 : \alpha_0 \neq \alpha_1\alpha_2$. The first partial derivatives of $\mathscr{L}$ with respect to θ, α_1, α_2, and α_0 evaluated under the restriction that $\alpha_0 = \alpha_1\alpha_2$ are given by

$$\left.\frac{\partial \mathscr{L}}{\partial \theta}\right|_{\alpha_0=\alpha_1\alpha_2} = \sum_{i=1}^{n}\left[\frac{2}{\theta} - \log\left(\left(\alpha_1 y_{i1}+1\right)\left(\alpha_2 y_{i2}+1\right)\right)\right] \tag{3.4}$$

$$\left.\frac{\partial \mathscr{L}}{\partial \alpha_1}\right|_{\alpha_0=\alpha_1\alpha_2} = \sum_{i=1}^{n} -\frac{\left(\theta+1\right)\left(\alpha_1\theta y_{i1}-1\right)}{\alpha_1\theta\left(\alpha_1 y_{i1}+1\right)\left(\alpha_2 y_{i2}+1\right)} \tag{3.5}$$

$$\left.\frac{\partial \mathscr{L}}{\partial \alpha_2}\right|_{\alpha_0=\alpha_1\alpha_2} = \sum_{i=1}^{n} -\frac{\left(\theta+1\right)\left(\alpha_2\theta y_{i2}-1\right)}{\alpha_2\theta\left(\alpha_1 y_{i1}+1\right)\left(\alpha_2 y_{i2}+1\right)} \tag{3.6}$$

$$\left.\frac{\partial \mathscr{L}}{\partial \alpha_0}\right|_{\alpha_0=\alpha_1\alpha_2} = \sum_{i=1}^{n} -\frac{\left(\alpha_1\theta y_{i1}-1\right)\left(\alpha_2\theta y_{i2}-1\right)}{\alpha_1\alpha_2\theta\left(\alpha_1 y_{i1}+1\right)\left(\alpha_2 y_{i2}+1\right)} \tag{3.7}$$

The null parameter space and gradient vector evaluated under such a space are $\boldsymbol{\theta_0} = \left[\widetilde{\theta}\ \widetilde{\alpha}_1\ \widetilde{\alpha}_2\ \widetilde{\alpha}_0\right]^T$ and $\nabla\mathscr{L}\left(\boldsymbol{\theta_0}\right)$, which is a 1×4 vector comprised of Equations 3.4, 3.5, 3.6, and 3.7. To calculate the expected Fisher information matrix, the second partial derivatives are needed along with their respective expectations evaluated under the restriction that $\alpha_0 = \alpha_1\alpha_2$. These are

$$-E\left[\frac{\partial^2 \mathscr{L}}{\partial\theta\partial\theta^T}\right]\Bigg|_{\alpha_0=\alpha_1\alpha_2} = \frac{2n}{\theta^2}$$

$$-E\left[\frac{\partial^2 \mathscr{L}}{\partial\alpha_1\partial\alpha_1^T}\right]\Bigg|_{\alpha_0=\alpha_1\alpha_2} = \frac{n(\theta+1)^2}{\alpha_1^2(\theta+2)^2}$$

$$-E\left[\frac{\partial^2 \mathscr{L}}{\partial\alpha_2\partial\alpha_2^T}\right]\Bigg|_{\alpha_0=\alpha_1\alpha_2} = \frac{n(\theta+1)^2}{\alpha_2^2(\theta+2)^2}$$

$$-E\left[\frac{\partial^2 \mathscr{L}}{\partial\alpha_0\partial\alpha_0^T}\right]\Bigg|_{\alpha_0=\alpha_1\alpha_2} = \frac{n}{\alpha_1^2\alpha_2^2(\theta+2)^2}$$

$$-E\left[\frac{\partial^2 \mathscr{L}}{\partial\theta\partial\alpha_1^T}\right]\Bigg|_{\alpha_0=\alpha_1\alpha_2} = -E\left[\frac{\partial^2 \mathscr{L}}{\partial\alpha_1\partial\theta^T}\right]\Bigg|_{\alpha_0=\alpha_1\alpha_2} = \frac{n}{\alpha_1\theta+\alpha_1}$$

$$-E\left[\frac{\partial^2 \mathscr{L}}{\partial\theta\partial\alpha_2^T}\right]\Big|_{\alpha_0=\alpha_1\alpha_2} = -E\left[\frac{\partial^2 \mathscr{L}}{\partial\alpha_2\partial\theta^T}\right]\Big|_{\alpha_0=\alpha_1\alpha_2} = \frac{n}{\alpha_2\theta+\alpha_2}$$

$$-E\left[\frac{\partial^2 \mathscr{L}}{\partial\theta\partial\alpha_0^T}\right]\Big|_{\alpha_0=\alpha_1\alpha_2} = -E\left[\frac{\partial^2 \mathscr{L}}{\partial\alpha_0\partial\theta^T}\right]\Big|_{\alpha_0=\alpha_1\alpha_2} = 0$$

$$-E\left[\frac{\partial^2 \mathscr{L}}{\partial\alpha_1\partial\alpha_2^T}\right]\Big|_{\alpha_0=\alpha_1\alpha_2} = -E\left[\frac{\partial^2 \mathscr{L}}{\partial\alpha_2\partial\alpha_1^T}\right]\Big|_{\alpha_0=\alpha_1\alpha_2} = \frac{n}{\alpha_1\alpha_2(\theta+2)^2}$$

$$-E\left[\frac{\partial^2 \mathscr{L}}{\partial\alpha_1\partial\alpha_0^T}\right]\Big|_{\alpha_0=\alpha_1\alpha_2} = -E\left[\frac{\partial^2 \mathscr{L}}{\partial\alpha_0\partial\alpha_1^T}\right]\Big|_{\alpha_0=\alpha_1\alpha_2} = \frac{-n}{\alpha_1^2\alpha_2(\theta+2)^2}$$

$$-E\left[\frac{\partial^2 \mathscr{L}}{\partial\alpha_2\partial\alpha_0^T}\right]\Big|_{\alpha_0=\alpha_1\alpha_2} = -E\left[\frac{\partial^2 \mathscr{L}}{\partial\alpha_0\partial\alpha_2^T}\right]\Big|_{\alpha_0=\alpha_1\alpha_2} = \frac{-n}{\alpha_1\alpha_2^2(\theta+2)^2}$$

The full score statistic for testing the null is then given by

$$\xi_{1A} = \frac{B\pi_1 + C\pi_2 + D\pi_3 + A\pi_4}{n} \tag{3.8}$$

where

1. $A = \sum_{i=1}^{n}\left[\frac{2}{\theta} - \log\left(\left(\alpha_1 y_{i1}+1\right)\left(\alpha_2 y_{i2}+1\right)\right)\right]$

2. $B = \sum_{i=1}^{n} -\frac{(\theta+1)(\alpha_1\theta y_{i1}-1)}{\alpha_1\theta(\alpha_1 y_{i1}+1)(\alpha_2 y_{i2}+1)}$

3. $C = \sum_{i=1}^{n} -\frac{(\theta+1)(\alpha_2\theta y_{i2}-1)}{\alpha_2\theta(\alpha_1 y_{i1}+1)(\alpha_2 y_{i2}+1)}$

4. $D = \sum_{i=1}^{n} -\frac{(\alpha_1\theta y_{i1}-1)(\alpha_2\theta y_{i2}-1)}{\alpha_1\alpha_2\theta(\alpha_1 y_{i1}+1)(\alpha_2 y_{i2}+1)}$

5. $\pi_1 = \frac{\theta+2}{2}\left[C\alpha_1\alpha_2(\theta+2) + \frac{B\left(\theta^2+2\theta+2\right)\alpha_1^2}{\theta} - A\theta(\theta+1)\alpha_1 + \frac{2D(\theta+1)^2\alpha_1^2\alpha_2}{\theta}\right]$

6. $\pi_2 = \frac{\theta+2}{2}\left[B\alpha_1\alpha_2(\theta+2) + \frac{C\left(\theta^2+2\theta+2\right)\alpha_2^2}{\theta} - A\theta(\theta+1)\alpha_2 + \frac{2D(\theta+1)^2\alpha_1\alpha_2^2}{\theta}\right]$

7. $\pi_3 = (\theta+2)\alpha_1\alpha_2\left[\pi_{31}+\pi_{32}\right]$

a) $\pi_{31} = \frac{D\left(3\theta^2+6\theta+2\right)\alpha_2\alpha_1}{\theta} + \frac{B(\theta+1)^2\alpha_1}{\theta}$

b) $\pi_{32} = \frac{C(\theta+1)^2\alpha_2}{\theta} - A\theta(\theta+1)$

8. $\pi_4 = \frac{1}{2}\left[\pi_{41} - \pi_{42}\right]$

   a) $\pi_{41} = A\theta^2(\theta+1)^2 - B\theta(\theta+2)\alpha_1(\theta+1)$

   b) $\pi_{42} = C\theta(\theta+2)\alpha_2(\theta+1) - 2D\theta(\theta+2)\alpha_1\alpha_2(\theta+1)$

It turns out that many of the terms above are zero when evaluated at $\alpha_0 = \alpha_1\alpha_2$, and so the full test statistic provided by Equation 3.8 can be reduced to

$$\xi_{1B} = \frac{D\pi_3}{n}. \tag{3.9}$$

Testing $\alpha_0 = 0$

I am interested in deriving a score test for testing whether the bivariate Pareto model introduced by Lindley–Singpurwalla can be used to analyze the data and thus testing $H_0 : \alpha_0 = 0$ versus. $H_0 : \alpha_0 > 0$. The first partial derivatives of $\mathscr{L}$ with respect to θ, α_1, α_2, and α_0 evaluated under the restriction that $\alpha_0 = 0$ are given by

$$\left.\frac{\partial \mathscr{L}}{\partial \theta}\right|_{\alpha_0=0} = \sum_{i=1}^{n}\left[\frac{1}{\theta} + \frac{1}{\theta+1} - \log\left(\alpha_1 y_{i1} + \alpha_2 y_{i2} + 1\right)\right] \tag{3.10}$$

$$\left.\frac{\partial \mathscr{L}}{\partial \alpha_1}\right|_{\alpha_0=0} = \sum_{i=1}^{n}\left[\frac{1}{\alpha_1} - \frac{(\theta+2)y_{i1}}{\alpha_1 y_{i1} + \alpha_2 y_{i2} + 1}\right] \tag{3.11}$$

$$\left.\frac{\partial \mathscr{L}}{\partial \alpha_2}\right|_{\alpha_0=0} = \sum_{i=1}^{n}\left[\frac{1}{\alpha_2} - \frac{(\theta+2)y_{i2}}{\alpha_1 y_{i1} + \alpha_2 y_{i2} + 1}\right] \tag{3.12}$$

$$\left.\frac{\partial \mathscr{L}}{\partial \alpha_0}\right|_{\alpha_0=0} = \sum_{i=1}^{n}\left[\frac{\alpha_1\theta y_{i1} - 1}{\alpha_1\alpha_2(\theta+1)} + y_{i2}\left(\frac{\theta}{\alpha_1(\theta+1)} - \frac{(\theta+2)y_{i1}}{\alpha_1 y_{i1} + \alpha_2 y_{i2} + 1}\right)\right] \tag{3.13}$$

The null parameter space and score vector evaluated under such a space are $\boldsymbol{\theta_0} = \left[\widetilde{\theta}\ \widetilde{\alpha}_1\ \widetilde{\alpha}_2\ \widetilde{\alpha}_0\right]^T$ and $\nabla\mathscr{L}\left(\boldsymbol{\theta_0}\right)$, which is a 1×4 vector comprised of Equations 3.10, 3.11, 3.12, and 3.13. To calculate the expected Fisher information matrix, the second partial derivatives are needed along with their respective expectations evaluated under the restriction that $\alpha_0 = 0$. These are

$$-E\left[\frac{\partial^2\mathscr{L}}{\partial\theta\partial\theta^T}\right]\Bigg|_{\alpha_0=0} = n\left(\frac{1}{\theta^2} + \frac{1}{(\theta+1)^2}\right)$$

$$-E\left[\frac{\partial^2\mathscr{L}}{\partial\alpha_1\partial\alpha_1^T}\right]\Bigg|_{\alpha_0=0} = \frac{n\left(\theta+1\right)}{\alpha_1^2(\theta+3)}$$

$$-E\left[\frac{\partial^2\mathscr{L}}{\partial\alpha_2\partial\alpha_2^T}\right]\Bigg|_{\alpha_0=0} = \frac{n\left(\theta+1\right)}{\alpha_2^2(\theta+3)}$$

$$-E\left[\frac{\partial^2\mathscr{L}}{\partial\alpha_0\partial\alpha_0^T}\right]\Bigg|_{\alpha_0=0} = \frac{n\left(\theta^2+\theta+2\right)}{\alpha_1^2\alpha_2^2(\theta-2)(\theta-1)(\theta+1)(\theta+3)}$$

$$-E\left[\frac{\partial^2\mathscr{L}}{\partial\theta\partial\alpha_1^T}\right]\Bigg|_{\alpha_0=0} = -E\left[\frac{\partial^2\mathscr{L}}{\partial\alpha_1\partial\theta^T}\right]\Bigg|_{\alpha_0=0} = \frac{n}{\alpha_1(\theta+2)}$$

$$-E\left[\frac{\partial^2\mathscr{L}}{\partial\theta\partial\alpha_2^T}\right]\Bigg|_{\alpha_0=0} = -E\left[\frac{\partial^2\mathscr{L}}{\partial\alpha_2\partial\theta^T}\right]\Bigg|_{\alpha_0=0} = \frac{n}{\alpha_2(\theta+2)}$$

$$-E\left[\frac{\partial^2\mathscr{L}}{\partial\theta\partial\alpha_0^T}\right]\Bigg|_{\alpha_0=0} = -E\left[\frac{\partial^2\mathscr{L}}{\partial\alpha_0\partial\theta^T}\right]\Bigg|_{\alpha_0=0} = \frac{-n}{\alpha_1\alpha_2\left(\theta^3+2\theta^2-\theta-2\right)}$$

$$-E\left[\frac{\partial^2\mathscr{L}}{\partial\alpha_1\partial\alpha_2^T}\right]\Bigg|_{\alpha_0=0} = -E\left[\frac{\partial^2\mathscr{L}}{\partial\alpha_2\partial\alpha_1^T}\right]\Bigg|_{\alpha_0=0} = \frac{-n}{\alpha_1\alpha_2(\theta+3)}$$

$$-E\left[\frac{\partial^2\mathscr{L}}{\partial\alpha_1\partial\alpha_0^T}\right]\Bigg|_{\alpha_0=0} = -E\left[\frac{\partial^2\mathscr{L}}{\partial\alpha_0\partial\alpha_1^T}\right]\Bigg|_{\alpha_0=0} = \frac{-n}{\alpha_1^2\alpha_2\left(\theta^2+4\theta+3\right)}$$

$$-E\left[\frac{\partial^2\mathscr{L}}{\partial\alpha_2\partial\alpha_0^T}\right]\Bigg|_{\alpha_0=0} = -E\left[\frac{\partial^2\mathscr{L}}{\partial\alpha_0\partial\alpha_2^T}\right]\Bigg|_{\alpha_0=0} = \frac{-n}{\alpha_1\alpha_2^2\left(\theta^2+4\theta+3\right)}$$

The full score statistic for testing the null is calculated to be

$$\xi_{2A} = \frac{\gamma_{2A}^{*} - \gamma_{2A}^{**}}{\gamma_{11}} \tag{3.14}$$

where

- $\gamma_{2A}^{*} = \alpha_1^2(\theta+3)\left(\gamma_1+\gamma_2+\gamma_3\right)$
- $\gamma_{2A}^{**} = 2\alpha_1(\theta+3)\left[\gamma_4+\left(\gamma_5-\gamma_6\right)\alpha_2+\gamma_7\right]+\theta\left(\gamma_8-\gamma_9-\gamma_{10}\right)$
- $F = \frac{\partial \mathscr{L}}{\partial \theta}\Big|_{\alpha_0=0} = \sum_{i=1}^{n}\left[\frac{1}{\theta}+\frac{1}{\theta+1}-\log\left(\alpha_1 y_{i1}+\alpha_2 y_{i2}+1\right)\right]$
- $G = \frac{\partial \mathscr{L}}{\partial \alpha_1}\Big|_{\alpha_0=0} = \sum_{i=1}^{n}\left[\frac{1}{\alpha_1}-\frac{(\theta+2)y_{i1}}{\alpha_1 y_{i1}+\alpha_2 y_{i2}+1}\right]$
- $H = \frac{\partial \mathscr{L}}{\partial \alpha_2}\Big|_{\alpha_0=0} = \sum_{i=1}^{n}\left[\frac{1}{\alpha_2}-\frac{(\theta+2)y_{i2}}{\alpha_1 y_{i1}+\alpha_2 y_{i2}+1}\right]$
- $J = \frac{\partial \mathscr{L}}{\partial \alpha_0}\Big|_{\alpha_0=0} = \sum_{i=1}^{n}\left[\frac{\alpha_1\theta y_{i1}-1}{\alpha_1\alpha_2(\theta+1)}+y_{i2}\left(\frac{\theta}{\alpha_1(\theta+1)}-\frac{(\theta+2)y_{i1}}{\alpha_1 y_{i1}+\alpha_2 y_{i2}+1}\right)\right]$
- $\gamma_1 = G^2\theta\left(\theta^8+8\theta^7+27\theta^6+50\theta^5+54\theta^4+20\theta^3-40\theta^2-56\theta-32\right)$
- $\gamma_2 = 2\alpha_2 G\left(\theta^9+3\theta^8-5\theta^7-23\theta^6-4\theta^5+48\theta^4+40\theta^3-12\theta^2-32\theta-16\right)J$
- $\gamma_3 = \alpha_2^2\theta\left(\theta^2-1\right)^2\left(3\theta^4+6\theta^3-8\theta^2-24\theta-16\right)J^2$
- $\gamma_4 = FG\theta^2(\theta+1)^2(\theta+2)^3\left(\theta^3-\theta^2+2\theta-2\right)$
- $\gamma_5 = F(\theta-1)\theta^2(\theta+1)^3\left(\theta^2-4\right)^2 J$
- $\gamma_6 = G\left(\theta^9+8\theta^8+25\theta^7+40\theta^6+36\theta^5+14\theta^4-16\theta^3-44\theta^2-32\theta-16\right)H$
- $\gamma_7 = \alpha_2^2\left(-\theta^9-3\theta^8+5\theta^7+23\theta^6+4\theta^5-48\theta^4-40\theta^3+12\theta^2+32\theta+16\right)HJ$
- $\gamma_8 = F^2\theta(\theta+1)^2(\theta+2)^3\left(\theta^5+\theta^4-\theta^3+7\theta^2-12\theta+4\right)$
- $\gamma_9 = 2\alpha_2 F\theta(\theta+1)^2(\theta+2)^3\left(\theta^4+2\theta^3-\theta^2+4\theta-6\right)H$

- $\gamma_{10} = \alpha_2^2 \left(\gamma_{10}^*\right) H^2$

    1. $\gamma_{10}^* = \theta^9 + 11\theta^8 + 51\theta^7 + 131\theta^6 + 204\theta^5 + 182\theta^4 + 20\theta^3 - 176\theta^2 - 200\theta - 96$

- $\gamma_{11} = 2(\theta + 2)\left(\theta^7 + 5\theta^6 + 9\theta^5 + 3\theta^4 - 12\theta^3 - 6\theta^2 + 8\right)$

It turns out that many of the terms above are zero when evaluated at $\alpha_0 = 0$, and so the full test statistic provided by Equation 3.14 can be reduced to a more simplified version, given as

$$\xi_{2B} = \frac{\alpha_1^2(\theta + 3)\gamma_3}{n\gamma_{11}}.$$

3.4 Statistical Inference

The maximum likelihood estimates can be obtained by maximizing (3.3) with respect to the unknown parameters $\alpha_1, \alpha_2, \theta$, and α_0, where $\alpha_1 > 0, \alpha_2 > 0, \theta > 0$ and $0 \leq \alpha_0 \leq (\theta + 1)\alpha_1\alpha_2$. There are many methods to go about computing these estimates. The most direct technique is to compute the estimating equations by setting the first derivative with respect to each parameter equal to zero and solving for the respective parameters. However, as in most cases, these don't exist in closed form and numerical methods need to be used. Sankaran and Kundu (2014) propose using a two-step estimation procedure, which essentially reduces a four-dimensional optimization problem into solving a two-dimensional optimization problem. The two-stage estimators are consistent (Sankaran and Kundu, 2014) and the asymptotic distribution of these estimators can be obtained as described in Joe (2005). For this chapter, a standard implementation of Nelder and Mead (Nelder and Mead, 1965) in R is utilized to obtain the maximum likelihood estimates.

As with obtaining the maximum likelihood estimates, there are multiple ways to generate correlated, bivariate Pareto data. Sankaran and Kundu (2014) generate such data via acceptance-rejection sampling. For this chapter, a common but easy to implement sampling technique via the inverse cumulative distribution function (CDF) is used. First, y_2 is generated via the inverse CDF of $y_2 = F^{-1}_{Y_2}(u_2)$. Given each value of y_2, $F^{-1}_{Y_1|Y_2=y_2}(u_1)$ is utilized to come up with the corresponding y_1 values. These paired observations comprise a correlated bivariate Pareto sample from two independently generated uniform(0,1) vectors u_1 and u_2.

3.5 Simulation Results

To check the efficiency of the proposed score test for testing $\alpha_0 = \alpha_1\alpha_2$, multiple simulation studies were run and the results were compared to the likelihood ratio and Wald tests. These procedures were repeated 5000 times independently for each value of α_1, α_2, θ, and α_0, and the significance level (when $\alpha_0 = \alpha_1\alpha_2$) and power (when $\alpha_0 > \alpha_1\alpha_2$) were estimated for sample sizes $n = 10, 15, 25, 50, 100, 250$ and 500. Tables 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, and 3.9 summarize the results for the nominal significance level 0.05.

In tables 3.2 and 3.3, our proposed score test is compared to the likelihood ratio and Wald tests when α_1 and α_2 are held constant at 0.50 and θ varies from 2.5 to 10.0. In tables 3.4 and 3.5, α_1 is held constant at 0.50, θ is held constant at 3.0, and α_2 varies from 1.0 to 6.0. In tables 3.6 and 3.7, α_1 is held constant at 0.50 and α_2 varies from 1.0 to 3.0 and θ varies from 2.5 to 10.0. In tables 3.8 and 3.9, α_1 varies from 1.0 to 4.0, α_2 varies from 1.0 to 6.0, and θ varies from 2.5 to 10.0. In all eight tables under each combination of α_1, α_2, θ, and α_0 the simulation results are relatively consistent in that the proposed score statistic, ξ_{1B}, remains close to the nominal size for large samples ($n = 100, 250, 500$). This performance matches that of the likelihood ratio and Wald tests. In a majority of cases where the sample sizes are small, our score statistic performs better than both the likelihood ratio and Wald tests. Under each combination of α_1, α_2 and θ, as α_0 and n increase, the power of the proposed test increased as well. However, as seen in tables 3.8 and 3.9 where

$\alpha_1 = 4.0$, $\alpha_2 = 6.0$, and $\theta = 10.0$, the power of each test under all sample sizes is much smaller compared to other combinations of α_1, α_2, θ, and α_0. It appears that as α_1, α_2, θ, and α_0 increase simultaneously, there is a reduction in the power of each test. In general, the more parameters that are fixed, the higher the power as sample size increases.

To check the efficiency of the proposed score test for testing $\alpha_0 = 0$, multiple simulation studies were run and the results were compared to the likelihood ratio test. These procedures were also repeated 5000 times independently for each value of α_1, α_2, θ, and α_0, and the significance level (when $\alpha_0 = 0$) and power (when $\alpha_0 > 0$) were estimated for sample sizes $n = 10, 15, 25, 50, 100, 250$ and 500. Tables 3.10, 3.11, 3.12, 3.13, 3.14, 3.15, 3.16, 3.17 summarize the results for the nominal significance level 0.025 (likelihood ratio) and 0.05 (score).

In tables 3.10 and 3.11, our proposed score test is compared to the likelihood ratio test when α_1 and α_2 are held constant at 0.50 and θ varies from 2.5 to 10.0. In tables 3.12 and 3.13, α_1 is held constant at 0.50, θ is held constant at 3.0, and α_2 varies from 1.0 to 6.0. In tables 3.14 and 3.15, α_1 is held constant at 0.50 and α_2 varies from 1.0 to 3.0 and θ varies from 2.5 to 10.0. In tables 3.16 and 3.17, α_1 varies from 1.0 to 4.0, α_2 varies from 1.0 to 6.0, and θ varies from 2.5 to 10.0. Both tests perform poorly in small samples ($n = 10, 15, 25, 50$) and noticeably better in larger samples. For all eight of these tables, as the sample size increases, the size of the likelihood ratio test approaches the nominal size of 0.025 and our score test, ξ_{2B}, approaches the nominal size of 0.05. As I saw when testing $\alpha_0 = \alpha_1\alpha_2$, under each combination of α_1, α_2 and θ, as α_0 and n increase, the power of the proposed test increased as well. However, in tables 3.16 and 3.17 where $\alpha_1 = 4.0$, $\alpha_2 = 6.0$, and $\theta = 10.0$, the power of each test under all sample sizes is much smaller compared to other combinations of α_1, α_2, θ, and α_0. Once again, as more parameters are varied together and increase simultaneously, power is lost regardless of sample size.

3.6 Illustration

Many scientists and evolutionary biologists are interested in investigating the process of evolution, specifically the characteristics that particular species need to withstand such a process (Ramsey and Schafer, 2012). In this regard, a variable of interest is brain size. One might conjecture that the larger the brain, the smarter the species and thus the species will be able to survive longer and have a better chance of evolving (Ramsey and Schafer, 2012). Many variables could effect brain size, including body weight, gestation period, and litter size (Ramsey and Schafer, 2012). The purpose of this illustration is not to determine which of these variables are associated with brain size (though this is an interesting question), but instead to explore the relationship between body weight, gestation period, and litter.

The following data set was originally reported in Sacher and Staffeldt (1974) and consists of the average brain size, average body weight, average gestation period, and average litter size for 96 species. Measurements made on the first five species in this data set can be seen in Table 3.1.

Table 3.1: Measurements Made on 5 Different Species

Species	Brain	Body	Gestation	Litter
Aardvark	9.60	2.20	31	5.0
Acouchis	9.90	0.78	98	1.2
African Elephant	4480.00	2800	655	1.0
Agoutis	20.30	2.80	104	1.3
Axis Deer	219.00	89.00	218	1.0

For this example, I analyzed the bivariate outcomes of the average body weight and the average gestation period. Furthermore, Y_1 is the average body weight and Y_2 is the average gestation period. A total of 96 paired observations are used in this example. From the aforementioned data, for testing $H_0 : \alpha_0 = 0$ versus $H_1 : \alpha_0 > 0$, the p-value is $p < 0.001$. Thus, if the significance level 0.05 is used, my test strongly rejects fitting Lindley and Singpurwalla's distribution to these data, and concludes the bivariate distribution given by Sankaran and Nair is more appropriate. Furthermore, since Sankaran and Nair's distribution better fits these data, I can conduct a test for independence using the test statistic as given by Equation 3.8 or Equation 3.9. The p-value for testing $H_0 : \alpha_0 = \alpha_1\alpha_2$ versus. $H_1 : \alpha_0 \neq \alpha_1\alpha_2$ is $p < 0.001$. Thus, at the 0.05 significance level, our test strongly rejects the independence of Y_1 and Y_2. This is not surprising as we would expect the body weight and gestation period made on the same species to be correlated.

3.7 Conclusions

I derived two score tests: one for testing independence based on Sankaran and Nair's bivariate Pareto distribution and one for testing whether Sankaran and Nair's parameterization reduces to the more popular bivariate Pareto distribution introduced by Lindley–Singpurwalla. The performace of both score tests were examined and compared to that of the likelihood ratio and Wald tests under a variety of different sample sizes and different values of α_1, α_2, θ, and α_0. For testing both $\alpha_0 = \alpha_1\alpha_2$, the score tests perform better in maintaining the nominal significance level in smaller samples. For testing $\alpha_0 = \alpha_1\alpha_2$, our test performs as well as the likelihood ratio and Wald tests in larger samples. Similarly, for testing $\alpha_0 = 0$, as sample size increases, the likelihood ratio test approaches the nominal 0.025 level and the score test approaches the nominal 0.05 level Both tests perform poorly in small samples. As sample size and α_0 increase, so does power. However, power tends to decrease as more parameters are simultaneously varied and increased.

Table 3.2: Estimated significance level and power of tests for testing independence in Sankaran and Nair's bivariate Pareto distribution at the nominal size $\alpha = 0.05$ for $n = 10, 15, 25$, $\alpha_1 = \alpha_2 = 0.50$, and increasing values of θ.

				$n = 10$			$n = 15$			$n = 25$		
α_1	α_2	θ	α_0	ξ_{1B}	LR	Wald	ξ_{1B}	LR	Wald	ξ_{1B}	LR	Wald
0.5	0.5	2.5	0.25	0.052	0.112	0.123	0.056	0.086	0.098	0.048	0.057	0.081
			0.75	0.103	0.371	0.189	0.165	0.375	0.249	0.302	0.441	0.438
			1.25	0.232	0.659	0.217	0.421	0.769	0.343	0.695	0.889	0.587
			1.75	0.357	0.831	0.217	0.603	0.927	0.328	0.857	0.984	0.545
			2.25	0.444	0.910	0.204	0.709	0.966	0.336	0.926	0.994	0.518
			5.25	0.687	0.975	0.176	0.904	0.997	0.251	0.991	0.999	0.376
			10.75	0.757	0.985	0.136	0.939	0.998	0.180	0.998	1	0.237
0.5	0.5	3.0	0.25	0.050	0.110	0.153	0.056	0.084	0.120	0.058	0.061	0.100
			0.75	0.083	0.306	0.207	0.135	0.312	0.252	0.245	0.376	0.410
			1.25	0.200	0.586	0.247	0.345	0.687	0.388	0.600	0.813	0.626
			1.75	0.312	0.792	0.254	0.517	0.881	0.388	0.820	0.967	0.604
			2.25	0.393	0.884	0.256	0.639	0.956	0.374	0.893	0.991	0.584
			5.25	0.638	0.973	0.223	0.872	0.994	0.308	0.986	0.999	0.465
			10.75	0.745	0.987	0.171	0.938	0.999	0.229	0.996	1	0.309
0.5	0.5	4.0	0.25	0.049	0.097	0.175	0.058	0.074	0.152	0.056	0.055	0.109
			0.75	0.058	0.261	0.255	0.091	0.242	0.272	0.172	0.288	0.351
			1.25	0.108	0.460	0.313	0.224	0.515	0.393	0.450	0.683	0.587
			1.75	0.209	0.675	0.334	0.388	0.779	0.444	0.685	0.909	0.630
			2.25	0.292	0.818	0.328	0.513	0.903	0.450	0.806	0.978	0.651
			5.25	0.554	0.973	0.287	0.820	0.993	0.407	0.971	0.998	0.596
			10.75	0.713	0.989	0.189	0.914	0.998	0.276	0.993	1	0.402
0.5	0.5	5.0	0.25	0.043	0.087	0.195	0.047	0.060	0.178	0.052	0.050	0.133
			0.75	0.045	0.208	0.284	0.068	0.208	0.298	0.124	0.242	0.336
			1.25	0.084	0.370	0.339	0.158	0.422	0.418	0.333	0.561	0.520
			1.75	0.146	0.567	0.394	0.282	0.652	0.497	0.541	0.817	0.673
			2.25	0.209	0.732	0.384	0.410	0.836	0.522	0.719	0.954	0.715
			5.25	0.490	0.963	0.291	0.760	0.993	0.432	0.951	0.998	0.642
			10.75	0.665	0.989	0.215	0.893	0.998	0.296	0.988	1	0.468
0.5	0.5	10.0	0.25	0.051	0.074	0.261	0.052	0.056	0.243	0.070	0.048	0.206
			0.75	0.044	0.125	0.311	0.045	0.106	0.297	0.060	0.123	0.274
			1.25	0.051	0.198	0.349	0.066	0.211	0.369	0.116	0.283	0.370
			1.75	0.061	0.280	0.388	0.106	0.319	0.420	0.214	0.450	0.460
			2.25	0.096	0.370	0.428	0.156	0.459	0.477	0.316	0.621	0.531
			5.25	0.250	0.861	0.418	0.451	0.941	0.571	0.744	0.981	0.750
			10.75	0.458	0.964	0.300	0.716	0.982	0.456	0.923	0.989	0.669

Table 3.3: Estimated significance level and power of tests for testing independence in Sankaran and Nair's bivariate Pareto distribution at the nominal size $\alpha = 0.05$ for $n = 50, 100, 250, 500$, $\alpha_1 = \alpha_2 = 0.50$, and increasing values of θ.

				$n = 50$			$n = 100$			$n = 250$			$n = 500$		
α_1	α_2	θ	α_0	ξ_{1B}	LR	Wald	ξ_{1B}	LR	Wald	ξ_{1B}	LR	Wald	ξ_{1B}	LR	Wald
0.5	0.5	2.5	0.25	0.057	0.081	0.048	0.051	0.066	0.053	0.055	0.056	0.046	0.046	0.042	0.049
			0.75	0.608	0.688	0.787	0.902	0.931	0.964	1	1	1	1	1	1
			1.25	0.953	0.989	0.887	1	1	0.990	1	1	1	1	1	1
			1.75	0.995	1	0.842	1	1	0.976	1	1	1	1	1	1
			2.25	0.998	1	0.816	1	1	0.967	1	1	0.999	1	1	1
			5.25	1	1	0.608	1	1	0.824	1	1	0.984	1	1	1
			10.75	1	1	0.361	1	1	0.573	1	1	0.857	1	1	0.979
0.5	0.5	3.0	0.25	0.048	0.052	0.067	0.049	0.052	0.058	0.042	0.044	0.041	0.051	0.052	0.048
			0.75	0.522	0.611	0.712	0.829	0.869	0.926	0.999	1	1	1	1	1
			1.25	0.924	0.972	0.932	0.999	1	0.998	1	1	1	1	1	1
			1.75	0.986	0.999	0.872	1	1	0.990	1	1	1	1	1	1
			2.25	0.996	1	0.864	1	1	0.982	1	1	1	1	1	1
			5.25	1	1	0.695	1	1	0.895	1	1	0.995	1	1	1
			10.75	1	1	0.485	1	1	0.716	1	1	0.939	1	1	0.995
0.5	0.5	4.0	0.25	0.054	0.049	0.076	0.053	0.051	0.059	0.049	0.050	0.047	0.049	0.052	0.047
			0.75	0.379	0.468	0.544	0.708	0.762	0.839	0.983	0.986	0.994	1	1	1
			1.25	0.827	0.911	0.871	0.988	0.996	0.988	1	1	1	1	1	1
			1.75	0.961	0.994	0.914	1	1	0.995	1	1	1	1	1	1
			2.25	0.989	0.999	0.898	1	1	0.992	1	1	1	1	1	1
			5.25	1	0.999	0.825	1	1	0.964	1	1	1	1	1	1
			10.75	1	1	0.639	1	1	0.845	1	1	0.985	1	1	1
0.5	0.5	5.0	0.25	0.050	0.046	0.084	0.055	0.053	0.057	0.054	0.054	0.056	0.050	0.051	0.047
			0.75	0.336	0.286	0.387	0.455	0.578	0.644	0.716	0.945	0.955	0.981	1	1
			1.25	0.520	0.700	0.829	0.770	0.964	0.982	0.952	1	1	0.999	1	1
			1.75	0.673	0.911	0.979	0.881	0.998	0.999	0.981	1	1	1	1	1
			2.25	0.715	0.968	0.997	0.921	0.999	0.999	0.991	1	1	1	1	1
			5.25	0.642	0.999	1	0.884	1	1	0.982	1	1	1	1	1
			10.75	1	1	0.723	1	1	0.918	1	1	0.998	1	1	1
0.5	0.5	10.0	0.25	0.055	0.042	0.144	0.051	0.038	0.096	0.050	0.043	0.063	0.054	0.052	0.058
			0.75	0.102	0.180	0.256	0.223	0.303	0.336	0.583	0.630	0.654	0.882	0.896	0.916
			1.25	0.287	0.458	0.424	0.622	0.734	0.609	0.977	0.988	0.908	1	1	0.988
			1.75	0.520	0.720	0.557	0.887	0.950	0.740	0.998	0.998	0.942	1	1	0.996
			2.25	0.697	0.896	0.657	0.955	0.987	0.792	0.999	0.999	0.955	1	1	0.997
			5.25	0.961	0.991	0.923	0.996	0.997	0.989	1	0.999	1	1	1	1
			10.75	0.996	0.997	0.900	0.999	0.999	0.991	1	1	1	1	1	1

Table 3.4: Estimated significance level and power of tests for testing independence in Sankaran and Nair's bivariate Pareto distribution at the nominal size $\alpha = 0.05$ for $n = 10, 15, 25$, $\alpha_1 = \theta = 0.50$, and increasing values of α_2.

				$n = 10$			$n = 15$			$n = 25$		
α_1	α_2	θ	α_0	ξ_{1B}	LR	Wald	ξ_{1B}	LR	Wald	ξ_{1B}	LR	Wald
0.5	1.0	3.0	0.5	0.049	0.105	0.139	0.057	0.085	0.113	0.057	0.063	0.092
			1.0	0.078	0.308	0.225	0.138	0.330	0.263	0.254	0.389	0.406
			1.5	0.078	0.308	0.225	0.138	0.330	0.263	0.254	0.389	0.406
			2.0	0.134	0.433	0.245	0.230	0.499	0.333	0.436	0.627	0.567
			2.5	0.194	0.586	0.261	0.346	0.684	0.394	0.607	0.818	0.624
			5.5	0.464	0.931	0.267	0.714	0.979	0.387	0.939	0.996	0.585
			11.0	0.636	0.975	0.217	0.878	0.996	0.323	0.991	1	0.493
0.5	1.5	3.0	0.75	0.045	0.105	0.140	0.050	0.077	0.121	0.053	0.059	0.096
			1.25	0.046	0.159	0.174	0.051	0.132	0.173	0.070	0.121	0.186
			1.75	0.059	0.229	0.211	0.084	0.220	0.226	0.130	0.221	0.292
			2.25	0.078	0.299	0.236	0.127	0.318	0.283	0.249	0.391	0.415
			2.75	0.113	0.400	0.256	0.190	0.442	0.351	0.376	0.556	0.527
			5.75	0.322	0.831	0.279	0.561	0.922	0.405	0.851	0.984	0.631
			11.25	0.547	0.955	0.241	0.807	0.989	0.371	0.971	1	0.574
0.5	2.0	3.0	1.0	0.043	0.111	0.151	0.054	0.077	0.126	0.052	0.063	0.093
			1.5	0.039	0.139	0.165	0.046	0.103	0.153	0.057	0.099	0.161
			2.0	0.045	0.188	0.195	0.064	0.172	0.199	0.098	0.166	0.242
			2.5	0.057	0.242	0.208	0.098	0.238	0.243	0.164	0.274	0.333
			3.0	0.076	0.310	0.224	0.136	0.319	0.285	0.254	0.388	0.431
			6.0	0.240	0.690	0.297	0.433	0.814	0.413	0.724	0.930	0.657
			11.5	0.471	0.930	0.256	0.720	0.980	0.393	0.940	0.998	0.590
0.5	3.0	3.0	1.0	0.044	0.108	0.152	0.046	0.078	0.129	0.053	0.061	0.103
			1.5	0.039	0.124	0.166	0.049	0.100	0.142	0.055	0.085	0.138
			2.0	0.039	0.149	0.183	0.048	0.128	0.167	0.073	0.120	0.189
			2.5	0.052	0.197	0.190	0.066	0.172	0.210	0.098	0.170	0.238
			3.0	0.050	0.235	0.195	0.082	0.220	0.235	0.140	0.231	0.303
			6.0	0.148	0.493	0.263	0.247	0.555	0.382	0.495	0.699	0.606
			11.5	0.323	0.830	0.260	0.575	0.929	0.399	0.848	0.981	0.613
0.5	6.0	3.0	3.0	0.048	0.104	0.137	0.049	0.086	0.119	0.043	0.056	0.090
			3.5	0.044	0.118	0.147	0.049	0.094	0.132	0.047	0.067	0.112
			4.0	0.042	0.136	0.149	0.050	0.103	0.144	0.050	0.080	0.131
			4.5	0.047	0.144	0.168	0.048	0.125	0.160	0.065	0.105	0.164
			5.0	0.047	0.157	0.170	0.057	0.138	0.165	0.068	0.119	0.176
			8.0	0.065	0.269	0.215	0.106	0.268	0.244	0.192	0.312	0.349
			13.5	0.140	0.481	0.251	0.266	0.563	0.375	0.504	0.705	0.599

Table 3.5: Estimated significance level and power of tests for testing independence in Sankaran and Nair's bivariate Pareto distribution at the nominal size $\alpha = 0.05$ for $n = 50, 100, 250, 500$, $\alpha_1 = \theta = 0.50$, and increasing values of α_2.

				$n = 50$			$n = 100$			$n = 250$			$n = 500$		
α_1	α_2	θ	α_0	ξ_{1B}	LR	Wald	ξ_{1B}	LR	Wald	ξ_{1B}	LR	Wald	ξ_{1B}	LR	Wald
0.5	1.0	3.0	0.5	0.048	0.050	0.068	0.050	0.052	0.056	0.051	0.052	0.045	0.054	0.053	0.043
			1.0	0.528	0.615	0.711	0.842	0.877	0.933	0.997	0.998	0.999	1	1	1
			1.5	0.528	0.615	0.711	0.842	0.877	0.933	0.997	0.998	0.999	1	1	1
			2.0	0.801	0.881	0.894	0.980	0.991	0.996	1	1	1	1	1	1
			2.5	0.925	0.970	0.924	0.999	1	0.997	1	1	1	1	1	1
			5.5	0.999	1	0.842	1	1	0.976	1	1	1	1	1	1
			11.0	1	1	0.717	1	1	0.909	1	1	0.996	1	1	1
0.5	1.5	3.0	0.75	0.052	0.059	0.070	0.046	0.048	0.047	0.051	0.050	0.050	0.052	0.051	0.045
			1.25	0.119	0.147	0.231	0.204	0.229	0.343	0.454	0.469	0.570	0.742	0.755	0.816
			1.75	0.301	0.359	0.486	0.553	0.594	0.726	0.926	0.939	0.966	0.998	0.998	0.999
			2.25	0.524	0.616	0.708	0.833	0.874	0.933	0.996	0.997	0.999	1	1	1
			2.75	0.714	0.796	0.837	0.952	0.973	0.986	1	1	1	1	1	1
			5.75	0.994	1	0.890	1	1	0.986	1	1	1	1	1	1
			11.25	0.999	1	0.807	1	1	0.962	1	1	1	1	1	1
0.5	2.0	3.0	1.0	0.048	0.054	0.068	0.050	0.052	0.061	0.053	0.054	0.051	0.050	0.049	0.048
			1.5	0.082	0.102	0.173	0.128	0.147	0.244	0.287	0.303	0.399	0.532	0.542	0.624
			2.0	0.194	0.236	0.358	0.371	0.408	0.551	0.756	0.775	0.849	0.965	0.967	0.978
			2.5	0.363	0.434	0.553	0.637	0.684	0.798	0.961	0.968	0.983	1	1	1
			3.0	0.522	0.606	0.709	0.841	0.878	0.935	0.999	0.999	1	1	1	1
			6.0	0.972	0.994	0.911	1	1	0.993	1	1	1	1	1	1
			11.5	0.999	1	0.844	1	1	0.976	1	1	1	1	1	1
0.5	3.0	3.0	1.0	0.053	0.051	0.065	0.051	0.052	0.060	0.048	0.051	0.049	0.054	0.054	0.047
			1.5	0.066	0.082	0.136	0.078	0.092	0.161	0.162	0.171	0.248	0.281	0.293	0.362
			2.0	0.110	0.142	0.232	0.198	0.227	0.340	0.455	0.479	0.589	0.749	0.764	0.819
			2.5	0.192	0.245	0.368	0.373	0.408	0.541	0.757	0.773	0.849	0.966	0.970	0.981
			3.0	0.296	0.360	0.485	0.556	0.599	0.727	0.927	0.940	0.966	0.999	0.999	0.999
			6.0	0.860	0.926	0.920	0.990	0.997	0.998	1	1	1	1	1	1
			11.5	0.994	1	0.882	1	1	0.989	1	1	1	1	1	1
0.5	6.0	3.0	3.0	0.051	0.050	0.068	0.050	0.057	0.059	0.048	0.049	0.047	0.055	0.055	0.050
			3.5	0.055	0.065	0.102	0.059	0.065	0.104	0.079	0.085	0.125	0.105	0.109	0.156
			4.0	0.059	0.075	0.132	0.089	0.103	0.171	0.159	0.171	0.238	0.287	0.299	0.369
			4.5	0.085	0.105	0.183	0.144	0.166	0.262	0.288	0.302	0.407	0.532	0.544	0.626
			5.0	0.116	0.148	0.250	0.213	0.231	0.347	0.456	0.477	0.586	0.750	0.761	0.818
			8.0	0.403	0.481	0.607	0.728	0.766	0.860	0.987	0.991	0.996	1	1	1
			13.5	0.851	0.922	0.907	0.991	0.996	0.996	1	1	1	1	1	1

Table 3.6: Estimated significance level and power of tests for testing independence in Sankaran and Nair's bivariate Pareto distribution at the nominal size $\alpha = 0.05$ for $n = 10, 15, 25$, $\alpha_1 = 0.50$, and increasing values of α_2 and θ.

				$n=10$			$n=15$			$n=25$		
α_1	α_2	θ	α_0	ξ_{1B}	LR	Wald	ξ_{1B}	LR	Wald	ξ_{1B}	LR	Wald
0.5	1.0	2.5	0.5	0.051	0.116	0.128	0.056	0.093	0.095	0.058	0.071	0.084
			1.0	0.060	0.226	0.161	0.079	0.202	0.172	0.120	0.192	0.241
			1.5	0.100	0.360	0.202	0.172	0.383	0.253	0.311	0.456	0.442
			2.0	0.165	0.507	0.215	0.295	0.589	0.317	0.523	0.722	0.594
			2.5	0.242	0.670	0.221	0.419	0.779	0.347	0.693	0.892	0.595
			5.5	0.529	0.941	0.216	0.768	0.983	0.331	0.958	0.998	0.511
			11.0	0.672	0.974	0.193	0.897	0.996	0.271	0.993	1	0.405
0.5	1.5	3.0	0.75	0.045	0.105	0.140	0.050	0.077	0.121	0.053	0.059	0.096
			1.25	0.046	0.159	0.174	0.051	0.132	0.173	0.070	0.121	0.186
			1.75	0.059	0.229	0.211	0.084	0.220	0.226	0.130	0.221	0.292
			2.25	0.078	0.299	0.236	0.127	0.318	0.283	0.249	0.391	0.415
			2.75	0.113	0.400	0.256	0.190	0.442	0.351	0.376	0.556	0.527
			5.75	0.322	0.831	0.279	0.561	0.922	0.405	0.851	0.984	0.631
			11.25	0.547	0.955	0.241	0.807	0.989	0.371	0.971	1	0.574
0.5	2.0	4.0	1.0	0.042	0.098	0.180	0.051	0.071	0.160	0.052	0.055	0.115
			1.5	0.039	0.116	0.192	0.040	0.086	0.188	0.049	0.086	0.166
			2.0	0.042	0.157	0.220	0.051	0.144	0.220	0.070	0.138	0.216
			2.5	0.045	0.194	0.237	0.071	0.190	0.242	0.114	0.211	0.295
			3.0	0.053	0.248	0.251	0.089	0.248	0.291	0.170	0.297	0.367
			6.0	0.161	0.562	0.326	0.296	0.671	0.431	0.577	0.825	0.643
			11.5	0.370	0.896	0.293	0.612	0.958	0.434	0.876	0.991	0.645
0.5	2.5	6.0	1.25	0.050	0.089	0.228	0.053	0.066	0.203	0.059	0.053	0.157
			1.75	0.046	0.099	0.231	0.050	0.078	0.205	0.054	0.069	0.178
			2.25	0.044	0.111	0.240	0.045	0.097	0.223	0.057	0.095	0.195
			2.75	0.052	0.136	0.259	0.047	0.119	0.240	0.066	0.123	0.224
			3.25	0.041	0.165	0.266	0.056	0.153	0.272	0.076	0.154	0.257
			6.25	0.069	0.310	0.324	0.111	0.345	0.389	0.257	0.471	0.456
			11.75	0.171	0.624	0.370	0.315	0.749	0.497	0.613	0.898	0.661
0.5	3.0	10.0	1.5	0.057	0.075	0.273	0.064	0.067	0.252	0.057	0.043	0.199
			2.0	0.055	0.076	0.272	0.058	0.070	0.247	0.059	0.055	0.215
			2.5	0.052	0.097	0.278	0.053	0.076	0.253	0.052	0.066	0.215
			3.0	0.059	0.103	0.272	0.053	0.086	0.264	0.055	0.083	0.220
			3.5	0.050	0.107	0.293	0.056	0.100	0.264	0.050	0.094	0.246
			6.5	0.046	0.177	0.335	0.059	0.177	0.328	0.092	0.222	0.322
			12.0	0.071	0.301	0.380	0.117	0.377	0.436	0.247	0.522	0.471

Table 3.7: Estimated significance level and power of tests for testing independence in Sankaran and Nair's bivariate Pareto distribution at the nominal size $\alpha = 0.05$ for $n = 50, 100, 250, 500$, $\alpha_1 = 0.50$, and increasing values of α_2 and θ.

				$n=50$			$n=100$			$n=250$			$n=500$		
α_1	α_2	θ	α_0	ξ_{1B}	LR	Wald	ξ_{1B}	LR	Wald	ξ_{1B}	LR	Wald	ξ_{1B}	LR	Wald
0.5	1.0	2.5	0.5	0.047	0.053	0.069	0.049	0.054	0.055	0.050	0.053	0.046	0.053	0.053	0.045
			1.0	0.236	0.282	0.403	0.437	0.465	0.598	0.832	0.846	0.900	0.985	0.987	0.993
			1.5	0.620	0.703	0.786	0.902	0.931	0.965	0.999	0.999	1	1	1	1
			2.0	0.867	0.934	0.923	0.992	0.997	0.998	1	1	1	1	1	1
			2.5	0.954	0.988	0.883	1	1	0.992	1	1	0.999	1	1	1
			5.5	0.999	1	0.775	1	1	0.947	1	1	0.999	1	1	1
			11.0	1	1	0.608	1	1	0.818	1	1	0.983	1	1	1
0.5	1.5	3.0	0.75	0.052	0.059	0.070	0.046	0.048	0.047	0.051	0.050	0.050	0.052	0.051	0.045
			1.25	0.119	0.147	0.231	0.204	0.229	0.343	0.454	0.469	0.570	0.742	0.755	0.816
			1.75	0.301	0.359	0.486	0.553	0.594	0.726	0.926	0.939	0.966	0.998	0.998	0.999
			2.25	0.524	0.616	0.708	0.833	0.874	0.933	0.996	0.997	0.999	1	1	1
			2.75	0.714	0.796	0.837	0.952	0.973	0.986	1	1	1	1	1	1
			5.75	0.994	1	0.890	1	1	0.986	1	1	1	1	1	1
			11.25	0.999	1	0.807	1	1	0.962	1	1	1	1	1	1
0.5	2.0	4.0	1.0	0.049	0.050	0.075	0.049	0.050	0.063	0.052	0.053	0.053	0.049	0.051	0.048
			1.5	0.067	0.087	0.148	0.095	0.116	0.199	0.213	0.230	0.327	0.399	0.414	0.509
			2.0	0.131	0.184	0.272	0.268	0.301	0.431	0.613	0.642	0.742	0.888	0.897	0.934
			2.5	0.258	0.336	0.433	0.484	0.539	0.665	0.888	0.903	0.947	0.996	0.997	0.998
			3.0	0.374	0.478	0.558	0.710	0.758	0.839	0.982	0.986	0.995	0.9997	1	1
			6.0	0.920	0.978	0.905	0.998	1	0.994	1	1	1	1	1	1
			11.5	0.995	0.999	0.893	1	1	0.990	1	1	1	1	1	1
0.5	2.5	6.0	1.25	0.059	0.043	0.101	0.046	0.043	0.074	0.050	0.049	0.056	0.052	0.051	0.054
			1.75	0.051	0.070	0.135	0.056	0.073	0.129	0.107	0.122	0.191	0.170	0.186	0.277
			2.25	0.072	0.104	0.183	0.120	0.152	0.226	0.276	0.305	0.423	0.504	0.525	0.640
			2.75	0.111	0.168	0.245	0.207	0.259	0.341	0.516	0.560	0.669	0.821	0.836	0.902
			3.25	0.155	0.238	0.306	0.333	0.395	0.471	0.739	0.767	0.857	0.965	0.969	0.984
			6.25	0.594	0.746	0.670	0.914	0.954	0.890	1	1	0.994	1	1	1
			11.75	0.927	0.988	0.879	0.998	0.997	0.978	1	1	0.999	1	1	1
0.5	3.0	10.0	1.5	0.061	0.045	0.152	0.058	0.042	0.108	0.047	0.040	0.085	0.053	0.053	0.070
			2.0	0.050	0.059	0.169	0.053	0.058	0.132	0.063	0.071	0.127	0.063	0.071	0.127
			2.5	0.049	0.068	0.170	0.061	0.084	0.157	0.111	0.132	0.189	0.204	0.223	0.317
			3.0	0.059	0.090	0.196	0.090	0.133	0.201	0.192	0.222	0.297	0.369	0.396	0.514
			3.5	0.071	0.118	0.215	0.131	0.181	0.237	0.304	0.350	0.422	0.579	0.612	0.705
			6.5	0.214	0.353	0.363	0.508	0.619	0.538	0.929	0.952	0.875	0.998	0.999	0.982
			12.0	0.580	0.795	0.584	0.914	0.968	0.754	0.998	0.999	0.942	1	1	0.994

Table 3.8: Estimated significance level and power of tests for testing independence in Sankaran and Nair's bivariate Pareto distribution at the nominal size $\alpha = 0.05$ for $n = 10, 15, 25$, and increasing values of α_1, α_2 and θ.

				$n = 10$			$n = 15$			$n = 25$		
α_1	α_2	θ	α_0	ξ_{1B}	LR	Wald	ξ_{1B}	LR	Wald	ξ_{1B}	LR	Wald
1.0	1.0	2.5	1.0	0.047	0.107	0.119	0.054	0.085	0.097	0.047	0.056	0.078
			1.5	0.043	0.167	0.145	0.054	0.125	0.135	0.063	0.109	0.155
			2.0	0.057	0.218	0.160	0.076	0.200	0.193	0.125	0.201	0.243
			2.5	0.066	0.273	0.188	0.118	0.288	0.230	0.202	0.308	0.353
			3.0	0.094	0.358	0.193	0.166	0.368	0.273	0.310	0.452	0.452
			6.0	0.293	0.776	0.252	0.527	0.878	0.375	0.798	0.959	0.578
			11.5	0.513	0.941	0.227	0.777	0.985	0.355	0.954	0.999	0.538
1.5	2.0	3.0	3.0	0.049	0.105	0.140	0.051	0.083	0.123	0.056	0.061	0.094
			3.5	0.048	0.114	0.144	0.051	0.090	0.123	0.052	0.069	0.115
			4.0	0.047	0.132	0.160	0.049	0.107	0.145	0.052	0.081	0.136
			4.5	0.046	0.153	0.160	0.049	0.118	0.161	0.061	0.102	0.161
			5.0	0.044	0.154	0.170	0.053	0.139	0.174	0.069	0.118	0.178
			8.0	0.064	0.263	0.208	0.099	0.268	0.248	0.193	0.307	0.356
			13.5	0.136	0.490	0.254	0.266	0.562	0.346	0.506	0.713	0.594
2.0	2.5	4.0	5.0	0.056	0.101	0.182	0.063	0.078	0.161	0.058	0.056	0.119
			5.5	0.052	0.100	0.184	0.055	0.074	0.151	0.053	0.059	0.124
			6.0	0.050	0.109	0.189	0.048	0.081	0.170	0.048	0.061	0.138
			6.5	0.047	0.110	0.182	0.045	0.089	0.169	0.051	0.067	0.140
			7.0	0.048	0.130	0.202	0.050	0.095	0.178	0.050	0.069	0.148
			10.0	0.052	0.168	0.217	0.054	0.146	0.218	0.080	0.142	0.223
			15.5	0.066	0.256	0.231	0.091	0.246	0.277	0.165	0.287	0.328
2.5	3.0	6.0	7.5	0.058	0.087	0.228	0.056	0.064	0.210	0.066	0.049	0.169
			8.0	0.052	0.097	0.221	0.064	0.068	0.206	0.058	0.059	0.177
			8.5	0.061	0.093	0.227	0.058	0.067	0.208	0.061	0.059	0.165
			9.0	0.054	0.100	0.233	0.059	0.073	0.208	0.062	0.061	0.178
			9.5	0.058	0.104	0.222	0.058	0.082	0.210	0.055	0.059	0.176
			12.5	0.047	0.109	0.237	0.047	0.091	0.223	0.054	0.086	0.194
			18.0	0.052	0.146	0.247	0.054	0.139	0.255	0.066	0.131	0.245
4.0	6.0	10.0	24.0	0.070	0.081	0.268	0.062	0.059	0.257	0.075	0.050	0.217
			24.5	0.066	0.082	0.265	0.069	0.057	0.250	0.066	0.043	0.202
			25.0	0.067	0.078	0.273	0.062	0.060	0.245	0.070	0.051	0.208
			25.5	0.061	0.080	0.268	0.066	0.067	0.259	0.066	0.049	0.218
			26.0	0.071	0.077	0.275	0.071	0.065	0.254	0.070	0.055	0.211
			29.0	0.063	0.081	0.268	0.064	0.064	0.259	0.066	0.059	0.213
			34.5	0.058	0.087	0.273	0.062	0.074	0.247	0.060	0.061	0.220

Table 3.9: Estimated significance level and power of tests for testing independence in Sankaran and Nair's bivariate Pareto distribution at the nominal size $\alpha = 0.05$ for $n = 50, 100, 250, 500$, and increasing values of α_1, α_2 and θ.

				$n=50$			$n=100$			$n=250$			$n=500$		
α_1	α_2	θ	α_0	ξ_{1B}	LR	Wald	ξ_{1B}	LR	Wald	ξ_{1B}	LR	Wald	ξ_{1B}	LR	Wald
1.0	1.0	2.5	1.0	0.049	0.051	0.065	0.053	0.055	0.055	0.046	0.046	0.042	0.049	0.047	0.045
			1.5	0.092	0.119	0.199	0.166	0.184	0.286	0.353	0.368	0.467	0.609	0.619	0.682
			2.0	0.238	0.276	0.397	0.452	0.488	0.620	0.831	0.849	0.896	0.988	0.990	0.993
			2.5	0.422	0.487	0.614	0.731	0.769	0.862	0.987	0.989	0.995	1	1	1
			3.0	0.615	0.693	0.783	0.905	0.931	0.965	0.999	1	1	1	1	1
			6.0	0.987	0.998	0.857	1	1	0.980	1	1	1	1	1	1
			11.5	1	1	0.798	1	1	0.955	1	1	1	1	1	1
1.5	2.0	3.0	3.0	0.048	0.052	0.064	0.049	0.052	0.058	0.042	0.044	0.041	0.051	0.052	0.048
			3.5	0.051	0.058	0.096	0.059	0.067	0.098	0.073	0.078	0.119	0.114	0.117	0.157
			4.0	0.062	0.076	0.131	0.089	0.104	0.168	0.165	0.174	0.248	0.284	0.293	0.368
			4.5	0.089	0.108	0.180	0.136	0.155	0.249	0.295	0.314	0.415	0.525	0.541	0.620
			5.0	0.102	0.138	0.228	0.207	0.227	0.344	0.458	0.475	0.578	0.763	0.772	0.827
			8.0	0.405	0.485	0.604	0.709	0.755	0.849	0.983	0.986	0.992	1	1	1
			13.5	0.843	0.910	0.913	0.991	0.997	0.997	1	1	1	1	1	1
2.0	2.5	4.0	5.0	0.054	0.050	0.087	0.053	0.051	0.061	0.049	0.050	0.047	0.049	0.052	0.047
			5.5	0.045	0.050	0.093	0.049	0.056	0.083	0.056	0.064	0.088	0.060	0.064	0.085
			6.0	0.050	0.057	0.104	0.056	0.064	0.112	0.070	0.079	0.117	0.107	0.114	0.161
			6.5	0.056	0.072	0.129	0.073	0.085	0.136	0.106	0.118	0.183	0.174	0.186	0.259
			7.0	0.056	0.078	0.142	0.076	0.090	0.151	0.152	0.164	0.246	0.284	0.299	0.387
			10.0	0.135	0.190	0.277	0.268	0.306	0.432	0.608	0.635	0.745	0.892	0.902	0.940
			15.5	0.373	0.469	0.548	0.704	0.750	0.828	0.983	0.987	0.994	1	1	1
2.5	3.0	6.0	7.5	0.053	0.046	0.111	0.056	0.049	0.075	0.055	0.053	0.064	0.051	0.052	0.051
			8.0	0.052	0.046	0.120	0.051	0.050	0.084	0.051	0.053	0.082	0.049	0.051	0.067
			8.5	0.051	0.050	0.128	0.050	0.052	0.088	0.054	0.057	0.091	0.066	0.070	0.100
			9.0	0.055	0.057	0.136	0.050	0.058	0.104	0.055	0.065	0.111	0.073	0.080	0.133
			9.5	0.046	0.057	0.132	0.054	0.065	0.119	0.068	0.076	0.126	0.105	0.113	0.173
			12.5	0.064	0.090	0.173	0.092	0.120	0.178	0.209	0.235	0.341	0.384	0.406	0.528
			18.0	0.118	0.183	0.259	0.246	0.300	0.390	0.585	0.622	0.738	0.881	0.895	0.937
4.0	6.0	10.0	24.0	0.058	0.043	0.150	0.051	0.041	0.127	0.050	0.043	0.097	0.054	0.052	0.079
			24.5	0.064	0.046	0.158	0.062	0.047	0.123	0.051	0.047	0.099	0.056	0.056	0.084
			25.0	0.056	0.046	0.159	0.051	0.041	0.114	0.052	0.046	0.094	0.053	0.054	0.085
			25.5	0.066	0.053	0.172	0.054	0.043	0.121	0.052	0.049	0.109	0.052	0.054	0.085
			26.0	0.059	0.041	0.155	0.058	0.047	0.130	0.051	0.046	0.103	0.054	0.053	0.095
			29.0	0.059	0.050	0.162	0.051	0.055	0.139	0.058	0.060	0.112	0.060	0.065	0.113
			34.5	0.058	0.061	0.175	0.060	0.066	0.147	0.073	0.086	0.142	0.101	0.114	0.183

Table 3.10: Estimated significance level and power of tests for testing the fit of Lindley and Singpurwalla's distribution to these data at the nominal size $\alpha = 0.025$ and $\alpha = 0.05$ for $n = 10, 15, 25$, $\alpha_1 = \alpha_2 = 0.50$, and increasing values of θ.

				$n = 10$		$n = 15$		$n = 25$	
α_1	α_2	θ	α_0	LR	ξ_{2B}	LR	ξ_{2B}	LR	ξ_{2B}
0.5	0.5	2.5	0.0	0.045	0.012	0.030	0.009	0.025	0.013
			0.5	0.311	0.121	0.373	0.208	0.522	0.375
			1.0	0.590	0.302	0.697	0.501	0.850	0.759
			1.5	0.780	0.457	0.874	0.700	0.954	0.902
			2.0	0.863	0.569	0.933	0.805	0.975	0.946
			5.0	0.954	0.760	0.973	0.892	0.982	0.965
			10.0	0.966	0.740	0.985	0.845	0.992	0.907
0.5	0.5	3.0	0.0	0.046	0.015	0.027	0.013	0.024	0.015
			0.5	0.262	0.109	0.304	0.189	0.447	0.353
			1.0	0.504	0.274	0.610	0.464	0.782	0.719
			1.5	0.714	0.418	0.832	0.660	0.929	0.882
			2.0	0.829	0.532	0.913	0.778	0.961	0.942
			5.0	0.945	0.764	0.974	0.924	0.976	0.981
			10.0	0.973	0.784	0.984	0.917	0.993	0.966
0.5	0.5	4.0	0.0	0.051	0.016	0.034	0.018	0.026	0.021
			0.5	0.199	0.084	0.217	0.136	0.319	0.273
			1.0	0.393	0.187	0.479	0.356	0.627	0.583
			1.5	0.592	0.315	0.707	0.546	0.864	0.816
			2.0	0.753	0.434	0.852	0.681	0.944	0.905
			5.0	0.946	0.738	0.964	0.921	0.974	0.983
			10.0	0.972	0.817	0.981	0.954	0.988	0.992
0.5	0.5	5.0	0.0	0.050	0.015	0.031	0.015	0.025	0.021
			0.5	0.160	0.060	0.181	0.112	0.235	0.208
			1.0	0.322	0.146	0.397	0.274	0.531	0.494
			1.5	0.496	0.251	0.592	0.444	0.775	0.717
			2.0	0.660	0.356	0.775	0.587	0.905	0.848
			5.0	0.948	0.695	0.971	0.892	0.977	0.985
			10.0	0.971	0.822	0.988	0.964	0.990	0.996
0.5	0.5	10.0	0.0	0.055	0.017	0.045	0.024	0.030	0.026
			0.5	0.108	0.037	0.099	0.056	0.115	0.105
			1.0	0.178	0.068	0.191	0.118	0.249	0.228
			1.5	0.251	0.111	0.305	0.197	0.428	0.380
			2.0	0.342	0.158	0.434	0.298	0.592	0.518
			5.0	0.812	0.453	0.891	0.714	0.936	0.940
			10.0	0.926	0.679	0.942	0.908	0.942	0.976

Table 3.11: Estimated significance level and power of tests for testing the fit of Lindley and Singpurwalla's distribution to these data at the nominal size $\alpha = 0.025$ and $\alpha = 0.05$ for $n = 50, 100, 250, 500$, $\alpha_1 = \alpha_2 = 0.50$, and increasing values of θ.

				$n = 50$		$n = 100$		$n = 250$		$n = 500$	
α_1	α_2	θ	α_0	LR	ξ_{2B}	LR	ξ_{2B}	LR	ξ_{2B}	LR	ξ_{2B}
0.5	0.5	2.5	0.0	0.023	0.015	0.027	0.029	0.019	0.033	0.026	0.041
			0.5	0.817	0.728	0.977	0.962	1	1	1	1
			1.0	0.979	0.965	0.998	0.999	1	1	1	1
			1.5	0.992	0.992	0.998	0.999	1	1	1	1
			2.0	0.991	0.995	0.996	1	1	1	1	1
			5.0	0.995	0.995	1	0.999	1	1	1	1
			10.0	0.998	0.958	1	0.994	1	1	1	1
0.5	0.5	3.0	0.0	0.021	0.021	0.023	0.028	0.025	0.037	0.022	0.041
			0.5	0.704	0.670	0.945	0.934	1	1	1	1
			1.0	0.951	0.942	0.997	0.996	1	1	1	1
			1.5	0.985	0.986	0.995	0.999	1	1	1	1
			2.0	0.982	0.990	0.995	0.998	0.998	0.999	1	1
			5.0	0.988	0.996	0.997	0.999	1	1	1	1
			10.0	0.998	0.993	1	1	1	1	1	1
0.5	0.5	4.0	0.0	0.019	0.025	0.024	0.040	0.023	0.043	0.023	0.041
			0.5	0.512	0.509	0.824	0.822	0.994	0.993	1	1
			1.0	0.887	0.880	0.987	0.988	0.999	0.999	1	1
			1.5	0.971	0.969	0.991	0.996	0.996	0.998	0.998	0.999
			2.0	0.979	0.990	0.985	0.994	0.993	0.998	0.998	0.999
			5.0	0.979	0.992	0.988	0.997	0.994	0.998	0.999	0.999
			10.0	0.992	0.997	0.997	0.999	1	1	1	1
0.5	0.5	5.0	0.0	0.021	0.035	0.021	0.033	0.024	0.044	0.023	0.052
			0.5	0.401	0.405	0.674	0.686	0.968	0.969	1	1
			1.0	0.801	0.806	0.961	0.965	0.997	0.999	0.999	0.999
			1.5	0.942	0.939	0.983	0.992	0.991	0.996	0.996	0.998
			2.0	0.975	0.982	0.981	0.993	0.988	0.995	0.989	0.996
			5.0	0.977	0.994	0.984	0.994	0.983	0.992	0.987	0.993
			10.0	0.991	0.995	0.994	0.998	0.997	0.999	0.999	1
0.5	0.5	10.0	0.0	0.023	0.038	0.023	0.043	0.023	0.041	0.023	0.048
			0.5	0.170	0.181	0.284	0.307	0.580	0.605	0.866	0.878
			1.0	0.419	0.433	0.677	0.716	0.949	0.963	0.982	0.992
			1.5	0.665	0.684	0.895	0.922	0.966	0.986	0.962	0.986
			2.0	0.844	0.850	0.949	0.973	0.957	0.983	0.956	0.985
			5.0	0.936	0.979	0.946	0.985	0.937	0.978	0.924	0.970
			10.0	0.939	0.980	0.936	0.979	0.929	0.977	0.931	0.974

Table 3.12: Estimated significance level and power of tests for testing the fit of Lindley and Singpurwalla's distribution to these data at the nominal size $\alpha = 0.025$ and $\alpha = 0.05$ for $n = 10, 15, 25$, $\alpha_1 = \theta = 0.50$, and increasing values of α_2.

				$n = 10$		$n = 15$		$n = 25$	
α_1	α_2	θ	α_0	LR	ξ_{2B}	LR	ξ_{2B}	LR	ξ_{2B}
0.5	1.0	3.0	0.0	0.041	0.011	0.031	0.014	0.020	0.014
			0.5	0.142	0.046	0.145	0.070	0.208	0.130
			1	0.255	0.100	0.298	0.180	0.433	0.344
			1.5	0.378	0.171	0.465	0.325	0.623	0.550
			2	0.498	0.248	0.610	0.450	0.768	0.705
			5	0.895	0.617	0.954	0.844	0.979	0.965
			10	0.955	0.759	0.980	0.927	0.986	0.984
0.5	1.5	3.0	0.0	0.041	0.001	0.028	0.009	0.022	0.015
			0.5	0.102	0.026	0.105	0.045	0.135	0.076
			1	0.182	0.062	0.195	0.100	0.285	0.197
			1.5	0.271	0.113	0.299	0.184	0.438	0.345
			2	0.330	0.146	0.420	0.269	0.570	0.486
			5	0.772	0.453	0.875	0.699	0.956	0.902
			10	0.938	0.697	0.974	0.891	0.986	0.981
0.5	2.0	3.0	0.0	0.044	0.013	0.032	0.014	0.023	0.019
			0.5	0.090	0.024	0.078	0.030	0.103	0.059
			1	0.137	0.040	0.143	0.071	0.203	0.129
			1.5	0.188	0.062	0.226	0.118	0.310	0.229
			2.0	0.259	0.107	0.301	0.173	0.423	0.336
			5.0	0.615	0.326	0.737	0.564	0.879	0.812
			10.0	0.901	0.616	0.960	0.832	0.990	0.969
0.5	3.0	3.0	0.0	0.048	0.012	0.030	0.012	0.024	0.015
			0.5	0.083	0.019	0.067	0.027	0.069	0.035
			1.0	0.110	0.029	0.104	0.043	0.126	0.072
			1.5	0.139	0.045	0.145	0.071	0.205	0.126
			2.0	0.176	0.061	0.201	0.105	0.289	0.201
			5.0	0.423	0.193	0.531	0.374	0.699	0.620
			10.0	0.781	0.465	0.879	0.698	0.962	0.909
0.5	6.0	3.0	0.0	0.044	0.012	0.029	0.010	0.024	0.017
			0.5	0.063	0.013	0.046	0.015	0.043	0.018
			1.0	0.075	0.015	0.067	0.024	0.067	0.033
			1.5	0.092	0.021	0.084	0.034	0.093	0.051
			2.0	0.109	0.028	0.105	0.048	0.131	0.077
			5.0	0.216	0.077	0.255	0.140	0.354	0.262
			10.0	0.430	0.197	0.518	0.367	0.686	0.609

Table 3.13: Estimated significance level and power of tests for testing the fit of Lindley and Singpurwalla's distribution to these data at the nominal size $\alpha = 0.025$ and $\alpha = 0.05$ for $n = 50, 100, 250, 500$, $\alpha_1 = \theta = 0.50$, and increasing values of α_2.

				$n = 50$		$n = 100$		$n = 250$		$n = 500$	
α_1	α_2	θ	α_0	LR	ξ_{2B}	LR	ξ_{2B}	LR	ξ_{2B}	LR	ξ_{2B}
0.5	1.0	3.0	0.0	0.024	0.021	0.025	0.029	0.022	0.032	0.026	0.044
			0.5	0.382	0.294	0.648	0.592	0.959	0.949	1	0.999
			1	0.698	0.655	0.944	0.937	1	1	1	1
			1.5	0.885	0.867	0.989	0.987	1	1	1	1
			2	0.951	0.942	0.997	0.997	1	1	1	1
			5	0.989	0.994	0.994	0.997	0.999	1	1	1
			10	0.993	0.997	0.997	0.999	1	1	1	1
0.5	1.5	3.0	0.0	0.023	0.023	0.022	0.028	0.022	0.033	0.022	0.038
			0.5	0.234	0.166	0.425	0.346	0.818	0.774	0.983	0.972
			1	0.488	0.417	0.800	0.759	0.994	0.991	1	1
			1.5	0.716	0.673	0.935	0.921	1	0.999	1	1
			2	0.844	0.821	0.984	0.980	1	1	1	1
			5	0.990	0.991	0.997	0.998	0.999	1	1	1
			10	0.989	0.997	0.996	0.999	0.999	0.999	1	1
0.5	2.0	3.0	0.0	0.020	0.021	0.021	0.027	0.023	0.034	0.020	0.038
			0.5	0.163	0.115	0.315	0.244	0.647	0.583	0.907	0.881
			1	0.371	0.291	0.654	0.583	0.956	0.940	0.999	0.999
			1.5	0.566	0.497	0.850	0.824	0.996	0.994	1	1
			2.0	0.714	0.673	0.937	0.927	1	0.927	1	1
			5.0	0.986	0.977	0.998	0.998	1	1	1	1
			10.0	0.995	0.997	0.997	0.998	1	1	1	1
0.5	3.0	3.0	0.0	0.020	0.021	0.023	0.029	0.024	0.036	0.021	0.037
			0.5	0.111	0.073	0.197	0.137	0.405	0.339	0.691	0.639
			1.0	0.231	0.158	0.443	0.353	0.807	0.763	0.978	0.972
			1.5	0.379	0.295	0.657	0.589	0.955	0.943	0.999	0.999
			2.0	0.506	0.431	0.797	0.759	0.991	0.988	1	1
			5.0	0.921	0.899	0.994	0.992	1	1	1	1
			10.0	0.997	0.993	0.999	0.999	1	1	1	1
0.5	6.0	3.0	0.0	0.019	0.020	0.021	0.027	0.023	0.035	0.023	0.039
			0.5	0.054	0.038	0.081	0.063	0.178	0.141	0.302	0.259
			1.0	0.110	0.068	0.179	0.131	0.426	0.361	0.706	0.641
			1.5	0.172	0.110	0.311	0.239	0.662	0.592	0.910	0.881
			2.0	0.218	0.151	0.432	0.354	0.811	0.770	0.979	0.970
			5.0	0.609	0.546	0.886	0.870	0.999	0.998	1	1
			10.0	0.920	0.909	0.994	0.993	1	1	1	1

Table 3.14: Estimated significance level and power of tests for testing the fit of Lindley and Singpurwalla's distribution to these data at the nominal size $\alpha = 0.025$ and $\alpha = 0.05$ for $n = 10, 15, 25$, $\alpha_1 = 0.50$, and increasing values of α_2 and θ.

				$n = 10$		$n = 15$		$n = 25$	
α_1	α_2	θ	α_0	LR	ξ_{2B}	LR	ξ_{2B}	LR	ξ_{2B}
0.5	1.0	2.5	0.0	0.044	0.012	0.031	0.009	0.025	0.013
			0.5	0.173	0.050	0.189	0.079	0.263	0.137
			1.0	0.310	0.118	0.383	0.209	0.527	0.382
			1.5	0.441	0.208	0.548	0.365	0.725	0.606
			2.0	0.576	0.304	0.700	0.505	0.845	0.753
			5.0	0.913	0.651	0.960	0.841	0.981	0.965
			10.0	0.955	0.747	0.981	0.903	0.990	0.970
0.5	1.5	3.0	0.0	0.044	0.011	0.026	0.011	0.023	0.014
			0.5	0.109	0.031	0.104	0.045	0.143	0.081
			1.0	0.186	0.065	0.198	0.102	0.271	0.186
			1.5	0.249	0.102	0.302	0.178	0.437	0.349
			2.0	0.341	0.151	0.415	0.269	0.569	0.489
			5.0	0.768	0.450	0.872	0.693	0.958	0.910
			10.0	0.934	0.685	0.974	0.900	0.986	0.984
0.5	2.0	4.0	0.0	0.052	0.014	0.034	0.016	0.025	0.019
			0.5	0.083	0.026	0.062	0.028	0.083	0.054
			1.0	0.118	0.037	0.114	0.058	0.136	0.098
			1.5	0.153	0.054	0.174	0.097	0.218	0.172
			2.0	0.206	0.073	0.234	0.146	0.308	0.262
			5.0	0.496	0.256	0.605	0.450	0.775	0.707
			10.0	0.851	0.537	0.924	0.779	0.973	0.952
0.5	2.5	6.0	0.0	0.052	0.014	0.035	0.019	0.030	0.024
			0.5	0.066	0.019	0.059	0.027	0.051	0.044
			1.0	0.085	0.028	0.075	0.043	0.077	0.067
			1.5	0.104	0.039	0.098	0.052	0.109	0.094
			2.0	0.131	0.040	0.128	0.077	0.145	0.127
			5.0	0.275	0.116	0.316	0.212	0.446	0.404
			10.0	0.546	0.288	0.678	0.507	0.829	0.760
0.5	3.0	10.0	0.0	0.053	0.013	0.044	0.020	0.030	0.023
			0.5	0.063	0.016	0.052	0.020	0.045	0.034
			1.0	0.076	0.018	0.061	0.029	0.051	0.040
			1.5	0.082	0.023	0.068	0.032	0.077	0.057
			2.0	0.090	0.023	0.083	0.044	0.082	0.062
			5.0	0.151	0.049	0.160	0.095	0.200	0.177
			10.0	0.275	0.118	0.346	0.231	0.480	0.427

Table 3.15: Estimated significance level and power of tests for testing the fit of Lindley and Singpurwalla's distribution to these data at the nominal size $\alpha = 0.025$ and $\alpha = 0.05$ for $n = 50, 100, 250, 500$, $\alpha_1 = 0.50$, and increasing values of α_2 and θ.

				$n=50$		$n=100$		$n=250$		$n=500$	
α_1	α_2	θ	α_0	LR	ξ_{2B}	LR	ξ_{2B}	LR	ξ_{2B}	LR	ξ_{2B}
0.5	1.0	2.5	0.0	0.023	0.015	0.027	0.022	0.018	0.039	0.026	0.042
			0.5	0.493	0.314	0.780	0.647	0.986	0.971	1	1
			1.0	0.816	0.734	0.978	0.961	1	1	1	1
			1.5	0.940	0.908	0.997	0.995	1	1	1	1
			2.0	0.980	0.963	0.999	0.999	1	1	1	1
			5.0	0.991	0.997	0.991	0.998	1	1	1	1
			10.0	0.996	0.995	1	1	1	1	1	1
0.5	1.5	3.0	0.0	0.021	0.021	0.023	0.028	0.025	0.037	0.022	0.041
			0.5	0.244	0.168	0.439	0.357	0.818	0.776	0.977	0.970
			1.0	0.504	0.434	0.802	0.766	0.994	0.991	1	1
			1.5	0.711	0.673	0.939	0.928	0.999	0.999	1	1
			2.0	0.837	0.812	0.978	0.973	1	1	1	1
			5.0	0.991	0.991	0.996	0.998	0.998	0.999	1	1
			10.0	0.992	0.997	0.995	0.998	1	1	1	1
0.5	2.0	4.0	0.0	0.019	0.025	0.024	0.040	0.023	0.043	0.022	0.042
			0.5	0.106	0.087	0.181	0.173	0.402	0.393	0.679	0.666
			1.0	0.235	0.215	0.425	0.418	0.820	0.811	0.983	0.981
			1.5	0.388	0.367	0.668	0.665	0.957	0.956	1	1
			2.0	0.535	0.523	0.821	0.820	0.994	0.994	1	1
			5.0	0.950	0.944	0.994	1	1	1	1	1
			10.0	0.987	0.994	0.990	0.996	0.996	0.998	0.998	0.999
0.5	2.5	6.0	0.0	0.021	0.036	0.021	0.034	0.024	0.044	0.022	0.052
			0.5	0.053	0.056	0.069	0.076	0.146	0.155	0.250	0.262
			1.0	0.105	0.107	0.169	0.181	0.374	0.386	0.646	0.652
			1.5	0.179	0.180	0.288	0.297	0.629	0.639	0.894	0.896
			2.0	0.247	0.243	0.426	0.443	0.810	0.818	0.978	0.977
			5.0	0.701	0.710	0.914	0.925	0.992	0.996	0.996	0.998
			10.0	0.949	0.964	0.968	0.988	0.972	0.991	0.978	0.993
0.5	3.0	10.0	0.0	0.024	0.035	0.024	0.041	0.022	0.041	0.023	0.049
			0.5	0.041	0.046	0.043	0.055	0.061	0.076	0.087	0.099
			1.0	0.056	0.062	0.077	0.087	0.135	0.144	0.230	0.243
			1.5	0.082	0.089	0.125	0.141	0.221	0.240	0.407	0.422
			2.0	0.112	0.116	0.168	0.183	0.341	0.365	0.585	0.604
			5.0	0.327	0.340	0.562	0.600	0.896	0.921	0.980	0.991
			10.0	0.735	0.746	0.910	0.951	0.951	0.985	0.944	0.982

Table 3.16: Estimated significance level and power of tests for testing the fit of Lindley and Singpurwalla's distribution to these data at the nominal size $\alpha = 0.025$ and $\alpha = 0.05$ for $n = 10, 15, 25$, and increasing values of α_1, α_2 and θ.

				$n = 10$		$n = 15$		$n = 25$	
α_1	α_2	θ	α_0	LR	ξ_{2B}	LR	ξ_{2B}	LR	ξ_{2B}
1.0	1.0	2.5	0.0	0.044	0.010	0.030	0.009	0.025	0.013
			0.5	0.105	0.024	0.096	0.032	0.130	0.055
			1.0	0.167	0.045	0.189	0.077	0.277	0.145
			1.5	0.233	0.076	0.285	0.143	0.411	0.252
			2.0	0.307	0.122	0.373	0.213	0.526	0.382
			5.0	0.704	0.391	0.809	0.607	0.927	0.853
			10.0	0.919	0.626	0.971	0.851	0.988	0.968
1.5	2.0	3.0	0.0	0.045	0.011	0.025	0.011	0.025	0.014
			0.5	0.063	0.013	0.051	0.021	0.046	0.023
			1.0	0.077	0.020	0.066	0.026	0.071	0.036
			1.5	0.093	0.024	0.081	0.033	0.097	0.053
			2.0	0.108	0.032	0.102	0.042	0.127	0.074
			5.0	0.225	0.078	0.256	0.142	0.359	0.261
			10.0	0.427	0.198	0.530	0.363	0.692	0.620
2.0	2.5	4.0	0.0	0.051	0.012	0.032	0.016	0.026	0.020
			0.5	0.061	0.014	0.038	0.017	0.033	0.024
			1.0	0.057	0.016	0.048	0.021	0.039	0.022
			1.5	0.068	0.015	0.057	0.028	0.047	0.030
			2.0	0.076	0.022	0.068	0.034	0.061	0.040
			5.0	0.120	0.041	0.128	0.060	0.126	0.092
			10.0	0.202	0.073	0.226	0.139	0.313	0.265
2.5	3.0	6.0	0.0	0.049	0.011	0.034	0.014	0.028	0.022
			0.5	0.057	0.010	0.041	0.019	0.034	0.028
			1.0	0.062	0.014	0.043	0.019	0.036	0.033
			1.5	0.062	0.019	0.050	0.018	0.038	0.031
			2.0	0.067	0.014	0.052	0.021	0.041	0.032
			5.0	0.084	0.020	0.062	0.032	0.068	0.053
			10.0	0.113	0.032	0.109	0.057	0.128	0.102
4.0	6.0	10.0	0.0	0.050	0.013	0.042	0.019	0.029	0.025
			0.5	0.057	0.012	0.047	0.015	0.031	0.029
			1.0	0.060	0.013	0.050	0.017	0.035	0.028
			1.5	0.064	0.015	0.040	0.018	0.042	0.031
			2.0	0.060	0.012	0.045	0.022	0.035	0.027
			5.0	0.065	0.012	0.048	0.020	0.038	0.031
			10.0	0.066	0.015	0.057	0.024	0.044	0.038

Table 3.17: Estimated significance level and power of tests for testing the fit of Lindley and Singpurwalla's distribution to these data at the nominal size $\alpha = 0.025$ and $\alpha = 0.05$ for $n = 50, 100, 250, 500$, and increasing values of α_1, α_2 and θ.

				$n=50$		$n=100$		$n=250$		$n=500$	
α_1	α_2	θ	α_0	LR	ξ_{2B}	LR	ξ_{2B}	LR	ξ_{2B}	LR	ξ_{2B}
1.0	1.0	2.5	0.0	0.023	0.015	0.027	0.022	0.019	0.038	0.026	0.043
			0.5	0.236	0.116	0.433	0.247	0.800	0.654	0.979	0.941
			1.0	0.483	0.312	0.784	0.652	0.992	0.979	1	1
			1.5	0.683	0.556	0.928	0.883	1	0.998	1	1
			2.0	0.809	0.729	0.978	0.957	1	1	1	1
			5.0	0.989	0.984	0.999	0.999	1	1	1	1
			10.0	0.992	0.996	0.998	1	1	1	1	1
1.5	2.0	3.0	0.0	0.021	0.020	0.023	0.028	0.025	0.037	0.022	0.041
			0.5	0.054	0.034	0.090	0.069	0.171	0.141	0.296	0.255
			1.0	0.108	0.065	0.190	0.137	0.411	0.347	0.686	0.633
			1.5	0.165	0.114	0.309	0.236	0.653	0.589	0.904	0.875
			2.0	0.244	0.170	0.445	0.367	0.827	0.776	0.979	0.972
			5.0	0.626	0.568	0.885	0.868	0.999	0.999	1	1
			10.0	0.914	0.898	0.996	0.993	1	1	1	1
2.0	2.5	4.0	0.0	0.019	0.025	0.024	0.040	0.023	0.043	0.022	0.042
			0.5	0.035	0.035	0.039	0.047	0.064	0.070	0.087	0.092
			1.0	0.047	0.042	0.071	0.071	0.129	0.129	0.201	0.199
			1.5	0.071	0.061	0.104	0.102	0.192	0.195	0.374	0.371
			2.0	0.083	0.074	0.142	0.135	0.301	0.297	0.543	0.531
			5.0	0.230	0.209	0.440	0.422	0.822	0.815	0.984	0.980
			10.0	0.543	0.527	0.813	0.811	0.995	0.994	1	1
2.5	3.0	6.0	0.0	0.020	0.036	0.022	0.033	0.024	0.044	0.022	0.052
			0.5	0.024	0.035	0.025	0.039	0.031	0.046	0.045	0.059
			1.0	0.037	0.041	0.041	0.053	0.050	0.066	0.068	0.075
			1.5	0.038	0.045	0.038	0.050	0.068	0.077	0.097	0.108
			2.0	0.041	0.046	0.058	0.066	0.090	0.101	0.148	0.155
			5.0	0.086	0.087	0.142	0.152	0.295	0.309	0.514	0.521
			10.0	0.180	0.189	0.347	0.365	0.701	0.707	0.937	0.937
4.0	6.0	10.0	0.0	0.023	0.036	0.023	0.043	0.023	0.042	0.023	0.049
			0.5	0.029	0.037	0.025	0.042	0.023	0.044	0.027	0.048
			1.0	0.028	0.038	0.030	0.046	0.027	0.051	0.032	0.053
			1.5	0.033	0.043	0.031	0.047	0.025	0.046	0.029	0.046
			2.0	0.030	0.040	0.033	0.045	0.028	0.048	0.032	0.054
			5.0	0.033	0.036	0.038	0.050	0.046	0.061	0.059	0.070
			10.0	0.048	0.053	0.050	0.061	0.076	0.090	0.116	0.125

Chapter 4

A Copula Approach for Testing Independence Using Poisson Cumulative Distribution Functions

4.1 Introduction

Suppose y_1 and y_2 are two counts whose marginal distributions $F_1(y_1)$ and $F_2(y_2)$ are known and parametrically specified. When the bivariate distribution of (y_1, y_2) is known, all the familiar and standard methods can be used to make inference on this distribution. However, an issue surfaces when the bivariate distribution can't be expressed (i.e., it is unavailable) or only available under restricted criteria. Furthermore, this type of problem commonly arises under specifications of bivariate Poisson and Binomial distributions since these types of distributions usually only allow for positive dependence between counts. However, in some practical settings, it is of interest that the dependence between two variables be positive or negative. To addresses these complications, copula functions are examined. Copula functions, introduced by Sklar (1959), are one useful way to take known marginal distributions and derive unknown joint distributions.

Let $U_1, U_2, \ldots, U_q$ be uniform random variables on the $[0, 1]$ interval. Their joint cdf can be defined as $P(U_1 \leq u_1, U_2 \leq u_2, \ldots, U_q \leq u_q) = C(u_1, u_2, \ldots, u_q)$, where the function $C(\cdot, \ldots, \cdot)$ is the copula and for $j = 1, 2, \ldots, q$ and $q \geq 2$, u_j is a specific realization of U_j. For $C(\cdot, \ldots, \cdot)$ to be considered a copula, certain properties must be satisfied. C must be increasing and have a domain on the q dimensional unit hypercube and C must be grounded. For more details regarding these assumptions, see Nelsen (2007). If $C(\cdot, \ldots, \cdot)$ is considered a copula, then for q marginal cdf's $F_1(\cdot), F_2(\cdot), \ldots, F_q(\cdot)$ and for arbitrary $x_1, x_2, \ldots, x_q$, we may write

$$\begin{aligned} C[F_1(\cdot), F_2(\cdot), \ldots, F_q(\cdot)] &= P[U_1 \leq F_1(y_1), U_2 \leq F_2(y_2), \ldots, U_q \leq F_q(y_q)] \\ &= P[F_1^{-1}(U_1) \leq y_1, F_2^{-1}(U_2) \leq y_2, \ldots, F_q^{-1}(U_q) \leq y_q] \\ &= F(Y_1, Y_2, \ldots, Y_q). \end{aligned}$$

That is, if the marginal distributions are known and take a parametric form, then a copula-based joint cdf can be generated. Note that $X_j = F_j^{-1}(U_j), j = 1, 2, \ldots, q$. Though the joint cumulative distribution function (cdf) is interesting, for developing a score test the joint probability mass function (pmf) is needed. As given in Cameron et al. (2004), in the case of two discrete random variables, the continuous derivatives are replaced with finite differences and so the bivariate pmf is given by

$$\begin{aligned} c_{12}(F_1(y_1), F_2(y_2); \theta) = C(F_1(y_1), F_2(y_2); \theta) - C(F_1(y_1 - 1), F_2(y_2); \theta) - \\ C(F_1(y_1), F_2(y_2 - 1); \theta) + C(F_1(y_1 - 1), F_2(y_2 - 1); \theta). \end{aligned}$$

Alternatively, we may write

$$c_{12}(u_i, v_i; \theta) = C(u_i, v_i; \theta) - C(u_{i-1}, v_i; \theta) - C(u_i, v_{i-1}; \theta) + C(u_{i-1}, v_{i-1}; \theta),$$

where I am specifically interested in cumulative Poissons. Define

1. $C(\cdot, \ldots, \cdot)$ is the copula of interest.
2. $u_i = F_1(y_{1i}, \lambda_1) = \sum_{z=0}^{y_{1i}} \frac{e^{-\lambda_1}\lambda_1^z}{z!} \equiv \sum_{z=0}^{y_{1i}} P_1(z; \lambda_1)$
3. $u_{i-1} = F_1(y_{1i} - 1; \lambda_1) = \sum_{z=0}^{y_{1i}-1} \frac{e^{-\lambda_1}\lambda_1^z}{z!} \equiv \sum_{z=0}^{y_{1i}-1} P_1(z; \lambda_1)$
4. $v_i = F_2(y_{2i}; \lambda_2) = \sum_{z=0}^{y_{2i}} \frac{e^{-\lambda_2}\lambda_2^z}{z!} \equiv \sum_{z=0}^{y_{2i}} P_2(z; \lambda_2)$

5. $v_{i-1} = F_2(y_{2i} - 1; \lambda_2) = \sum_{z=0}^{y_{2i}-1} \frac{e^{-\lambda_2}\lambda_2^z}{z!} \equiv \sum_{z=0}^{y_{2i}-1} P_2(z; \lambda_2)$

The three most common Copula functions are the Frank (F), Normal (N), and Kimeldorf & Sampson (KS), respectively. They are given by

$$\begin{aligned} C^F &= \frac{-1}{\theta} \log \left[\frac{1 - e^{-\theta} - (1 - e^{-\theta u})(1 - e^{-\theta v})}{1 - e^{-\theta}} \right], \text{ where } \theta \in \Re \setminus \{0\} \\ C^N &= \Phi_2 \left[\Phi^{-1}(u), \Phi^{-1}(v); \theta \right], \text{ where } \theta \in [-1, 1] \\ C^{KS} &= \left(u^{-\theta} + v^{-\theta} - 1 \right)^{\frac{-1}{\theta}}, \text{ where } \theta \in [-1, \infty) \setminus \{0\} \end{aligned}$$

For these copulas, θ is known as the dependency parameter and hence measures the dependence between the marginal distributions. If no dependence is detected, then the joint CDF is written as the product of the marginal CDF's. In this situation, estimation procedures can be performed on each variable separately. In the case where it can't be shown that the two marginals are independent, then dependence is present and a copula function can be used to represent a bivariate distribution. In Section 4.2, this test of independence, i.e. $H_0 : \theta = 0$ versus $H_1 : \theta \neq 0$, is carried out via the Normal copula and a score test. Moreover, this test will determine if the two independent Poisson outcomes should be modeled as correlated observations through the Normal Copula function. The Normal copula was chosen because it specifically allows $\theta = 0$ for our test. Another reason the Normal copula is the primary copula of interest is due to the flexibility in the dependency parameter. As discussed in Cameron et al. (2004), most copulas do not require θ to be bounded below by -1 and bounded above by 1. However, by allowing θ to be both positive and negative, we do not restrict ourselves to variables that are only positively correlated. It should also be noted that θ can be converted to Kendall's tau or Spearman's rho for continuous responses only. For the case of discrete responses, as we have here, Kendall's tau and Spearman's rho are not used (Cameron et al., 2004).

In Section 4.3 a simulation study is presented to illustrate the estimated significance level and power of the score test when compared to the likelihood ratio and

Wald tests. Due to the complexity of the bivariate probability mass function and the second derivatives, the expected Fisher information matrix is unattainable, even with software. This issue is avoided by introducing the observed Fisher information matrix, which is common practice when the expected Fisher information matrix is difficult to calculate. In Section 4.4 a real world example is used to demonstrate this test. I will conclude in Section 4.5.

4.2 The Score Test

Before introducing the derivation of the score test for this problem, appropriate notation is defined. Let

1. $Q_{u_i} = \Phi^{-1}(u_i)$
2. $Q_{u_{i-1}} = \Phi^{-1}(u_{i-1})$
3. $Q_{v_i} = \Phi^{-1}(v_i)$
4. $Q_{v_{i-1}} = \Phi^{-1}(v_{i-1})$
5. $f_{i,i}(a,b) = f(a_i, b_i)$

Use the H function to denote

$$H(f(a,b)) = \begin{cases} f_{i,i}(a,b) - f_{i-1,i}(a,b) - f_{i,i-1}(a,b) + f_{i-1,i-1}(a,b); & \text{for } y_{1i} \neq 0, y_{2i} \neq 0 \\ f_{i,i}(a,b) - f_{i-1,i}(a,b); & \text{for } y_{1i} \neq 0, y_{2i} = 0 \\ f_{i,i}(a,b) - f_{i,i-1}(a,b); & \text{for } y_{1i} = 0, y_{2i} \neq 0 \\ f_{i,i}(a,b); & \text{for } y_{1i} = 0, y_{2i} = 0 \end{cases}$$

It should be noted that by choosing the Normal copula function, u_{i-1} and v_{i-1} only exist for when $y_1 > 0$ and $y_2 > 0$, respectively.

Consider independent observations $y_{11}, y_{12}, y_{13}, \ldots$ and $y_{21}, y_{22}, y_{23}, \ldots$ from Poisson marginals $F_1(y_{1i}; \lambda_1)$ and $F_2(y_{2i}; \lambda_2)$, respectively. The log-likelihood in terms of the normal copula function and the defined H function is given by

$$\mathscr{L} = \log\left\{H\left[\Phi_2(Q_u, Q_v; \theta)\right]\right\},$$

where $\Phi_2(\cdot,\cdot;\theta)$ is the bivariate normal distribution function of two standard normally distributed random variables with correlation $\theta \in [-1, 1]$, Φ is the cdf of the standard normal distribution, $N(0, 1)$, and Φ^{-1} (the quantile function) is its functional inverse.

The Score Test (Radhakrishna Rao, 1948) is a common large sample approach to hypothesis testing and will be used to assess $H_0 : \theta = 0$ versus $H_1 : \theta \neq 0$. In the multivariate setting, the score statistic can be defined as $\xi_S = [\nabla\mathscr{L}(\boldsymbol{\theta_0})]^T I_n^{-1}(\boldsymbol{\theta_0})\,[\nabla\mathscr{L}(\boldsymbol{\theta_0})]$, evaluated under the null parameter space $\boldsymbol{\theta_0} \in \Re^k$. It can be shown that $\xi_S \xrightarrow{d} \chi^2_1$ as $n \to \infty$ and the rejection region is $R = \{\xi_S > \chi^2_{1,1-\alpha}\}$. However, as discussed earlier, $I_n^{-1}(\boldsymbol{\theta_0})$ is computationally challenging and so the observed Fisher information matrix will used to attempt to remedy this problem. Denote the observed Fisher information by $O_n^{-1}(\boldsymbol{\theta_0})$. To calculate the score statistic, the first and second derivatives of the log-likelihood function with respect to u_i, v_i, and θ are needed.

Gradient Vector

Three components comprise the gradient vector. They are

$$\frac{\partial\mathscr{L}}{\partial\lambda_1} = H\left[\frac{\partial\Phi_2(Q_u, Q_v;\theta)}{\partial Q_u}\frac{\partial Q_u}{\partial u}\frac{\partial u}{\partial\lambda_1}\right]\frac{1}{H\left[\Phi_2(Q_u, Q_v;\theta)\right]}$$

$$\frac{\partial\mathscr{L}}{\partial\lambda_2} = H\left[\frac{\partial\Phi_2(Q_u, Q_v;\theta)}{\partial Q_v}\frac{\partial Q_v}{\partial v}\frac{\partial v}{\partial\lambda_2}\right]\frac{1}{H\left[\Phi_2(Q_u, Q_v;\theta)\right]}$$

$$\frac{\partial\mathscr{L}}{\partial\theta} = H\left[\frac{\partial\Phi_2(Q_u, Q_v;\theta)}{\partial\theta}\right]\frac{1}{H\left[\Phi_2(Q_u, Q_v;\theta)\right]},$$

where elements of each component are given by

1. $\frac{\partial \Phi_2(Q_u,Q_v;\theta)}{\partial Q_u} = H\left[\phi(Q_u)\Phi\left(\frac{Q_v-\theta Q_u}{\sqrt{1-\theta^2}}\right)\right]$

2. $\frac{\partial \Phi_2(Q_u,Q_v;\theta)}{\partial Q_v} = H\left[\phi(Q_v)\Phi\left(\frac{Q_u-\theta Q_v}{\sqrt{1-\theta^2}}\right)\right]$

3. $\frac{\partial \Phi_2(Q_u,Q_v;\theta)}{\partial \theta} = H\left[\phi(Q_u)\phi\left(\frac{Q_v-\theta Q_u}{\sqrt{1-\theta^2}}\right)\left(\frac{1}{\sqrt{1-\theta^2}}\right)\right] = H\left[\phi(Q_v)\phi\left(\frac{Q_u-\theta Q_v}{\sqrt{1-\theta^2}}\right)\left(\frac{1}{\sqrt{1-\theta^2}}\right)\right]$

4. $\frac{\partial Q_u}{\partial u} = H\left[\frac{1}{\phi(Q_u)}\right]$

5. $\frac{\partial Q_v}{\partial v} = H\left[\frac{1}{\phi(Q_v)}\right]$

6. $\frac{\partial u}{\partial \lambda_1} = H\left[-e^{-\lambda_1}\lambda_1^{y_1}/y_1!\right] = H\left[-P_1(y_1;\lambda_1)\right]$

7. $\frac{\partial v}{\partial \lambda_2} = H\left[-e^{-\lambda_2}\lambda_2^{y_2}/y_2!\right] = H\left[-P_2(y_2;\lambda_2)\right]$

After simplifying, the components of the gradient vector are given by

$$\begin{aligned}
\frac{\partial \mathscr{L}}{\partial \lambda_1} &= H\left[-\Phi\left(\frac{Q_v-\theta Q_u}{\sqrt{1-\theta^2}}\right)P_1(y_1;\lambda_1)\right]\frac{1}{H\left[\Phi_2(Q_u,Q_v;\theta)\right]} \\
\frac{\partial \mathscr{L}}{\partial \lambda_2} &= H\left[-\Phi\left(\frac{Q_u-\theta Q_v}{\sqrt{1-\theta^2}}\right)P_2(y_2;\lambda_2)\right]\frac{1}{H\left[\Phi_2(Q_u,Q_v;\theta)\right]} \\
\frac{\partial \mathscr{L}}{\partial \theta} &= H\left[\phi(Q_u)\phi\left(\frac{Q_v-\theta Q_u}{\sqrt{1-\theta^2}}\right)\left(\frac{1}{\sqrt{1-\theta^2}}\right)\right]\frac{1}{H\left[\Phi_2(Q_u,Q_v;\theta)\right]}
\end{aligned}$$

It can be shown that $\tilde{\lambda}_1 = \bar{y}_1$ and $\tilde{\lambda}_2 = \bar{y}_2$. Evaluated under the restriced maximum likelihood estimates and under $\theta = 0$, the components of the gradient vector become

$$\begin{aligned}
\frac{\partial \mathscr{L}}{\partial \lambda_1} &= g_{\lambda_1} = 0 \\
\frac{\partial \mathscr{L}}{\partial \lambda_2} &= g_{\lambda_2} = 0 \\
\frac{\partial \mathscr{L}}{\partial \theta} &= g_\theta = \frac{H\left[\phi(Q_u)\phi(Q_v)\right]}{H\left[\Phi_2(Q_u,Q_v;0)\right]}
\end{aligned}$$

The gradient vector is $\nabla\mathscr{L}\left(\boldsymbol{\theta_0}\right) = \left[g_{\lambda_1}\ g_{\lambda_2}\ g_\theta\right]^T$.

Observed Fisher Information Matrix

Recall the observed Fisher information matrix is given by

$$O_n^{-1}(\boldsymbol{\theta_0}) = \begin{bmatrix} -\frac{\partial^2 \mathscr{L}}{\partial\lambda_1\partial\lambda_1}\Big|_{\theta=0} & -\frac{\partial^2 \mathscr{L}}{\partial\lambda_1\partial\lambda_2}\Big|_{\theta=0} & -\frac{\partial^2 \mathscr{L}}{\partial\lambda_1\partial\theta}\Big|_{\theta=0} \\ -\frac{\partial^2 \mathscr{L}}{\partial\lambda_2\partial\lambda_1}\Big|_{\theta=0} & -\frac{\partial^2 \mathscr{L}}{\partial\lambda_2\partial\lambda_2}\Big|_{\theta=0} & -\frac{\partial^2 \mathscr{L}}{\partial\lambda_2\partial\theta}\Big|_{\theta=0} \\ -\frac{\partial^2 \mathscr{L}}{\partial\theta\partial\lambda_1}\Big|_{\theta=0} & -\frac{\partial^2 \mathscr{L}}{\partial\theta\partial\lambda_2}\Big|_{\theta=0} & -\frac{\partial^2 \mathscr{L}}{\partial\theta\partial\theta}\Big|_{\theta=0} \end{bmatrix}^{-1}.$$

The components of the observed Fisher information matrix are given by

$$\frac{\partial^2 \mathscr{L}}{\partial\lambda_1\partial\lambda_1} = H\Bigg[-\frac{\partial}{\partial Q_u}\Phi\left(\frac{Q_v-\theta Q_u}{\sqrt{1-\theta^2}}\right)\frac{\partial Q_u}{\partial u}\frac{\partial u}{\partial\lambda_1}P_1(y_1;\lambda_1) -\Phi\left(\frac{Q_v-\theta Q_u}{\sqrt{1-\theta^2}}\right)\frac{\partial}{\partial\lambda_1}P_1(y_1;\lambda_1)\Bigg]\frac{1}{H\left[\Phi_2(Q_u,Q_v,\theta)\right]} - g_{\lambda_1}^2$$

$$\frac{\partial^2 \mathscr{L}}{\partial\lambda_2\partial\lambda_2} = H\Bigg[-\frac{\partial}{\partial Q_v}\Phi\left(\frac{Q_u-\theta Q_v}{\sqrt{1-\theta^2}}\right)\frac{\partial Q_v}{\partial v}\frac{\partial v}{\partial\lambda_1}P_2(y_2;\lambda_2) -\Phi\left(\frac{Q_u-\theta Q_v}{\sqrt{1-\theta^2}}\right)\frac{\partial}{\partial\lambda_1}P_2(y_2;\lambda_2)\Bigg]\frac{1}{H\left[\Phi_2(Q_u,Q_v;\theta)\right]} - g_{\lambda_2}^2$$

$$\frac{\partial^2 \mathscr{L}}{\partial\theta\partial\theta} = -\left(\frac{H\left[\phi(Q_u)\phi(Q_v)\right]}{H\left[\Phi_2(Q_u,Q_v;\theta)\right]}\right)^2 = -g_\theta^2$$

$$\frac{\partial^2 \mathscr{L}}{\partial\lambda_1\partial\lambda_2} = H\left[-\frac{\partial}{\partial Q_v}\Phi\left(\frac{Q_v-\theta Q_u}{\sqrt{1-\theta^2}}\right)\frac{\partial Q_v}{\partial v}\frac{\partial v}{\partial\lambda_2}P_1(y_1;\lambda_1)\right]\frac{1}{H\left[\Phi_2(Q_u,Q_v;\theta)\right]} - g_{\lambda_1}g_{\lambda_2}$$

$$\frac{\partial^2 \mathscr{L}}{\partial\lambda_1\partial\theta} = H\left[-\frac{\partial}{\partial\theta}\Phi\left(\frac{Q_v-\theta Q_u}{\sqrt{1-\theta^2}}\right)P_1(y_1;\lambda_1)\right]\frac{1}{H\left[\Phi_2(Q_u,Q_v;\theta)\right]} - g_{\lambda_1}g_\theta$$

$$\frac{\partial^2 \mathscr{L}}{\partial\lambda_2\partial\theta} = H\left[-\frac{\partial}{\partial\theta}\Phi\left(\frac{Q_u-\theta Q_v}{\sqrt{1-\theta^2}}\right)P_2(y_2;\lambda_2)\right]\frac{1}{H\left[\Phi_2(Q_u,Q_v;\theta)\right]} - g_{\lambda_2}g_\theta,$$

where

$$\frac{\partial^2 \mathscr{L}}{\partial\lambda_1\partial\theta} = \frac{\partial^2 \mathscr{L}}{\partial\theta\partial\lambda_1},\ \frac{\partial^2 \mathscr{L}}{\partial\lambda_2\partial\theta} = \frac{\partial^2 \mathscr{L}}{\partial\theta\partial\lambda_2},\ \text{and}\ \frac{\partial^2 \mathscr{L}}{\partial\lambda_1\partial\lambda_2} = \frac{\partial^2 \mathscr{L}}{\partial\lambda_2\partial\lambda_1}.$$

The elements of each of these components are given by

$$\frac{\partial}{\partial Q_u}\Phi\left(\frac{Q_v-\theta Q_u}{\sqrt{1-\theta^2}}\right) = H\left[-\frac{\theta}{\sqrt{1-\theta^2}}\phi\left(\frac{Q_v-\theta Q_v}{\sqrt{1-\theta^2}}\right)\right]$$

$$\frac{\partial}{\partial Q_v}\Phi\left(\frac{Q_v-\theta Q_u}{\sqrt{1-\theta^2}}\right) = H\left[\frac{1}{\sqrt{1-\theta^2}}\phi\left(\frac{Q_v-\theta Q_v}{\sqrt{1-\theta^2}}\right)\right]$$

$$\frac{\partial}{\partial\theta}\Phi\left(\frac{Q_v-\theta Q_u}{\sqrt{1-\theta^2}}\right) = H\left[\frac{Q_v+Q_u(\theta^2-\theta-1)}{(1-\theta^2)^{3/2}}\phi\left(\frac{Q_v-\theta Q_v}{\sqrt{1-\theta^2}}\right)\right]$$

After simplifying, the components of the components of the observed Fisher information matrix are given by

$$\frac{\partial^2 \mathscr{L}}{\partial\lambda_1\partial\lambda_1} = H\Bigg[-\frac{\theta}{\sqrt{1-\theta^2}}\left(\frac{Q_v-\theta Q_u}{\sqrt{1-\theta^2}}\right)\frac{1}{\phi(Q_u)} - (P_1(y_1;\lambda_1)^2 - \Phi\left(\frac{Q_v-\theta Q_u}{\sqrt{1-\theta^2}}\right)P_1(y_1;\lambda_1)\left(\frac{y_1}{\lambda_1}-1\right)\Bigg]\frac{1}{H[\Phi_2(Q-u,Q_v;\theta)]} - \left(\frac{\partial\mathscr{L}}{\partial\lambda_1}\right)^2$$

$$\frac{\partial^2 \mathscr{L}}{\partial\lambda_2\partial\lambda_2} = H\Bigg[-\frac{\theta}{\sqrt{1-\theta^2}}\left(\frac{Q_u-\theta Q_v}{\sqrt{1-\theta^2}}\right)\frac{1}{\phi(Q_v)} - (P_2(y_2;\lambda_2) - \Phi\left(\frac{Q_u-\theta Q_v}{\sqrt{1-\theta^2}}\right)P_2(y_2;\lambda_2)\left(\frac{y_2}{\lambda_2}-1\right)\Bigg]\frac{1}{H[\Phi_2(Q-u,Q_v;\theta)]} - \left(\frac{\partial\mathscr{L}}{\partial\lambda_2}\right)^2$$

$$\frac{\partial^2 \mathscr{L}}{\partial\theta\partial\theta} = H\Bigg[\phi(Q_u)\phi\left(\frac{Q_v-\theta Q_u}{\sqrt{1-\theta^2}}\right)\left(-\frac{Q_v-\theta Q_u}{\sqrt{1-\theta^2}}\right)\left(-\frac{Q_u}{\sqrt{1-\theta^2}}+\frac{\theta(Q_v-\theta Q_u)}{(1-\theta^2)^{3/2}}\right)\left(\frac{1}{\sqrt{1-\theta^2}}\right)\Bigg]\frac{1}{H[\Phi_2(Q_u,Q_v;\theta)]} - \left(\frac{\partial\mathscr{L}}{\partial\theta}\right)^2$$

$$\frac{\partial^2 \mathscr{L}}{\partial\lambda_1\partial\lambda_2} = H\Bigg[-\frac{\theta}{\sqrt{1-\theta^2}}\left(\frac{Q_v-\theta Q_u}{\sqrt{1-\theta^2}}\right)\frac{1}{\phi(Q_u)} - P_2(y_2;\lambda_2)P_1(y_1;\lambda_1)\Bigg]\frac{1}{H[\Phi_2(Q-u,Q_v;\theta)]} - \left(\frac{\partial\mathscr{L}}{\partial\lambda_1}\right)\left(\frac{\partial\mathscr{L}}{\partial\lambda_2}\right)$$

$$\frac{\partial^2 \mathscr{L}}{\partial\lambda_1\partial\theta} = H\left[-\frac{\theta}{\sqrt{1-\theta^2}}\left(\frac{Q_v-\theta Q_u}{\sqrt{1-\theta^2}}\right)P_1(y_1;\lambda_1)\right]\frac{1}{H[\Phi_2(Q_u,Q_v;\theta)]} - \left(\frac{\partial\mathscr{L}}{\partial\lambda_1}\right)\left(\frac{\partial\mathscr{L}}{\partial\theta}\right)$$

$$\frac{\partial^2 \mathscr{L}}{\partial\lambda_2\partial\theta} = H\left[-\frac{\theta}{\sqrt{1-\theta^2}}\left(\frac{Q_u-\theta Q_v}{\sqrt{1-\theta^2}}\right)P_2(y_2;\lambda_2)\right]\frac{1}{H[\Phi_2(Q_u,Q_v;\theta)]} - \left(\frac{\partial\mathscr{L}}{\partial\lambda_2}\right)\left(\frac{\partial\mathscr{L}}{\partial\theta}\right)$$

Evaluated under the restricted maximum likelihood estimates and under $\theta = 0$, the components of the observed Fisher information matrix become

$$\frac{\partial^2 \mathscr{L}}{\partial\lambda_1\partial\lambda_1} = h_{\lambda_1\lambda_1} = \frac{H\left[vP_1(y_1;\lambda_1)\left(1-\frac{y_1}{\lambda_1}\right)\right]}{H\left[\Phi_2(Q_u,Q_v;0)\right]}$$

$$\frac{\partial^2 \mathscr{L}}{\partial\lambda_2\partial\lambda_2} = h_{\lambda_2\lambda_2} = \frac{H\left[uP_2(y_2;\lambda_2)\left(1-\frac{y_2}{\lambda_2}\right)\right]}{H\left[\Phi_2(Q_u,Q_v;0)\right]}$$

$$\frac{\partial^2 \mathscr{L}}{\partial\theta\partial\theta} = h_{\theta\theta} = \frac{H\left[\phi(Q_u)\phi(Q_v)\right]}{H\left[\Phi_2(Q_u,Q_v;0)\right]} - g_\theta^2$$

$$\frac{\partial^2 \mathscr{L}}{\partial\lambda_1\partial\lambda_2} = h_{\lambda_1\lambda_2} = 0$$

$$\frac{\partial^2 \mathscr{L}}{\partial\lambda_1\partial\theta} = h_{\lambda_1\theta} = \frac{H\left[Q_u\phi(Q_v)P_1(y_1;\lambda_1)\right]}{H\left[\Phi_2(Q_u,Q_v;0)\right]}$$

$$\frac{\partial^2 \mathscr{L}}{\partial\lambda_2\partial\theta} = h_{\lambda_2\theta} = \frac{H\left[Q_v\phi(Q_u)P_2(y_2;\lambda_2)\right]}{H\left[\Phi_2(Q_u,Q_v;0)\right]}$$

The observed Fisher information can now be re-written as

$$O_n^{-1}(\boldsymbol{\theta_0}) = \begin{bmatrix} -h_{\lambda_1\lambda_1} & -h_{\lambda_1\lambda_2} & -h_{\lambda_1\theta} \\ -h_{\lambda_1\lambda_2} & -h_{\lambda_2\lambda_2} & -h_{\lambda_2\theta} \\ -h_{\lambda_1\theta} & -h_{\lambda_2\theta} & -h_{\theta\theta} \end{bmatrix}^{-1}.$$

The score statistic is then given by

$$\begin{aligned}
\xi_S &= \left[\nabla \mathscr{L}\left(\boldsymbol{\theta_0}\right)\right]^T O_n^{-1}(\boldsymbol{\theta_0}) \left[\nabla \mathscr{L}\left(\boldsymbol{\theta_0}\right)\right] \\
&= \left[g_{\lambda_1}\ g_{\lambda_2}\ g_{\theta}\right] \begin{bmatrix} -h_{\lambda_1\lambda_1} & -h_{\lambda_1\lambda_2} & -h_{\lambda_1\theta} \\ -h_{\lambda_1\lambda_2} & -h_{\lambda_2\lambda_2} & -h_{\lambda_2\theta} \\ -h_{\lambda_1\theta} & -h_{\lambda_2\theta} & -h_{\theta\theta} \end{bmatrix}^{-1} \left[g_{\lambda_1}\ g_{\lambda_2}\ g_{\theta}\right]^T \\
&= \frac{g_\theta^2}{h_{\lambda_1\theta}^2/h_{\lambda_1\lambda_1} + h_{\lambda_2\theta}^2/h_{\lambda_2\lambda_2} - h_{\theta\theta}}.
\end{aligned}$$

4.3 Simulation Results

To check the efficiency of the score test, a simulation study compared the likelihood ratio, Wald, and the proposed score test. In every iteration, the score, likelihood ratio, and Wald tests were calculated to test $\theta = 0$. These procedures were repeated 1000 times independently for each combination of values of λ_1 and λ_2, and the significance level (when $\theta = 0$) and power (when $\theta > 0$) were estimated for sample sizes $n = 10, 15, 25, 50, 100$ and 200. To calculate power, correlated Poisson data was generated via rejection sampling from Lakshminarayana's bivariate Poisson distribution (Lakshminarayana et al., 1999) given by

$$f(y_1, y_2) = \frac{\theta_1^{y_1}\lambda_2^{y_2}e^{-\lambda_1-\lambda_2}\left[1+\lambda\left(e^{-y_1}-e^{-(1-e^{-1})\lambda_1}\right)\left(e^{-y_2}-e^{-(1-e^{-1})\lambda_2}\right)\right]}{y_1!y_2!}.$$

Furthermore, the correlation coefficient is also reported for each combination of λ_1 and λ_2. The simulation results can be seen in tables 4.1, 4.2, 4.3, and 4.4.

The covariance between Y_1 and Y_2 is $\lambda\lambda_1\lambda_2(1-e^{-1})^2e^{-(1-e^{-1})(\lambda_1+\lambda_2)}$. Thus, the correlation coefficient is $\rho = \lambda\sqrt{\lambda_1\lambda_2}(1-e^{-1})^2e^{-(1-e^{-1})(\lambda_1+\lambda_2)}$, which can take on both positive and negative values depending on λ (Famoye, 2010). Also, as described in Lakshminarayana et al. (1999), λ should lie in the range $|\lambda| \leq \frac{1}{(1-A)(1-B)}$ and ρ should lie in the range $|\rho| \leq \frac{\sqrt{\lambda_1\lambda_2}AB(1-e^{-1})^2}{(1-A)(1-B)}$, where $A = e^{-\lambda_1(1-e^{-1})}$ and $B = e^{-\lambda_2(1-e^{-1})}$. It should be noted that an explicit relationship between λ from Lakshminarayana's bivariate Poisson distribution and θ from the Normal copula function is unclear. However, it is clear that as λ and ρ increase in magnitude, θ approaches 1.

For small samples (i.e. $n = 10, 15, 25$), the score test outperformed the likelihood ratio and Wald tests in maintaining the nominal 0.05 significance level. Specifically, the Wald test badly overestimates the type 1 error rate for small sample sizes. The likelihood ratio test does a better job at maintaining the nominal 0.05 significance level, but in most cases for small sample sizes, doesn't do as good of a job as the score test. For large sample sizes (i.e. $n = 50, 100, 200$), the Wald, likelihood ratio, and score tests maintain the nominal 0.05 significance level in a majority of cases.

For each sample size and each pair of λ_1 and λ_2, the power of the proposed test increases as ρ increases. Furthermore, for large samples and for each pair of λ_1 and λ_2, as ρ increases, the power of the score test is close to that of the likelihood ratio and Wald tests.

4.4 Illustration

To compare the score test derived in this chapter to that derived in Chapter 2, the same real world example is used. As a refresher, the abundances of myctophid larvae from samples taken from spatially-independent sides of a paired bongo net are compared. Y_1 is the count of myctophid larvae in the left side of the bongo net and Y_2 is the count of myctophid larvae sampled in the right. A total of 261 paired samples are used.

From the aforementioned data, the p-value corresponding to the score test is $p < 0.001$. Thus, if the significance level 0.05 is used, our test strongly rejects the independence of Y_1 and Y_2, as did our test from chapter 2. Once again, this result is not surprising.

4.5 Conclusions

I derived the score test for testing whether two independent Poisson outcomes should be modeled as correlated observations through the use of the Normal Copula function. The performance of the proposed score test was examined and compared to that of the likelihood ratio and Wald tests under a variety of sample sizes and λ_1 and λ_2 values. The score test performs better in maintaining the nominal significance level as compared to the likelihood ratio and Wald tests in smaller samples ($n = 10, 15, 25$). In larger samples, the Wald, likelihood ratio, and score tests perform equally well in most cases. Larger values of the correlation result in higher power for all sample sizes and all combinations of λ_1 and λ_2.

Table 4.1: Estimated significance level and power of tests for testing independence using the Normal copula function and Poisson CDF's at the nominal size $\alpha = 0.05$ for $n = 10$, $n = 15$, $n = 25$, and varying values of λ_1 and λ_2

			$n=10$			$n=15$			$n=25$		
λ_1	λ_2	ρ	LR	ξ_S	Wald	LR	ξ_S	Wald	LR	ξ_S	Wald
0.50	0.50	0.00	0.080	0.075	0.146	0.080	0.062	0.132	0.065	0.054	0.106
		0.05	0.087	0.072	0.178	0.093	0.069	0.149	0.082	0.066	0.128
		0.10	0.094	0.079	0.198	0.091	0.075	0.183	0.090	0.079	0.160
		0.15	0.114	0.094	0.225	0.104	0.075	0.212	0.135	0.117	0.214
		0.20	0.165	0.122	0.289	0.164	0.134	0.269	0.187	0.168	0.285
		0.25	0.180	0.151	0.333	0.211	0.177	0.356	0.270	0.241	0.394
		0.50	0.475	0.414	0.604	0.621	0.563	0.749	0.805	0.774	0.891
0.50	1.00	0.00	0.096	0.077	0.217	0.082	0.063	0.171	0.071	0.064	0.127
		0.05	0.110	0.086	0.239	0.069	0.058	0.182	0.077	0.069	0.137
		0.10	0.121	0.084	0.232	0.097	0.079	0.200	0.098	0.088	0.167
		0.15	0.114	0.084	0.250	0.128	0.113	0.238	0.136	0.124	0.209
		0.20	0.174	0.136	0.304	0.166	0.144	0.300	0.219	0.199	0.304
		0.25	0.197	0.158	0.375	0.225	0.188	0.374	0.302	0.280	0.423
		0.50	0.450	0.388	0.653	0.590	0.557	0.765	0.805	0.796	0.890
0.50	1.50	0.00	0.092	0.087	0.234	0.082	0.064	0.187	0.065	0.057	0.120
		0.05	0.108	0.084	0.226	0.071	0.058	0.162	0.077	0.070	0.139
		0.10	0.119	0.101	0.268	0.083	0.068	0.181	0.103	0.093	0.173
		0.15	0.141	0.099	0.296	0.136	0.107	0.266	0.144	0.127	0.231
		0.20	0.123	0.111	0.273	0.145	0.124	0.270	0.205	0.178	0.324
		0.25	0.178	0.165	0.347	0.231	0.204	0.397	0.301	0.288	0.418
		0.50	0.395	0.315	0.614	0.504	0.466	0.685	0.752	0.745	0.844
0.50	2.00	0.00	0.094	0.081	0.250	0.064	0.052	0.162	0.066	0.057	0.132
		0.05	0.100	0.073	0.253	0.059	0.044	0.173	0.085	0.074	0.141
		0.10	0.113	0.088	0.260	0.097	0.080	0.187	0.120	0.101	0.197
		0.15	0.100	0.082	0.259	0.125	0.106	0.236	0.120	0.113	0.214
		0.20	0.137	0.112	0.308	0.161	0.138	0.289	0.198	0.184	0.308
		0.25	0.159	0.113	0.357	0.197	0.153	0.344	0.253	0.239	0.368
		0.50	0.316	0.253	0.508	0.433	0.379	0.611	0.618	0.603	0.734
1.00	1.00	0.00	0.090	0.089	0.216	0.066	0.061	0.159	0.055	0.051	0.107
		0.05	0.090	0.079	0.218	0.073	0.077	0.174	0.057	0.052	0.116
		0.10	0.095	0.092	0.265	0.110	0.089	0.212	0.095	0.083	0.164
		0.15	0.122	0.104	0.282	0.127	0.122	0.253	0.167	0.161	0.246
		0.20	0.138	0.120	0.324	0.156	0.140	0.307	0.225	0.219	0.315
		0.25	0.193	0.166	0.367	0.212	0.201	0.394	0.317	0.306	0.429
		0.50	0.502	0.454	0.741	0.708	0.703	0.869	0.921	0.920	0.969

Table 4.2: Estimated significance level and power of tests for testing independence using the Normal copula function and Poisson CDF's at the nominal size $\alpha = 0.05$ for $n = 50$, $n = 100$, $n = 200$, and varying values of λ_1 and λ_2

			$n = 50$			$n = 100$			$n = 200$		
λ_1	λ_2	ρ	LR	ξ_S	Wald	LR	ξ_S	Wald	LR	ξ_S	Wald
0.50	0.50	0.00	0.057	0.049	0.082	0.059	0.055	0.073	0.056	0.056	0.059
		0.05	0.070	0.064	0.098	0.096	0.094	0.110	0.125	0.122	0.128
		0.10	0.115	0.104	0.157	0.160	0.151	0.184	0.312	0.304	0.329
		0.15	0.203	0.187	0.264	0.349	0.340	0.389	0.603	0.596	0.619
		0.20	0.346	0.331	0.404	0.545	0.533	0.586	0.846	0.844	0.856
		0.25	0.483	0.467	0.566	0.776	0.762	0.804	0.968	0.966	0.972
		0.50	0.984	0.983	0.991	1.000	1.000	1.000	1.000	1.000	1.000
0.50	1.00	0.00	0.056	0.052	0.082	0.067	0.063	0.086	0.054	0.054	0.061
		0.05	0.059	0.056	0.090	0.086	0.082	0.095	0.111	0.110	0.120
		0.10	0.113	0.105	0.148	0.246	0.232	0.281	0.316	0.312	0.334
		0.15	0.214	0.196	0.274	0.376	0.367	0.406	0.659	0.656	0.674
		0.20	0.330	0.316	0.386	0.606	0.601	0.647	0.878	0.878	0.881
		0.25	0.525	0.509	0.582	0.801	0.795	0.829	0.978	0.978	0.982
		0.50	0.978	0.978	0.987	1.000	1.000	1.000	1.000	1.000	1.000
0.50	1.50	0.00	0.059	0.054	0.088	0.053	0.049	0.065	0.059	0.056	0.066
		0.05	0.078	0.072	0.112	0.092	0.090	0.107	0.110	0.108	0.115
		0.10	0.121	0.117	0.170	0.191	0.186	0.210	0.367	0.362	0.385
		0.15	0.228	0.216	0.294	0.416	0.411	0.452	0.668	0.665	0.688
		0.20	0.374	0.364	0.451	0.630	0.620	0.667	0.929	0.928	0.934
		0.25	0.544	0.531	0.614	0.842	0.840	0.868	0.988	0.988	0.989
		0.50	0.954	0.954	0.974	1.000	1.000	1.000	1.000	1.000	1.000
0.50	2.00	0.00	0.056	0.055	0.083	0.042	0.039	0.056	0.053	0.053	0.058
		0.05	0.082	0.076	0.108	0.101	0.098	0.125	0.128	0.126	0.137
		0.10	0.137	0.124	0.173	0.192	0.183	0.227	0.343	0.340	0.360
		0.15	0.234	0.221	0.295	0.399	0.392	0.443	0.707	0.703	0.724
		0.20	0.364	0.353	0.453	0.627	0.620	0.667	0.921	0.918	0.930
		0.25	0.504	0.486	0.580	0.795	0.793	0.829	0.981	0.981	0.982
		0.50	0.909	0.902	0.931	0.995	0.995	0.998	1.000	1.000	1.000
1.00	1.00	0.00	0.041	0.038	0.061	0.054	0.054	0.063	0.046	0.046	0.048
		0.05	0.071	0.066	0.105	0.097	0.096	0.111	0.145	0.143	0.158
		0.10	0.127	0.122	0.171	0.190	0.187	0.220	0.390	0.384	0.413
		0.15	0.223	0.217	0.287	0.389	0.384	0.416	0.679	0.674	0.695
		0.20	0.386	0.378	0.451	0.626	0.624	0.675	0.885	0.884	0.894
		0.25	0.545	0.535	0.608	0.817	0.815	0.837	0.987	0.987	0.990
		0.50	0.998	0.998	0.999	1.000	1.000	1.000	1.000	1.000	1.000

Table 4.3: Estimated significance level and power of tests for testing independence using the Normal copula function and Poisson CDF's at the nominal size $\alpha = 0.05$ for $n = 10$, $n = 15$, $n = 25$, and varying values of λ_1 and λ_2

			$n = 10$			$n = 15$			$n = 25$		
λ_1	λ_2	ρ	LR	ξ_S	Wald	LR	ξ_S	Wald	LR	ξ_S	Wald
1.00	1.50	0.00	0.097	0.087	0.229	0.085	0.075	0.174	0.056	0.058	0.110
		0.05	0.088	0.088	0.269	0.087	0.083	0.182	0.075	0.073	0.146
		0.10	0.101	0.087	0.243	0.103	0.099	0.213	0.100	0.090	0.150
		0.15	0.114	0.103	0.280	0.132	0.117	0.256	0.142	0.135	0.223
		0.20	0.167	0.153	0.332	0.158	0.142	0.307	0.244	0.234	0.350
		0.25	0.205	0.179	0.385	0.225	0.211	0.393	0.325	0.319	0.450
		0.50	0.459	0.418	0.703	0.589	0.584	0.782	0.831	0.831	0.912
1.00	2.00	0.00	0.093	0.077	0.233	0.081	0.074	0.175	0.062	0.055	0.111
		0.05	0.069	0.076	0.203	0.084	0.079	0.173	0.073	0.075	0.136
		0.10	0.110	0.093	0.258	0.086	0.084	0.191	0.114	0.104	0.182
		0.15	0.130	0.119	0.290	0.131	0.113	0.250	0.164	0.150	0.265
		0.20	0.156	0.130	0.325	0.195	0.189	0.333	0.240	0.235	0.343
		0.25	0.177	0.162	0.378	0.215	0.203	0.375	0.318	0.317	0.441
		0.50	0.432	0.390	0.645	0.552	0.541	0.715	0.774	0.774	0.861
1.50	1.50	0.00	0.066	0.061	0.206	0.079	0.074	0.166	0.068	0.064	0.130
		0.05	0.092	0.077	0.246	0.089	0.077	0.182	0.066	0.063	0.117
		0.10	0.107	0.091	0.256	0.112	0.107	0.216	0.112	0.104	0.189
		0.15	0.150	0.125	0.314	0.143	0.151	0.276	0.158	0.154	0.247
		0.20	0.153	0.137	0.341	0.177	0.163	0.291	0.244	0.235	0.344
		0.25	0.190	0.162	0.389	0.235	0.219	0.372	0.335	0.326	0.450
		0.50	0.469	0.442	0.717	0.630	0.613	0.806	0.870	0.869	0.940
1.50	2.00	0.00	0.075	0.080	0.207	0.083	0.083	0.172	0.062	0.058	0.125
		0.05	0.092	0.082	0.246	0.077	0.071	0.181	0.086	0.081	0.146
		0.10	0.106	0.097	0.264	0.099	0.091	0.203	0.102	0.094	0.174
		0.15	0.136	0.130	0.302	0.138	0.123	0.244	0.153	0.139	0.227
		0.20	0.161	0.148	0.326	0.175	0.165	0.308	0.251	0.246	0.359
		0.25	0.197	0.170	0.397	0.255	0.245	0.393	0.365	0.349	0.457
		0.50	0.444	0.407	0.683	0.617	0.611	0.786	0.806	0.803	0.892
2.00	2.00	0.00	0.084	0.076	0.214	0.071	0.070	0.170	0.062	0.058	0.117
		0.05	0.098	0.084	0.251	0.077	0.070	0.172	0.079	0.075	0.144
		0.10	0.125	0.106	0.297	0.103	0.100	0.216	0.103	0.100	0.185
		0.15	0.141	0.116	0.311	0.141	0.126	0.264	0.180	0.173	0.255
		0.20	0.155	0.136	0.342	0.174	0.156	0.324	0.248	0.239	0.338
		0.25	0.201	0.184	0.398	0.228	0.212	0.391	0.350	0.341	0.478
		0.50	0.423	0.392	0.663	0.549	0.523	0.724	0.784	0.787	0.869

Table 4.4: Estimated significance level and power of tests for testing independence using the Normal copula function and Poisson CDF's at the nominal size $\alpha = 0.05$ for $n = 50$, $n = 100$, $n = 200$, and varying values of λ_1 and λ_2

			$n = 50$			$n = 100$			$n = 200$		
λ_1	λ_2	ρ	LR	ξ_S	Wald	LR	ξ_S	Wald	LR	ξ_S	Wald
1.00	1.50	0.00	0.051	0.048	0.066	0.062	0.060	0.073	0.053	0.053	0.064
		0.05	0.073	0.073	0.100	0.101	0.096	0.125	0.137	0.136	0.144
		0.10	0.146	0.141	0.177	0.195	0.191	0.224	0.393	0.389	0.406
		0.15	0.245	0.239	0.295	0.418	0.414	0.455	0.702	0.702	0.719
		0.20	0.358	0.349	0.436	0.662	0.658	0.698	0.912	0.912	0.916
		0.25	0.598	0.594	0.659	0.836	0.835	0.862	0.989	0.989	0.991
		0.50	0.997	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1.00	2.00	0.00	0.053	0.049	0.077	0.062	0.060	0.073	0.045	0.044	0.053
		0.05	0.079	0.077	0.097	0.089	0.085	0.106	0.137	0.136	0.150
		0.10	0.140	0.135	0.188	0.201	0.200	0.222	0.380	0.375	0.402
		0.15	0.242	0.235	0.297	0.418	0.415	0.464	0.714	0.711	0.734
		0.20	0.392	0.382	0.454	0.683	0.681	0.709	0.920	0.920	0.930
		0.25	0.571	0.559	0.640	0.839	0.836	0.858	0.989	0.989	0.991
		0.50	0.966	0.967	0.986	1.000	1.000	1.000	1.000	1.000	1.000
1.50	1.50	0.00	0.068	0.063	0.086	0.054	0.052	0.067	0.051	0.049	0.056
		0.05	0.075	0.073	0.096	0.093	0.086	0.110	0.139	0.137	0.149
		0.10	0.117	0.111	0.137	0.215	0.212	0.246	0.364	0.362	0.387
		0.15	0.236	0.231	0.291	0.432	0.428	0.463	0.724	0.720	0.736
		0.20	0.402	0.393	0.472	0.658	0.655	0.692	0.932	0.932	0.940
		0.25	0.594	0.588	0.663	0.838	0.837	0.862	0.999	0.999	0.999
		0.50	0.991	0.991	0.998	1.000	1.000	1.000	1.000	1.000	1.000
1.50	2.00	0.00	0.049	0.046	0.066	0.052	0.049	0.064	0.059	0.059	0.061
		0.05	0.077	0.071	0.102	0.086	0.084	0.101	0.137	0.133	0.149
		0.10	0.156	0.153	0.193	0.224	0.217	0.249	0.392	0.388	0.409
		0.15	0.248	0.241	0.313	0.413	0.410	0.451	0.728	0.723	0.744
		0.20	0.395	0.390	0.477	0.675	0.668	0.713	0.928	0.927	0.936
		0.25	0.583	0.581	0.658	0.859	0.858	0.881	0.993	0.993	0.994
		0.50	0.983	0.983	0.989	1.000	1.000	1.000	1.000	1.000	1.000
2.00	2.00	0.00	0.065	0.065	0.090	0.059	0.057	0.075	0.052	0.052	0.060
		0.05	0.095	0.091	0.124	0.094	0.092	0.111	0.155	0.152	0.168
		0.10	0.166	0.164	0.213	0.228	0.225	0.253	0.382	0.380	0.402
		0.15	0.270	0.261	0.320	0.478	0.477	0.515	0.735	0.730	0.755
		0.20	0.428	0.420	0.487	0.696	0.695	0.738	0.920	0.920	0.930
		0.25	0.591	0.585	0.662	0.870	0.868	0.894	0.986	0.986	0.989
		0.50	0.963	0.965	0.982	1.000	1.000	1.000	1.000	1.000	1.000

Chapter 5

Conclusion

For this dissertation, four bonified score tests were derived for testing independence of two outcome variables. In one case, the bivariate distribution of interest was a probability mass function, in another case, the bivariate distribution of interest was a probability density function, in another case a score test was derived to determine which bivariate Pareto probability density function is more appropriate, in the last case the bivariate function was unknown and had to be estimated via the Normal copula function. In particular, the bivariate Poisson probability mass function as given by Lakshminarayana and the bivariate Pareto probability density functions as given by Sankaran–Nair and Lindley–Singpurwalla were chosen for this dissertation. For each test, a real world example was given to show the implementation of the test and interpretation of the results. Simulation studies were also performed to show each score test constructed achieved the appropriate size under a variety of different parameters and sample sizes. The power of each score test was also examined. As the correlation and sample size increased, so did power. In the simulation studies, the likelihood ratio and Wald tests were also compared to each of the score tests constructed. Overall, most score tests performed better in small samples as compared to the likelihood ratio and Wald tests. As sample size increased, the score tests and likelihood ratio tests performed equally well in most cases. The Wald test, on the other hand, didn't perform as well in some cases as compared to the likelihood ratio and score tests across sample size.

In conclusion, large sample hypothesis testing plays an important role in many statistical studies performed in the real world. It is to our advantage to compare the efficiency of these three tests in order to make a determination as to which method is most suitable for the study being conducted and the type of data being analyzed. Perhaps more importantly, assuming that certain criteria are met (e.g., independent outcomes) when in reality this is not the case, needs to be seriously addressed. Modeling random variables as independent when this in fact may not be true will most likely lead to misleading results. This dissertation has made an attempt to remedy this problem.

Chapter 6

Future Work

For this dissertation, no covariates were included in any of the bivariate models being used. This could be addressed in future work as real world application of these tests will probably include other variables of interest in addition to the responses. For each score test that was derived, a more general score test could also be derived with an inclusion of covariates into the models. For the score tests derived in chapter 2 and chapter 4, this is very doable and won't require much work beyond what was already done. However, an issue arises in chapter 3. If Sankaran and Nair's bivariate Pareto distribution is re-parameterized and reconstructed to include covariates, I discovered that there is no obvious and direct score test for testing independence of the response variables. Thus, a new bivariate distribution could be constructed (using multiple methods) and a test for independence could possibly be developed, or, once again, a copula approach could be implemented. Also, though the tests developed in chapter 3 are interesting and statistically valid (as shown in the simulation section), future work could be done to extend these results. Moreover, the bivariate Pareto distribution given by Sankaran and Nair is difficult to work with because it contains four unknown parameters that need to be estimated. An alternative and possibly attractive approach to test for independence of the response variables may be to use the Lomax copula, given in Balakrishnan and Lai (2009). Using this method, the bivariate distribution is defined by the Lomax copula and Pareto type II marginals, and Sankaran and Nair's parameterization is not needed. The copula approach proposed here is equivalent to testing $H_0 : \alpha_0 = \alpha_1\alpha_2$ versus $H_1 : \alpha_0 \neq \alpha_1\alpha_2$, which is the centerpiece of this chapter. Finally, the score test derived in chapter 4 presents many opportunities for future work. Instead of using the Normal copula, a different copula function could be used. Likewise, instead of using marginal Poisson CDF's, other discrete CDF's could potentially be substituted. Lastly, in chapter 4, discrete marginal CDF's were chosen and a score test was derived. However, what if the marginal CDF's were continuous? This is something that could easily be explored in future work.

Bibliography

Balakrishnan, N. and Lai, C.-D. (2009). *Continuous bivariate distributions*. Springer Science & Business Media.

Cameron, A. C., Li, T., Trivedi, P. K., and Zimmer, D. M. (2004). Modelling the differences in counted outcomes using bivariate copula models with application to mismeasured counts. *The Econometrics Journal*, 7(2):566–584.

Casella, G. and Berger, R. L. (2002). *Statistical inference*, volume 2. Duxbury Pacific Grove, CA.

Dykstra, R. and El Barmi, H. (1997). Chi-bar-square distributions. *Encyclopedia of Statistical Sciences*.

Famoye, F. (2010). A new bivariate generalized Poisson distribution. *Statistica Neerlandica*, 64(1):112–124.

Habtes, S., Muller-Karger, F. E., Roffer, M. A., Lamkin, J. T., and Muhling, B. A. (2014). A comparison of sampling methods for larvae of medium and large epipelagic fish species during spring SEAMAP ichthyoplankton surveys in the gulf of Mexico. *Limnology and Oceanography: Methods*, 12(2):86–101.

Joe, H. (2005). Asymptotic efficiency of the two-stage estimation method for copula-based models. *Journal of Multivariate Analysis*, 94(2):401–419.

Lakshminarayana, J., Pandit, S., and Srinivasa Rao, K. (1999). On a bivariate Poisson distribution. *Communications in Statistics-Theory and Methods*, 28(2):267–276.

Lindley, D. V. and Singpurwalla, N. D. (1986). Multivariate distributions for the life lengths of components of a system sharing a common environment. *Journal of Applied Probability*, pages 418–431.

Lyczkowski-Shultz, J., Hanisko, D. S., Sulak, K. J., Konieczna, M., and Bond, P. J. (2013). Characterization of ichthyoplankton in the northeastern gulf of Mexico from SEAMAP plankton surveys, 1982-1999. Gulf and Caribbean Research 25 (1): 43-98.

Muhling, B. A., Lamkin, J. T., and Richards, W. J. (2012). Decadal-scale responses of larval fish assemblages to multiple ecosystem processes in the northern gulf of Mexico. *Marine Ecology Progress Series*, 450:37–53.

Nelder, J. A. and Mead, R. (1965). A simplex method for function minimization. *The computer journal*, 7(4):308–313.

Nelsen, R. B. (2007). *An introduction to copulas*. Springer Science & Business Media.

Radhakrishna Rao, C. (1948). Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. *Mathematical Proceedings of the Cambridge Philosophical Society*, 44:50–57.

Ramsey, F. and Schafer, D. (2012). *The statistical sleuth: a course in methods of data analysis*. Cengage Learning.

Sacher, G. A. and Staffeldt, E. F. (1974). Relation of gestation time to brain weight for placental mammals: implications for the theory of vertebrate growth. *American Naturalist*, pages 593–615.

Sankaran, P. and Kundu, D. (2014). A bivariate Pareto model. *Statistics*, 48(2):241–255.

Sankaran, P. and Nair, N. U. (1993). A bivariate Pareto model and its applications to reliability. *Naval Research Logistics (NRL)*, 40(7):1013–1020.

Sklar, M. (1959). *Fonctions de répartition à n dimensions et leurs marges*. Université Paris 8.

Smith, P. E. and Moser, H. G. (2003). Long-term trends and variability in the larvae of pacific sardine and associated fish species of the California current region. *Deep Sea Research Part II: Topical Studies in Oceanography*, 50(14):2519–2536.

Yahav, I. and Shmueli, G. (2007). An elegant method for generating multivariate Poisson random variables. *ArXiv e-prints*.