GOODNESS-OF-FIT: A COMPARISON OF PARAMETRIC BOOTSTRAP AND EXACT CONDITIONAL TESTS

by

Wei Qian

Bachelor of Arts, Concordia University, 2006

Bachelor of Engineering, Harbin Institute of Technology, 1997

A PROJECT SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
in the Department
of
Statistics and Actuarial Science

SIMON FRASER UNIVERSITY
Summer 2008

APPROVAL

Name: Wei Qian

Degree: Master of Science

Title of project: Goodness-of-fit: a comparison of parametric bootstrap and exact conditional tests

Examining Committee: Dr. Jiguo Cao
Chair

____

Dr. Richard A. Lockhart
Senior Supervisor
Simon Fraser University

____

Dr. Michael A. Stephens
Supervisor
Simon Fraser University

____

Dr. Joan Hu
External Examiner
Simon Fraser University

Date Approved: August 26, 2008

Abstract

We study goodness of fit tests for exponential families. We compare, via Monte Carlo simulations, the powers of exact conditional tests based on co-sufficient samples (samples from the conditional distribution given the sufficient statistic) and approximate unconditional tests based on the parametric bootstrap. We use the Gibbs sampler to generate the co-sufficient samples. The gamma and von Mises families are investigated, and the Cramér-von Mises and Watson test statistics are applied. The results of this study show that those two tests have very similar powers even for samples of very small size, such as $n = 5$ for the gamma family and $n = 10$ for the von Mises family.

Acknowledgments

I would like to express my gratitude to my senior supervisor, Dr. Richard A. Lockhart, for his encouragement, patience, and guidance. In addition to his expertise and wisdom, he provided me vast assistance in writing reports. I also appreciate his generous support in each semester.

A special thanks goes out to Dr. Michael A. Stephens who is my co-supervisor. I am lucky to work with him. He spent much time and energy on my report and slides. I enjoyed every time talking with him.

I would like to thank Dr. Joan Hu, the external examiner in my committee. She taught me the first course at Simon Fraser University and provided me research assistantship in my first semester.

I would also like to thank other professors and staff in the department for teaching me courses and offering me help.

I acknowledge NSERC for the financial assistance during my graduate study.

Especially, thanks to my family for their understanding and support.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Suppose $X_1, ..., X_n$ is a random sample of size n from an unknown distribution F, and $\mathcal{F} = \{F_\theta : \theta \in \Theta\}$ is a parametric distribution family. We are going to determine whether the sample is from a distribution in $\mathcal{F}$. The general test of goodness of fit is a test of

$$H_0: \ F \in \mathcal{F}, \text{ or } F = F_\theta; \ \theta \in \Theta. \tag{1.1}$$

If the parametric vector θ is fully specified, the null hypothesis H_0 is simple. We call this Case 0, following Stephens (1986). Otherwise, with one or more elements of θ being unknown, H_0 is composite.

In goodness-of-fit, tests based on the empirical distribution function (EDF) are highly recommended, in which the EDF statistics are applied. For Case 0, those tests of fit could be simplified to a test of uniformity by the probability integral transformation (PIT). Tables of the critical values for those statistics are available, even for finite samples. When θ contains unknown element(s), the EDF statistics use the same calculation formulas except replacing F_θ by a suitable estimate. However, different from those of Case 0, their distributions will depend on the null distribution family tested, the method of estimation, and the sample size. That means different tables are required for the different families. Asymptotic tables of the critical values have been provided for some families in Stephens (1986). If the unknown parameters are location or scale parameters, and appropriate estimates are applied, the distributions of the EDF statistics will not depend on the unknown parameters but on the null distribution family and the sample size. However, if there is an unknown shape parameter present, the distributions of EDF statistics depend on the true value of the unknown shape parameter. When the dependence is slight, the tables are still useful as a reference. But, when the dependence is strong, it is risky to use these tables.

Instead of using the tables, a Monte Carlo method is suggested. It works as follows. First, find an estimate $\hat{\theta}$ of θ for the given sample, and fit a parametric model $F_{\hat{\theta}}$ to the given sample. If H_0 is true, then the fitted model $F_{\hat{\theta}}$ will be a close approximation to the unknown true model. Second, independent Monte Carlo samples are derived from the fitted model. Third, calculate the values of the test statistic, say S, for the original sample and derived samples. The p-value is evaluated by comparing the value of the test statistic from the original sample to those from the derived samples. That is, the EDF of those values of S from the derived samples simulates the true distribution of S. Finally, if the p-value is less than the significance level α, we reject H_0; otherwise, we accept it. This is the test based on the *parametric bootstrap*. See Cheng (2006) for more information. The Monte Carlo sample size should be as large as possible to make the EDF of the values of S as close as possible to its true distribution. Usually, the number is limited by computing ability and time. When the sample size is finite, increasing the number of Monte Carlo samples only makes the EDF of the test statistic values approach the distribution of S under $F_{\hat{\theta}}$, instead of F_{θ}. Thus, the test based on the parametric bootstrap is an approximate unconditional test.

Recently, it has been proposed to generate the samples from the conditional distribution of $X_1, ..., X_n$ given the sufficient statistic T_n, $F(X_1, ..., X_n|T_n)$, which leads to an exact conditional test. Those samples are called "*look-alike* " samples or "*samples" in O'Reilly & Gracia-Medrano (2006) or "*co-sufficient*" samples in Lockhart *et al.* (2007) and Lockhart *et al.* (2008). In this project, we will use the term "*co-sufficient*". Because of the sufficiency, the conditional distribution of any statistic calculated from the co-sufficient samples will not depend on the unknown parameters. In the composite case of goodness of fit, if H_0 is true, the co-sufficient samples given the value of the sufficient statistic have exactly the same distributional properties as the original sample. Therefore, by increasing the number of the Monte Carlo samples, tests based on the co-sufficient samples can be made as exact as required.

This project is a complement to Lockhart *et al.* (2007) and Lockhart *et al.* (2008). It compares powers of the exact conditional test based on co-sufficient samples and the approximate unconditional test based on the parametric bootstrap. We will use the same notations as those in the above two papers, as well as their methods of generating the co-sufficient samples. Large-sample theory suggests that the exact test should have the same power as the bootstrap test when the sample size is large enough. In this project, we will answer the question of "how large is large enough". Chapter 2 gives a general set-up for our study. Chapter 3 summarizes the methods so far available for generating the co-sufficient samples. In Chapters 4 and 5, the gamma and von Mises families are investigated, and the powers of the exact conditional tests and the approximate unconditional tests against different alternatives are studied. Finally, we summarize our work, and give some discussion of possible future work.

Chapter 2

General set-up

We investigate tests for the gamma and von Mises families in this project. These are members of the exponential families, whose sufficient statistics can be found quickly.

We review exponential families and quadratic EDF statistics in Section 2.1 and define the exact conditional test in Section 2.2. Finally, we illustrate the Monte Carlo realization of both exact and approximate tests, as well as their power study.

2.1 Review

2.1.1 Exponential families

Generally, a k-parameter exponential family of full rank has density or mass function of the form

$$f(x;\theta) = g(x)\exp\{\eta_1(\theta)h_1(x) + \dots + \eta_k(\theta)h_k(x) + c(\theta)\}\,;\theta \in \Theta, \tag{2.1}$$

where θ is a k-dimensional parameter vector, $c(\cdot)$, $g(\cdot)$, $\eta_i(\cdot)$, and $h_i(\cdot)$ are known functions, and $(\eta_1, \dots, \eta_k)$ has dimension k.

For a random sample $X_1, \dots, X_n$ from the above distribution with θ unknown, the minimal sufficient statistic for θ is $T_n = (T_{n1}, \dots, T_{nk})$, where $T_{ni} = \sum_{j=1}^{n} h_i(X_j)$.

2.1.2 Quadratic EDF statistics

Suppose $X_1, \dots, X_n$ is a random sample from a continuous distribution F_θ, which may be abbreviated as $F_\theta(\cdot)$ or F_θ. Let $X_{(1)} < X_{(2)} < \dots < X_{(n)}$ be the order statistics. Based on

$X_1, ..., X_n$, the empirical distribution function (EDF) F_n is a step function defined by

$$F_n(u) = \begin{cases} 0 & \text{if } u < X_{(1)}; \\ \frac{i}{n} & \text{if } X_{(i)} \leq u < X_{(i+1)}, \quad i = 1, ..., n-1; \\ 1 & \text{if } X_{(n)} \leq u. \end{cases}$$

As $n \to \infty$, $F_n \to F_\theta$.

The statistics based on the EDF measure the discrepancy between F_θ and F_n. Only quadratic EDF statistics are applied in our study. The general form of a quadratic EDF statistic Q, for the simple case, is

$$Q = n \int_{-\infty}^{\infty} \{F_n(u) - F_\theta(u)\}^2 \psi(F_\theta(u)) \mathrm{d}F_\theta(u),$$

where $\psi(F_\theta(u))$ is a weight function. Different weights give different statistics. The choice $\psi(x) = 1$ gives the Cramér-von Mises statistic W^2. The choice $\psi(x) = [x\{1-x\}]^{-1}$ gives the Anderson-Darling statistic A^2.

A related statistic is Watson's statistic U^2 given by

$$U^2 = n \int_{-\infty}^{\infty} \left\{ F_n(u) - F_\theta(u) - \int_{-\infty}^{\infty} [F_n(u) - F_\theta(u)] \mathrm{d}F_\theta(u) \right\}^2 \mathrm{d}F_\theta(u).$$

This statistic is applicable for circular data because it is invariant under rotations.

For case 0, if H_0 is true, the EDF statistic has a distribution independent of the null distribution F_θ. Let $Z = F(X;\theta)$ (the Probability Integral Transformation PIT), then $Z \sim U(0,1)$. Statistic Q could be written as

$$Q = n \int_0^1 \{F_n^*(z) - z\}^2 \psi(z) \mathrm{d}z;$$

where $F_n^*(z)$ is the EDF of the $z_i = F(x_i;\theta)$. Thus, all the simple hypothesis cases will be a test of uniformity.

Computing Formulas

Let $z_i = F(x_i;\theta)$, and let $z_{(i)}$, $i = 1, ..., n$ be the order statistics; practical computing formulas for W^2, A^2, and U^2 are given by

$$W^2 = \sum_i \left\{ Z_{(i)} - \frac{2i-1}{2n} \right\}^2 + \frac{1}{12n};$$

$$
\begin{aligned}
A^2 &= -n - \frac{1}{n}\sum_i (2i-1)[\log Z_{(i)} + \log\{1 - Z_{(n+1-i)}\}] \text{ or,} \\
&= -n - \frac{1}{n}\sum_i [(2i-1)\log Z_{(i)} + (2n+1-2i)\log\{1 - Z_{(i)}\}]; \\
U^2 &= W^2 - n(\bar{Z} - 0.5)^2.
\end{aligned}
$$

In the case of unknown parameters, the same formulas are used with $\hat{z}_i = F(x_i; \hat{\theta})$, where $\hat{\theta}$ is an approriate estimate (usually the MLE).

In this project, W^2 and A^2 are used to test the gamma family, U^2 to test the von Mises family.

2.2 Exact conditional test

When considering the problem of testing H_0 defined in 1.1 where θ is unknown, we wish for a powerful unbiased test with the critical function Φ, where

$$
\Phi = \begin{cases} 1 & \text{if } H_0 \text{ is rejected;} \\ 0 & \text{if } H_0 \text{ is accepted.} \end{cases} \tag{2.2}
$$

Denote the power function by $\beta_\Phi(F)$, which is the probability that the test rejects H_0 when the sample is from F; then,

$$
\beta_\Phi(F) = E_F[\Phi].
$$

A test of H_0 is said to be unbiased if the power function satisfies

$$
\begin{aligned}
\beta_\Phi(F) &\leq \alpha, \text{ if } F \in \mathcal{F}; \\
\beta_\Phi(F) &\geq \alpha, \text{ if } F \notin \mathcal{F}.
\end{aligned}
$$

That is, for no alternative should the rejection probability be less than the size of the test; otherwise, there will exist an alternative distribution which is more likely to be accepted than is some distribution in the null hypothesis.

Suppose the power function $\beta_\Phi(F)$ is a continuous function of F, unbiasedness implies

$$
\beta_\Phi(F) = E_F[\Phi] = \alpha; \text{ for all } F \in \mathcal{F}. \tag{2.3}
$$

Tests satisfying Condition 2.3 are said to be similar.

Suppose that there exists a sufficient statistic T for $\mathcal{F}$, and $\mathcal{T}$ is the family of distributions of T as F varies over $\mathcal{F}$. If $\mathcal{T}$ is complete, Condition 2.3 is equivalent to

$$E[\Phi|t] \equiv \alpha, \text{ for all } t. \tag{2.4}$$

Because of the sufficiency, the conditional power function over H_0 is independent of θ, and Condition 2.4 reduces the problem to that of testing a simple hypothesis for each value of t. Thus, any good unbiased test must be some exact conditional test, when there is a complete sufficient statistic to the null distribution.

Let S be an appropriate test statistic defined in Subsection 2.1.2. The distribution of S is $G(s;\theta)$ which depends on the true value of the unknown θ. For an observed sample, suppose $T = t$, we denote the conditional distribution of S given $T = t$ by $G_c(s|t)$; then, $G_c(s|t)$ is independent of θ.

We construct a nonrandomized test, conditioned on $T = t$, with the critical function $\Phi(S, t)$, satisfying

$$\Phi(S,t) = \begin{cases} 1 & \text{if } S > c(t), \\ 0 & \text{otherwise;} \end{cases} \tag{2.5}$$

where the function $c(t)$ are found by solving

$$E_{F_\theta}[\Phi(S,T)|t] = \alpha, \text{ for all } t. \tag{2.6}$$

This conditional test is exact because its rejection probability over H_0 is exactly α. For this given sample, suppose $S = s_0$, then its exact p-value is $p_c(s_0) = 1 - G_c(s_0|t)$.

If an alternative distribution, say K, is specified, the conditional power function of the exact conditional test given $T = t$ is

$$\beta_\Phi(K|t) = E_K[\Phi(S,T)|t].$$

The overall power is

$$\begin{aligned} \beta_\Phi(K) &= E_K[\beta_\Phi(K|T)] \\ &= E_K[E_K[\Phi(S,T)|T]] \\ &= E_K[\Phi(S,T)]. \end{aligned} \tag{2.7}$$

2.3 The Monte Carlo methods

Our power study is a two-level Monte Carlo study to give powers of two methods. Each of a sequence of samples is tested by both of these Monte Carlo test methods.

2.3.1 Monte Carlo implementation of the tests

Given an observed sample $x_1, ..., x_n$, we calculate the sufficient statistic $t_n = T_n(x_1, ..., x_n)$, the maximum likelihood estimate $\hat{\theta}$ of θ, and the test statistic $s_0 = S(x_1, ..., x_n)$. The procedures of tests based on both the co-sufficient and parametric bootstrap samples are as follows.

Exact conditional test based on the co-sufficient samples

Given $T_n = t_n$, we calculate the conditional p-value by $p_c(s_0) = \Pr(S > s_0|T_n = t_n) = 1 - G_c(s_0|t_n)$, if $p_c(s_0) < \alpha$, reject H_0; otherwise, accept it. Since $G_c(s|t_n)$ is hard or impossible to calculate, we use co-sufficient samples to simulate $G_c(s|t_n)$. Steps are as follows.

1. *Generate* $x^*_{1,1}, ..., x^*_{1,n}$ *from* $F(X_1, ..., X_n|t_n)$, *calculate* $s^*_1 = S(x^*_{1,1}, ..., x^*_{1,n})$;

   repeat the above step until B;

B. *generate* $x^*_{B,1}, ..., x^*_{B,n}$ *from* $F(X_1, ..., X_n|t_n)$, *calculate* $s^*_B = S(x^*_{B,1}, ..., x^*_{B,n})$;

   return $s^*_1, ..., s^*_B$.

We estimate the exact conditional p-value by $p_c(s_0) \approx \frac{1}{B}\sum_{j=1}^{B} I(s^*_j > s_0)$. As B increases, the estimate will become as accurate as required. The generation of co-sufficient samples is described in 3.2.

Approximate unconditional test based on the parametric bootstrap

The approximate p-value for the given sample is calculated by $p_{\hat{\theta}}(s_0) = 1 - G(s_0; \hat{\theta})$. We simulate $G(S; \hat{\theta})$ by using parametric bootstrap samples. Steps are as follows.

1. *Generate* $x_{1,1}, ..., x_{1,n}$ *from* $F_{\hat{\theta}}$, *calculate* $s_1 = S(x_{1,1}, ..., x_{1,n})$;

2. *generate* $x_{2,1}, ..., x_{2,n}$ *from* $F_{\hat{\theta}}$, *calculate* $s_2 = S(x_{2,1}, ..., x_{2,n})$;

repeat the above step until B;

B. generate $x_{B,1}, ..., x_{B,n}$ from $F_{\hat{\theta}}$, calculate $s_B = S(x_{B,1}, ..., x_{B,n})$;

return $s_1, ..., s_B$.

The approximate p-value $p_{\hat{\theta}}(s_0)$ is then estimated by $p_{\hat{\theta}}(s_0) \approx \frac{1}{B}\sum_{j=1}^{B} I(s_j > s_0)$. As B increases, the estimate will be as accurate as required. However, since the parametric bootstrap is based on the $F_{\hat{\theta}}$, and not on F_θ, the test is approximate.

2.3.2 Monte Carlo integration for the power function

When the alternative K is fully specified, we are interested in the powers of the approximate unconditional test and the exact conditional test. M independent samples are generated from K; for ith sample, we define Φ_i as in Equation 2.2.

The power function for the alternatives K will be estimated by Monte Carlo integration,

$$\beta_\Phi(K) \approx \frac{1}{M}\sum_{i=1}^{M} \Phi_i; \text{where } M \to \infty. \tag{2.8}$$

That is, the overall power is estimated by the average of the values of Φ_i. As M increases, the accuracy will be improved.

In our study, we set $M = 500$, and $B = 1,000$. For different M, standard errors of the power estimates are listed in Table 2.1. The same standard errors apply to an individual simulated p-value (a p-value for a single sample) when M is replaced by B, the number of bootstrap or co-sufficient samples used.

Power (p-value)	0.05	0.10	0.20	0.30	0.40	0.50
M =500	0.0097	0.0134	0.0179	0.0205	0.0219	0.0224
M =1,000	0.0069	0.0095	0.0126	0.0145	0.0155	0.0158
M =10,000	0.0022	0.0030	0.0040	0.0046	0.0049	0.0050

Table 2.1: Standard errors of the power estimates when $M = 500, 1,000$, and $10,000$. These standard errors also apply to individual p-value when M is replaced by B.

Chapter 3

Generating the co-sufficient samples

In this chapter, we will discuss two methods of generating the co-sufficient samples: use of Rao-Blackwell estimates in Section 3.1 and use of the Gibbs sampler in Section 3.2.

Suppose $X_1, ..., X_n$ is a random sample from F_θ defined in Equation 2.1, where θ is not specified, and T_n is the sufficient statistic, with k components, for θ in F_θ. Given $T_n = t_n$, a co-sufficient sample $x_1^*, ..., x_n^*$ is a sample from $F(X_1, ..., X_n|t_n)$. The conditioning reduces the effective dimension of the sample by k. That is, to obtain a co-sufficient sample $x_1^*, ..., x_n^*$, we only need generate $x_{k+1}^*, ..., x_n^*$, and then $x_1^*, ..., x_k^*$ could be found by solving a system of equations $T_n(x_1^*, ..., x_n^*) = t_n$. Moreover, given $T_n = t_n$, $X_1, ..., X_n$ are exchangeable, so $x_{k+1}^*, ..., x_n^*$ could be any $n-k$ entries of the co-sufficient sample.

3.1 Use of the Rao-Blackwell estimate

O'Reilly & Gracia-Medrano (2006) applied the Rao-Blackwell estimate of F_θ to generate the co-sufficient sample directly for inverse Gaussian family. Given $T_n = t_n$, the Rao-Blackwell estimate of F_θ, denoted by $\tilde{F}_n$, is defined by

$$\tilde{F}_n(x) = \Pr(X_i \leq x | T_n = t_n),$$

where i could be $1, ..., n$. We denote the inverse $\tilde{F}_n$ of by $\tilde{F}_n^{-1}$. If $\tilde{F}_n^{-1}$ can be computed, the following algorithm may be used to generate independent co-sufficient samples:

Algorithm

1. *Generate a* u_n *from* $U(0,1)$*, then* $x_n^* = \tilde{F}_n^{-1}(u_n)$*;*

2. *recalculate* t_{n-1}^* *from* t_n *and* x_n^**;*

3. *define* $\tilde{F}_{n-1}(x) = P(X_i \leq x | T_{n-1} = t_{n-1}^*)$*, where* i *could be* $1, ..., n-1$*, and let* $\tilde{F}_{n-1}^{-1}$ *be the inverse; generate another* u_{n-1} *from* $U(0,1)$ *independent of* u_n*, then* $x_{n-1}^* = \tilde{F}_{n-1}^{-1}(u_{n-1})$*;*

4. *continue computing* t_{n-2}^**,* x_{n-2}^* *and so on, we obtain* $x_n^*, ..., x_{k+1}^*$*;*

5. *solve the system of equations* $T_n(x_1^*, ..., x_n^*) = t_n$ *to get* $x_1^*, ..., x_k^*$*;*

6. *return* $(x_1^*, ..., x_n^*)$ *as a complete co-sufficient sample.*

The above algorithm is justified as follows. Assume F_θ is a continuous distribution. We write the conditional joint density of $X_{k+1}, ..., X_n$ given t_n by $f(x_{k+1}, ..., x_n | t_n)$; then,

$$\begin{aligned}
f(x_{k+1}, ..., x_n | t_n) &= \frac{f(x_{k+1}, ..., x_n, t_n)}{f(t_n)} \\
&= \frac{f(x_{k+1}, ..., x_n, t_n)}{f(t_n)} \frac{f(x_{k+2}, ..., x_n, t_n)}{f(x_{k+2}, ..., x_n, t_n)} \cdots \frac{f(x_n, t_n)}{f(x_n, t_n)} \\
&= \frac{f(x_{k+1}, ..., x_n, t_n)}{f(x_{k+2}, ..., x_n, t_n)} \frac{f(x_{k+2}, ..., x_n, t_n)}{f(x_{k+3}, ..., x_n, t_n)} \cdots \frac{f(x_n, t_n)}{f(t_n)} \\
&= f(x_{k+1} | x_{k+2}, ..., x_n, t_n) f(x_{k+2} | x_{k+3}, ..., x_n, t_n) \cdots f(x_n | t_n) \\
&= f(x_n | t_n) f(x_{n-1} | x_n, t_n) \cdots f(x_{k+1} | x_{k+2}, ..., x_n, t_n).
\end{aligned}$$

The terms in the right side of the identity correspond to the steps in the algorithm.

3.2 Use of the Gibbs sampler

Generally, the Rao-Blackwell estimate $\tilde{F}_n$ cannot be evaluated. The Gibbs sampler has been proposed to generate co-sufficient samples for the gamma family in Lockhart *et al.* (2007) and the von Mises family in Lockhart *et al.* (2008).

The Gibbs sampler is a special case of the Metropolis-Hastings algorithm where the random candidate is always accepted. It was first devised by Geman & Geman (1984) for image processing and popularized by Casella & George (1992). Chib & Greenberg (1995) and Walsh (2004) gave good introductions to MCMC and Gibbs sampler. The

Gibbs sampler is used to generate samples when the target distribution is a multivariate distribution. The point is that given the joint density function, it is simpler to sample from a univariate conditional distribution than to integrate over a joint distribution to get a marginal distribution. For example, when asked to generate samples of a bivariate vector (X, Y) with the joint density $p(x, y)$, we begin with a value y_0 of Y and for $i = 1, 2, ...$ and so on, sample x_i from $p(x|Y = y_{i-1})$; once x_i is found, sample y_i from $p(y|X = x_i)$. Then, a sequence of samples (x_i, y_i) is obtained. The Gibbs sampler is applicable when the joint distribution is not known explicitly, but it requires a set of full conditional distributions.

In our case, the target distribution is the conditional distribution of $X_{k+1}, ..., X_n$ given $T_n = t_n$. When using the Gibbs sampler, the set of full conditional densities is

$f_c(x_{k+1}|x_{k+2}, ..., x_n, t_n)$,

...,

$f_c(x_i|x_{k+1}, ..., x_{i-1}, x_{i+1}, ..., x_n, t_n)$,

...,

$f_c(x_n|x_{k+1}, ..., x_{n-1}, t_n)$.

Denote the full conditional distribution of x_i by $f_c(x_i)$. Then, $f_c(x_i)$ is derived from the joint density $f_1(x_{k+1}, ..., x_n, t_n)$, and

$$f_c(x_i) \propto f_1(x_{k+1}, ..., x_n, t_n), \tag{3.1}$$

where $f_1(x_{k+1}, ..., x_n, t_n)$ is computed from the joint density $f_s(x_1, ..., x_n)$ using the change of variables formula. The algorithm for using the Gibbs sampler is as follows.

3.2.1 Algorithm

1. *Begin with the original sample* $x_1^{(0)}, ..., x_n^{(0)}$.

2. *Repeat for* $j = 0, ..., M$,

   generate $x_{k+1}^{(j+1)}$ *from* $f_c(x_{k+1}|x_{k+2}^{(j)}, ..., x_n^{(j)}, t_n)$,

   generate $x_{k+2}^{(j+1)}$ *from* $f_c(x_{k+2}|x_{k+1}^{(j+1)}, x_{k+3}^{(j)}, ..., x_n^{(j)}, t_n)$,

   ...,

generate $x_n^{(j+1)}$ *from* $f_c(x_n|x_{k+1}^{(j+1)}, ..., x_{n-1}^{(j+1)}, t_n)$,

solve $T_n = t_n$ *to find* $x_1^{(j+1)}, ..., x_k^{(j+1)}$,

obtain the sample $x_1^{(j+1)}, ..., x_n^{(j+1)}$.

3. *Return a sequence of M samples where the jth sample is* $\left\{x_1^{(j)}, ..., x_n^{(j)}\right\}$.

The sequence of samples generated is a Markov chain whose statitionary initial distribution is the conditional distribution of $X_1, ..., X_n$ given $T_n = t_n$. If $x_1^{(0)}, ..., x_n^{(0)}$ is known to be from the null distribution, all the samples generated are co-sufficient samples; otherwise, a burn-in period is required so that the Markov chain converges to its statitionary distribution. The first 100 samples are discarded here. Moreover, in order to have the samples approximately independent of each other, we take only every 10th sample.

In Lockhart *et al.* (2007), the values are always drawn from $f_c(x_n)$. A new value x_n^* is drawn from $f_c(x_n)$. The variables are rotated so that x_{k+2} becomes x_{k+1}, x_{k+3} becomes x_{k+2}, and so on x_n^* becomes x_{n-1}, and x_3 becomes the new x_n with no asterisk. This is equivalent to the above algorithm since given $T_n = t_n$, $X_{k+1}, ..., X_n$ are exchangeable, and all the $f_c(x_i)$'s share a common computational formula. In the remaining parts of this document, we will discuss only $f_c(x_n)$.

3.2.2 Draw a value from $f_c(x_n)$

Write $f_c(x_n) = Ch(x_n)$, where C is a normalizing constant. Let F_c be the corresponding CDF. If F_c could be written explicitly, we could use the inversion method as follows:

1. Generate u from $U(0, 1)$,

2. return $x_n = F_c^{-1}(u)$, where F_c^{-1} is the inverse of F_c.

However, this method, applied to gamma and von Mises distributions, involves complicated special functions, and has not been used here. We use an acceptance-rejection (A-R) method.

Acceptance-rejection (A-R) method

Suppose there exists a known constant c such that $h(x_n) \leq cp(x_n)$ for all x_n, where p is some so-called candidate density.

1. Generate a candidate z from $p(z)$, and u from $U(0, 1)$.

2. If $u \leq \frac{h(z)}{cp(z)}$, return $x_n^* = z$;

3. else, go to (1).

In the A-R method, c is the expected number of iterations to accept z, and it determines the rate of acceptance. An optimal value of c is $c = \sup_x \frac{h(x)}{p(x)}$.

3.3 Comments

There is another more direct method that we are not going to discuss here because it is only applicable for those special cases with only unknown location or scale parameters. For them, if appropriate estimates are applied, the distribution of the test statistic will not depend on the unknown parameters, and we can utilize the asymptotic tables in Stephens (1986). It is also possible in these cases to use the parametric bootstrap to give exact tests; see Lillegard & Engen (1997) and Lindqvist & Taraldsen (2005).

Chapter 4

The Gamma Family

In this chapter, the gamma family is investigated. A distribution of the gamma family has the probability density function (PDF),

$$f(x;\alpha,\beta) = \frac{x^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)}\exp(-x/\beta), x \geq 0; \beta > 0, \alpha > 0,$$

where α is the shape parameter, and β is the scale parameter. Graphs of the PDF and CDF are shown in Figure 4.1, for various shapes and scales.

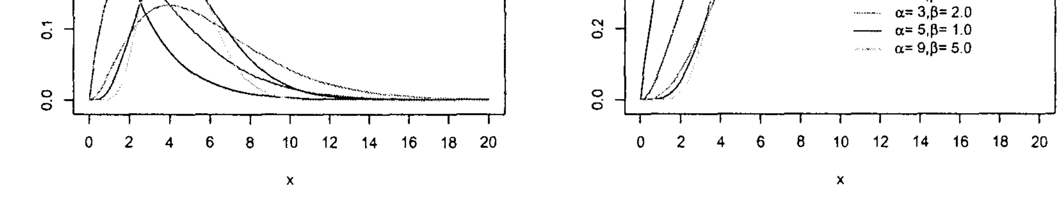


Figure 4.1: PDF and CDF of gamma distribution

When $\alpha > 1$, the Gamma distribution assumes a mounded, but skewed shape. The skewness decreases as the value of α increases. When $\alpha = 1$, it is a exponential distribution with the scale parameter β. When $\alpha < 1$, it is exponentially shaped but the density approaches ∞ as $x \to 0$.

In Section 4.1, we derive the full conditional density $f_c(x_n)$ and discuss the difficulty of using the A-R method to draw a value x_n^* from $f_c(x_n)$. Section 4.2 presents Monte Carlo computations of the powers of both the approximate and exact tests against different gamma alternatives.

4.1 Generating the co-sufficient samples

When the null distribution is a gamma distribution whose shape parameter α is unknown, no direct method is available to generate the co-sufficient samples, and we use the Gibbs sampler. Two cases are discussed here.

Case 1: α is unknown, and $\beta = \beta_0$ is known,

Case 2: both α and β are unknown.

4.1.1 Case 1: α is unknown, and $\beta = \beta_0$ is known

Given a sample data $y_1, ..., y_n$, we want to know whether the sample is from a distribution of the gamma family with a fixed scale β_0. Consider the test of

$$H_0 : (y_1, ..., y_n) \text{ comes from Gamma}(\alpha, \beta = \beta_0), \alpha \text{ is unknown.}$$

Let $x_i = y_i/\beta_0$; then H_0 can be simplified to

$$H_0' : (x_1, ..., x_n) \text{ comes from Gamma}(\alpha, \beta = 1), \alpha \text{ is unknown.}$$

The likelihood function is

$$L(\alpha) = \frac{1}{\Gamma^n(\alpha)} \left(\prod_{i=1}^{n} x_i \right)^{\alpha - 1} \exp\left(-\sum_{i=1}^{n} x_i \right).$$

Then, the sufficient statistic of α is $t_n = \prod_{i=1}^{n} x_i$ (or $\sum_{i=1}^{n} \log(x_i)$). The maximum likelihood estimate $\hat{\alpha}$ of α, is found by solving the equation,

$$\psi(\hat{\alpha}) = \sum_{i=1}^{n} \log(x_i)/n,$$

where $\psi(.)$ is the digamma function defined as

$$\psi(t) = \frac{\Gamma'(t)}{\Gamma(t)}.$$

The joint density function of $x_2, ..., x_n, t_n$ is as follows

$$f_1(x_2, ..., x_n, t_n) = ||J(x_2, ..., x_n, t_n)||^{-1} f_s(x_1, ..., x_n, \alpha);$$

where $||A||$is the absolute determinant of A, and the Jacobian is calculated as

$$\begin{aligned} ||J(x_2, ..., x_n, t_n)|| &= ||\frac{\partial(x_2, ..., x_n, t_n)}{\partial(x_1, ..., x_n)}|| \\ &= \prod_{i=2}^{n} x_i. \end{aligned}$$

Thus,

$$\begin{aligned} f_1(x_2, ..., x_n, t_n) &= \frac{1}{\prod_{i=2}^{n} x_i} f_s(x_1, ..., x_n; \alpha) \\ &= \frac{1}{\prod_{i=2}^{n} x_i} \left(\frac{1}{\Gamma(\alpha)}\right)^n t_n^{\alpha-1} \exp\left(-\sum_{i=2}^{n-1} x_i - \frac{t_n}{\prod_{i=2}^{n} x_i}\right) \end{aligned}$$

where $0 < x_2, ..., x_n, t_n < +\infty$.

The full conditional distribution of x_n given $x_2, ..., x_{n-1}, t_n$ is

$$\begin{aligned} f_c(x_n) &= f_c(x_n | x_2, ..., x_{n-1}, t_n) \\ &= \frac{f_1(x_2, ..., x_n, t_n)}{\int_0^{+\infty} f_1(x_2, ..., x_n, t_n) \mathrm{d}x_n} \\ &\propto \frac{1}{x_n} \exp(-x_n - \frac{a}{x_n}) \\ &\equiv h(x_n);\ 0 < x_n, \end{aligned}$$

where $a = t_n / \left(\prod_{j=2}^{n-1} x_j\right)$. This quantity a must be updated every time a new x_n is generated.

The function $h(\cdot)$ does not have a closed form integral. Thus we cannot use the inversion method to generate a value, so the acceptance-rejection method is applied to draw a new x_n. The value of a affects the shape of $h(\cdot)$ dramatically (see Figure 4.2). To make the A-R method efficient, we choose the candidate generation distribution $p(x)$ according to the value of a.

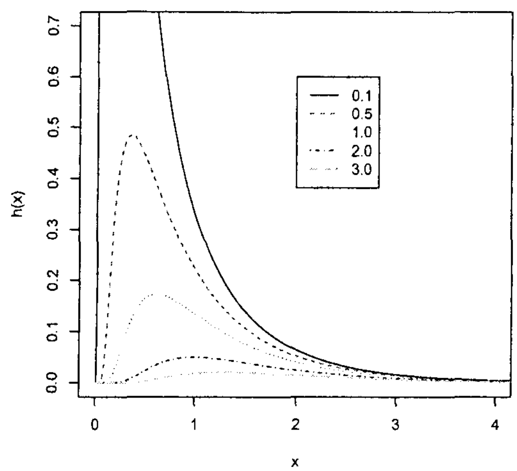


Figure 4.2: Graph of $h(\cdot)$ under different values of a

When $a < 1$, we use

$$p(x) = \begin{cases} \frac{C_1}{C}\frac{\exp(-a/x)}{x} & x \le b; \\ \frac{C_2}{C}\frac{\exp(-x)}{x} & x > b, \end{cases}$$

where $b = \frac{-1+\sqrt{1+4a}}{2}$, $C_1 = \exp(-b)$, $C_2 = \exp(-a/b)$ and $C = C_1\mathrm{Ei}(1, \frac{a}{b}) + C_2\mathrm{Ei}(1, b)$. Here, Ei is the exponential integral function; see Abramowitz & Stegun (1972). The CDF of $p(x)$ is

$$P(x) = \begin{cases} \frac{C_1}{C}\mathrm{Ei}(1, \frac{a}{x}) & x \le b; \\ 1 - \frac{C_2}{C}\mathrm{Ei}(1, x) & x > b. \end{cases}$$

Therefore, the probability of acceptance α is calculated by

$$\alpha = \frac{h(x)}{p(x)C_{max}} = \frac{C}{C_1 C_{max}}\exp(-x)\mathbf{I}_{(x\le b)} + \frac{C}{C_2 C_{max}}\exp(-a/x)\mathbf{I}_{(x>b)},$$

where $C_{max} = \sup_{x>0}\frac{h(x)}{p(x)} = C \times \max(\frac{1}{C_1}, \frac{1}{C_2})$.

When $a \ge 1$, the candidate distribution could be Gamma(shape=κ, scale=1), where $\kappa = \frac{1+\sqrt{1+4a}}{2}$. The probability of acceptance α is given by

$$\alpha = \frac{h(x)}{p'(x)C'_{max}} = \frac{\exp(-a/x)}{x^{\kappa}},$$

where $C'_{max} = \exp(-\kappa)(\kappa/a)^{\kappa}$.

4.1.2 Case 2: both α and β are unknown

For an observed sample $x_1, ..., x_n$, where both α and β in the null distribution are unknown, the likelihood function will be

$$\begin{aligned} L(\alpha,\beta) &= \prod_{i=1}^{n} f(x_i;\alpha,\beta) \\ &= \frac{1}{\beta^{n\alpha}}\frac{1}{\Gamma^n(\alpha)}(\prod_{i=1}^{n} x_i)^{\alpha-1}\exp(-\sum_{i=1}^{n} x_i/\beta); \end{aligned}$$

Therefore, the sufficient statistic is $T_n = (s_n, p_n) = (\sum_{i=1}^n x_i, \prod_{i=1}^n x_i)$. The MLEs $\hat{\alpha}$ and $\hat{\beta}$ are found by solving the equations

$$\begin{aligned} \log(p_n)/n - \log(s_n/n) &= \psi(\hat{\alpha}) - \log(\hat{\alpha}), \\ \hat{\beta} &= \frac{s_n/n}{\hat{\alpha}}. \end{aligned}$$

Similar to Case 1, the joint density function of $x_3, ..., x_n, s_n, p_n$ is given by

$$f_1(x_3, ..., x_n, s_n, p_n) = 2\frac{p_n^{\alpha-1}\exp(-s_n/\beta)}{\beta^{n\alpha}\Gamma^n(\alpha)}\|J(x_3, ..., x_n, s_n, p_n)\|^{-1},$$

where the Jacobian $J(x_3, ..., x_n, s_n, p_n)$ is given by

$$\begin{aligned} \|J(x_3, ..., x_n, s_n, p_n)\| &= \|\frac{\partial(x_3, ..., x_n, s_n, p_n)}{\partial(x_1, ..., x_n)}\| \\ &= (\prod_{i=3}^{n} x_i)\sqrt{\left\{(s_n - \sum_{i=3}^{n} x_i)^2 - 4p_n/\prod_{i=3}^{n} x_i\right\}}. \end{aligned}$$

Define $\tilde{s} = \sum_{i=1}^{n-1} x_i$, $\tilde{p} = \prod_{i=1}^{n-1} x_i$, $C = s_n - \tilde{s}$, and $D = p_n/\tilde{p}$, The full conditional distribution of x_n is given by

$$\begin{aligned} f_c(x_n) &= f_c(x_n|x_3, ..., x_{n-1}, s_n, p_n) \\ &\propto \frac{1}{x_n}\frac{1}{\sqrt{(s_n - \tilde{s} - x_n)^2 - 4(p_n/\tilde{p})/x_n}} \\ &= \frac{1}{x_n}\frac{1}{\sqrt{(C - x_n)^2 - 4D/x_n}} \\ &= \frac{1}{x_n}\frac{1}{\sqrt{(C - x_n)^2 - 4D/x_n}}. \end{aligned}$$

Let $c = 4D/C^3$, and define a new variable $V = X_n/C$. The PDF of V, $y(\nu)$, is then

$$\begin{aligned} y(\nu) &\propto \left[\sqrt{\nu}\sqrt{\nu(1-\nu)^2 - c}\right]^{-1} \\ &= \frac{1}{\sqrt{\nu}\sqrt{(\nu-a)(b-\nu)(d-\nu)}} \\ &\equiv h(\nu);\ \nu \in (a,b), \end{aligned} \tag{4.1}$$

where $a < b < d$ and a, b, and d are the three solutions of the equation $\nu(1-\nu)^2 = c$. They can be found numerically with $a \in (0, 1/3)$, $b \in (1/3, 1)$, and $d > 1$. Since $h(\nu)$ is a U-shaped function, a scaled Beta(0.5, 0.5) in (a, b) is used as the candidate generation function. The density is given by

$$p(\nu) = \frac{1}{\pi}\frac{1}{\sqrt{(\nu-a)(b-\nu)}};\ \nu \in (a,b).$$

The probability of acceptance is $\alpha = h(\nu)/\{p(\nu)C_{max}\}$, where the constant C_{max} is calculated by

$$\begin{aligned} C_{max} &= \sup_{\nu\in(a,b)}\frac{h(\nu)}{p(\nu)} \\ &= \sup_{\nu\in(a,b)}\frac{\pi}{\sqrt{\nu(d-\nu)}} \\ &= \max\left(\frac{\pi}{\sqrt{a(d-a)}}, \frac{\pi}{\sqrt{b(d-b)}}\right). \end{aligned} \tag{4.2}$$

After we derive a value ν of V, we let $x_n^* = C\nu$.

Once $x_3^*, ...x_n^*$ have been found, we can find x_1^* and x_2^* by solving the equations

$$\begin{aligned} x_1 + x_2 &= s_n - \sum_{i=3}^{n} x_i^* \\ x_1 x_2 &= p_n / \prod_{i=3}^{n} x_i^*. \end{aligned}$$

The values x_1 and x_2 are the two solutions of the quadratic

$$x^2 - (s_n - \sum_{i=3}^{n} x_i^*)x - p_n/\prod_{i=3}^{n} x_i^* = 0.$$

Discussion:

As $c \to 0$ in Equation 4.1, $a \to 0$, $b \to 1$, and $d \to 1$. In this limit, the A-R method fails because $C_{max} \to \infty$; that is, the candidate is always rejected. To avoid this problem, we study the limiting conditional distribution of V as $c \to 0$.

For any $\nu \in (a+\epsilon, b-\epsilon)$, $\epsilon \in (0, (a+b)/2)$, the limiting density of V is, $\lim_{c\to 0} y(\nu) = \lim_{c\to 0} \frac{h(\nu)}{\int h(\nu)\mathrm{d}\nu} = 0$, which means that $\Pr(V = a \text{ or } b | c = 0) = 1$.

In practice, when $c > 10^{-9}$, we used the regular A-R method. On very few occasions, $c \leq 10^{-9}$, in which case we took ν to be a value close to b or a ($b - 10^{-5}$ or $a + 10^{-5}$) with equal probability. When $c = 10^{-9}$, $C_{max} \approx 99336$.

4.2 Power study

The Anderson-Darling statistic A^2 and the Cramér-von Mises statistic W^2 are used for tests at the 0.05 and 0.10 levels. The alternatives studied are Weibull, log-normal and half-normal distributions. From each alternative, 500 samples are drawn; for each sample, we use 1000 co-sufficient samples and 1000 bootstrap samples to evaluate the exact and parametric p-values by the methods shown in Section 2.3.1. The powers for both tests are estimated as in Equation 2.8.

4.2.1 Weibull alternatives

The Weibull distribution is a continuous probability distribution. The probability density function is

$$f(x; \kappa, \lambda) = \frac{\kappa}{\lambda}\left(\frac{x}{\lambda}\right)^{\kappa-1} \exp(-(x/\lambda)^{\kappa}); x > 0; \kappa > 0, \lambda > 0,$$

where κ is the shape parameter, and λ is the scale parameter. Plots of the PDF are given in Figure 4.3.

For the Weibull alternatives, we discuss three cases (denote Case 1, 2.a, and 2.b). For Case 1, in the null hypothesis distribution, α is unknown, but β is known, and we use a Weibull alternative with $\kappa = 1.2$ and $\lambda = 1$. For Case 2.a and 2.b, both α and β in the null hypothesis distribution are unknown. The alternative in Case 2.a is a Weibull distribution with $\kappa = 0.5$ and $\lambda = 1$; that in Case 2.b is a Weibull distribution with $\kappa = 2$ and $\lambda = 1$.

Table 4.1 shows the powers of the two tests for Case 1 at levels 0.05 and 0.10 for sample size 5, 10, 20, 30, 40, and 50. The scale parameter in the null distribution is known, $\beta = 1$.

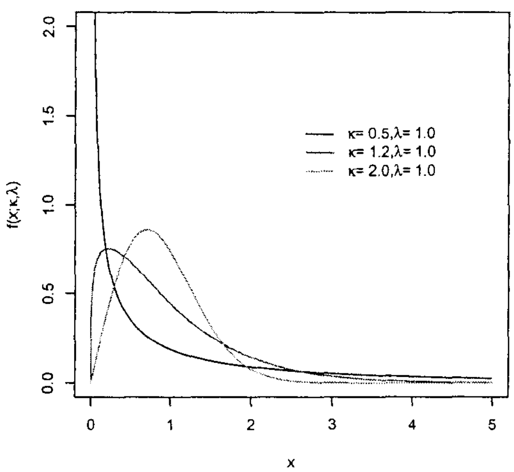


Figure 4.3: PDF of Weibull distribution

Sample Size	Exact				Parametric Bootstrap			
	A^2		W^2		A^2		W^2	
	0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
5	0.034	0.078	0.056	0.122	0.042	0.092	0.056	0.112
10	0.044	0.118	0.064	0.142	0.044	0.108	0.068	0.142
20	0.096	0.174	0.112	0.186	0.082	0.180	0.104	0.190
30	0.116	0.196	0.114	0.228	0.122	0.198	0.114	0.210
40	0.148	0.244	0.162	0.226	0.154	0.242	0.160	0.232
50	0.190	0.300	0.174	0.272	0.192	0.294	0.178	0.276

Table 4.1: Powers of exact and bootstrap tests at levels 0.05 and 0.10 based on 500 samples from the alternative Weibull(κ=1.2,λ=1) in Case 1

Figure 4.4 gives plots of corresponding p-values given by the two tests. The correlation was calculated from the 500 samples. This was found to be very high for all sample sizes; for n=5 the value with test statistic W^2 was 0.9977, and that with A^2 was 0.9976. The correlations increase with sample size.

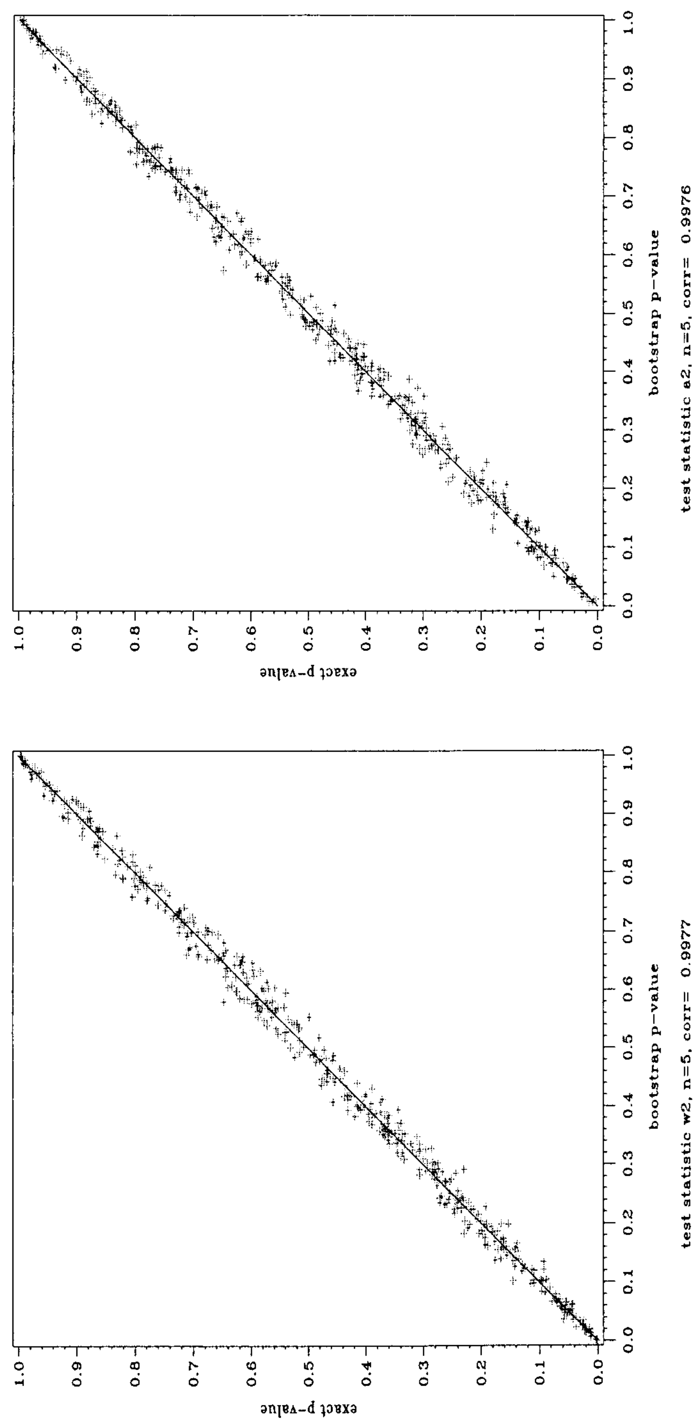


Figure 4.4: p-values of exact and bootstrap tests based on 500 samples from Weibull(κ=1.2,λ=1) in Case 1; corr: correlation between exact p-value and bootstrap p-value

Table 4.2 shows the powers of the two tests for the Case 2.a at levels 0.05 and 0.10 for sample size 5, 10, 20, 30, 40, and 50. Both α and β in the null distribution are unknown.

Sample Size	Exact				Parametric Bootstrap			
	A^2		W^2		A^2		W^2	
	0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
5	0.044	0.084	0.046	0.096	0.048	0.090	0.046	0.100
10	0.070	0.130	0.076	0.138	0.072	0.126	0.084	0.144
20	0.126	0.194	0.132	0.192	0.128	0.188	0.130	0.186
30	0.200	0.296	0.198	0.294	0.198	0.296	0.198	0.300
40	0.236	0.346	0.226	0.332	0.224	0.336	0.220	0.324
50	0.282	0.362	0.266	0.352	0.278	0.366	0.272	0.350

Table 4.2: Powers of exact and bootstrap tests at levels 0.05 and 0.10 level based on 500 samples from the alternative Weibull(κ=0.5,λ=1) in Case 2.a

Table 4.3 shows the powers of the two tests for Case 2.b at levels 0.05 and 0.10 for sample size 5, 10, 20, 30, 40, and 50. Both α and β in the null distribution are unknown.

Sample Size	Exact				Parametric Bootstrap			
	A^2		W^2		A^2		W^2	
	0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
5	0.046	0.084	0.048	0.084	0.042	0.094	0.048	0.098
10	0.066	0.126	0.064	0.126	0.064	0.128	0.064	0.132
20	0.088	0.166	0.076	0.174	0.086	0.168	0.076	0.174
30	0.108	0.174	0.098	0.156	0.108	0.172	0.092	0.158
40	0.136	0.192	0.120	0.196	0.128	0.194	0.118	0.196
50	0.162	0.268	0.148	0.252	0.164	0.262	0.152	0.240

Table 4.3: Powers of exact and bootstrap tests at levels 0.05 and 0.10 based on 500 samples from the alternative Weibull(κ=2,λ=1) in Case 2.b

4.2.2 Half-normal alternatives (HN)

The half-normal distribution is a normal distribution with mean $\mu = 0$ and scale σ, limited to the domain $x \in [0, \infty)$. It has probability density given by

$$f(x; \mu, \sigma) = \frac{1}{\sigma}\sqrt{\frac{2}{\pi}}\exp(-x^2/(2\sigma^2)), x > 0;\ \sigma > 0.$$

Two cases (called Case 1 and 2) are discussed. For Case 1, in the null hypothsis distribuion, α is unknown, but β is known; the alternative is a half-normal distribution with $\mu = 0$ and $\sigma = 1$. For Case 2, both α and β are unknown, and the alternative is same.

Table 4.4 shows the powers of the two tests for Case 1 at levels 0.05 and 0.10 for sample size 5, 10, 20, 30, 40, and 50. The scale parameter in the null distribution is known, $\beta = 1$.

Sample Size	Exact				Parametric Bootstrap			
	A^2		W^2		A^2		W^2	
	0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
5	0.056	0.124	0.086	0.168	0.058	0.114	0.090	0.168
10	0.078	0.150	0.106	0.204	0.076	0.154	0.098	0.206
20	0.158	0.292	0.180	0.290	0.156	0.304	0.182	0.300
30	0.266	0.432	0.262	0.388	0.280	0.434	0.260	0.390
40	0.390	0.562	0.338	0.482	0.396	0.566	0.346	0.492
50	0.504	0.650	0.442	0.568	0.516	0.652	0.444	0.564

Table 4.4: Powers of exact and bootstrap tests at levels 0.05 and 0.10 based on 500 samples from the alternative HN(μ=0,σ=1) in Case 1

Table 4.5 shows the power of the two tests for Case 2 at levels 0.05 and 0.10 for sample size 5, 10, 20, 30, 40, and 50. The scale parameter β in the null distribution is unknown.

Sample Size	Exact				Parametric Bootstrap			
	A^2		W^2		A^2		W^2	
	0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
5	0.064	0.130	0.070	0.124	0.066	0.126	0.068	0.124
10	0.060	0.114	0.068	0.106	0.062	0.108	0.060	0.106
20	0.088	0.156	0.090	0.154	0.086	0.166	0.082	0.154
30	0.112	0.192	0.112	0.182	0.112	0.190	0.112	0.186
40	0.154	0.248	0.142	0.228	0.156	0.250	0.146	0.224
50	0.194	0.302	0.182	0.276	0.194	0.300	0.182	0.284

Table 4.5: Powers of exact and bootstrap tests at levels 0.05 and 0.10 based on 500 samples from the alternative HN(μ=0,σ=1) in Case 2

4.2.3 Log-normal alternatives (LN)

If Y is a random variable from the normal distribution, then $X = \exp(Y)$ has a log-normal distribution. The probability density function of X is given by

$$f(x; \mu, \sigma) = \frac{1}{(\sigma\sqrt{2\pi x})} \exp\left\{ -\frac{(\ln x - \mu)^2}{2\sigma^2} \right\}; x > 0,$$

where μ and σ are the mean and standard deviation of the normal distribution respectively. This density has a mounded shape as in Figure 4.5.

Three cases (denote Case 1, 2.a, and 2.b) are discussed. For Case 1, in the null distribution, α is unknown, but β is known; the alternative is a log-normal distribuion with $\mu = 0$ and $\sigma = 1$. For Case 2.a and 2.b, both α and β in the null distribuion are unknown; the alternative in Case 2.a is a log-normal distribution with $\mu = 0$ and $\sigma = 1$, and that in Case 2.b is a log-normal distribution with $\mu = 0$ and $\sigma = 2$.

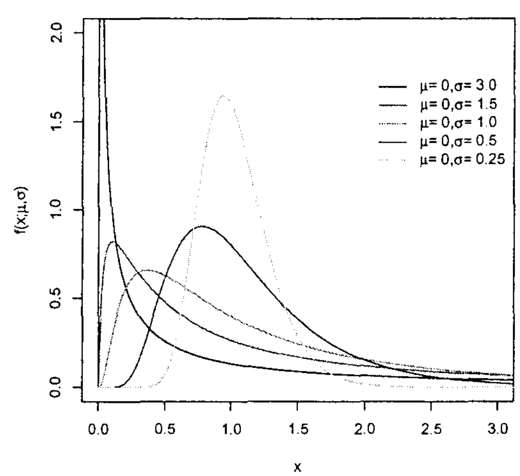


Figure 4.5: PDF of log-normal distribution

Table 4.6 shows the power of the two tests for Case 1 at levels 0.05 and 0.10 for sample size 5, 10, 20, 30, 40, and 50. The scale parameter in the null distribution is known, $\beta = 1$.

Sample Size	Exact				Parametric Bootstrap			
	A^2		W^2		A^2		W^2	
	0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
5	0.152	0.208	0.082	0.152	0.148	0.214	0.080	0.152
10	0.184	0.260	0.112	0.186	0.180	0.268	0.114	0.184
20	0.236	0.330	0.166	0.264	0.242	0.326	0.156	0.254
30	0.316	0.422	0.208	0.316	0.310	0.424	0.208	0.308
40	0.348	0.476	0.228	0.334	0.350	0.474	0.220	0.348
50	0.424	0.544	0.280	0.414	0.424	0.562	0.284	0.418

Table 4.6: Powers of exact and bootstrap tests at levels 0.05 and 0.10 based on 500 samples from the alternative LN(μ=0,σ=1) in Case 1

Table 4.7 shows the power of the two tests for Case 2.a at levels 0.05 and 0.10 for sample size 5, 10, 20, 30, 40, and 50; Both α and β in the null distribution are unknown.

Sample Size	Exact				Parametric Bootstrap			
	A^2		W^2		A^2		W^2	
	0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
5	0.066	0.114	0.068	0.124	0.068	0.120	0.072	0.128
10	0.098	0.154	0.102	0.154	0.104	0.156	0.108	0.158
20	0.234	0.322	0.226	0.306	0.230	0.326	0.218	0.308
30	0.350	0.460	0.332	0.438	0.350	0.458	0.336	0.428
40	0.464	0.564	0.426	0.528	0.460	0.564	0.428	0.530
50	0.512	0.628	0.486	0.594	0.514	0.624	0.490	0.598

Table 4.7: Powers of exact and bootstrap tests at levels 0.05 and 0.10 level based on 500 samples from the alternative LN(μ=0,σ=1) in Case 2.a

Table 4.8 shows the power of the two tests for Case 2.b at levels 0.05 and 0.10 for sample size 5, 10, 20, 30, 40, and 50; Both α and β in the null distribution are unknown.

Sample Size	Exact				Parametric Bootstrap			
	A^2		W^2		A^2		W^2	
	0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
5	0.082	0.176	0.102	0.192	0.086	0.178	0.108	0.196
10	0.196	0.290	0.204	0.302	0.190	0.300	0.198	0.312
20	0.516	0.600	0.516	0.606	0.514	0.610	0.514	0.600
30	0.676	0.754	0.658	0.740	0.680	0.744	0.658	0.740
40	0.836	0.882	0.822	0.876	0.836	0.882	0.826	0.878
50	0.916	0.944	0.912	0.938	0.912	0.942	0.910	0.938

Table 4.8: Powers of exact and bootstrap tests at levels 0.05 and 0.10 based on 500 samples from the alternative LN(μ=0,σ=2) in Case 2.b

4.3 Comments

In this chapter, we have studied powers for testing the gamma family with unknown shape and with scale known or unknown. Several alternatives from families close to the gamma family were investigated, such as the Weibull, half-normal, and log-normal distributions. The tables presented lead to the following conclusions:

1. The exact conditional tests and the approximate bootstrap tests have very similar powers even for samples of very small size, such as $n = 5$. This conclusion is also verified by the closeness of two p-values for each sample, and their high correlation based on 500 Monte Carlo samples from each alternative.

2. As expected, the powers increase as the sample size increases.

3. When the sample size is large enough, both tests appear to be unbiased.

Chapter 5

The von Mises Family

In this chapter, we will discuss the von Mises family, which is an important distribution family in circular statistics. Circular data measure directions and can be represented as angles or as points on the circumference of a unit circle. The left panel of Figure 5.1 shows a unit circle centered at the origin O. It also shows a typical point P_i from a sample $P_1, ..., P_n$ of n points on the circumference. Each P_i is related to an angle θ_i, the angle of a vector $\bar{OP_i}$ from O to P_i. We represent the vector, $\bar{OP_i}$, by $(1, \theta_i)$ in the polar co-ordinate system or $(\cos(\theta_i), \sin(\theta_i))$ in the rectangular co-ordinate system. The right hand panel of Figure 5.1 shows the resultant vector $t = (\sum_{i=1}^{n} \cos(\theta_i), \sum_{i=1}^{n} \sin(\theta_i))$ which is the vector sum of the $\bar{OP_i}$. This vector plays an important role later in this chapter.

The von Mises family (VM) is a continuous distributional family on the circle. A random variable Θ from the generalized l-modal von Mises distribution has the probability density function:

$$f(\theta; \kappa, \mu) = \frac{1}{2\pi I_0(\kappa)} \exp\{\kappa \cos l(\theta - \mu)\}; \theta \in (0, 2\pi]; \mu \in (0, 2\pi], \kappa > 0,$$

where $I_0(\kappa)$ is the modified Bessel function of order 0, and l indicates the number of modes. In this project, we only work on the von Mises family where $l = 1$. Plots of the PDF and CDF of a von Mises distribution are shown in Figure 5.2. The density is symmetric about the direction μ, and a larger value of κ indicates a higher concentration around μ. When $\kappa = 0$, the distribution is a uniform distribution over $(0, 2\pi)$. As $\kappa \to \infty$, we have $\sqrt{\kappa}(\Theta - \mu) \to N(0, 1)$.

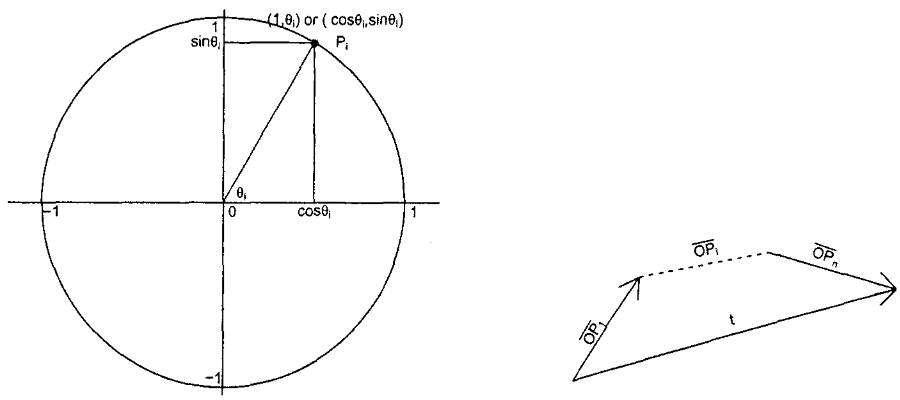


Figure 5.1: Circular data P_i, and resultant vector t

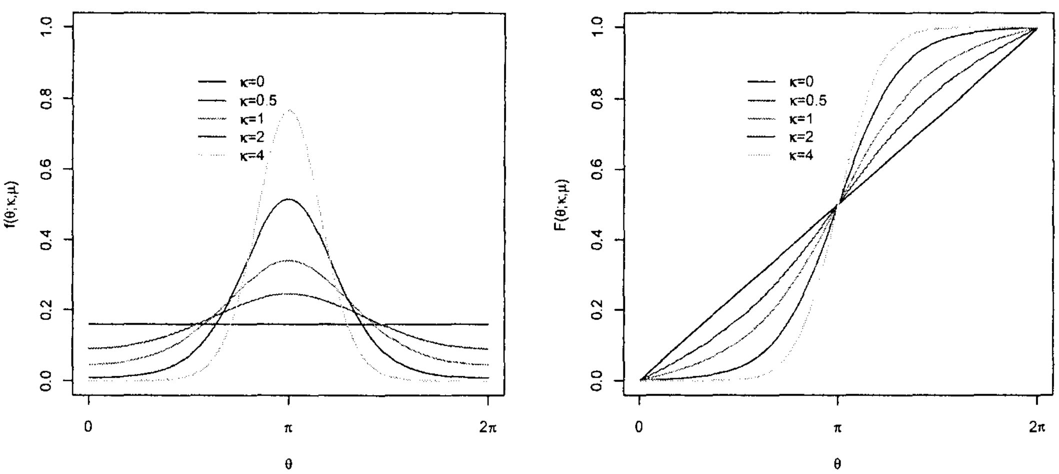


Figure 5.2: PDF and CDF of von Mises distribution

In Section 5.1 are discussed the methods of generating the co-sufficient samples for the von Mises distribution test (Lockhart *et al.* (2008)). In Section 5.2 the results of a power study on the tests against different alternatives are presented.

5.1 Generating the co-sufficient samples

There is no direct method to generate co-sufficient samples when κ is unknown. Instead we apply the Gibbs sampler. We will discuss two cases:

Case 1: $\mu = \mu_0$ is known, and κ is unknown,

Case 2: both μ and κ are unknown.

5.1.1 Case 1: μ is known, and κ is unknown

Given a sample of data $\alpha_1, ..., \alpha_n$, consider a test of H_0,

$$H_0 : \alpha_1, ..., \alpha_n \text{ comes from } VM(\kappa, \mu); \kappa \text{ is unknown, and } \mu = \mu_0.$$

Let $\theta_i = \alpha_i - \mu_0$. Then, H_0 simplifies to H_0'

$$H_0' : \theta_1, ..., \theta_n \text{ comes from } VM(\kappa, \mu); \kappa \text{ is unknown, and } \mu = 0.$$

The likelihood function is

$$\begin{aligned} L(\kappa) &= \prod_{i=1}^{n} f(\theta_i, \kappa, \mu = 0) \\ &= \left\{ \frac{1}{2\pi I_0(\kappa)} \right\}^n \exp\left\{ \kappa \sum_{i=1}^{n} \cos(\theta_i) \right\}. \end{aligned}$$

Therefore, the sufficient statistic for κ is $t_n = \sum_{i=1}^{n} \cos(\theta_i)$. The maximum likelihood estimate of κ can be obtained by solving the equation

$$\frac{I_1(\kappa)}{I_0(\kappa)} = \frac{\sum_{i=1}^{n} \cos(\theta_i)}{n},$$

where $I_1(\kappa)$ is the modified Bessel function of order 1; also $I_1(\kappa)$ is the derivative of $I_0(\kappa)$.

The joint density function of $\Theta_2, ..., \Theta_n, T_n$ is

$$f_1(\theta_2, ..., \theta_n, t_n) = \|J\|^{-1} f_s(\theta_1, ..., \theta_n),$$

where the Jacobian is

$$\begin{aligned} \|J(\theta_2, ..., \theta_n, t_n) &= \|\frac{\partial(\theta_2, ..., \theta_n, t_n)}{\partial(\theta_1, ..., \theta_n)}\| \\ &= \sqrt{1 - (t_n - \sum_{i=2}^{n} \cos(\theta_i))^2}. \end{aligned}$$

So, we have

$$\begin{aligned} f_1(\theta_2, ..., \theta_n, t_n) &= \frac{1}{\sqrt{1 - (t_n - \sum_{i=2}^{n} \cos(\theta_i))^2}} (\frac{1}{2\pi I_0(\kappa)})^n \exp(\kappa t_n) \\ &\propto \frac{1}{\sqrt{1 - (t_n - \sum_{i=2}^{n} \cos(\theta_i))^2}}. \end{aligned}$$

The full conditional distribution of Θ_n given $\Theta_2, ..., \Theta_{n-1}, t_n$ is

$$\begin{aligned} f_c(\theta_n) &= f_c(\theta_n | \theta_2, ..., \theta_{n-1}, t_n) \\ &\propto \frac{1}{\sqrt{1 - (t_n - \sum_{i=2}^{n} \cos(\theta_i))^2}}. \end{aligned}$$

Define $Z = \cos(\Theta_n)$, and let $h = t_n - \sum_2^{n-1} \cos(\theta_i)$. Then the conditional density $y(z)$ of Z given h is

$$\begin{aligned} y(z) &\propto (\sqrt{1 - (z - h)^2} \sqrt{1 - z^2})^{-1} \\ &\equiv g(z); \;\; z \in (a, b); \end{aligned} \tag{5.1}$$

where $a = \max(h-1, -1)$, and $b = \min(h+1, 1)$. The A-R method can be used to generate z. Since $g(\cdot)$ has a "U" shape, a scaled Beta(0.5, 0.5) in (a, b) can be the candidate generation distribution. This density is given by

$$p(z) = \frac{1}{\pi} \frac{1}{\sqrt{(z - a)(b - z)}}; \;\; z \in (a, b).$$

The probability α of acceptance is $\alpha = \frac{g(z)}{p(z) C_{max}}$, where C_{max} is calculated by

$$\begin{aligned} C_{max} &= \sup_{z \in (a,b)} \frac{g(z)}{p(z)} \\ &= \begin{cases} \sup_{z \in (h-1,1)} \frac{\pi}{\sqrt{(1-z+h)(1+z)}} & \text{if } h > 0, \\ \sup_{z \in (-1,1+h)} \frac{\pi}{\sqrt{(1-z)(z-h+1)}} & \text{if } h \leq 0. \end{cases} \end{aligned} \tag{5.2}$$

Once z is obtained, take $\theta_n^* = \arccos(z)$ or $-\arccos(z)$ with equal probability.

After obtaining a cycle of $\theta_2^*, ..., \theta_n^*$, we can find θ_1^* by solving $\cos(\theta_1) = t_n - \sum_{i=2}^n \cos\theta_i^*$ to set $\theta_1^* = \arccos(t_n - \sum_{i=2}^n \cos\theta_i^*)$ or $-\arccos(t_n - \sum_{i=2}^n \cos\theta_i^*)$ with equal probability.

Discussion

As $h \to 0$ in Equation 5.1, $a \to -1$, and $b \to 1$. The A-R method then fails because $C_{max} \to \infty$; as we did in the case of the gamma distribution, we set a bound on h to keep it away from 0. When h is close to 0, we sample from the limiting distribution of Z.

For any $\epsilon \in (0, (a+b)/2)$, the limiting density function

$$\lim_{h\to 0} \frac{g(z)}{\int_a^b g(z)\mathrm{d}z}\mathrm{d}z = 0; \text{ for } z \in (a+\epsilon, b-\epsilon),$$

which means $\Pr(Z = a \text{ or } b | h = 0) = 1$.

As defined, h is the sum of two i.i.d objects, say Z_1 and Z_2. The desired Z could be either of them. When $h = 0$, $Z_1 = -Z_2$. Thus, $\Pr(Z = a|h = 0) = \Pr(Z = b|h = 0) = 1/2$.

In practice, a value $z \in (a, b)$ close to a or b can be drawn with equal probability when h is very small.

5.1.2 Case 2: Both μ and κ are unknown

Given a random sample $\theta_1, ..., \theta_n$, the likelihood function for Case 2 is

$$\begin{aligned} L(\kappa,\mu) &= \prod_{i=1}^n f(\theta_i;\kappa,\mu) \\ &= \frac{1}{(2\pi)^n I_0^n(\kappa)} \exp\left\{\kappa\sum_{i=1}^n \cos(\theta_i-\mu)\right\} \\ &= \frac{1}{(2\pi)^n I_0^n(\kappa)} \exp\left\{\kappa\cos(\mu)\sum_{i=1}^n \cos(\theta_i) + \kappa\sin(\mu)\sum_{i=1}^n \sin(\theta_i)\right\}, \end{aligned}$$

and the sufficient statistic for (κ, μ) is the resultant vector $t_n = (t_1, t_2)$, where $t_1 = \sum_{i=1}^n \cos(\theta_i)$ and $t_2 = \sum_{i=1}^n \sin(\theta_i)$. Let $|t_n|$ and $\bar{\alpha}_0$ be the norm and angle of t_n respectively.

The maximum likelihood estimates $\hat{\kappa}$ and $\hat{\mu}$, can be found by solving the equations

$$\begin{aligned} \frac{I_1(\kappa)}{I_0(\kappa)} &= \frac{|t_n|}{n} \\ \hat{\mu} &= \bar{\alpha}_0. \end{aligned}$$

The joint density function of $\Theta_3, ..., \Theta_n, T_1, T_2$ is given by

$$f_1(\theta_3, ..., \theta_n, t_1, t_2)\| = 2\|J(\theta_3, ..., \theta_n, t_1, t_2)\|^{-1} f_s(\theta_1, ..., \theta_n),$$

where the Jacobian is

$$\begin{aligned} \|J(\theta_3, ..., \theta_n, t_1, t_2)\| &= \|\frac{\partial(\theta_3, ..., \theta_n, t_1, t_2)}{\partial(\theta_1, ..., \theta_n)}\| \\ &= \sqrt{1 - \left\{1 - \frac{(t_1 - \sum_{i=3}^{n}\cos(\theta_i))^2 + (t_2 - \sum_{i=3}^{n}\sin(\theta_i))^2}{2}\right\}^2}. \end{aligned}$$

The full conditional distribution of Θ_n, $f_c(\theta_n)$ is given by

$$\begin{aligned} f_c(\theta_n) &= f_c(\theta_n|\theta_3, ..., \theta_{n-1}, t_1, t_2) \\ &\propto \frac{1}{\sqrt{1 - \left\{1 - \frac{(h_1 - \cos(\theta_n))^2 + (h_2 - \sin(\theta_n))^2}{2}\right\}^2}}, \end{aligned}$$

where $h_1 = t_1 - \sum_{i=3}^{n-1}\cos(\theta_i)$ and $h_2 = t_2 - \sum_{i=3}^{n-1}\sin(\theta_i)$.

Define a vector h by $h = (h_1, h_2)$. Let $|h|$ and α_h denote the norm and angle of h respectively. Define a new variable $Z = \cos(\Theta_n - \alpha_h)$. Then Z has the conditional PDF $y(z)$ such that

$$\begin{aligned} y(z) &\propto \left(\sqrt{1 - \{1 - (|h|^2 - 2|h|z + 1)/2\}^2}\sqrt{1 - z^2}\right)^{-1} \\ &\propto \frac{1}{\sqrt{(1 - z)(z - a)(\frac{|h|^2+1}{2|h|} - z)(z - b)}} \\ &\equiv g(z);\ z \in (a, 1)\ , \end{aligned} \tag{5.3}$$

where $a = \max\left\{(|h|^2 - 3)/2|h|, -1\right\}$, and $b = \min\left\{(|h|^2 - 3)/2|h|, -1\right\}$. When $|h| > 1$, $a = \frac{|h|^2-3}{2|h|}$, and $b = -1$; else, $a = -1$, and $b = \frac{|h|^2-3}{2|h|}$.

To draw a value z from $y(z)$, we apply the acceptance and rejection method. Since $g(\cdot)$ is a "U" shape function, a scaled Beta(0.5,0.5) in $(a, 1)$ will be the candidate distribution, whose density is given by

$$p(z) = \frac{1}{\pi}\frac{1}{\sqrt{(z - a)(1 - z)}};\ z \in (a, 1).$$

The probability of acceptance $\alpha = \frac{g(z)}{C_{max}p(z)}$, where C_{max} is calculated by

$$\begin{aligned} C_{max} &= \sup_{z\in(a,1)}\frac{g(z)}{p(z)} \\ &= \max\left(\frac{\pi}{\sqrt{(\frac{|h|^2+1}{2|h|} - a)(a - b)}}, \frac{\pi}{\sqrt{(\frac{|h|^2+1}{2|h|} - 1)(1 - b)}}\right). \end{aligned} \tag{5.4}$$

Once z is obtained, $\theta_n^* = \alpha_h + \arccos(z)$ or $\alpha_h - \arccos(z)$ with equal probability. After a cycle of $\theta_3^*, ..., \theta_n^*$ is found, find θ_1^* and θ_2^* by solving the equations

$$\begin{aligned}\cos(\theta_1) + \cos(\theta_2) &= t_1 - \sum_{i=3}^{n} \cos(\theta_i^*) \\ \sin(\theta_1) + \sin(\theta_2) &= t_2 - \sum_{i=3}^{n} \sin(\theta_i^*).\end{aligned}$$

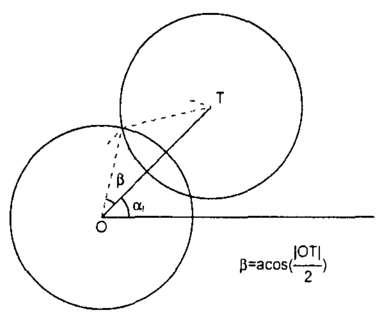


Figure 5.3: θ_1^* and θ_2^*

Define $\bar{OT}$ in Figure 5.3 as the vector of $(t_1 - \sum_{i=3}^{n} \cos(\theta_i^*), t_2 - \sum_{i=3}^{n} \sin(\theta_i^*))$. Let l be the norm of $\bar{OT}$, and α_l be the angle. Then, $\bar{OT}$ is the sum of two unit vectors $(\cos(\theta_1), \sin(\theta_1))$ and $(\cos(\theta_2), \sin(\theta_2))$. As Figure 5.3 shows, we have $\theta_1^* = \alpha_l + \arccos(l/2)$, and $\theta_2^* = \alpha_l - \arccos(l/2)$, where θ_1^* and θ_2^* are exchangeable.

Discussion

We have the same difficulty as that in Case 1. When $|h| \to 1$ in Equation.5.3, $a \to -1$. The A-R iterations become stuck, so we consider using the limiting distribution of Z.

For any $z \in (a + \epsilon, 1 - \epsilon)$, where $\epsilon \in (0, (1 + a)/2)$, the limiting density function

$$\lim_{|h| \to 1} \frac{g(z)}{\int_a^1 g(\mu) \mathrm{d}\mu} \mathrm{d}z = 0,$$

which means $\Pr(Z = -1 \text{ or } 1 | |h| = 1) = 1$.

From the definition, we see that h is the resultant vector of three unit random vectors, which are exchangeable The desired Z could belong to any one of them. The fact that $\Pr(Z = -1 \text{ or } 1) = 1$ means that the directions of those three vectors must be identical or opposite to that of h. Geometrically, this happens only when some two of them are identical to h and the third one is opposite. Thus, $\Pr(Z = -1 | |h| = 1) = 1/3$, and $\Pr(Z = 1 | |h| = 1) = 2/3$.

In practice, if $|h| > 1.0001$ or $|h| < 0.9999$, the regular A-R method is applied. Otherwise, we draw a value from the limiting distribution. Fewer than 0.02% of the values were drawn from this limiting distribution, so the impact on the power estimates is negligible. When $|h| = 1.0001$ or 0.9999, $C_{max} \approx 31410$.

5.2 Power study

The Watson's statistic U^2 is used for the power study for tests at the 0.05 and 0.10 levels. In the following subsections, we present a variety of alternative distributions and compare the powers of the parametric bootstrap and exact conditional tests in a Monte Carlo study. In each subsection we consider a different family of alternatives. Alternatives from the offset normal family are studied in Subsection 5.2.1, the cardioid family in Subsection 5.2.2, the wrapped skew-normal family in Subsection 5.2.3, and the asymmetric circular family in Subsection 5.2.4. We generated 500 Monte Carlo samples from each alternative. Again, for each sample, we use 1000 co-sufficient samples and 1000 bootstrap samples to evaluate the exact and parametric p-value. The powers of both tests for the alternative are evaluated by the percentages of the samples with p-value less than α.

We only work on Case 2 (both μ and κ are unknown).

5.2.1 Offset normal alternatives (ON)

Suppose (X, Y) is a random vector from a bivariate normal (BN) distribution, say with $X, Y \sim \text{BN}(\mu, \nu, \sigma_1, \sigma_2, \rho)$. Let Θ be the angle of (X, Y). Then we say Θ is a random variable from the offset normal (ON) distribution. The PDF for Θ is

$$f(\theta) = \frac{1}{C(\theta)} \left\{ \phi(\mu, \nu; 0, \Sigma) + aD(\theta)\Phi[D(\theta)]\phi\left[\frac{a(\mu \sin\theta - \nu\cos\theta)}{\sqrt{C(\theta)}}\right] \right\},$$

where

$$\begin{aligned} a &= \frac{1}{\sigma_1\sigma_2\sqrt{1-\rho^2}}, \\ C(\theta) &= a^2(\sigma_2^2\cos^2\theta - \rho\sigma_1\sigma_2\sin(2\theta) + \sigma_1^2\sin^2\theta), \\ D(\theta) &= \frac{a^2}{\sqrt{C(\theta)}}[\mu\sigma_2(\sigma_2\cos\theta - \rho\sigma_1\sin\theta) + \nu\sigma_1(\sigma_1\sin\theta - \rho\sigma_2\cos\theta)], \end{aligned}$$

and where $\phi(\cdot)$ and $\Phi(\cdot)$ are the PDF and CDF of $N(0,1)$ respectively. See Figure 5.4 for examples of the probability density function of the ON distribution.

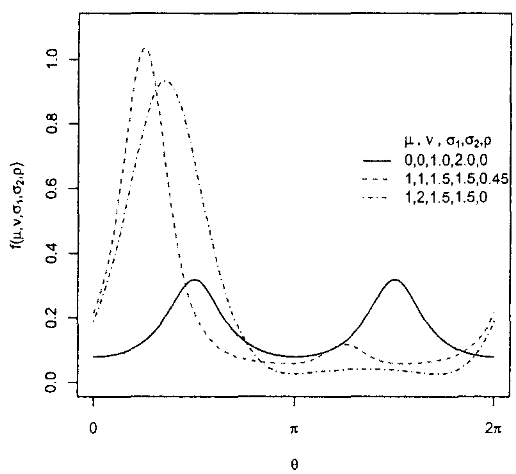


Figure 5.4: PDF of offset normal distribution

We discuss three cases (denote Case 2.a, 2.b, and 2.c). For all the cases, κ and μ are unknown. For Case 2.a, the alternative is a transformed offset normal distribution whose density function has only one mode. The alternatives in the other two cases have a density function with two modes.

In Case 2.a, the alternative is ON(0,0,$\sigma_1 = 1$,$\sigma_2 = 2$, $\rho = 0$) which we transform into a uni-modal density by setting $\Theta' = 2\Theta$. The density function of Θ' is

$$f(\theta') = \frac{\sqrt{1-b^2}}{2\pi(1-b\cos(\theta'))}; \theta' \in (0, 2\pi),$$

where $b = \frac{\sigma_1^2-\sigma_2^2}{\sigma_1^2+\sigma_2^2} = -0.6$. Table 5.1 shows the power of the two tests for Case 2.a at levels 0.05 and 0.10 for sample size 5, 10, 20, 30, 40, and 50.

Sample Size	Exact U^2		Bootstrap U^2	
	0.05	0.10	0.05	0.10
5	0.046	0.114	0.020	0.072
10	0.050	0.104	0.030	0.082
20	0.058	0.116	0.048	0.090
30	0.052	0.106	0.042	0.100
40	0.084	0.136	0.076	0.128
50	0.082	0.152	0.078	0.144

Table 5.1: Powers of exact and bootstrap tests at 0.05 and 0.10 levels based on 500 samples from the alternative of transformed ON$(0, 0, 1.0, 2.0, 0)$ in Case 2.a

Table 5.2 shows the power of the two tests for Case 2.b, in which the alternative distribution is ON(1,1,$\sigma_1 = 1.5$,$\sigma_2 = 1.5$, $\rho = 0.45$), at levels 0.05 and 0.10 for sample size 5, 10, 20, 30, 40, and 50.

Sample Size	Exact U^2		Bootstrap U^2	
	0.05	0.10	0.05	0.10
5	0.062	0.128	0.046	0.108
10	0.128	0.208	0.110	0.188
20	0.208	0.298	0.196	0.290
30	0.240	0.344	0.234	0.334
40	0.292	0.412	0.278	0.412
50	0.364	0.494	0.354	0.480

Table 5.2: Powers of exact and bootstrap tests at 0.05 and 0.10 levels based on 500 samples from the alternative ON(1,1,1.5,1.5, 0.45) for Case 2.b

Table 5.3 shows the power of the two tests for Case 2.c, in which the alternative is ON(1,2,σ_1 = 1.5,σ_2 = 1.5, $\rho = 0$), at levels 0.05 and 0.10 for sample size 5, 10, 20, 30, 40, and 50.

Sample Size	Exact		Bootstrap	
	U^2		U^2	
	0.05	0.10	0.05	0.10
5	0.066	0.136	0.062	0.136
10	0.090	0.144	0.088	0.152
20	0.060	0.138	0.072	0.130
30	0.092	0.166	0.084	0.164
40	0.094	0.174	0.092	0.168
50	0.108	0.192	0.112	0.202

Table 5.3: Powers of exact and bootstrap tests at 0.05 and 0.10 levels based on 500 samples from the alternative ON(1,2,1.5,1.5, 0) for Case 2.c

5.2.2 Cardioid alternative (CA)

The Cardioid distribution, derived from the cardioid curve, has the probability density function:

$$f(\theta; \mu, \rho) = \frac{1}{2\pi}\{1 + 2\rho\cos(\theta - \mu)\},$$

where $0 \leq \mu < 2\pi$, $-\frac{1}{2} < \rho < \frac{1}{2}$. It has a symmetric and uni-modal shape as in Figure 5.5.

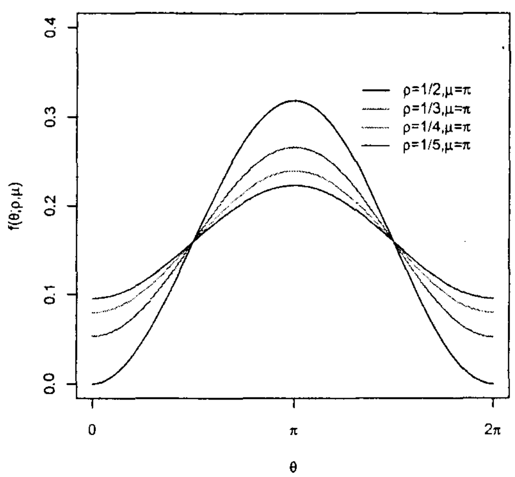


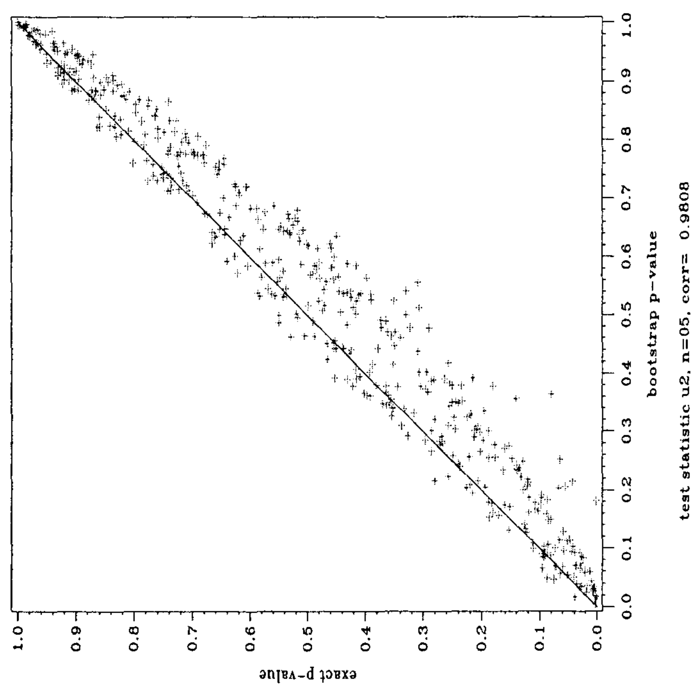
Figure 5.5: PDF of cardioid distribution

Table 5.4 shows the power of the two tests for the case, in which the alternative is Cardioid($\mu = \pi, \rho = 1/3$), at levels 0.05 and 0.10 for sample size 5, 10, 20, 30, 40, and 50.

Sample Size	Exact		Bootstrap	
	U^2		U^2	
	0.05	0.10	0.05	0.10
5	0.046	0.100	0.036	0.060
10	0.040	0.108	0.022	0.070
20	0.066	0.116	0.062	0.100
30	0.062	0.112	0.048	0.104
40	0.076	0.134	0.066	0.124
50	0.070	0.146	0.068	0.130

Table 5.4: Powers of exact and bootstrap tests at 0.05 and 0.10 levels based on 500 samples from the alternative CA(π, 1/3)

Figures 5.6 and 5.7 present the exact p-values and approximate p-values for the 500 samples drawn from the alternative CA($\mu = \pi$,$\rho = 1/3$) for sample sizes n=5, 10, 20, and 30. Points in the plots lie close to the diagonal line, which shows the closeness between the two kinds of p-values. The correlations between the p-values of two tests based on 500 samples are corr=0.9808 when $n = 5$, corr=0.9958 when $n = 10$, corr=0.9975 when $n = 20$, and corr=0.9978 when $n = 30$. As the sample size increases, the correlation between two kinds of p-values gets stronger. The plots also show there are more samples whose exact p-values are smaller than their bootstrap p-values, which implies the exact tests are more powerful.



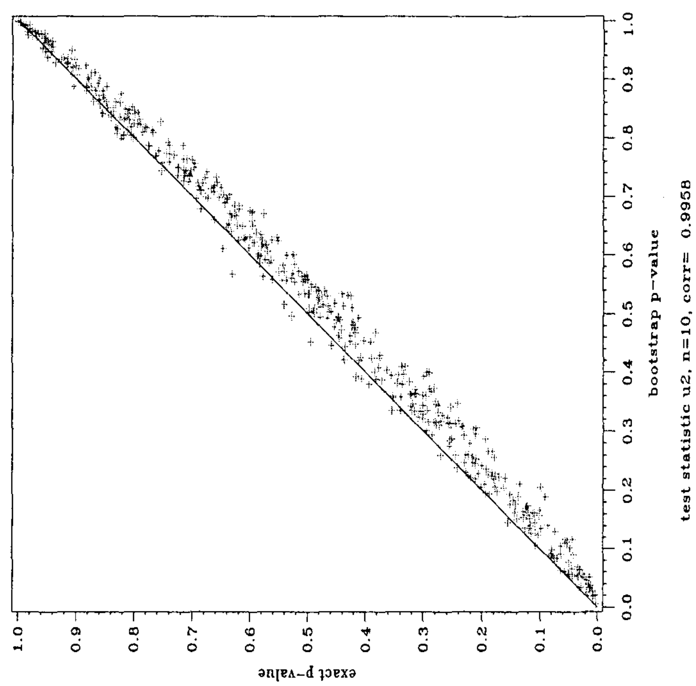
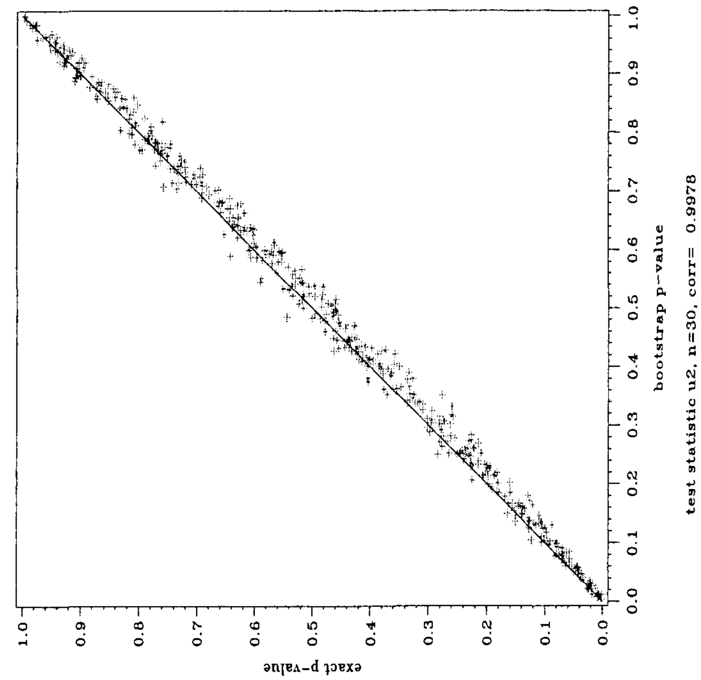


Figure 5.6: p-values of exact and bootstrap tests based on 500 samples from the alternative CA($\mu = \pi, \rho = 1/3$), corr: correlation between exact p-value and bootstrap p-value



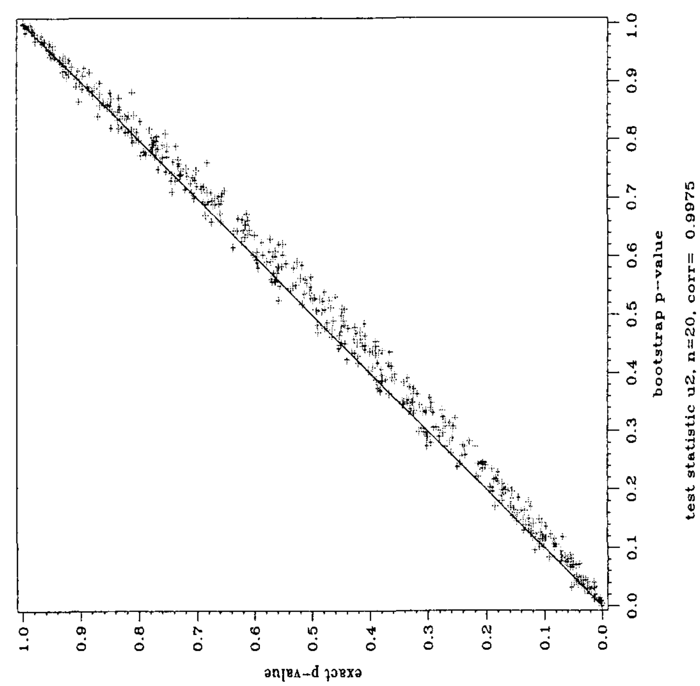


Figure 5.7: Continued

5.2.3 Wrapped skew-normal alternatives (WSN)

Suppose Z is a random variable from the skew-normal distribution, which we denote by $Z \sim SN(\mu, \sigma^2, \lambda)$, where μ, σ, and λ are the location, scale, and shape, respectively. The PDF of $SN(\mu, \sigma^2, \lambda)$ is given by

$$f(z; \mu, \sigma, \lambda) = \frac{2}{\sigma}\phi\left(\frac{z-\mu}{\sigma}\right)\Phi\left(\lambda\frac{z-\mu}{\sigma}\right),$$

$-\infty < z < \infty$, $-\infty < \mu < \infty$, $\sigma > 0$, and $-\infty < \lambda < \infty$.

Define a circular random variable $\Theta = Z \text{ (mod)} 2\pi$. We say Θ is from a wrapped skew-normal distribution having the density function

$$f(\theta; \mu, \sigma, \lambda) = \frac{2}{\sigma}\sum_{r=-\infty}^{\infty}\phi\left(\frac{\theta + 2\pi r - \mu}{\sigma}\right)\Phi\left(\lambda\frac{\theta + 2\pi r - \mu}{\sigma}\right),$$

$\theta \in (0, 2\pi)$. Denote $\Theta \sim WSN(\mu, \sigma^2, \lambda)$. See Pewsey (2000) for more about the WSN distribution.

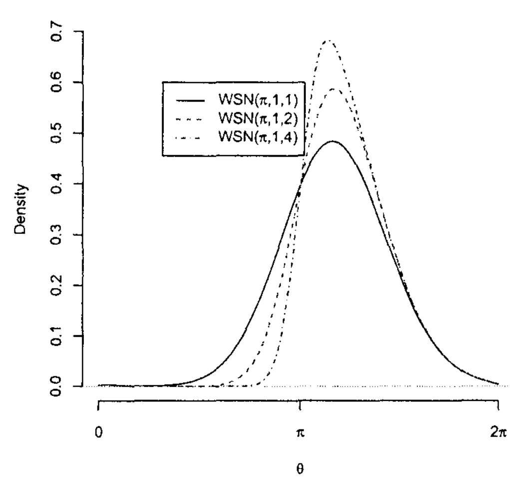


Figure 5.8: PDF of wrapped skew-normal distribution

Two cases (denote Case 2.a and 2.b) are discussed. For Case 2.a, the alternative is WSN($\mu = \pi, \sigma = 1, \lambda = 2.0$); for Case 2.b, the alternative is WSN($\mu = \pi, \sigma = 1, \lambda = 1.0$). The density function of the alternative in Case 2.a is more skewed than that in Case 2.b.

Table 5.5 shows the power of the two tests for Case 2.a, in which the alternative is WSN($\mu = \pi, \sigma = 1, \lambda = 2.0$), at levels 0.05 and 0.10 for sample size 5, 10, 20, 30, 40, and 50.

Sample Size	Exact		Bootstrap	
	U^2		U^2	
	0.05	0.10	0.05	0.10
5	0.048	0.102	0.052	0.094
10	0.036	0.098	0.038	0.104
20	0.058	0.116	0.054	0.112
30	0.072	0.130	0.076	0.126
40	0.082	0.148	0.082	0.156
50	0.114	0.196	0.120	0.204

Table 5.5: Powers of exact and bootstrap tests at 0.05 and 0.10 levels based on 500 samples from the alternative WSN(π, 1, 2.0) in Case 2.a

Table 5.6 shows the power of the two tests for Case 2.b, in which the alternative is WSN($\mu = \pi, \sigma = 1, \lambda = 1.0$), at levels 0.05 and 0.10 for sample size 5, 10, 20, 30, 40, and 50.

Sample Size	Exact		Bootstrap	
	U^2		U^2	
	0.05	0.10	0.05	0.10
5	0.044	0.088	0.048	0.088
10	0.042	0.090	0.038	0.092
20	0.056	0.098	0.054	0.098
30	0.056	0.108	0.056	0.106
40	0.072	0.130	0.064	0.134
50	0.070	0.140	0.062	0.138

Table 5.6: Powers of exact and bootstrap tests at 0.05 and 0.10 levels based on 500 samples from the alternative WSN(π,1,1.0) for Case 2.b

5.2.4 Asymmetric circular alternatives (AC)

We are also interested in a family of circular densities given by

$$f(\theta; \alpha, \beta) = C \exp(\alpha \cos(\theta - \mu_1) + \beta \cos(2(\theta - \mu_2))); \theta \in (0, 2\pi]; \beta > 0, \alpha > 0,$$

where C is a normalizing factor. Since we do not know C, only the shape of the density function is provided in Figure 5.9. This is an asymmetric circular distribution (AC) discussed in Jammalamadaka & SenGupta (2001). In our study, we consider only the symmetric shape with two modes, that is we take $\mu_1 = \mu_2 = \pi$.

Two cases (denote Case 2.a and 2.b) are discussed here. For Case 2.a, the alternative is AC($\alpha = 1, \beta = 1$); for Case 2.b, the alternative is AC($\alpha = 1, \beta = 1/2$).

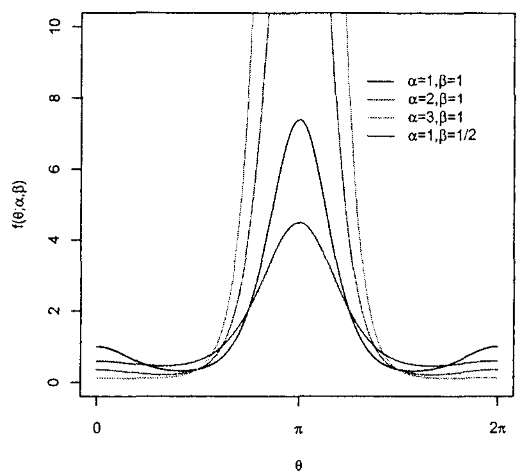


Figure 5.9: PDF of asymmetric circular distribution with $\mu_1 = \mu_2 = \pi$

Table 5.7 shows the power of the two tests for Case 2.a, in which the alternative is $AC(\alpha = 1, \beta = 1)$, at levels 0.05 and 0.10 for sample size 5, 10, 20, 30, 40, and 50.

Sample Size	Exact		Bootstrap	
	U^2		U^2	
	0.05	0.10	0.05	0.10
5	0.110	0.190	0.076	0.166
10	0.198	0.308	0.162	0.288
20	0.422	0.558	0.408	0.538
30	0.570	0.678	0.560	0.666
40	0.722	0.818	0.722	0.818
50	0.838	0.890	0.836	0.894

Table 5.7: Powers of exact and bootstrap tests at 0.05 and 0.10 levels based on 500 samples from the alternative AC(1,1) for Case 2.a

Table 5.8 shows the power of the two tests for Case 2.b, in which the alternative $AC(\alpha =$

$1, \beta = 1/2)$, at levels 0.05 and 0.10 for sample size 5, 10, 20, 30, 40, and 50.

Sample Size	Exact U^2		Bootstrap U^2	
	0.05	0.10	0.05	0.10
5	0.074	0.136	0.048	0.090
10	0.132	0.204	0.110	0.170
20	0.126	0.212	0.120	0.190
30	0.202	0.304	0.192	0.280
40	0.266	0.386	0.252	0.382
50	0.358	0.476	0.348	0.474

Table 5.8: Powers of exact and bootstrap tests at 0.05 and 0.10 levels based on 500 samples from the alternative AC(1,1/2) for Case 2.b

5.3 Comments

In this chapter we investigated the von Mises family in which both concentration and location parameters are unknown, and studied the powers of both the exact and bootstrap tests against different alternatives: offset normal distribution, cardioid distribution, wrapped skew-normal distribution, and asymmetric circular distribution. The tables presented lead to the following conclusions:

1. The estimates of the powers of exact tests are slightly larger than those of the bootstrap tests, but it is hard to say which test is more powerful because the estimates are subject to variability.

2. The difference between the powers of two tests gets smaller as the sample size n increases. When $n \geq 10$, both tests have very similar powers. This conclusion is also verified by the strong correlations between the exact and bootstrap p-values.

3. Both kinds of tests appear to be unbiased when the sample size is large enough.

Chapter 6

Summary

In this project, we studied the powers of both the exact conditional tests and the bootstrap tests. The gamma and von Mises families were investigated against different alternatives. Our study leads to the following conclusions.

1. Both kinds of tests have very similar powers even the sample size is very small.

2. The exact conditional tests are easier in terms of the computation of test statistics, which involves the calculation of the maximum likelihood estimates (MLE) for the unknown parameters. The MLE's depend on the value of sufficient statistics T. For co-sufficient samples from the same conditional distribution provided $T = t$, we need compute the MLE only once. However, the parametric bootstrap samples are not subject to the constraint $T = t$, we have to calculate the MLE for every parametric bootstrap sample.

3. The bootstrap tests are faster to implement, but the difference of time spent is acceptable if we conduct tests on 1 data set. For example, when a data set of size 15, Example 1 in Lockhart *et al.* (2008), is tested for von Mises family (both κ and μ unknown), and the same number (10,000) of co-sufficient samples and parametric bootstrap samples are used, it takes 13 minutes for the exact test and 2 minites for the bootstrap test.

Future work

There is still more work that could be done in the future.

1. We expected exact tests to be always slightly more powerful than bootstrap tests before this project. However, we saw the powers of bootstrap tests were larger than those of exact tests occasionally. This happened because both the powers and p-values are estimates, subject to variability. Therefore, the numbers of Monte Carlo samples (both M and B) could be increased for a future study to improve the accuracy.

2. In Subsection 3.2.2, the inverse of the conditional distribution for x_n is mentioned. This could be used in future work. For Case 2 of von Mises family, this also requires the distribution of $|h|$ for any 3 unit vectors, which are uniformally distributed on the circle. This was discussed in Stephens (1962). For $|h| = 1$, the density is infinite. We do not explore the inversion method in this project.

3. The application of the Gibbs sampler makes exact tests possible for other distribution families in which the Rao-Blackwell estiamte is not available, so we can extend our work to other distribution families.

Bibliography

ABRAMOWITZ, M., & STEGUN, I. A. (eds). 1972. *Handbook of Mathematical Functions.* New York: Dover.

CASELLA, G., & GEORGE, E. I. 1992. Explaining the Gibbs sampler. *The American Statistician*, **46(3)**, 167–174.

CHENG, R. C. H. 2006. Validating and comparing simulation models using resampling. *Journal of Simulation*, **1**, 53–63.

CHIB, S., & GREENBERG, E. 1995. Understanding the Metropolis-hastings Algorithm. *The American Statistician*, **49(4)**, 327–335.

GEMAN, S., & GEMAN, D. 1984. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6**, 721–741.

JAMMALAMADAKA, R. S., & SENGUPTA, A. 2001. *Topics in Circular Statistics.* Singapore: World Scientific Publishing Co.

LILLEGARD, M., & ENGEN, S. 1997. Stochastic simulations conditioned on sufficient statistics. *Biometrika*, **84**, 235–240.

LINDQVIST, B. H., & TARALDSEN, G. 2005. Monte carlo conditional on a sufficient statistic. *Biometrika*, **92**, 451–464.

LOCKHART, R. A., O'REILLY, F., & STEPHENS, M. A. 2007. Use of the Gibbs sampler to obtain conditional tests, with applications. *Biometrika*, **94(4)**, 992–998.

LOCKHART, R. A., O'REILLY, F., & STEPHENS, M. A. 2008. On testing fit of the von Mises distribution: conditional p-values. *Unpublished.*

O'REILLY, F., & GRACIA-MEDRANO, L. 2006. On the Conditional Distribution of Goodness-of-Fit Tests. *Communications in Statistics-Theory and Methods*, **35**, 541–549.

PEWSEY, A. 2000. The wrapped skew-normal distribution on the circle. *Communications in Statistics - Theory and Methods*, **29(11)**, 2459–2472.

STEPHENS, M. A. 1962. Exact and approximate tests for directions. *Biometrika*, **49(4)**, 463–477.

STEPHENS, M. A. 1986. Tests based on EDF statistics. *Pages 97–193 of:* D'AGOSTINO, R.B., & STEPHENS, M.A. (eds), *Goodness of Fit Techniques.* New York: Marcel Dekker.

WALSH, B. 2004 (April). *Markov Chain Monte Carlo and Gibbs Sampling.* Unpublished lecture notes, available at *http://www.stat.columbia.edu/~liam/teaching/neurostat-spr07/papers/mcmc/mcmc-gibbs-intro.pdf.*