© Grace Scientific Publishing

Journal of
Statistical Theory and Practice
Volume 3, No. 3, September 2009

Exact Conditional Tests and Approximate Bootstrap Tests for the von Mises Distribution

Richard A. Lockhart, *Department of Statistics and Actuarial Sciences, Simon Fraser University, Burnaby BC V5A 1S6, Canada.*
Email: lockhart@stat.sfu.ca

Federico O'Reilly, *Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México.*
Email: federico@sigma.iimas.unam.mx

Michael Stephens, *Department of Statistics and Actuarial Sciences, Simon Fraser University, Burnaby BC V5A 1S6, Canada.*
Email: stephens@stat.sfu.ca

Received: April 1, 2009 *Revised:* April 7, 2009

Abstract

Exact and approximate tests of fit are compared for testing that a given sample comes from the von Mises distribution. For the exact test, Gibbs sampling is used to generate samples from the conditional distribution of sample data, given the values of the sufficient statistics. The samples, called co-sufficient samples, are used to estimate the distribution of Watson's statistic, and hence to find the exact p-value for the given sample. The test is compared to the approximate test using the parametric bootstrap. Two examples are analyzed, and the p-values of the two tests are compared. When more examples are examined, an unexpectedly high correlation is discovered between the two sets of p-values, suggesting a strong mathematical connection.

AMS Subject Classification: 62G09; 62G10.

Key-words: Circular data; Gibbs sampler; Simulation; Watson test.

1. Introduction

The introduction of the concept of sufficiency by Fisher launched a vast amount of research to exploit the idea that all the information about parameters was contained in the sufficient statistic. The main emphasis was on parameter estimation, and less attention was paid to how best to test fit to a distribution. However, Lehmann, in his celebrated book on hypothesis testing, states that tests of fit should be made using the conditional distribution of

* 1559-8608/09-3/$5 + $1pp – see inside front cover

the data, given the sufficient statistic. The point is made again in the recent edition (Lehmann and Romano, 2005).

In this paper these ideas are applied to testing the fit of a sample to the von Mises distribution. The Gibbs sampler is used to generate samples from the conditional distribution given the sufficient statistic; these will be called *co-sufficient* samples, following the terminology used in Lockhart, O'Reilly and Stephens (2007), where similar tests were developed for the gamma distribution. The samples may be obtained from the univariate conditional distributions, which can be found analytically in terms of elliptic integrals. From the samples, the distribution of Watson's U^2 statistic is estimated and used to find the p-value of the sample U^2. The test is elegant and exact, and can be made as accurate as required by increasing the number of co-sufficient samples.

This exact test is compared with that given using the necessarily approximate parametric bootstrap. Remarkably, a very high correlation is found between the two p-values. We return to this discovery in Section 7.

In the next Section the general theory of conditional tests is given. The von Mises distribution is next described and conditional tests considered for this distribution. In Section 4 the Gibbs sampler procedure is outlined and in Section 5 the simulation methods are given to produce von Mises co-sufficient samples. In Section 6 the methods are applied to two data sets. Comparisons with the parametric bootstrap are described in Section 7 and some final remarks are in Section 8.

2. Conditional tests and sufficiency

2.1. General considerations

Suppose, in the general setting, that we wish to test H_o: a given random sample $X_1, X_2, \cdots, X_n$ with values $x_1, x_2, \cdots, x_n$ comes from distribution $F(x;\eta)$, where η is a vector of unknown parameters. Suppose also that there exists a minimal sufficient statistic T for η, which takes the value $T = t$ in the given sample. A statistic for testing fit, say S, is calculated from the sample, and the value is s_0. This value should be compared with the conditional distribution, say $F_S(s \mid t)$, of S given t.

The distribution $F_S(s \mid t)$ will depend on the conditional distribution $F_c(x \mid t)$ of $X_1, X_2, \cdots, X_n$ given t, but it will not depend on the true value of η.

In general, $F_c(x \mid t)$ is difficult to find, and therefore $F_S(s \mid t)$ is also difficult to find. However, if, say, M independent co-sufficient samples of size n can be generated from $F_c(x \mid t)$, and S is calculated from each, the distribution $F_S(s \mid t)$ may be estimated and the test of H_o made. Tests based on co-sufficient samples will be exact tests; the estimated p-value may be made as accurate as desired by increasing M.

Several authors have studied the problem of generating co-sufficient samples; see for example Engen and Lillegård (1997), Lindqvist and Taraldsen (2005), O'Reilly and Gracia-Medrano (2006). For tests for the gamma distribution, where $F_c(x \mid t)$ cannot be written explicitly, Lockhart, O'Reilly and Stephens (2007) used the Gibbs sampler to generate co-sufficient samples. The same approach will be used here for testing the von Mises distribution.

2.2. The parametric bootstrap

For tests based on the parametric bootstrap, the parameters of the von Mises distribution are first estimated from the given sample; then M Monte Carlo samples are drawn from this distribution to give an estimated distribution of the test statistic. This method cannot be made more accurate by increasing M since the samples are always drawn from only an estimate of the true parent distribution.

3. The von Mises distribution and conditional tests

Suppose vectors OP_i, $i = 1, \cdots, n$ are drawn from the centre O of a circle with radius 1, to points P_i on the circumference of the circle. Let x_i be the angular co-ordinate of OP_i : a vector might be referred to by its x-value, or by its components $\{\cos(x), \sin(x)\}$. The von Mises density is

$$f(x;\theta,\kappa) = \frac{1}{2\pi I_0(\kappa)} \exp\{\kappa \cos(x-\theta)\}$$

where θ is the location in $[0, 2\pi)$ and κ is a positive shape, and both are unknown. The functions $I_0(\kappa)$, and later, $I_1(\kappa)$, are modified Bessel functions of order 0 and 1, and $I_1(\kappa)$ is the derivative of $I_0(\kappa)$. The sample above is then $x_1, x_2, \ldots, x_n$ and it is desired to test H_o: the sample comes from the von Mises distribution with unknown shape parameter κ and unknown modal angle θ. The minimal sufficient statistic for the parameters is the resultant vector $t = (t_1, t_2)$, where

$$t_1 = \sum_{j=1}^{n} \cos(x_i); \quad t_2 = \sum_{j=1}^{n} \sin(x_i). \tag{3.1}$$

The vector t is the resultant (vector sum) of the n given vectors OP_i; suppose its length (norm) is $|t|$. Vector t, given here by its components, is often called R. Maximum likelihood estimates of the parameters are respectively $\hat{\theta}$, the angle of t, and $\hat{\kappa}$ satisfying the equation $I_1(\kappa)/I_0(\kappa) = |t|/n$.

The alternative hypothesis considered is the omnibus hypothesis that the sample is drawn from a density which is not in the von Mises family.

Classical ideas of hypothesis testing may be applied to this problem; see Lehmann and Romano (2005). Any unbiased level α test must be similar: the power function must be identically equal to α on the boundary between the null and alternative. In our case the boundary is the null so the power function must be identically α for all choices of shape κ and modal angle θ.

Lehmann and Romano (2005, p.115) show that whenever there is a complete sufficient statistic on the null hypothesis, a similar test must have conditional level exactly equal to α in order to be unbiased.

The von Mises model has such a complete sufficient statistic, and in this article we implement exact α-level conditional tests for this model. The resulting tests are therefore similar, but we do not know if they are unbiased.

The goodness-of-fit statistic used to test H_o will be Watson's (1961, 1962) U^2, an adaptation for the circle of the Cramér-von Mises W^2; U^2 does not depend on the origin used for the angles x. The calculation of U^2 is described at the end of this section.

Thus the ultimate goal is to find the conditional distribution of U^2 given t. There are two cases to consider.

Case 1. both κ and θ are unknown.

Case 2. κ is unknown but θ is known, a case which sometimes arises in biological applications.

Case 1 is most common and will be discussed first.

3.1. Watson's U^2

Suppose an ordered random sample $x_{(1)}, x_{(2)}, \cdots, x_{(n)}$ is given. Make the probability integral transformation $z_i = F(x_{(i)}; \theta, \kappa), i = 1, \cdots, n$, and let $\bar{z} = \sum_1^n z_i/n$. When parameters are unknown, they are replaced by their maximum likelihood estimates. Watson's U_n^2 will here be called simply U^2; it is given by

$$U^2 = \sum_1^n (z_i - (2i-1)/2n)^2 - n(\bar{z} - 0.5)^2.$$

4. The Gibbs Sampler

As stated earlier, the Gibbs sampler will be used to generate co-sufficient samples. The steps in the Gibbs sampling procedure are as follows. The details of the simulation will be given in Section 5.

(1) First transform the vector of variables $x_1, \ldots, x_n$ to the variables t_1, t_2, $x_3, \ldots, x_n$. Then, assuming H_o is true: the sample is from the von Mises distribution, the conditional density of, say, X_3, given t and the variables $x_4, \ldots, x_n$ can be found to within a constant of proportionality.

(2) A new value of X_3, based on this conditional distribution can then be simulated by a procedure to be described in the next section. Let the new value be x_3^*.

(3) Using the simulation procedure again, simulate a value x_4^* based on the conditional distribution of X_4 given t and $x_3^*, x_5, \ldots, x_n$.

(4) Similarly simulate x_5^*, having conditioned on t, and $x_3^*, x_4^*, x_6, \ldots, x_n$; continue until a set of $n-2$ new values $x_3^*, \ldots, x_n^*$ has been generated.

(5) The last two values (x_1^* and x_2^*) are obtained by solving (3.1). The co-sufficient sample (the x^*-sample) then has sufficient statistic t. The details of this solution are given in Section 5.5.

(6) This procedure is repeated many times with each newly generated sample of x^* values replacing the original sample. The generated samples form a Markov chain. If, as was assumed above, the original sample distribution was von Mises, the co-sufficient samples will be generated from the required conditional distribution of the data, given t.

However, when the original sample is not von Mises, the Gibbs sampler procedure ensures that the Markov Chain converges to a stationary distribution which is the conditional distribution, and then the generated samples are co-sufficient. In general, therefore, when H_o cannot be assumed true, a burn-in period is required where say M_0 samples are generated but discarded before keeping the remaining successive samples as the co-sufficient samples. These might be serially correlated, but the simulated distribution of a statistic (for example U^2) can nevertheless be estimated if all the generated samples are used after the initial burn-in period.

If it is desired to maintain independence amongst generated samples some may be discarded between retained samples. For example, for the values of U^2, the serial correlations were negligible for lags larger than about 4 and so about 4 samples would be discarded between retained samples if independence is wanted. Clearly, this makes the generation much slower.

5. Simulation procedure

5.1. General Considerations

From the given sample, as stated above, the resultant vector $t = (t_1, t_2)$ is the minimal sufficient statistic. Define also

$$\tilde{t}_1 = \sum_4^n \cos(x_i), \quad \tilde{t}_2 = \sum_4^n \sin(x_i), \tag{5.1}$$

and calculate components h_1 and h_2 of vector h by

$$h = (h_1, h_2) \equiv (t_1 - \tilde{t}_1, t_2 - \tilde{t}_2). \tag{5.2}$$

The vector h is the resultant of three unit vectors; suppose its angular co-ordinate is α_h. Define z by

$$z = \cos(x_3^* - \alpha_h). \tag{5.3}$$

Remark 5.1. The conditional distribution of z given t and $x_4, \ldots, x_n$ is the same as the conditional distribution of z given h, a result to be used below.

A value of z is to be generated in order to obtain x_3^*. The simulation procedure depends on the following Proposition.

Proposition 5.1. *The conditional density of z given t and $x_4, x_5, \ldots, x_n$ is proportional to the kernel*

$$g(z) = \frac{1}{\sqrt{\left[1 - \{1 - (|h|^2 - 2|h|z + 1)/2\}^2\right](1 - z^2)}} \tag{5.4}$$

where the argument z is in the range:

$$\max\{(|h|^2 - 3)/2|h|, -1\} < z < 1\,.$$

The proposition is proved in two steps. First use the remark above to simplify the problem to that of computing the conditional density of z given h. Second compute the joint density of z and h via the change of variables formula from the joint density of x_1, x_2, x_3. It will then be seen that the kernel g arises as the Jacobian of the transformation.

Using the proposition a z-value may then be generated in one of two ways – using an acceptance-rejection procedure similar to that used by Lockhart, O'Reilly and Stephens (2007) for the Gamma distribution, or analytically.

5.2. Method 1: Acceptance-Rejection

In generating an observation by acceptance-rejection, it is not necessary to know the multiplying constant which relates the kernel $g(z)$ to its conditional density; it is sufficient to majorize (5.4).

The denominator of $g(z)$ tends to zero as z approaches the extremes of its interval, yielding a "U" shaped kernel. This may be majorized with an appropriately scaled beta density in order to generate a z value by acceptance-rejection. Recall that $z = \cos(x_3^* - \alpha_h)$. The sign of the corresponding angle $\alpha_z = x_3^* - \alpha_h$ is selected with equal probability and x_3^* is then $x_3^* = \alpha_h + \alpha_z$.

5.3. Method 2: Analytic Generation

In implementing the Gibbs sampler, new x^* values may be generated by using the inverse Probability Integral Transform. The kernel $g(z)$ in (5.4) may be integrated explicitly in terms of elliptic integrals; it is then normalized to obtain the corresponding cumulative distribution function denoted by

$$G(\cdot \mid t, x_4, \ldots, x_n) = G(\cdot \mid h)$$

where, as in (5.2), h is the (vector) resultant for the first three data points.

This distribution can then be inverted so that, if $G\{z \mid h\} = u$, for $0 < u < 1$, the inverse may be written

$$z = Q(u \mid h). \tag{5.5}$$

This is the inverse probability integral transform and will be used to give a value $\tilde{z}$ from a value u, uniformly generated in the interval (0,1). The calculation of $Q(u \mid h)$ will be described in the next section.

When $\tilde{z}$ has been generated a new x^* value is computed as described in Section 5 for the acceptance rejection method, with $\tilde{z}$ replacing z. At each step h is modified by calculating $\tilde{t}_1$ and $\tilde{t}_2$ using new values x^* as generated.

5.4. The Inverse Probability Transform

In order to calculate the function $Q(u \mid h)$, elliptic integrals are needed. There are several notational conventions for elliptic integrals, so definitions will be given for the versions used here.

The incomplete elliptic integral of the first kind is, for $0 < k < 1$ and $0 \leq u \leq 1$,

$$F(u,k) = \int_0^u \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}}.$$

The notation $F(u,k)$ is standard, and will be used here. There should be no confusion with cumulative distribution functions as used in other Sections.

The corresponding complete integral is $K(k) = F(1,k)$. The Jacobi amplitude function $\mathrm{am}(u,k)$ satisfies $F(\mathrm{am}(u,k),k) \equiv u$ for $0 \leq u \leq F(1,k)$ and $0 < k < 1$. Finally the Jacobi sine function is $\mathrm{sn}(u,k) = \sin(\mathrm{am}(u,k))$. Software to compute these functions is available in many packages; for instance, in `R`, the package `elliptic` provides all the needed functions.

With these definitions in hand, we can return to the calculation of $Q(u \mid h)$. Recall from Section 5 that we need z defined by (5.3). The first step is to find the marginal density of the length $|h|$ of the resultant of a sample of size $n = 3$ from the uniform distribution on the circle, (that is, from the von Mises distribution with $\kappa = 0$). Stephens (1962, formulas 23a and 23b) has given this marginal density: the density of $|h|$ has a discontinuity at $|h| = 1$ and this makes the conditional density of z have a form which differs according to the sign of $|h| - 1$.

To state the result put

$$r_1(z,|h|) = \begin{cases} \sqrt{\dfrac{(3-|h|^2+2|h|)(1+z)}{2(3-|h|^2+2|h|z)}}, & |h| < 1, \\ \sqrt{\dfrac{2(3-|h|^2+2|h|z)}{(3-|h|^2+2|h|)(1+z)}}, & |h| > 1, \end{cases}$$

and

$$r_2(|h|) = \begin{cases} 4|h|^{1/2}(|h|+1)^{-3/2}(3-|h|)^{-1/2}, & |h| < 1, \\ (|h|+1)^{3/2}(3-|h|)^{1/2}|h|^{-1/2}/4, & |h| > 1. \end{cases}$$

We then find

Proposition 5.2. *The conditional cumulative distribution of z given*

$$h = \left(\sum_1^3 \cos(x_i), \sum_1^3 \sin(x_i) \right)$$

is

$$G(z \mid h) = \frac{F\{r_1(z,|h|), r_2(|h|)\}}{K\{r_2(|h|)\}}$$

for the interval

$$\max\left\{-1, \frac{|h|^2-3}{2|h|}\right\} \leq z \leq 1.$$

The corresponding conditional inverse function $Q(u \mid h)$, satisfying (5.5)*, becomes*

$$Q(u \mid h) = \begin{cases} \dfrac{(2|h|^2-6)J\{u,r_2(|h|)\}+(|h|^2-2|h|-3)}{4|h|J\{u,r_2(|h|)\}-(|h|^2-2|h|-3)}, & |h|<1, \\ \dfrac{2|h|^2-6-J\{u,r_2(|h|)\}(|h|^2-2|h|-3)}{4|h|+(|h|^2-2|h|-3)J\{u,r_2(|h|)\}}, & |h|>1, \end{cases}$$

where

$$J(u,r_2) = sn^2(uK(r_2),r_2).$$

5.5. The final two x^*-values

After generating the $n-2$ new variables $x_3^*,\ldots,x_n^*$, the two values x_1^*,x_2^* are found as the solutions to the system:

$$\cos(x_1^*)+\cos(x_2^*) = t_1 - \sum_3^n \cos(x_i^*),$$

$$\sin(x_1^*)+\sin(x_2^*) = t_2 - \sum_3^n \sin(x_i^*).$$

If the components on the right hand sides of these two equations form the vector l, whose angle is α_l and norm is $|l|$, expressed in polar coordinates, the solution for the unordered pair $\{x_1^*,x_2^*\}$ is $(\alpha_l+\cos^{-1}(|l|/2), \alpha_l-\cos^{-1}(|l|/2))$.

With this procedure the system always has a unique solution, except for the permutation of the indices 1 and 2.

5.6. The von Mises test: location known

In some applications, the practitioner might know the modal direction $\theta=\theta_0$ and still want to test the fit of the von Mises as a model for the data. For this, create the new sample y by $y_i = x_i-\theta_0$ and follow the steps in the case where both parameters were unknown; the minimal sufficient statistic is now the component $t_1' = \sum_1^n \cos(y_i)$ of the resultant of the vectors with angles y_i. In this case, start the Gibbs sampler by simulating a value y_2^* from the conditional for Y_2 given t_1' and $y_3,y_4,\ldots,y_n$; then, as before, continue by simulating a value y_3^* from the conditional distribution given t_1', y_2^* and $y_4,y_5,\ldots,y_n$, and so on. The required conditional density for Y_2 depends on a density of, say, $z_0=\cos(Y_2-y_0)$, where y_0 is the angle of the resultant vector whose first coordinate is $t_1'-\sum_3^n\cos(y_i^*)$. This density is proportional to a kernel similar to that of Proposition 5.1; this simpler kernel (for known θ_0) is

$$g_0(z) = \frac{1}{\sqrt{\left[1-\left\{z-t_1'+\sum_3^n\cos(y_i^*)\right\}^2\right](1-z^2)}}$$

for z in the interval

$$\max\left\{-1-t_1'+\sum_3^n \cos(y_i^*),-1\right\} < z < \min\left\{1-t_1'+\sum_3^n \cos(y_i^*),1\right\}.$$

As in the previous case, the angle whose cosine is z is then randomly assigned a sign and added to y_0 to obtain a new y_2^*. The sample generated (the y_i^* set) may then be changed to $x_i^* = y_i^* + \theta_0$, for $i = 1,\ldots,n$. Since the test statistic U^2 is invariant to location changes, this change makes no difference in the resulting p-value, but for the purpose of making plots it may be best to express the angles with the original orientation.

6. Examples

In this Section two data sets will be used to compare the p-values of the exact conditional test with those given by the parametric bootstrap. The p-values will also be estimated from the asymptotic points given in Lockhart and Stephens (1985).

Example 6.1. The first example comes from Stephens (1972), where two sets of geological data are given. The second set has 34 observations, given in Table 6.1. When tested to come from the von Mises distribution, estimates are $\hat{\theta}_0 = 3.771$ radians (equals 216.0 degrees), and $\hat{\kappa} = 1.326$; then $U^2 = 0.0948$.

For the exact test, using $M = 100,000$ co-sufficient samples with burn-in $M_0 = 100$, the p-value for U^2 is 0.03646. For the bootstrap, $M = 100,000$ again, and $p = 0.03904$. These results both suggest a poor fit to the von Mises distribution.

Example 6.2. The data in this example are from biology, and are directions taken by nematodes when subjected to a light source. They are given in Table 6.2. The estimates of θ_0 and κ are 0.712 radians (= 40.8 degrees) and 0.706. The value of U^2 is 0.0251.

With 100,000 good co-sufficient samples and 100 burn-in samples omitted, $p = 0.7290$ from the exact test, and 0.7353 from 100,000 bootstrap samples. The von Mises distribution is clearly acceptable.

7. Correlation between Exact and Approximate p-values

The close values of p given by the exact and approximate methods for the above examples suggested a closer investigation. It is interesting to base this on a specific example, so for Examples 6.1 and 6.2, 200 samples were generated with the given parameter estimates. For each of these samples, parameters were estimated and U^2 calculated, and the p-values were found by the conditional and bootstrap methods. The 200 pairs of p-values have been plotted in Figures 7.1 and 7.2. They show a remarkable correlation ($R^2 = 0.997$ for the geological data of Example 6.1 and $R^2 = 0.9997$ for the nematode data of Example 6.2) between the two sets. A similar high correlation between p-values was observed by Lockhart, O'Reilly and Stephens (2007) in connection with tests for the gamma distribution. The bootstrap values are slightly larger, on the whole, than the exact values (as they were in Examples 6.1 and 6.2), so that the tests based on these will give a size α slightly smaller than the exact.

Table 6.1. Geological Data: Example 6.1.

90	100	115	115	130	135	145	160	165	170	180	190
190	196	200	205	210	225	230	245	250	250	253	254
254	255	256	261	270	277	280	290	290	305		

Table 6.2. Nematode Data: Example 6.2.

2.75	17.32	29.19	31.90	34.50	41.81	41.86	45.56
45.70	47.27	58.62	59.32	62.57	69.92	71.88	73.84
74.33	74.68	76.91	86.91	91.53	105.68	108.13	108.61
117.03	119.05	132.08	134.19	135.76	168.51	177.11	179.85
188.19	196.63	198.91	228.69	251.96	255.68	290.99	291.53
294.30	305.42	308.88	321.66	323.20	328.32	335.93	341.80
351.11	353.24	353.65	354.91	355.00	355.83	359.79	

Similar plots were made for several other data sets and all show the same pattern and very high correlation.

8. Final Remarks

Remark 8.1. An exact test of fit has been given for testing the von Mises distribution around a circle. The test is straightforward to implement: computer code is available from the authors.

Remark 8.2. The existence of the exact test has enabled calibration of the bootstrap approximate test, a result not usually available. In the tests for von Mises or gamma, the bootstrap tends to give a conservative test; the bootstrap p is higher than the true value.

Remark 8.3. The calibration is made possible by a very high correlation existing between exact p-values and those given by the parametric bootstrap test for both gamma and von Mises tests. These unexpected results bring up wider issues; they suggest a strong mathematical connection between the two procedures. This has been investigated by Lockhart (2008a, unpublished), who shows that the two methods have the same large sample theory, so this good agreement between p-values can be explained for large n. However, plots like Figures 7.1 and 7.2 arise even for small samples. Also, Lockhart (2008b, unpublished) has shown that conditional tests using quadratic functionals of the empirical distribution function (such as Watson's U^2, the Cramér-von Mises W^2 and the Anderson-Darling A^2 statistics) approximately maximize the average power with respect to a suitable prior distribution on the alternative hypothesis in these problems. Thus U^2 used as indicated here is optimal in this sense of maximizing a certain average power. Neyman smooth tests, studied, for example, by Rayner and Best (1989, 1990), when implemented with the conditional method given above, are similarly optimal, although for somewhat different priors. These issues will be investigated further.

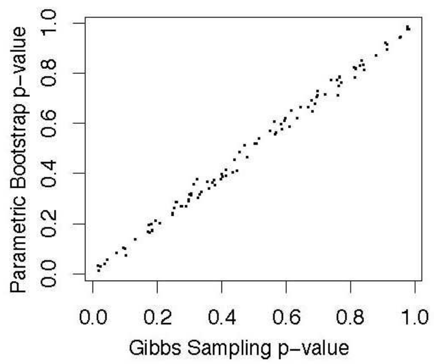

Figure 7.1. Plot of bootstrap p-values against exact values: Example 6.1.

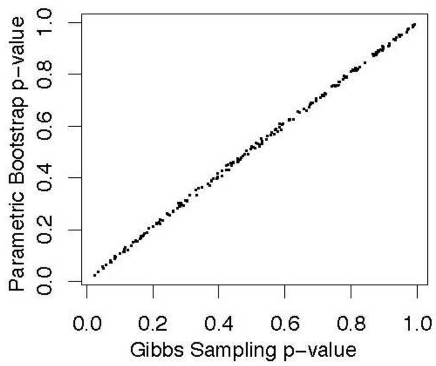

Figure 7.2. Plot of bootstrap p-values against exact values: Example 6.2.

Remark 8.4. Other Markov Chain Monte Carlo methods are available for this problem. In particular, investigations are under way of a Metropolis Hastings algorithm which would replace the acceptance rejection step of our first algorithm with a Metropolis Hastings step.

Acknowledgements

The authors thank Professor John Rayner for his helpful comments. They also thank Jay Burr of Simon Fraser University Department of Biological Sciences for providing the data used in Example 6.2, and Zheng Sun for making Figures 7.1 and 7.2. This research was supported by the Natural Sciences and Engineering Research Council of Canada, and the Council is thanked for this support. Also support from Consejo Nacional de Ciencia y Tecnología is acknowledged for the second author while on sabbatical leave.

References

Engen, S., Lillegård, M., 1997. Stochastic Simulation conditioned on sufficient statistics. *Biometrika*, 84, 235–240.

Lehmann, E.L., Romano, J.P., 2005. *Testing Statistical Hypotheses*. 3rd edition. Springer, New York.

Lindqvist, B.H., Taraldsen, G., 2005. Monte Carlo conditioning on a sufficient statistic. *Biometrika*, 92, 451–464.

Lockhart, R.A., 2008a. Conditional limit laws for goodness-of-fit tests. Unpublished manuscript.

Lockhart, R.A., 2008b. Bayesian assessment of goodness-of-fit tests. Unpublished manuscript.

Lockhart, R.A., O'Reilly, F.J., Stephens, M.A., 2007. Use of the Gibbs sampler to obtain conditional tests, with applications. *Biometrika*, 94, 992–998.

Lockhart, R.A., Stephens, M.A., 1985. Tests of fit for the von Mises distribution. *Biometrika*, 72, 647–652.

O'Reilly, F.J., Gracia-Medrano, L., 2006. On the conditional distribution of goodness-of-fit tests. *Commun. Statist. (Theory and Methods)*, 35, 541–549.

Rayner, J.C.W., Best, D.J., 1989. *Smooth Tests of Goodness-of-Fit*. Oxford University Press, New York.

Rayner, J.C.W., Best, D.J., 1990. Smooth tests of goodness of fit: an overview. *Int. Statist. Rev.*, 58, 9–17

Stephens, M.A., 1962. Exact and approximate tests for directions I. *Biometrika*, 49, 463–477.

Stephens, M.A., 1972. Multisample tests for the von Mises distribution. *J. Am. Statist. Assoc.*, 67, 456–461.

Watson, G.S., 1961. Goodness-of-fit tests on a circle. *Biometrika*, 48, 109–114.

Watson, G.S., 1962. Goodness-of-fit tests on a circle. 2. *Biometrika*, 49, 57–63.