ANOTHER CALCULATION OF THE $p$-VALUE FOR THE PROBLEM OF REGIONS USING THE SCALED BOOTSTRAP RESAMPLINGS

by

Hidetoshi Shimodaira

SEPTEMBER 2000

Department of Statistics
STANFORD UNIVERSITY
Stanford, California
ANOTHER CALCULATION OF THE p-VALUE FOR THE PROBLEM OF REGIONS USING THE SCALED BOOTSTRAP RESAMPLINGS

by

HIDETOSHI SHIMODAIRA
Visiting at
Department of Statistics
Stanford University

SEPTEMBER 2000

Department of Statistics
Sequoia Hall
STANFORD UNIVERSITY
Stanford, California 94305-4065

http://www-stat.stanford.edu
Another calculation of the $p$-value for the problem of regions using the scaled bootstrap resamplings

Hidetoshi Shimodaira

The Institute of Statistical Mathematics
4-6-7 Minami-Azabu, Minatoku, Tokyo 106-8569, JAPAN

Visiting at
Department of Statistics, Stanford University
shimo@ism.ac.jp, shimo@stat.stanford.edu

September 26, 2000

Abstract

An approximately unbiased test is considered for the null hypothesis represented as a region with smooth boundaries. This problem is discussed previously in Efron and Tibshirani (1998), and our argument is based on their results. We give another calculation of the corrected $p$-value without finding the "nearest point" on the boundary to the observation, which is required in the calculation of Efron, Halloran and Holmes (1996). We generate sets of bootstrap replicates with several sample sizes which may differ from that of the observed data. For each set of replicates, the frequency that the replicates fall in the region is counted. Only these frequencies are used to estimate the signed distance and the curvature of the boundary in the calculation of the $p$-value. Our calculation will be useful for applications with correlated data structure where complicated bootstrap methods are used.

Keywords: problem of regions; curvature; similar on the boundary; unbiased test; scaled bootstrap
1 Introduction

We consider the problem to test the null hypothesis represented as a region in a space under the assumption that we only determine 1/0 value indicating whether the observation is in the region or not. Given a large number of nonparametric bootstrap replicates of the data, the naive $p$-value is obtained as the frequency that the replicates fall in the region. However, this $p$-value is biased either upward or downward depending on the curvature of the boundary of the region, and a bootstrap algorithm to calculate the bias corrected $p$-value is proposed in Efron, Halloran and Holmes (1996) and Efron and Tibshirani (1998). In this paper, we propose another calculation of the corrected $p$-value using a set of bootstrap replicates with several sample sizes which may differ from the sample size of the observed data. An advantage of this idea is that there is no need to calculate the “nearest point” on the boundary to the observation, which is required in the calculation of Efron et al. (1996).

We follow the argument of Efron and Tibshirani (1998), and consider the simple normal model. Let the observation $y$ be distributed as $K$-dimensional multivariate normal with unknown mean $\mu$ and covariance matrix the identity,

$$y \sim N_K(\mu, I). \tag{1}$$

We are interested in the null hypothesis that $\mu$ is included in a region $R$ which is a nondegenerated subset of the $K$-dimensional Euclidean space with smooth boundaries. The alternative hypothesis is represented as the region that $\mu$ is not included in $R$. The indicator function $I_R(y) = 1$ for $y \in R$, and zero otherwise.

Note that the normal model (1) is a simplification of the practical problem where the unknown parameter vector is estimated from the observation of sample size $n$, and the estimate may have the asymptotic normal distribution with the covariance matrix proportional to $n^{-1/2}$. The estimate will be converted to $y$ in (1) if it is rescaled by the factor $n^{1/2}$ and the appropriate linear transformation is applied to it.

The bootstrap replicate $y^*$ of $y$ will be distributed as

$$y^* \sim N_K(y, \sigma^2 I), \tag{2}$$

where $\sigma$ is the scale parameter. In the practical problem, we can easily specify $\sigma$ by changing the resample size $m$ which may differ from $n$; $\sigma = \sqrt{n/m}$. We
will use $k \geq 2$ set of bootstrap replicates, where each set consists of $B_i$ samples from
\[ y_{i,j}^* \sim N_K(y, \sigma_i^2 I), \quad j = 1, \ldots, B_i \]  
for $i = 1, \ldots, k$.

The naive $p$-value obtained from each set of the replicates is
\[ \tilde{p}_i = B_i^{-1} \sum_{j=1}^{B_i} I_R(y_{i,j}^*). \]  

For $\sigma_i = 1$, a small value of $\tilde{p}_i$ may indicate the rejection of the null hypothesis, but it is biased when the boundary is curved. In the following sections, we give a bias corrected $p$-value $\hat{p}$ in the sense that $\hat{p}$ is approximately distributed as the uniform on $(0, 1)$, yet non-randomized, when the true mean vector $\mu$ lies on the boundary of the region $\mathcal{R}$; i.e., approximately similar on the boundary (Lehmann p. 135, 1986).

## 2 Signed Distance and Curvature

The signed distance $x_0$ is defined to be the distance from point $y$ to the nearest point $\hat{\mu}_0$ on the boundary, taken positive if $y \not\in \mathcal{R}$ and negative if $y \in \mathcal{R}$. Let $\hat{d}$ be the $(K - 1) \times (K - 1)$ matrix of the curvature of the boundary at $\hat{\mu}_0$, and define
\[ \hat{d}_1 = \text{trace}(\hat{d}), \quad \hat{d}_2 = \text{trace}(\hat{d}^2). \]

Let $\Phi(\cdot)$ denote the distribution function of the standard normal distribution. Efron and Tibshirani (1998) show that $p$-value defined by
\[ \hat{p} = 1 - \Phi \left( \frac{x_0 - \hat{d}_1}{1 - \hat{d}_2} \right) \]  
is third-order accurate when we consider the asymptotics such that the sample size $n$ goes to infinity, i.e.,
\[ \Pr\{\hat{p} < \alpha\} = \alpha + O(n^{-3/2}), \quad 0 < \alpha < 1 \]
when $\mu$ is on the boundary. Since $\hat{d} = O_p(n^{-1/2})$, (5) is rewritten as
\[ \hat{Z} \approx x_0 - (\hat{d}_1 - x_0 \hat{d}_2), \]  

3
where $\tilde{Z} = \Phi^{-1}(1 - \hat{p})$, and "≈" indicates the third order accuracy. An estimate of the right hand side on (6) is given by the following argument.

Let $\tilde{p}_\sigma$ be the naive p-value obtained from $B$ samples from (2);

$$\tilde{p}_\sigma = B^{-1} \sum_{j=1}^{B} I_{R}(y_j^*) .$$  \hspace{1cm} (7)

It follows from (2.19) of Efron and Tibshirani (1998) that

$$\tilde{p}_\sigma = 1 - \Phi \left( \frac{x_0/\sigma + \sigma \tilde{d}_1}{1 + \sigma^2 \tilde{d}_2} \right) + O_p(n^{-3/2}) + O_p(B^{-1/2}),$$  \hspace{1cm} (8)

because $x_0$ and $\tilde{d}$ in the eq. (2.19) are replaced by $x_0/\sigma$ and $\sigma \tilde{d}$, respectively. Then we have

$$\tilde{Z}_\sigma \approx \sigma^{-1} x_0 + \sigma (\tilde{d}_1 - x_0 \tilde{d}_2) + O_p(B^{-1/2}),$$  \hspace{1cm} (9)

where $\tilde{Z}_\sigma = \Phi^{-1}(1 - \tilde{p}_\sigma)$. Therefore, $x_0$ and $c := \tilde{d}_1 - x_0 \tilde{d}_2$ will be estimated by plotting points $(\sigma^{-1}, \tilde{Z}_\sigma)$ for several $\sigma$ values, and then $\hat{p}$ will be obtained by $1 - \Phi(x_0 - c)$.

### 3 Implementation

The proposed algorithm to calculate $\hat{p}$ is quite simple.

(Step 1) Specify the scale parameter $\sigma_i$ and the number of replicates $B_i$ for $i = 1, \ldots, k$.

(Step 2) Generate the replicates $y_{i,j}$, $i = 1, \ldots, k$, $j = 1, \ldots, B_i$ according to (3).

(Step 3) Count the number of times that $y_{i,1}, \ldots, y_{i,B_i}$ are in the region $R$ to calculate $\tilde{p}_i$ according to (4). If the count is 0 or $B_i$, then do not use it.

(Step 4) Plot the points $(\sigma_i^{-1}, \tilde{Z}_i)$, $i = 1, \ldots, k$, where $\tilde{Z}_i = \Phi^{-1}(1 - \tilde{p}_i)$. Estimate $x_0$ and $c$ by minimizing

$$\text{RSS} = \sum v_i^{-1} \left( \sigma_i^{-1} x_0 + \sigma_i c - \tilde{Z}_i \right)^2,$$

where $v_i = \tilde{p}_i(1 - \tilde{p}_i)/(\phi(\tilde{Z}_i)^2 B_i)$ is the approximate variance of $\tilde{Z}_i$. $\phi(\cdot)$ denotes the density function of the standard normal distribution.
(Step 5) Calculate the bias corrected $p$-value by

$$\hat{p} = 1 - \Phi(x_0 - c).$$  \hspace{1cm} (10)

The approximate variance of $\hat{p}$ is given by

$$\phi \left( \Phi^{-1}(\hat{p}) \right)^2 (1, -1) \left( \sum_{i=1}^{k} u_i^{-1} \begin{pmatrix} \sigma_i^{-2} & 1 \\ 1 & \sigma_i^2 \end{pmatrix} \right)^{-1} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$  \hspace{1cm} (11)

(Step 6) Check the validity of the approximation employed in Section 2. RSS is approximately distributed as $\chi^2$ with $k - 2$ degrees of freedom when the argument of Section 2 is justified. If RSS is significantly large, then we should not depend on this algorithm to calculate $\hat{p}$.

(Step 7) The standard method of bias correction may be applied to the estimated $\hat{p}$, because the distribution of $\hat{p}$ is skewed for small value of $B_i$. Step 2 is often expensive in computation, so the replicates of the counts in Step 3 are directly obtained from the multinomial distribution using the observed values of the counts. This “further” bias corrected $p$-value is denoted $\hat{p}$.

In Step 1 of the algorithm, it is important to specify good values of $B_i$ to reduce (11). We will use a heuristic algorithm to minimize (11) by fixing $B = \sum_{i=1}^{k} B_i$, whereas it might be more practical to fix $\sum_{i=1}^{k} \sigma_i^{-2} B_i$ in terms of the computational cost.

4 Example

We take the same example as given in Section 2 of Efron and Tibshirani (1998), which is simple enough so that we can calculate the exact value of $\hat{p}$, denoted $p$.

Consider the region $\mathcal{R} = \{ \mu : \| \mu \|^2 \leq r^2 \}$ for $r > 0$ and $\mathcal{R} = \{ \mu : \| \mu \|^2 \geq r^2 \}$ for $r < 0$. When $\mu$ is on the boundary, i.e. $\| \mu \|^2 = r^2$, $\| y \|^2$ is distributed as $\chi^2$ with $K$ degrees of freedom and the noncentrality parameter $r^2$; we denote it by $F_{K,r^2}(s) = \Pr\{ \| y \|^2 \leq s \}$. Thus $p = 1 - F_{K,r^2}(\| y \|^2)$ for $r > 0$ and $p = F_{K,r^2}(\| y \|^2)$ for $r < 0$. Also $\hat{d}_1 = (K - 1)/2r$ and $\hat{d}_2 = (K - 1)/(2r)^2$ are used to calculate the true value of $c$.

First we consider the case that $K = 4$, $r = 5$, and $\| y \| = 7$. I chose $k = 16$ scale parameters $(\sigma_1, \ldots, \sigma_{16}) = (0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4,$
1.5, 1.6, 1.7, 1.8, 1.9, 2.0), and used \( B_i = 1000 \) for all \( i \), thus the total number of replicates is \( B = 16,000 \). For the sixteen sets of replicates, we obtained the counts of falling in the region as \((0, 0, 1, 6, 9, 9, 24, 25, 30, 49, 38, 46, 56, 59, 73, 67)\). Then \( x_0 \) and \( c \) are estimated as 1.94 and 0.26 from the plot of \((\sigma_i^{-1}, \tilde{Z}_i)\), \( i = 1, \ldots, k \) as shown in Fig. 1, where the true values of them are 2 and 0.24, respectively. The fitting of the curve is quite good as indicated by the \( \chi^2 \) statistic with 12 degrees of freedom \( \text{RSS} = 9.4 \). The corrected p-value is \( \hat{p} = 1 - \Phi(1.94 - 0.26) = 0.046 \). This agrees nicely with the exact value \( p = 0.040 \). On the other hand, the naive p-value for \( \sigma = 1 \) is \( 9/1000 = 0.009 \), and its theoretical value is \( \tilde{p} = F_{K, \sigma^2}(5^2) = 0.012 \) for \( B = \infty \). The difference between \( \hat{p} \) and \( \tilde{p} \) is quite large, and \( \tilde{p} \) is a very poor estimate of \( p \). In Step 7 of the algorithm, we applied the bootstrap-t method to \( \hat{p} \) with the number of replicates 1000, and obtained 95% confidence interval (0.028, 0.063), and \( \hat{p} = 0.042 \), which is closer to the exact value than that obtained in Step 5.

Next we consider the case that \( K = 4 \), \( r = -5 \), and \( ||y|| = 3 \). The same \( k \), \( \sigma_i \), and \( B_i \) are used as the previous case. We got the counts \((0, 1, 5, 11, 27, 59, 83, 109, 155, 174, 233, 245, 303, 327, 362, 426)\), and the corrected p-value is obtained as \( \hat{p} = 1 - \Phi(2.01 - (-0.39)) = 0.0083 \). The fitting of the curve is quite good as indicated by the \( \chi^2 \) statistic with 13 degrees of freedom \( \text{RSS} = 7.5 \). The obtained \( \hat{p} \) agrees well with the exact value \( p = 0.0085 \). Again, the naive p-value \( \tilde{p} = 1 - F_{K, \sigma^2}(5^2) = 0.053 \) is very different from \( \hat{p} \). The bootstrap-t gives 95% confidence interval (0.0066, 0.0100) and \( \hat{p} = 0.0079 \). This deviates away from the exact value \( p = 0.0085 \), but becomes closer to \( \hat{p} = 0.0080 \), the corrected p-value with \( B = \infty \). Note that \( \tilde{p} \) with \( B_i = \infty \) is calculated analytically for our simple example. We will see the difference between \( p \) and \( \hat{p} \) with \( B = \infty \) in the next section.

The values of \( B_i \) used in the above example are not very efficient to reduce the variance of \( \hat{p} \). The square root of (11) gives an estimate of the standard deviance of \( \hat{p} \), and it was 0.009 for the first case, and 0.0013 for the second case. A stochastic searching algorithm to obtain optimal \( B_i \) values with \( B = 16,000 \) was applied to the two cases, and it gave \((B_5, B_{16}) = (11565, 4435)\) and \((B_5, B_{16}) = (11405, 4595)\), respectively, while the other \( B_i \) values were zero. The estimated standard deviances then became 0.006 and 0.0009, respectively, and they were about 2.4 times efficient in saving the number of replicates.

6
5 Simulation Study

We show the results of two sets of simulations in Table 1 and Table 2. The same scale parameters as the previous section are used throughout.

In the first set of simulations, the boundary of the null region is sphere, and p-values are calculated for fixed points of $y$ located $x_0 = 2$ far away from the null region. The influence of the magnitude of the curvature on the calculation of p-values is shown in Table 1. It is easy to see that all the p-values become closer to $p$ as the absolute value of the curvature $d_1$ becomes smaller, which verifies the asymptotic argument. In fact, $\hat{p}$ and $\tilde{p}$ with $B = \infty$ give the same value as $p$ when $d_1 = 0$. $\hat{p}$ and $p$ are similar to each other and they increase as $d_1$ becomes larger, whereas $\tilde{p}$ changes in the other way. This is because $\tilde{p}$ is "bias corrected" in the wrong direction.

The influence of the magnitude of $B$ on $\hat{p}$ is also seen in Table 1. $\hat{p}$ has a relatively large bias when $B_1 = 1000$, and it is adjusted reasonably by the further correction of Step 7. For $B_1 = 10000$, $\hat{p}$ gives a very good estimate of the $\hat{p}$ with $B = \infty$, and the further correction becomes almost marginal.

In the second set of simulations, the boundary of the null region is oval in $K = 2$ dimensions defined by $(\mu_1/rs)^2 + (\mu_2/rs^{-1})^2 = 1$ for $s = 1, 2, 3$. In each repetition of the simulations, $y$ is generated from $N(\mu, I)$ where $\mu$ is located on the boundary. $\mu$ is chosen so that the curvature at $\mu$ is $d_1 = 1/2r$; $(0.71r, 0.71r)$, $(1.79r, 0.22r)$, $(2.85r, 0.11r)$ for $s = 1, 2, 3$, respectively. The curvature of the boundary ranges from $s^{-3}/2r$ at $(0, s^{-1}r)$ to $s^3/2r$ at $(sr, 0)$.

If the test is similar on the boundary, i.e. unbiased, then the rejection probability shown in Table 2 should be equal to $\alpha$. This is almost the case when the absolute value of the curvature $|d_1|$ is not very large and the shape is sphere $s = 1$. The test becomes biased, or inaccurate, as $|d_1|$ and/or $s$ increase. This is because the asymptotic theory of Section 2 is valid only when the curvature as well as its change is not very large. $\hat{p}$ is not much better than $\hat{p}$, since only the effect of small $B$ is adjusted in Step 7 of the algorithm. The break-down of the asymptotic theory is indicated as high probabilities of significantly large $\chi^2$ values.
6 Concluding Remarks

The other method (Efron et al. 1996) to calculate \( \hat{p} \) uses the bootstrap samples from \( \hat{\mu}_0 \). This should be efficient in the sense that the number of bootstrap replicates needed to have the same standard deviance in \( \hat{p} \) is generally smaller than that of our method. However, the bootstrap samples from \( \hat{\mu}_0 \) can be difficult to obtain in some situations in practical applications; e.g., the block bootstrap for time-series analysis. On the other hand, changing the scale parameter is often easy to implement in a large class of problems.

Another advantage of our method is the diagnostic given by Step 6 of the algorithm to detect the break-down of the asymptotic theory.

Tests based on (5) are regarded as the likelihood ratio tests (LRT) adapted locally to the neighborhood of \( y \) or \( \hat{\mu}_0 \). The rejection region of the genuine LRT should be the form \( x_0 > h \), where \( h \) is a global constant determined by the distribution function of \( x_0 \) at the least favorable configuration. Our test as well as that of Efron et al. (1996) determines \( h \) locally depending on the curvature of the boundary using the statistics other than \( x_0 \); \( h \) will be the critical constant of the standard normal distribution when \( c = 0 \), while \( h \) becomes larger for \( c > 0 \) and \( h \) becomes smaller for \( c < 0 \).

Approximately unbiased tests have been discussed in the literature for the regions with nonsmooth boundaries such as cones. Perlman and Wu (1999) gave a good critical review on such “New Tests” for the concave regions, e.g., \( \theta(\mu) := \max\{\mu_1, \ldots, \mu_K\} \geq 0 \), while Shimodaira (2000) has discussed the convex region \( \theta(\mu) \leq 0 \). These new tests are less biased than LRT, but often have critical regions in intuitively unreasonable shape with “bumps.” However, the bumps of the rejection regions resemble those obtained by smoothing the boundaries of the null hypothesis and applying (5) to it. Thus the new tests may not be completely senseless. One is tempted to develop a theory to cover both smooth and nonsmooth boundaries.

The computer programs used in this paper are available from the author by request.
Acknowledgement

I would like to thank for Ziheng Yang, Satoshi Kuriki, Michael Perlman, and Brad Efron for helpful comments and suggestions. Most part of the research was done during my visit to Stanford University.

References


Figure 1: The points \((\sigma_i^{-1}, \tilde{Z}_i)\) are plotted for the two cases of the example. The left panel corresponds to the positive curvature and the right corresponds to the negative curvature. The signed distance and the curvature are estimated by fitting the curve (9) to the points.
Table 1: \( p \)-values calculated for \( K = 4, x_0 = 2 \) and several \( r \) values. \( \hat{d}_1 \) ranges from -0.5 to 1.5. The averages and the standard deviations of the corrected \( p \)-values are obtained from the simulation with hundred repetitions.

<table>
<thead>
<tr>
<th>( r )</th>
<th>( \hat{d}_1 )</th>
<th>( B = 1000 \times 16 )</th>
<th>( B = 10,000 \times 16 )</th>
<th>( B = \infty )</th>
</tr>
</thead>
<tbody>
<tr>
<td>-3</td>
<td>-0.5</td>
<td>0.0033</td>
<td>0.0032</td>
<td>0.0032</td>
</tr>
<tr>
<td>-5</td>
<td>-0.3</td>
<td>0.0084</td>
<td>0.0080</td>
<td>0.0080</td>
</tr>
<tr>
<td>-7.5</td>
<td>-0.2</td>
<td>0.0134</td>
<td>0.0128</td>
<td>0.0127</td>
</tr>
<tr>
<td>-15</td>
<td>-0.1</td>
<td>0.0176</td>
<td>0.0165</td>
<td>0.0176</td>
</tr>
<tr>
<td>( \infty )</td>
<td>0</td>
<td>0.0239</td>
<td>0.0224</td>
<td>0.0229</td>
</tr>
<tr>
<td>15</td>
<td>0.1</td>
<td>0.0307</td>
<td>0.0285</td>
<td>0.0286</td>
</tr>
<tr>
<td>7.5</td>
<td>0.2</td>
<td>0.0381</td>
<td>0.0353</td>
<td>0.0345</td>
</tr>
<tr>
<td>5</td>
<td>0.3</td>
<td>0.0465</td>
<td>0.0430</td>
<td>0.0416</td>
</tr>
<tr>
<td>3</td>
<td>0.5</td>
<td>0.0655</td>
<td>0.0600</td>
<td>0.0580</td>
</tr>
<tr>
<td>1.5</td>
<td>1</td>
<td>0.1248</td>
<td>0.1113</td>
<td>0.1104</td>
</tr>
<tr>
<td>1</td>
<td>1.5</td>
<td>0.1884</td>
<td>0.1734</td>
<td>0.1677</td>
</tr>
</tbody>
</table>
Table 2: The probability of rejecting the null hypothesis of oval shapes when \( \mu \) is on the boundary. \( \Pr\{\hat{p} < \alpha\} \) and \( \Pr\{\hat{p} < \alpha\} \) for \( \alpha = 0.05 \) are calculated from the simulation of 10,000 repetitions. The probability of significantly large \( \chi^2 \) \( (p < 0.01) \) in Step 6 of the algorithm is tabulated as well.

<table>
<thead>
<tr>
<th>( r )</th>
<th>( d_1 )</th>
<th>( \hat{p} )</th>
<th>( \hat{p} )</th>
<th>( \chi^2 )</th>
<th>( \hat{p} )</th>
<th>( \hat{p} )</th>
<th>( \chi^2 )</th>
<th>( \hat{p} )</th>
<th>( \hat{p} )</th>
<th>( \chi^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>-0.500</td>
<td>0.000</td>
<td>0.129</td>
<td>0.000</td>
<td>0.259</td>
<td>0.000</td>
<td>0.000</td>
<td>0.395</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-2</td>
<td>-0.250</td>
<td>0.055</td>
<td>0.054</td>
<td>0.023</td>
<td>0.000</td>
<td>0.248</td>
<td>0.000</td>
<td>0.000</td>
<td>0.380</td>
<td></td>
</tr>
<tr>
<td>-4</td>
<td>-0.125</td>
<td>0.050</td>
<td>0.051</td>
<td>0.009</td>
<td>0.011</td>
<td>0.137</td>
<td>0.000</td>
<td>0.000</td>
<td>0.310</td>
<td></td>
</tr>
<tr>
<td>-8</td>
<td>-0.062</td>
<td>0.046</td>
<td>0.047</td>
<td>0.010</td>
<td>0.062</td>
<td>0.063</td>
<td>0.030</td>
<td>0.002</td>
<td>0.002</td>
<td>0.221</td>
</tr>
<tr>
<td>-16</td>
<td>-0.031</td>
<td>0.047</td>
<td>0.049</td>
<td>0.011</td>
<td>0.054</td>
<td>0.055</td>
<td>0.011</td>
<td>0.071</td>
<td>0.071</td>
<td>0.079</td>
</tr>
<tr>
<td>16</td>
<td>0.031</td>
<td>0.048</td>
<td>0.050</td>
<td>0.010</td>
<td>0.047</td>
<td>0.049</td>
<td>0.011</td>
<td>0.044</td>
<td>0.046</td>
<td>0.080</td>
</tr>
<tr>
<td>8</td>
<td>0.062</td>
<td>0.049</td>
<td>0.050</td>
<td>0.009</td>
<td>0.040</td>
<td>0.041</td>
<td>0.029</td>
<td>0.033</td>
<td>0.036</td>
<td>0.221</td>
</tr>
<tr>
<td>4</td>
<td>0.125</td>
<td>0.046</td>
<td>0.048</td>
<td>0.010</td>
<td>0.043</td>
<td>0.045</td>
<td>0.132</td>
<td>0.034</td>
<td>0.037</td>
<td>0.316</td>
</tr>
<tr>
<td>2</td>
<td>0.250</td>
<td>0.043</td>
<td>0.046</td>
<td>0.021</td>
<td>0.039</td>
<td>0.043</td>
<td>0.248</td>
<td>0.039</td>
<td>0.047</td>
<td>0.376</td>
</tr>
<tr>
<td>1</td>
<td>0.500</td>
<td>0.029</td>
<td>0.033</td>
<td>0.125</td>
<td>0.040</td>
<td>0.047</td>
<td>0.256</td>
<td>0.044</td>
<td>0.054</td>
<td>0.392</td>
</tr>
</tbody>
</table>
```r
# R program

##
## bctr: bootstrap corrected test of regions
##
## ss: vector of scales
## bb: vector of the numbers of bootstrap replicates
## cc: vector of the numbers of "success"
##
## returns a list of
## p: hat p --- corrected p-value
## vp: the variance of it
## xc=c(x0,c): signed distance and the curvature
## used: logical vector to indicate used "cc"
##

bctr <- function(ss,bb,cc,quiet=F) {
  if(length(ss) != length(cc)) stop("cc and ss have different lengths")
  bb <- rep(bb,length.out=length(cc))
  pp <- cc/bb
  pp[is.na(pp)] <- NA
  used <- (pmin(pp,1-pp)>1e-6) & !is.na(pp)
  zz <- ss
  zz[used] <- -qnorm(pp[used])
  zz[!used] <- NA
  vv <- (1-pp)*pp/(dnorm(zz)^2*bb)

  if(!quiet & sum(used)<length(cc))
    if(!quiet) warning(paste(sum(used),"out of",
                           length(cc),"points are used"))

  df <- sum(used)-2

  if(df < 0) {
    p <- vp <- xc <- NA;
    return(p,vp,xc,df,used,ss,bb,cc,pp,zz,vv)
  }
}
```
X <- cbind(1/ss[used],ss[used])
fit <- lsfit(X,zz[used],1/vv[used],intercept=F)
xc <- fit$coef; names(xc) <- c("x0","c");

a <- c(1,-1); names(a) <- names(xc);
z <- drop(a %% xc) # hat Z
p <- 1-pnorm(z) # hat p

vxc <- solve(crossprod(X,X/vv[used]))
vz <- drop(a %% vxc %% a)
vp <- dnorm(z)^2 * vz

if(df > 0) {
  rss <- sum((fit$residuals)^2/vv[used])
pchi <- 1-pchisq(rss,df=df)
  if(!quiet & & pchi<0.01)
    warning(paste("model inconsistency with pvalue=" ,pchi))
} else {
  rss<-0; pchi<-1;
}

return(p,vp,xc,df,used,ss,bb,cc,pp,zz,vv,
       z,vz,vxc,rss,pchi)