A new goodness of fit test: the reversed Berk-Jones statistic

Leah Jager[1] and Jon A. Wellner[2]

University of Washington

January 23, 2004; revised July 22 and 29, 2005

Abstract

Owen inverted a goodness of fit statistic due to Berk and Jones to obtain confidence bands for a distribution function using Noé's recursion. As argued by Owen, the resulting bands are narrower in the tails and wider in the center than the classical Kolmogorov-Smirnov bands and have certain advantages related to the optimality theory connected with the test statistic proposed by Berk and Jones.

In this article we introduce a closely related statistic, the "reversed Berk-Jones statistic" which differs from the Berk and Jones statistic essentially because of the asymmetry of Kullback-Leibler information in its two arguments. We parallel the development of Owen for the new statistic, giving a method for constructing the confidence bands using the recursion formulas of Noé to compute rectangle probabilities for order statistics. Along the way we uncover some difficulties in Owen's calculations and give appropriate corrections. We also compare the exclusion probabilites (corresponding to the power of the tests) of our new bands with the (corrected version of) Owen's bands for a simple Lehmann type alternative considered by Owen and show that our bands are preferable over a certain range of alternatives.

[1] Research supported by the University of Washington VIGRE Grant, NSF DMS-9810726

[2] Research supported in part by National Science Foundation grant DMS-0203320, NIAID grant 2R01 AI291968-04

AMS 2000 *subject classifications.* 62G30, 62E15, 62G15

Key words and phrases. Berk-Jones statistic, confidence bands, coverage probability, empirical distribution, goodness of fit test, Noé's recursion, power

1 Introduction

Consider the classical goodness-of-fit testing problem: based on $X_1, \ldots, X_n$ i.i.d. F, test

$$H_0 : F(x) = F_0(x) \qquad \text{for all } x \in I\!R \tag{1}$$

versus

$$H_1 : F(x) \neq F_0(x) \qquad \text{for some } x \in I\!R \tag{2}$$

where F_0 is a fixed continuous distribution function.

Berk and Jones (1979) introduced the test statistic R_n, which is defined as

$$R_n = \sup_{-\infty < x < \infty} K(\mathbb{F}_n(x), F_0(x)), \tag{3}$$

where

$$K(x, y) = x \log \frac{x}{y} + (1 - x) \log \frac{1 - x}{1 - y}, \tag{4}$$

and $\mathbb{F}_n$ is the empirical distribution function of the X_i's, given by

$$\mathbb{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} 1_{[X_i \leq x]}. \tag{5}$$

The statistic R_n has several remarkable optimality properties, among which is the fact that it has greater power than any weighted Kolmogorov statistic. Wellner and Koltchinskii (2003) present a proof of the limiting null distribution of the Berk-Jones statistic, and Owen (1995) computes exact quantiles under the null distribution for finite n. Using these quantiles, Owen constructs confidence bands for F by inverting the Berk and Jones test, and then calculates the power associated with the Berk-Jones test statistic for fixed alternatives of the form $F(x) = F_0(x)^\alpha$.

An alternative test statistic, $\tilde{R}_n$, which we will call the *reversed Berk-Jones statistic*, is defined by

$$\tilde{R}_n = \sup_{X_{(1)} \leq x < X_{(n)}} K(F_0(x), \mathbb{F}_n(x)) \tag{6}$$

where $X_{(1)}$ and $X_{(n)}$ are the first and last order statistics, respectively.

The motivation behind this statistic comes from examination of the functions $K(F_0(x), F(x))$ and $K(F(x), F_0(x))$ for an alternative distribution function F. When F is stochastically smaller than F_0, we expect the Berk-Jones test to be more powerful than the reversed Berk-Jones test, since $\sup_x K(F(x), F_0(x)) > \sup_x K(F_0(x), F(x))$ in this case. However, in the case where F is stochastically larger than F_0, we have $\sup_x K(F(x), F_0(x)) < \sup_x K(F_0(x), F(x))$, and so we expect the reversed test statistic to be more powerful. This topic is more specifically addressed in section 4 of this paper, dealing with heuristics and calculations of power.

For the reversed statistic, $\tilde{R}_n$, we present an analysis similar to that in Owen (1995) for the Berk-Jones statistic. We calculate exact finite quantiles and the related confidence bands. Additionally, we present an analysis of the power of this reversed statistic when compared to the Berk-Jones statistic in the case of alternatives of the form $F(x) = F_0(x)^\alpha$, $\alpha \in (0, \infty)$. Finally, we compare our new proposed bands to those based on the Berk-Jones statistic for the galaxy data analyzed by Roeder (1990) and Owen (1995). Along the way we comment briefly on and give corrections to the development in Owen (1995).

2 Exact quantiles of the null distribution of $\tilde{R}_n$ for finite n

2.1 Exact null distribution of $\tilde{R}_2$

Proposition 1. Under the null hypothesis,

$$P(\tilde{R}_2 \leq x) = r_x^2, \qquad 0 \leq x \leq \log 2 \tag{7}$$

where $0 \leq r_x \leq 1$ is the unique solution of

$$(1 - r_x) \log (1 - r_x) + (1 + r_x) \log (1 + r_x) = 2x. \tag{8}$$

Proof. Without loss of generality we can take F_0 to be the uniform distribution on $[0, 1]$, $F_0(x) = x$, $0 \leq x \leq 1$. Note that

$$\begin{aligned} \tilde{R}_2 &= \sup_{X_{(1)} \leq x < X_{(2)}} K(x, \mathbb{F}_2(x)) \\ &= \max \{K(X_{(1)}, \frac{1}{2}), K(X_{(2)}, \frac{1}{2})\} \\ &= \max \{K(X_1, \frac{1}{2}), K(X_2, \frac{1}{2})\} \end{aligned}$$

where $X_1, X_2 \sim$ Uniform[0,1] are independent. Thus we calculate

$$P(K(U, \frac{1}{2}) \leq x) = P(l_x \leq U \leq u_x) \tag{9}$$

where $u_x \geq \frac{1}{2}$ solves $K(u_x, \frac{1}{2}) = x$ and $l_x \leq \frac{1}{2}$ solves $K(l_x, \frac{1}{2}) = x$. But since $K(\frac{1}{2} - t, \frac{1}{2}) = K(\frac{1}{2} + t, \frac{1}{2})$ for $0 \leq t \leq \frac{1}{2}$, it is clear that $u_x - \frac{1}{2} = \frac{1}{2} - l_x$, or $u_x = 1 - l_x$. Hence it follows that (9) equals

$$P(l_x \leq U \leq 1 - l_x) = P(\frac{1}{2} - v_x \leq U \leq \frac{1}{2} + v_x) = 2v_x \tag{10}$$

where $v_x = \frac{1}{2} - l_x \in (0, \frac{1}{2}]$ satisfies $K(\frac{1}{2} - v_x, \frac{1}{2}) = x$. But this means that $r_x = 2v_x$ satisfies (8). The conclusion follows since

$$\begin{aligned} P(\tilde{R}_2 \leq x) &= P(\max \{K(X_1, \frac{1}{2}), K(X_2, \frac{1}{2})\} \leq x) \\ &= P(K(U, \frac{1}{2}) \leq x)^2 = (2v_x)^2 = r_x^2. \end{aligned}$$

□

Knowing the exact distribution of $\tilde{R}_2$ allows us to calculate the exact quantiles for $n = 2$. We do this by solving

$$P(\tilde{R}_2 \leq \tilde{\lambda}_2^{1-\alpha}) = 1 - \alpha \tag{11}$$

for $\tilde{\lambda}_2^{1-\alpha}$ given a value of $1 - \alpha$. This yields

$$\tilde{\lambda}_2^{1-\alpha} = \frac{1}{2}((1 - \sqrt{1-\alpha}) \log (1 - \sqrt{1-\alpha}) + (1 + \sqrt{1-\alpha} \log (1 + \sqrt{1-\alpha})). \tag{12}$$

The 0.95 quantile is then $\tilde{\lambda}_2^{0.95} = 0.625251$. The 0.99 quantile is $\tilde{\lambda}_2^{0.99} = 0.675634$.

2.2 Quantiles of the null distribution of $\tilde{R}_n$ for $n > 2$

Owen (1995) computed exact quantiles of the Berk-Jones statistic under the null distribution for finite n using a recursion of Noé (1972). Using an analagous method, we compute exact quantiles of the reversed statistic using this recursion.

We want to calculate $\tilde{\lambda}_n^{1-\alpha}$ such that $P(\tilde{R}_n \leq \tilde{\lambda}_n^{1-\alpha}) = 1 - \alpha$. With this $\tilde{\lambda}_n^{1-\alpha}$ we can form $1 - \alpha$ confidence bands for F by finding $\tilde{L}_n(x)$ and $\tilde{H}_n(x)$ (depending on the data) such that $P(\tilde{R}_n \leq \tilde{\lambda}_n^{1-\alpha}) = P(\tilde{L}_n(x) \leq F(x) \leq \tilde{H}_n(x),\ x \in \mathbb{R})$. We can rewrite this probability in terms of the order statistics, and then use the recursions due to Noé (1972) to compute it.

Our procedure is as follows. We want to find $\tilde{\lambda}_n^{1-\alpha}$ for a given confidence interval corresponding to $1 - \alpha$. Given this $\tilde{\lambda}_n^{1-\alpha}$, we calculate a confidence band of the form $\{\tilde{a}_i, i = 1, \cdots, n\}$ and $\{\tilde{b}_i, i = 1, 2, \cdots, n\}$ such that $P(\tilde{R}_n \leq \tilde{\lambda}_n^{1-\alpha}) = P(\tilde{a}_i < X_{(i)} \leq \tilde{b}_i, i = 1, 2, \cdots, n)$.

To see how to calculate $\{\tilde{a}_i\}$ and $\{\tilde{b}_i\}$, we look at the reversed statistic, $\tilde{R}_n$, itself. Similarly to the $n = 2$ case above, we can separate the event $[\tilde{R}_n \leq \tilde{\lambda}_n]$ into parts associated with each order statistic. Now

$$\begin{aligned} \tilde{R}_n &= \sup_{X_{(1)} \leq x < X_{(n)}} K(x, \mathbb{F}_n(x)) \\ &= K(X_{(1)}, \frac{1}{n}) \vee \max_{2 \leq i \leq n-1} \{K(X_{(i)}, \frac{i-1}{n}), K(X_{(i)}, \frac{i}{n})\} \vee K(X_{(n)}, \frac{n-1}{n}). \end{aligned}$$

So the event $[\tilde{R}_n \leq \tilde{\lambda}_n]$ is equivalent to the intersection of the events

$$K(X_{(1)}, \frac{1}{n}) \leq \tilde{\lambda}_n, \tag{13}$$

$$\max\{K(X_{(i)}, \frac{i-1}{n}), K(X_{(i)}, \frac{i}{n})\} \leq \tilde{\lambda}_n, \quad 2 \leq i \leq n-1, \tag{14}$$

$$K(X_{(n)}, \frac{n-1}{n}) \leq \tilde{\lambda}_n. \tag{15}$$

Here we have managed to divide the event into smaller events relating to each order statistic separately.

In order to compute the finite sample quantiles, we are looking for $\{\tilde{a}_i\}_{i=1}^n$ and $\{\tilde{b}_i\}_{i=1}^n$ such that $P(\tilde{R}_n \leq \tilde{\lambda}_n) = P(\tilde{a}_i < X_{(i)} \leq \tilde{b}_i, 1 \leq i \leq n)$. (Note that we have deliberately chosen somewhat different notation from Owen (1995).) Splitting the event $[\tilde{R}_n \leq \tilde{\lambda}_n]$ into events (13), (14), and (15), we can define $\{\tilde{a}_i\}$ and $\{\tilde{b}_i\}$ in terms of these smaller events.

From (13), $K(X_{(1)}, \frac{1}{n}) \leq \tilde{\lambda}_n$, we obtain bounds on $X_{(1)}$, and we see that

$$\tilde{b}_1 = \max\{x | K(x, \frac{1}{n}) \leq \tilde{\lambda}_n\}, \tag{16}$$

$$\tilde{a}_1 = \min\{x | K(x, \frac{1}{n}) \leq \tilde{\lambda}_n\}. \tag{17}$$

Similarly, from (15) we find that bounds on $X_{(n)}$ are given by

$$\begin{aligned} \tilde{b}_n &= \max\{x | K(x, \frac{n-1}{n}) \leq \tilde{\lambda}_n\}, \\ \tilde{a}_n &= \min\{x | K(x, \frac{n-1}{n}) \leq \tilde{\lambda}_n\}. \end{aligned}$$

Finally, the event (14) yields that for $2 \le i \le n-1$,

$$\begin{aligned}
\tilde{b}_i &= \max\{x|\max\{K(x,\frac{i-1}{n}),K(x,\frac{i}{n})\}\le\tilde{\lambda}_n\} \\
&= \max\{x|K(x,\frac{i-1}{n})\le\tilde{\lambda}_n, K(x,\frac{i}{n})\le\tilde{\lambda}_n\}\,, \qquad (18) \\
\tilde{a}_i &= \min\{x|\max\{K(x,\frac{i-1}{n}),K(x,\frac{i}{n})\}\le\tilde{\lambda}_n\} \\
&= \min\{x|K(x,\frac{i-1}{n})\le\tilde{\lambda}_n, K(x,\frac{i}{n})\le\tilde{\lambda}_n\}\,. \qquad (19)
\end{aligned}$$

The above equations for $2 \le i \le n-1$ are not as easy to deal with as those for the cases where $i = 1$ and $i = n$. However, by noticing the relationship between $K(x, \frac{i-1}{n})$ and $K(x, \frac{i}{n})$, we can simplify further. To do this we use the following claim.

Claim 1. Let $\tilde{\lambda}_n > 0$. Then for any fixed y_1 and y_2 such that $0 < y_1 < y_2 < 1$,

$$\begin{aligned}
&\text{(i)} \quad \max\{x|K(x,y_1)\le\tilde{\lambda}_n, K(x,y_2)\le\tilde{\lambda}_n\} = \max\{x|K(x,y_1)\le\tilde{\lambda}_n\}\,, \\
&\text{(ii)} \quad \min\{x|K(x,y_1)\le\tilde{\lambda}_n, K(x,y_2)\le\tilde{\lambda}_n\} = \min\{x|K(x,y_2)\le\tilde{\lambda}_n\},
\end{aligned}$$

provided $\{x|K(x,y_1)\le\tilde{\lambda}_n, K(x,y_2)\le\tilde{\lambda}_n\}$ is not empty.

Proof. (i) First, note that $\frac{\partial}{\partial x}K(x,y) = \log\frac{x}{y} - \log\frac{1-x}{1-y}$. So K decreases in x on the interval $[0, y)$, has a minimum of 0 at $x = y$, and increases in x on the interval $(y, 1]$. This means that $\max\{x|K(x,y)\le\tilde{\lambda}_n\}$ will occur on the interval $(y, 1]$.

Now fix y_1 and y_2 such that $y_1 < y_2$. Then $K(x, y_1)$ and $K(x, y_2)$ will have a point of intersection at c in the interval (y_1, y_2). That is, $K(c, y_1) = K(c, y_2)$. Now $K(x, y_1)$ is increasing in x on the interval $(c, y_2]$, while $K(x, y_2)$ is decreasing in x on this same interval. Thus $K(x, y_2) < K(x, y_1)$ on $(c, y_2]$. But $\frac{\partial}{\partial x}K(x,y_1) > \frac{\partial}{\partial x}K(x,y_2)$ for all x. So $K(x, y_2) < K(x, y_1)$ for all x in $(c, 1]$.

There are three cases to consider. First, suppose $\tilde{\lambda}_n > K(c, y_1)$, where again, c is the point of intersection. Then $\max\{x|K(x,y_1)\le\tilde{\lambda}_n, K(x,y_2)\le\tilde{\lambda}_n\} > c$. Since we have shown that $K(x, y_2) < K(x, y_1)$ for all $x > c$, the maximum x value where both $K(x, y_1) \le \tilde{\lambda}_n$ and $K(x, y_2) \le \tilde{\lambda}_n$ is the same as the maximum x for which $K(x, y_1) \le \tilde{\lambda}_n$. So $\max\{x|K(x,y_1)\le\tilde{\lambda}_n, K(x,y_2)\le\tilde{\lambda}_n\} = \max\{x|K(x,y_1)\le\tilde{\lambda}_n\}$.

Now suppose $\tilde{\lambda}_n = K(c, y_1)$. Then $\max\{x|K(x,y_1)\le\tilde{\lambda}_n, K(x,y_2)\le\tilde{\lambda}_n\} = c = \max\{x|K(x,y_1)\le\tilde{\lambda}_n\}$.

Finally, suppose $\tilde{\lambda}_n < K(c, y_1)$. Then there is no x in the interval $[0, 1]$ that satisfies $\max\{x|K(x,y_1)\le\tilde{\lambda}_n, K(x,y_2)<\tilde{\lambda}_n\}$, since $K(x, y_1) > \tilde{\lambda}_n$ for $x \ge c$ and $K(x, y_2) > \tilde{\lambda}_n$ for $x \le c$.

(ii) The result for the minimum is proved in a similar way.

□

The above result allows (18) and (19) to be written simply (for $2 \le i \le n-1$) as

$$\tilde{b}_i = \max\{x|K(x,\frac{i-1}{n})\le\tilde{\lambda}_n\}, \qquad (20)$$

$$\tilde{a}_i = \min\{x|K(x,\frac{i}{n})\le\tilde{\lambda}_n\}. \qquad (21)$$

We can further simplify the calculation by noticing that $\tilde{a}_i = 1 - \tilde{b}_{n-i+1}$ for $1 \leq i \leq n$. This is shown in the following two claims.

Claim 2. Let $\lambda > 0$. Then for any fixed y,

$$\max\{x|K(x,y) \leq \tilde{\lambda}\} = 1 - \min\{x|K(x,1-y) \leq \tilde{\lambda}\} \tag{22}$$

Proof. Fix y. Now $\frac{\partial}{\partial x}K(x,y) = \log\frac{x}{y} - \log\frac{1-x}{1-y}$. So K decreases in x on the interval $[0,y)$, has a minimum at $x = y$, and increases in x on the interval $(y,1]$. This means that $\max\{x|K(x,y) \leq \tilde{\lambda}\}$ will occur on the interval $(y,1)$, while $\min\{x|K(x,y) \leq \tilde{\lambda}\}$ will occur on the interval $(0,y)$. Also, notice that for all x and y, $K(x,y) = K(1-x,1-y)$. We break this proof into two cases.

First, suppose there exists a $x^* > y$ such that $K(x^*,y) = \tilde{\lambda}$. Then $x^* = \max\{x|K(x,y) \leq \tilde{\lambda}\}$. But $K(1-x^*,1-y) = \tilde{\lambda}$ as well. Now $1-x^*$ is in the interval $[0,1-y)$. So $1-x^* = \min\{x|K(x,1-y) \leq \tilde{\lambda}\}$. Thus the result is proved.

Now, suppose there does not exist a $x^* > y$ such that $K(x^*,y) = \tilde{\lambda}$. Since $x \mapsto K(x,y)$ is continuous and increasing on $(y,1]$, this means that $K(x,y) < \tilde{\lambda}$ for all $x > y$. So $\max\{x|K(x,y) \leq \tilde{\lambda}\} = 1$, since this is the maximum value of x that is possible. Now, since $K(x,y) < \tilde{\lambda}$ for all $x > y$, we also have $K(1-x,1-y) < \tilde{\lambda}$ for all $1-x < 1-y$. So $\min\{x|K(x,1-y) \leq \tilde{\lambda}\} = 0$. Again, the result is proved.

□

Claim 3. For $1 \leq i \leq n$,

$$\tilde{a}_i = 1 - \tilde{b}_{n-i+1}. \tag{23}$$

Proof. From (20) and (21) it follows that, for $i \in \{2,\ldots,n-1\}$,

$$\begin{aligned}
\tilde{a}_i &= \min\{x|K(x,\frac{i}{n}) \leq \tilde{\lambda}_n\} \\
&= 1 - \max\{x|K(x,1-\frac{i}{n}) \leq \tilde{\lambda}_n\} \qquad \text{by Claim 2} \\
&= 1 - \max\{x|K(x,\frac{n-i}{n}) \leq \tilde{\lambda}_n\} \\
&= 1 - \tilde{b}_{n-i+1}.
\end{aligned}$$

The cases $i = 1$ and $i = n$ are trivial.

□

Now that we have defined $\{\tilde{b}_i\}$ for all i, and thus defined $\{\tilde{a}_i\}$, we can calculate $P(\tilde{R}_n \leq \tilde{\lambda}_n) = P(\tilde{a}_i < X_{(i)} \leq \tilde{b}_i,\ 1 \leq i \leq n)$ by a recursion due to Noé (1972). The computing process involved in finding the values of $\{\tilde{a}_i\}$, $\{\tilde{b}_i\}$, and $\tilde{\lambda}_n$ follows the method outlined in Owen (1995), using the Van Wijngaarden-Decker-Brent method to first find the $\{\tilde{a}_i\}$ and $\{\tilde{b}_i\}$ corresponding to a particular $\tilde{\lambda}_n$ and then reapplying the same Van Wijngaarden-Decker-Brent method again to solve for the $\tilde{\lambda}_n^{1-\alpha}$ associated with the $1-\alpha$ quantile.

There is, however, one slight complication in these calculations compared to the method outlined in Owen (1995). When calculating confidence bands by inverting the Berk-Jones statistic, we look at $K(x,y)$ as a function of y for a fixed x. For each $x \in (0,1)$, $K(x,y)$ is a continuous function of y with a minimum of 0 at $y = x$ that tends to ∞ as $y \to 0$ or $y \to 1$. Therefore for any $\lambda > 0$ there exists a y^* such that $K(x, y^*) = \tilde{\lambda}$.

For the reversed statistic, however, we look at $K(x,y)$ as a function of x for a fixed y. Again, $K(x,y)$ is a continuous function of x with a minimum of 0 at $x = y$. But $K(0,y) = \log \frac{1}{1-y} < \infty$ and $K(1,y) = \log \frac{1}{y} < \infty$. So we are not guaranteed that for each $\tilde{\lambda} > 0$ there exists an x^* such that $K(x^*, y) = \tilde{\lambda}$. Thus care must be taken when looking at $\tilde{b}_i$ as defined in (20). If there is no x^* satisfying $K(x^*, \frac{i-1}{n}) = \tilde{\lambda}_n$ then $\tilde{b}_i = 1$ necessarily.

Now we determine the confidence bands

$$\tilde{L}_n(x) = \sum_{i=0}^{n} \tilde{l}_i 1_{(X_{(i)}, X_{(i+1)}]}(x)\,,$$

and

$$\tilde{H}_n(x) = \sum_{i=0}^{n} \tilde{h}_i 1_{[X_{(i)}, X_{(i+1)})}(x)\,,$$

where $X_{(0)} \equiv -\infty$ and $X_{(n+1)} = \infty$ by convention. For $x \in (X_{(i)}, X_{(i+1)})$ we have $\mathbb{F}_n(x) = i/n$ and it is clear that the event $[\tilde{R}_n \le \tilde{\lambda}_n^{1-\alpha}]$ restricts $F(x)$ only by $K(F(x), i/n) \le \tilde{\lambda}_n^{1-\alpha}$. Hence

$$\tilde{h}_i = \max\{p | K(p, i/n) \le \tilde{\lambda}_n^{1-\alpha}\}\,,$$

while

$$\tilde{l}_i = \min\{p | K(p, i/n) \le \tilde{\lambda}_n^{1-\alpha}\}\,.$$

From (16) and (20) it follows that $\tilde{h}_i = \tilde{b}_{i+1}$, $i \in \{0, \ldots, n-1\}$, and from (17) and (21) we have $\tilde{l}_i = a_i$, $i \in \{1, \ldots, n\}$. Furthermore $\tilde{h}_n = 1$ and $\tilde{l}_0 = 0$ (trivially). (Note that the statistic $\tilde{R}_n$ gives no constraints on the values of F on the corresponding sets since the supremum is only taken over $X_{(1)} \le x < X_{(n)}$.)

Although we have experimented with several different approximations for the 0.95 and 0.99 quantiles of the reversed statistic analogous to the formulas given by Owen (1995) for the Berk-Jones statistic itself (and with the corrections developed here in the next section), none of our attempts so far have been sufficiently accurate over the range $10 < n \le 1000$ to recommend in practice. Our current recommendation is to use the exact values as computed via Noé's recursion in the programs (which are available at the second author's web site).

3 Some comments on Owen's quantiles for the Berk-Jones statistic

We can apply the same approach detailed above to the problem of finding the exact quantiles of for the Berk-Jones statistic, as Owen (1995) does.

3.1 Exact null distribution of R_1

Proposition 2. Under the null hypothesis,

$$P(R_1 \le x) = (1 - 2e^{-x}) 1_{[\log 2, \infty)}(x). \qquad (24)$$

Proof. Note that

$$\begin{aligned} R_n &= \sup_{0 \le x \le 1} K(\mathbb{F}_n(x), x) \\ &= \max_{1 \le i \le n} \{K(\frac{i-1}{n}, X_{(i)}) \vee K(\frac{i}{n}, X_{(i)})\}. \end{aligned}$$

Thus for $n = 1$ we have (interpreting $0 \log 0 = 0$)

$$R_1 = \log \frac{1}{1 - X_1} \vee \log \frac{1}{X_1}. \tag{25}$$

It follows that

$$\begin{aligned} P(R_1 \le x) &= P(\log \frac{1}{1 - X_1} \vee \log \frac{1}{X_1} \le x) \\ &= P(e^{-x} \le X_1 \le 1 - e^{-x}) \\ &= (1 - 2e^{-x}) 1_{[\log 2, \infty)}(x). \end{aligned}$$

□

Knowing the exact distribution of R_1 allows us to calculate the exact quantiles for $n = 1$. We do this by solving

$$P(R_1 \le \lambda_1^{1-\alpha}) = 1 - \alpha \tag{26}$$

for $\lambda_1^{1-\alpha}$ given a value of $1 - \alpha$. This implies that the $1 - \alpha$ quantile $\lambda_1^{1-\alpha}$ of R_1 is given by $\lambda_1^{1-\alpha} = -\log \frac{\alpha}{2}$; in particular $\lambda_1^{0.95} = 3.68888$ and $\lambda_1^{0.99} = 5.29832$. [Owen (1995), page 518, second column, claims both that $\lambda_1^{1-\alpha} = -\log(1 - \alpha)$, and $\lambda_1^{0.95} = -\log(1 - 0.95) = 2.9957$.]

3.2 Quantiles of the null distribution of R_n for $n > 1$

Now we look at the problem of finding the $1 - \alpha$ quantile of R_n. Again, we want to find λ_n such that $P(R_n \le \lambda_n) = 1 - \alpha$. If we break down the problem as we did for the reversed statistic, we have that

$$\begin{aligned} R_n &= \sup_{0 \le x \le 1} K(\mathbb{F}_n(x), x) \\ &= \max_{1 \le i \le n} \max\{K(\frac{i-1}{n}, X_{(i)}), K(\frac{i}{n}, X_{(i)})\} \\ &= \max\{\log \frac{1}{1 - X_{(1)}}, K(\frac{1}{n}, X_{(1)})\} \vee \max_{2 \le i \le n-1} \{K(\frac{i-1}{n}, X_{(i)}), K(\frac{i}{n}, X_{(i)})\} \\ &\qquad \vee \max\{K(\frac{n-1}{n}, X_{(n)}), \log \frac{1}{X_{(n)}}\}. \end{aligned}$$

Now breaking the event $[R_n \le \lambda_n]$ into separate events involving each of the order statistics separately gives us that the event $[R_n \le \lambda_n]$ is equivalent to the intersection of the events

$$\max\{\log \frac{1}{1 - X_{(1)}}, K(\frac{1}{n}, X_{(1)})\} \le \lambda_n\,, \tag{27}$$

$$\max\{K(\frac{i-1}{n}, X_{(i)}), K(\frac{i}{n}, X_{(i)})\} \le \lambda_n, \quad 2 \le i \le n-1, \tag{28}$$

$$\max\{K(\frac{n-1}{n}, X_{(n)}), \log \frac{1}{X_{(n)}}\} \le \lambda_n. \tag{29}$$

Again we are looking for numbers $\{a_i\}_{i=1}^n$ and $\{b_i\}_{i=1}^n$ (which depend also on n and λ_n, dependence suppressed in the notation) such that

$$P(R_n \le \lambda_n) = P(a_i < X_{(i)} \le b_i,\ 1 \le i \le n)\,.$$

Splitting the event $[R_n \le \lambda_n]$ into events (27), (28), and (29), we can define $\{a_i\}$ and $\{b_i\}$ in terms of these smaller events, as we did in the case of the reversed statistic.

From (27), we see that

$$\begin{aligned}
b_1 &= \max\{x|\max\{\log\frac{1}{1-x}, K(\frac{1}{n}, x)\} \le \lambda_n\} \\
&= \max\{x|\log\frac{1}{1-x} \le \lambda_n, K(\frac{1}{n}, x) \le \lambda_n\}, \\
a_1 &= \min\{x|\max\{\log\frac{1}{1-x}, K(\frac{1}{n}, x)\} \le \lambda_n\} \\
&= \min\{x|\log\frac{1}{1-x} \le \lambda_n, K(\frac{1}{n}, x) \le \lambda_n\}.
\end{aligned}$$

Similarly, because of (29), we have

$$\begin{aligned}
b_n &= \max\{x|\max\{K(\frac{n-1}{n}, x), \log\frac{1}{x}\} \le \lambda_n\} \\
&= \max\{x|K(\frac{n-1}{n}, x) \le \lambda_n, \log\frac{1}{x} \le \lambda_n\}, \\
a_n &= \min\{x|\max\{K(\frac{n-1}{n}, x), \log\frac{1}{x}\} \le \lambda_n\} \\
&= \min\{x|K(\frac{n-1}{n}, x) \le \lambda_n, \log\frac{1}{x} \le \lambda_n\}.
\end{aligned}$$

Finally, event (28) gives us that for $2 \le i \le n-1$,

$$\begin{aligned}
b_i &= \max\{x|\max\{K(\frac{i-1}{n}, x), K(\frac{i}{n}, x)\} \le \lambda_n\} \\
&= \max\{x|K(\frac{i-1}{n}, x) \le \lambda_n, K(\frac{i}{n}, x) \le \lambda_n\} && (30) \\
a_i &= \min\{x|\max\{K(\frac{i-1}{n}, x), K(\frac{i}{n}, x)\} \le \lambda_n\} \\
&= \min\{x|K(\frac{i-1}{n}, x) \le \lambda_n, K(\frac{i}{n}, x) \le \lambda_n\}. && (31)
\end{aligned}$$

Again, we can simplify these expressions for a_i and b_i by noticing a few things. We begin with the following claim.

Claim 4. Let $\lambda_n > 0$. Then for any fixed x_1 and x_2 such that $0 < x_1 < x_2 < 1$,

$$\begin{aligned}
&\text{(i)} \quad \max\{y|K(x_1, y) \le \lambda_n, K(x_2, y) \le \lambda_n\} = \max\{y|K(x_1, y) \le \lambda_n\}\,, \\
&\text{(ii)} \quad \min\{y|K(x_1, y) \le \lambda_n, K(x_2, y) \le \lambda_n\} = \min\{y|K(x_2, y) \le \lambda_n\},
\end{aligned}$$

provided $\{y|K(x_1, y) \le \lambda_n, K(x_2, y) \le \lambda_n\}$ is not empty.

Proof. (i) First, note that $\frac{\partial}{\partial y}K(x,y) = \frac{y-x}{y(1-y)}$. So K decreases in y on the interval $[0,x)$, has a minimum of 0 at $y=x$, and increases in y on the interval $(x,1]$. This means that $\max\{y|K(x,y) \leq \lambda_n\}$ will occur on the interval $(x,1]$.

Now fix x_1 and x_2 such that $x_1 < x_2$. Then $K(x_1,y)$ and $K(x_2,y)$ will have a point of intersection at c in the interval (x_1,x_2). That is, $K(x_1,c) = K(x_2,c)$. Now $K(x_1,y)$ is increasing in y on the interval $(c,x_2]$, while $K(x_2,y)$ is decreasing in y on this same interval. So $K(x_2,y) < K(x_1,y)$ on $(c,x_2]$. But $\frac{\partial}{\partial y}K(x_1,y) > \frac{\partial}{\partial y}K(x_2,y)$ for all y. So $K(x_2,y) < K(x_1,y)$ for all y in $(c,1]$.

There are three cases to consider. First, suppose $\lambda_n > K(x_1,c)$, where again, c is the point of intersection. Then $\max\{y|K(x_1,y) \leq \lambda_n, K(x_2,y) \leq \lambda_n\} > c$. Since we have shown that $K(x_2,y) < K(x_1,y)$ for all $y > c$, the maximum y value where both $K(x_1,y) \leq \lambda_n$ and $K(x_2,y) \leq \lambda_n$ is the same as the maximum y for which $K(x_1,y) \leq \lambda_n$. So $\max\{y|K(x_1,y) \leq \lambda_n, K(x_2,y) \leq \lambda_n\} = \max\{y|K(x_1,y) \leq \lambda_n\}$.

Now suppose $\lambda_n = K(x_1,c)$. Then $\max\{y|K(x_1,y) \leq \lambda_n, K(x_2,y) \leq \lambda_n\} = c = \max\{y|K(x_1,y) \leq \lambda_n\}$.

Finally, suppose $\lambda_n < K(x_1,y)$. Then there is no y in the interval $[0,1]$ that satisfies $\max\{y|K(x_1,y) \leq \lambda_n, K(x_2,y) \leq \lambda_n\}$, since $K(x_1,y) > \lambda_n$ for $y \geq c$ and $K(x_2,y) > \lambda_n$ for $y \leq c$.

(ii) The result for the minimum is proved in a similar way.

$\square$

This claim allows us to simplify (30) and (31) to be (for $2 \leq i \leq n-1$)

$$b_i = \max\{x|K(\frac{i-1}{n},x) \leq \lambda_n\}\,, \tag{32}$$

$$a_i = \min\{x|K(\frac{i}{n},x) \leq \lambda_n\}. \tag{33}$$

We can also simplify the cases where $i=1$ and $i=n$. These cases can be written as

$$b_1 = \max\{x|\log\frac{1}{1-x} \leq \lambda_n\} = 1 - e^{-\lambda_n}\,, \tag{34}$$

$$a_1 = \min\{x|K(\frac{1}{n},x) \leq \lambda_n\}\,, \tag{35}$$

$$b_n = \max\{x|K(\frac{n-1}{n},x) \leq \lambda_n\}\,, \tag{36}$$

$$a_n = \min\{x|\log\frac{1}{x} \leq \lambda_n\} = e^{-\lambda_n}. \tag{37}$$

The rationale for this is as follows. First the $i=1$ case. Notice that $\log\frac{1}{1-y}$ has value 0 at $y=0$ and then is increasing to ∞ over the interval $(0,1)$. And $K(x,y)$ decreases in y on $(0,x)$, has a minimum value of 0 at $x=y$, and increases in y on $(x,1)$. Since both functions are continuous, this means there is a point of intersection, c, on the interval $(0,x)$. Since $\log\frac{1}{1-y}$ is increasing on $(0,c)$ and $K(x,y)$ is decreasing on this interval, $\log\frac{1}{1-y} < K(x,y)$ on this interval. This gives the result for a_1. But $\frac{\partial}{\partial y}\log\frac{1}{1-y} = \frac{1}{1-y} > \frac{1}{1-y}(1-\frac{x}{y}) = \frac{\partial}{\partial y}K(x,y)$ for all y. So $\log\frac{1}{1-y} > K(x,y)$ on the interval $(c,1)$. This gives the result for b_1.

The case for $i = n$ is similar, except that $\log \frac{1}{y}$ decreases from ∞ to 0 over the interval $(0, 1]$ rather than increasing from 0 to ∞ over the interval $[0, 1)$, as in the $\log \frac{1}{1-y}$ case.

Finally, as in the case of the reversed statistic described above, we see that we can once again define the $\{a_i\}$ in terms of the $\{b_i\}$ as $a_i = 1 - b_{n-i+1}$ for $1 \leq i \leq n$. The following two claims (analogous to claims 2 and 3 in the case of the reversed statistic) show this.

Claim 5. Let $\lambda > 0$. Then for any fixed x,

$$\max\{y|K(x, y) \leq \lambda\} = 1 - \min\{y|K(1 - x, y) \leq \lambda\} \tag{38}$$

Proof. Fix x. Now $\frac{\partial}{\partial y} K(x, y) = \frac{y-x}{y(1-y)}$. So K decreases in y on the interval $[0, x)$, has a minimum of 0 at $y = x$, and increases in y on the interval $(x, 1]$. This means that $\max\{y|K(x, y) \leq \lambda\}$ will occur on the interval $(x, 1]$, while $\min\{y|K(x, y) \leq \lambda\}$ will occur on the interval $[0, x)$. Also, notice that for all x and y, $K(x, y) = K(1 - x, 1 - y)$.

Now, since $K(x, y) \to \infty$ as $x \to 1$ for a fixed y, there exists a y^* in $(x, 1]$ such that $K(x, y^*) = \lambda$. So $y^* = \max\{y|K(x, y) \leq \lambda\}$. But $K(1 - x, 1 - y^*) = \lambda$ as well. Now $1 - y^*$ is in the interval $[0, 1 - x)$. So $1 - y^* = \min\{y|K(1 - x, y) \leq \lambda\}$. Thus the result is proved.

□

Claim 6. For $1 \leq i \leq n$,

$$a_i = 1 - b_{n-i+1}. \tag{39}$$

Proof. From (32) and (33) we have that for $2 \leq i \leq n - 1$

$$\begin{aligned}
a_i &= \min\{x|K(\frac{i}{n}, x) \leq \lambda_n\} \\
&= 1 - \max\{x|K(1 - \frac{i}{n}, x) \leq \lambda_n\} \qquad \text{by Claim 5} \\
&= 1 - \max\{x|K(\frac{n-i}{n}, x) \leq \lambda_n\} \\
&= 1 - b_{n-i+1}.
\end{aligned}$$

The cases for $i = 1$ and $i = n$ are trivial.

□

From here we can calculate $P(R_n \leq \lambda_n) = P(a_i < X_{(i)} \leq b_i,\ 1 \leq i \leq n)$ by the same recursion due to Noé, computing the actual values of $\{a_i\}$, $\{b_i\}$, and λ_n following Owen's algorithm as we did for the reversed statistic.

This is where we notice a difference from Owen. Owen (1995), page 517, finds his $1 - \alpha$ confidence bands based on the Berk-Jones statistic by taking

$$L_n(x) \equiv L(x) = \min\{p|K(\mathbb{F}_n(x), p) \leq \lambda_n\} \tag{40}$$

and

$$H_n(x) \equiv H(x) = \max\{p|K(\mathbb{F}_n(x), p) \leq \lambda_n\}. \tag{41}$$

Because L and H are step functions, however, they need only to be calculated at the order statistics themselves. That is, we only need to find $\{L_i\}$ and $\{H_i\}$ such that $P(R_n \leq \lambda_n) =$

$P(L_{i-1} < X_{(i)} \leq H_i, i = 1, 2, \cdots, n)$. As defined by Owen, in the two displays below his formula (7), page 517, $\{H_i\}$ and $\{L_i\}$ become

$$\begin{aligned} H_i &= \max\{x | K(\frac{i}{n}, x) \leq \lambda_n\}, \quad 1 \leq i \leq n-1, \\ H_n &= 1\,, \\ L_i &= \min\{x | K(\frac{i}{n}, x) \leq \lambda_n\}, \quad 1 \leq i \leq n-1, \\ L_0 &= 0\,, \end{aligned}$$

and then Owen (1995) *claims* in his formula (9) page 518 that these are linked to the event $\{R_n \leq \lambda_n\}$ by

$$\{R_n \leq \lambda_n\} = \{L_{i-1} \leq X_{(i)} \leq H_i : \quad i = 1, \ldots, n\}\,. \tag{42}$$

But in fact, (42) is false. Note that the event on the right side of (42) implies that

$$R_n \geq K((n-1)/n, H_n - \epsilon) = K((n-1)/n, 1 - \epsilon) \to \infty$$

as $\epsilon \downarrow 0$. Thus

$$\{R_n \leq \lambda_n\} \subset \{L_{i-1} \leq X_{(i)} \leq H_i : \quad i = 1, \ldots, n\} \tag{43}$$

with strict inclusion since the event on the right side allows $R_n = \infty$. Moreover, note that $R_n = R_n(X_1, \ldots, X_n) > \lambda_n^{.95}$ also at the points $X_{(i)} = H_i$ with $i < n$ since $K((i-1)/n, H_i) > \lambda_n^{.95}$.

Our definition of $\{b_i\}$ is

$$\begin{aligned} b_1 &= \max\{x | \log \frac{1}{1-x} \leq \lambda_n\} = 1 - e^{-\lambda_n}\,, \\ b_i &= \max\{x | K(\frac{i-1}{n}, x) \leq \lambda_n\}, \quad 2 \leq i \leq n\,. \end{aligned}$$

Note that Owen's H_i's involve the maximum x such that $K(\frac{i}{n}, x) \leq \lambda_n$, while our b_i's are defined as the maximum x such that $K(\frac{i-1}{n}, x) \leq \lambda_n$. Thus it follows that $b_i = H_{i-1}$ for $i = 2, \ldots, n-1$, and similarly $a_i = 1 - b_{n-i+1} = 1 - H_{n-i} = L_i$, $i = 2, \ldots, n-1$ by virtue of Owen's relation $L_i = 1 - H_{n-i}$. Thus we claim the correct event identity is:

$$\{R_n \leq \lambda_n\} = \{a_i < X_{(i)} \leq b_i, \ i = 1, \ldots, n\} \tag{44}$$

where $a_i = L_i$, $i = 1, \ldots, n-1$, $a_n = e^{-\lambda_n}$, $b_i = H_{i-1}$, $i = 2, \ldots, n$, $b_1 = 1 - e^{-\lambda_n}$ The following table illustrates the situation numerically for $n = 4$.

Table 1: Numerical illustration of order statistic bounds, $n = 4$

Owen's claim $\lambda_4^{.95} = .914983$		Jager-Wellner $\lambda_4^{.95} = 1.092493$	
Owen bounds	J-W bounds	Owen bounds	J-W bounds
$L_0 = 0$	$a_1 = .002737$	$L_0 = 0$	$a_1 = .001340$
$L_1 = .002737$	$a_2 = .041857$	$L_1 = .001340$	$a_2 = .028958$
$L_2 = .041857$	$a_3 = .147088$	$L_2 = .028958$	$a_3 = .114653$
$L_3 = .147088$	$a_4 = .400523$	$L_3 = .114653$	$a_4 = .335379$
$H_1 = .852912$	$b_1 = .599477$	$H_1 = .885347$	$b_1 = .664621$
$H_2 = .958143$	$b_2 = .852912$	$H_2 = .971042$	$b_2 = .885347$
$H_3 = .997263$	$b_3 = .958143$	$H_3 = .998660$	$b_3 = .971042$
$H_4 = 1$	$b_4 = .997263$	$H_4 = 1$	$b_4 = .998660$
Actual coverage = .99841 wrong λ wrong bounds	Actual coverage = .901771 wrong λ correct bounds	Actual coverage = .9995 correct λ wrong bounds	Actual coverage = .95 correct λ correct bounds

The first column of table 1 gives the constants L_i and H_i involved in the right side of Owen's claimed event identity (42) corresponding to his $\lambda_4^{.95} = .9149...$. The "actual coverage" in this column is the probability of the event on the right side of (42) and (43). [Note that this probability, .99841..., is not .95. It seems that in his computer program Owen used $\{a_1, \ldots, a_n\} = \{L_1, \ldots, L_{n-1}, e^{-\lambda_n}\}$ in place of $L_0, \ldots, L_{n-1}$. When we make this replacement we find $P(M_4) \equiv P(\cap_{i=1}^4 [a_i \le X_{(i)} \le H_i]) = .95$ when a_i and H_i are determined by Owen's $\lambda_4^{.95}$. But this latter event M_4 also satisfies $[R_4 \le \lambda_4^{.95}] \subset M_4$.] The second column of table 1 gives the constants $\{a_i\}$ and $\{b_i\}$ corresponding to Owen's $\lambda_4^{.95} = .9149...$. The resulting probability of the two equal events in (44) is $0.9017... < .95$. (This makes sense since the four dimensional rectangle involved on the right side of (44) is strictly contained in the rectangle on the right side of (42), even with the L_i's replaced by a_i's as in Owen's program.) The third column of table 1 gives the constants L_i and H_i corresponding to our $\lambda_4^{.95} = 1.092...$. This is the correct $\lambda_4^{.95}$, but Owen's bounds L_i and H_i are still wrong, so equality in (42) fails and the resulting probability of the event on right side of (43) is .9995... (or .9747... if we use the a_i's in place of $L_0, ..., L_4$ as the lower bounds as in Owen's program). Finally, column 4 of table 1 gives our $\{a_i\}$ and $\{b_i\}$ corresponding to our $\lambda_4^{.95}$, and the probability of the event on the left side of (43) and both terms in (44) is .95.

Table 2 compares the $\lambda_n^{0.95}$ values calculated using Owen's definition of the $\{H_i\}$'s and those calculated using our $\{b_i\}$'s with exact results for $n = 2$, and simulation results for $3 \le n \le 20$ and selected larger values of n. The Monte Carlo simulations were carried out by simulating the Berk-Jones statistic $100,000$ times and taking the 0.95 quantile (or 95000^{th} order statistic) of the simulated values.

In each case, the finite sample quantile calculated according to the $\{b_i\}$ found by our method (which is similar to the method used for the reversed statistic) agrees more closely with the simulated result.

Finally, we determine the confidence bands

$$L_n(x) = \sum_{i=0}^{n} l_i 1_{(X_{(i)}, X_{(i+1)}]}(x)\,,$$

and

$$H_n(x) = \sum_{i=0}^{n} h_i 1_{[X_{(i)}, X_{(i+1)})}(x)\,,$$

where $X_{(0)} \equiv -\infty$ and $X_{(n+1)} = \infty$ by convention. For $x \in (X_{(i)}, X_{(i+1)})$ we have $\mathbb{F}_n(x) = i/n$ and it is clear that the event $[R_n \leq \lambda_n^{1-\alpha}]$ restricts $F(x)$ only by $K(i/n, F(x)) \leq \lambda_n^{1-\alpha}$. Hence

$$h_i = \max\{p | K(i/n, p) \leq \lambda_n^{1-\alpha}\}\,,$$

while

$$l_i = \min\{p | K(i/n, p) \leq \lambda_n^{1-\alpha}\}\,.$$

From (32) and (36) it follows that $h_i = b_{i+1}$, $i \in \{0, \ldots, n-1\}$, and from (35) and (33) we have $l_i = a_i$, $i \in \{1, \ldots, n\}$. Furthermore $h_n = 1$ and $l_0 = 0$ (trivially).

Table 2: Comparison of 0.95 quantiles of the Berk-Jones statistic with simulation

n	Owen's $\lambda_n^{.95}$	Estimated Coverage	Owen's approximate $\lambda_n^{.95}$	Estimated Coverage	Our $\lambda_n^{.95}$	Simulation
2	1.67031	.90032	1.67117	.90058	2.024950	2.02769
3	1.176631	.90122	1.17665	.90122	1.414108	1.41362
4	0.914983	.90195	0.915054	.90198	1.092493	1.08907
5	0.751753	.90184	0.751894	.90263	0.892788	0.891337
6	0.639718	.90385	0.639889	.90397	0.756251	0.755725
7	0.557816	.90350	0.557992	.90359	0.656788	0.659785
8	0.495200	.90595	0.495369	.90603	0.580990	0.578748
9	0.445698	.90767	0.445852	.90776	0.521242	0.519253
10	0.405531	.90660	0.405670	.90650	0.472895	0.473739
11	0.372252	.90646	0.372376	.90661	0.432943	0.433429
12	0.344209	.90594	0.344319	.90605	0.399358	0.402910
13	0.320240	.90865	0.320337	.90876	0.370718	0.370418
14	0.299506	.90903	0.299592	.90908	0.345995	0.344839
15	0.281384	.91073	0.281461	.91041	0.324432	0.322943
16	0.265404	.90926	0.265473	.90935	0.305456	0.306919
17	0.251203	.91145	0.251265	.91156	0.288622	0.287871
18	0.238495	.91172	0.238551	.91180	0.273585	0.272886
19	0.227054	.91212	0.227104	.91219	0.260069	0.258733
20	0.216696	.91142	0.216742	.91134	0.247853	0.249022
50	0.093344	.91985	0.0933644	.91988	0.104239	0.103634
100	0.048899	.92186	0.0489062	.92333	0.053766	0.053617
500	0.010631	.93043	0.0106328	.93029	0.011381	0.011379
1000	0.005466	.93251	0.0054659	.93145	0.005804	.00580896

As in OWEN (1995) and OWEN (2001), we give approximation formulas for the 0.95 and 0.99 quantiles of the Berk-Jones statistic which are polynomial in $\log n$. These formulas compare to (10)-(13) in OWEN (1995) and to Table 7.1, page 159 in OWEN (2001). We find that Owen's exact and approximate critical values for bands with claimed confidence coefficient .95 have true coverage ranging from about .90 to .93 for sample sizes between 3 and 1000; see Table 2 for estimated coverage probabilities (with 10^5 monte-carlo samples).

$$\lambda_n^{0.95} \doteq \frac{1}{n}(3.6792 + 0.5720 \log n - 0.0567(\log n)^2 + 0.0027(\log n)^3), \qquad 1 < n \leq 100. \tag{45}$$

$$\lambda_n^{0.95} \doteq \frac{1}{n}(3.7752 + 0.5062 \log n - 0.0417(\log n)^2 + 0.0016(\log n)^3), \qquad 100 < n \leq 1000. \tag{46}$$

$$\lambda_n^{0.99} \doteq \frac{1}{n}(5.3318 + 0.5539 \log n - 0.0370(\log n)^2), \qquad 1 < n \leq 100. \tag{47}$$

$$\lambda_n^{0.99} \doteq \frac{1}{n}(5.6392 + 0.4018 \log n - 0.0183(\log n)^2), \qquad 100 < n \leq 1000. \tag{48}$$

4 Power considerations

4.1 Power heuristics

Here we more specifically address issues of power. We are able to get some qualitative ideas of the behavior of these test statistics against different alternatives by considering the functions $K(F_0(x), F(x))$ and $K(F(x), F_0(x))$ pointwise in x, rather than taking the supremum over all x.

Consider the distribution functions

$$F_1(x) = \frac{1}{1 + \log \frac{1}{x}} \tag{49}$$

and

$$F_2(x) = e^{-(\frac{1}{x} - 1)}. \tag{50}$$

The distribution function F_1 is an "extreme" upward alternative to F_0 in a neighborhood of zero, while F_2 is an "extreme" downward alternative to F_0 (Figure 1). So F_1 is an example of a distribution function with high density near zero, while F_2 is an example of a distribution function with low density near zero.

Using the distribution functions F_1 and F_2, we can define the functions

$$\begin{aligned} g_1(x) &= K(F_1(x), x)\,, \\ g_2(x) &= K(F_2(x), x)\,, \\ \tilde{g}_1(x) &= K(x, F_1(x))\,, \\ \tilde{g}_2(x) &= K(x, F_2(x)). \end{aligned}$$

Figure 2 suggests that the Berk-Jones statistic (R_n) will be more powerful against alternatives of the type F_1, while the reversed statistic ($\tilde{R}_n$) will be more powerful against alternatives of the form F_2.

Although these plots suggest that the reversed statistic may be more powerful than the Berk-Jones statistic for the situation where the alternative distribution function is extremely different

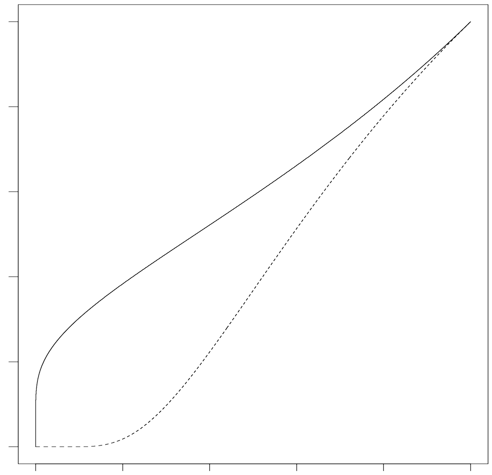


Figure 1: Extreme distribution functions F_1 (solid line) and F_2 (dashed line).

than the null, we are actually more interested in the power behavior for alternatives which are slightly different from the null distribution. For example, alternatives which are more moderately stochastically larger or smaller than F_0.

Natural alternatives to consider are those of the form F_0^c, for different values of $c \in (0, \infty)$. For values of $c > 1$, this distribution is stochastically larger than F_0. For values of $c < 1$, this distribution is stochastically smaller than F_0.

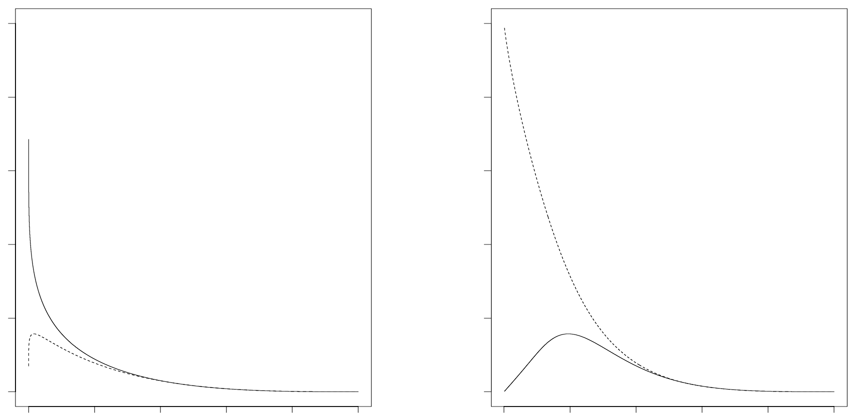
Based on the behavior of the functions g_1, g_2, $\tilde{g}_1$, and $\tilde{g}_2$, we would guess that in this case as well, the dual statistic would be more powerful against stochastically larger alternatives ($c > 1$), while the Berk-Jones statistic would remain more powerful for stochastically smaller alternatives ($c < 1$).

4.2 Power calculations

To test our conjectures about power, we use the same algorithm by Noé (1972) to calculate the probability that F_1, F_2, and F^c are contained in the 95% confidence band for F_0. Figure 3 plots these probabilities for F^c against c for different sample sizes. The curves for sample size $n = 20$ can be compared to Figure 5 in Owen (1995). The line representing the Berk-Jones statistic is the same as that which Owen calls the curve for the nonparametric likelihood bands. We see that the reverse Berk-Jones statistic has greater power than both the Berk-Jones statistic and the Kolmogorov-Smirnov statistic for values of $c > 1$.

5 Examples

Figure 4 shows the empirical distribution function of the velocities of 82 galaxies from the Corona Borealis region along with 95% confidence intervals generated by inverting both the Berk-Jones



(a) (b)

Figure 2: (a) The functions g_1 (solid line) and $\tilde{g}_1$ (dashed line). (b) The functions g_2 (solid line) and $\tilde{g}_2$ (dashed line).

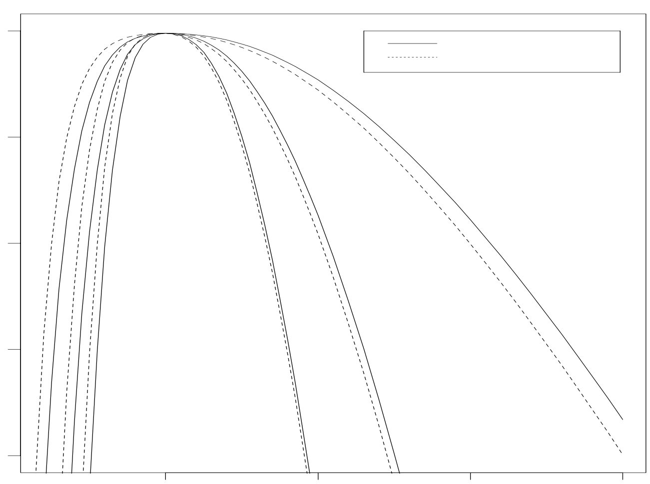


Figure 3: The probability that the alternative distribution F_0^c is included in the 95% confidence bands for F_0 (vertical axis), based on samples of $n = 20, 50, 100$ from a continuous distribution F_0.

statistic and the reversed statistic. This data appears in Table 1 of Roeder (1990). This figure can be compared to Figures 1 and 2 in Owen (1995). Comparison shows that the confidence band based on the reversed Berk-Jones statistic are narrower at the tails than the one based on the Kolmogorov-Smirnov statistic. Also, there are slight differences between the confidence band based on the reversed statistic compared to the Berk-Jones statistic. In the region of the lower tail, the band based on the reversed statistic is shifted slightly downward, while in the region of the upper tail, this band is shifted slightly upward. This behavior is more noticable when looking at equally spaced data points. Figure 5 shows these same 95% confidence bands for $n = 20$ equispaced observations.

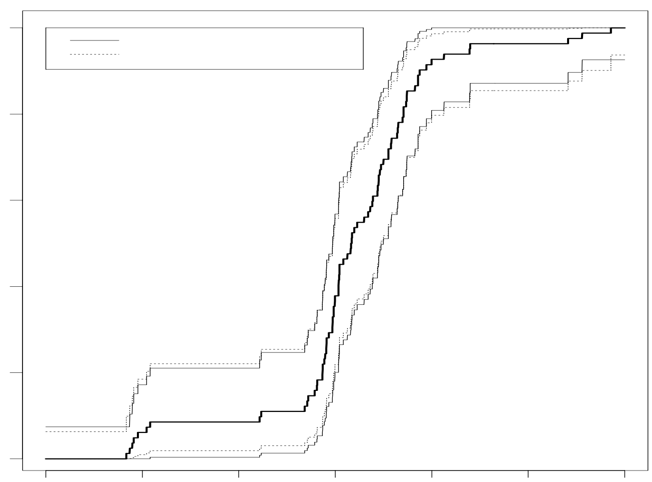


Figure 4: The empirical CDF of the velocities of 82 galaxies in the Corona Borealis Region (dark solid line) and 95% confidence bands obtained by inverting the Berk-Jones statistic (dashed line) and the reversed Berk-Jones statistic (solid line)

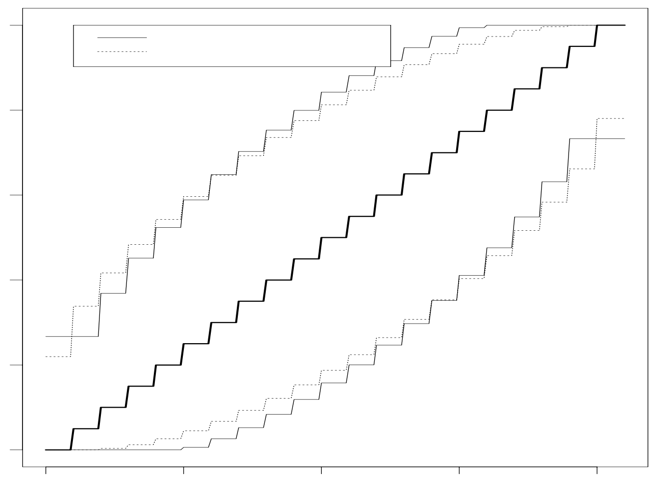


Figure 5: The empirical CDF of 20 equally spaced data points (dark solid line) and 95% confidence bands obtained by inverting the Berk-Jones statistic (dashed line) and the reversed Berk-Jones statistic (solid line)

Finally, Figure 6 gives a comparison of Owen's bands based on the Berk-Jones statistic to our bands based on the same statistic; as argued above, Owen's bands do not have the correct coverage probability.

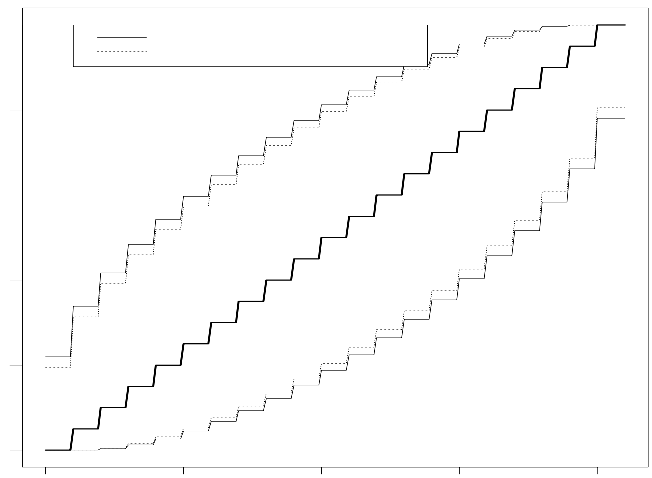


Figure 6: Comparison of Owen's bands based on the Berk-Jones statistic (dashed line) to our bands based on the Berk-Jones statistic (solid line) for 20 equally spaced data points

The C and R programs used to carry out the computations presented here are available (in several forms) at

http://www.stat.washington.edu/jaw/RESEARCH/SOFTWARE/software.list.html.

Acknowledgements: We owe thanks to Art Owen for sharing his C programs used to carry out the computations for his 1995 paper. Those programs were used as a starting point for the programs used here. We also owe thanks to Art for several helpful discussions. Mame Astou Diouf pointed out several typographical errors in the first version.

References

Berk, R. H. and Jones, D. H. (1979). Goodness-of-fit test statistics that dominate the Kolmogorov statistics. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **47**, 47-59.

Noé, M. (1972). The calculation of distributions of two-sided Kolmogorov-Smirnov type statistics. *Annals of Mathematical Statistics* **43**, 58-64.

Owen, A. B. (1995). Nonparametric likelihood confidence bands for a distribution function. *Journal of the American Statistical Association* **90**, 516-521.

Owen, A. B. (2001). *Empirical Likelihood*. Chapman & Hall/CRC, Boca Raton.

Roeder, K. (1990). Density estimation with confidence sets exemplified by superclusters and voids in the galaxies. *Journal of the American Statistical Association* **85**, 617-624.

Wellner, J. A. and Koltchinskii, V. (2003). A note on the asymptotic distribution of Berk-Jones type statistics under the null hypothesis. *High Dimensional Probability III*, 321-332. Birkhäuser, Basel (2003).

University of Washington
Statistics
Box 354322
Seattle, Washington 98195-4322
e-mail: leah@stat.washington.edu

University of Washington
Statistics
Box 354322
Seattle, Washington 98195-4322
e-mail: jaw@stat.washington.edu