PLOTS AND TESTS FOR SYMMETRY

by

Kjell A. Doksum
University of California, Berkeley

Grete Fenstad and Rolf Aaberge
University of Oslo
PLOTS AND TESTS FOR SYMMETRY

by

Kjell A. Doksum
University of California, Berkeley

Grete Fenstad and Rolf Aaberge
University of Oslo

SUMMARY

We consider plots which can be used to check whether a distribution is symmetric about some point and which give information about possible deviations from symmetry. Simultaneous confidence bands for checking the accuracy of the plots are derived and compared. Tests that are related to the plots and confidence bands are considered and their power functions are compared using Monte Carlo methods. As tests of normality, the power of two of these tests compare favorably with the power of the Shapiro-Wilk statistic.

Some key words: Symmetry, skewness, confidence bands, Butler statistic, tests for normality.
1. INTRODUCTION

Many of the robust methods recently developed rely to a certain extent on the assumption of symmetry. Therefore it is important to have methods which check this assumption and indicate possible deviations from symmetry.

Let the observations, before or after a transformation, be denoted by \( X_1, \ldots, X_n \) and let \( X_1(1) \leq X_2(2) \leq \cdots \leq X_n(n) \) denote the order statistics. We assume that the distribution \( F \) of \( X_i \) is continuous and let \( G \) denote the distribution of \( -X_i \), thus

\[
G(x) = 1 - F(-x)
\]

\( F_n \) and \( G_n \) will denote the empirical distributions of the \( X_i \) and \( -X_i \), respectively.

Three closely related plots for assessing symmetry are those of Gnanadesikan and Wilk (1968), Tukey (given by Gnanadesikan and Wilk) and Doksum (1975). These plots, which are introduced in Section 2, are all based on

\[
h_n(x) = G_n^{-1}(F_n(x))
\]

where

\[
G_n^{-1}(u) = \inf\{ t : G_n(t) \geq u \}
\]

is the left inverse of \( G_n \).

Note that for fixed \( x \), \( h_n(x) \) converges a.s. to

\[
h(x) = G^{-1}(F(x))
\]

The plots based on \( h_n \) are reliable only for "large \( n \)". One gets a good idea of how reliable and how large \( n \) must be by considering simultaneous confidence bands for \( h \) based on \( F_n \) and \( G_n \).

In Section 3, we introduce bands based on test statistics of the form
where \( a \) and \( b \) are given constants and \( \Psi \) is a weight function. \( \Psi = 1 \) and \( a = 1 - b = 0 \) yields the Butler (1969) statistic \( B_n \), in which case the confidence band for \( h \) is

\[
\left[ G_n^{-1}(F_n(x)-\varepsilon), G_n^{-1}(F_n(x)+\varepsilon) \right]
\]

where \( G_n^{-1} \) is the right inverse

\[
G_n^{-1}(u) = \sup\{t : G_n(t) \leq u\}
\]

and \( \varepsilon \) is the level \( \alpha \) critical value of \( B_n \). This band is preferable to the band given by Doksum (1975) which uses a larger \( \varepsilon \).

We also consider a statistic \( A_n \) obtained by choosing the weight function

\[
\Psi_1^2(u,v) = u + v, \quad 0 < u + v \leq 1
\]

\[
= 2 - u - v, \quad 1 < u + v < 2
\]

in \( T_n \). \( \Psi_1 \) is chosen so as to make the variance of

\[
\frac{|F_n(x)-G_n(x)|}{\Psi_1(F_n(x),G_n(x))}
\]

approximately constant in \( x \).

The asymptotic distribution of \( T_n \) is given and the bands derived from the Butler and the \( A_n \) statistics are compared in terms of their asymptotic widths. It turns out that the latter is narrower in the "tails" and is preferable on this basis.

Next, some tests for symmetry related to the plots are considered and their power functions compared for common skew parametric models. The two best statistics are found to be

\[
B_n(\hat{\eta}) = \sup x\{F_n(x+\hat{\eta})-G_n(x-\hat{\eta})\}
\]

and the David-Johnson (1956) statistic

\[
J = \frac{\hat{\eta} - M}{\sigma}
\]
where $M$ is the sample median, $\hat{\sigma}$ is the quartile range and 

$$\hat{\eta} = \frac{1}{2}(x(1) + x(n)).$$

As tests for normality, $B_n(\hat{\eta})$ and $J$ are better than the Kolmogorov statistic with the parameters estimated and they compare favorably with the Shapiro-Wilk statistic. The Monte Carlo results slightly favors $J$ over all other competitors. It is by far the easiest test to carry out.

2. THE PLOTS AND FIRST PROPERTIES

The three plots are

(A) The Wilk-Gnanadesikan (1968) plot $w_n(x)$.

This plot is $(x(i), x(n+1-i))$, $i=1, \ldots, \lfloor \frac{1}{2}(n+1) \rfloor$ where $[t]$ is the greatest integer $\leq t$. It is the function

$$w_n(x) = - G_n^{-1}(F_n(x))$$

evaluated at the order statistics $x(i)$, $i \leq \lfloor \frac{1}{2}(n+1) \rfloor$. For fixed $x$, $w_n(x)$ converges a.s. to

$$w(x) = - G^{-1}(F(x)) = F^{-1}(1-F(x)).$$

Let $m$ denote the median of $F$, then

$$F \text{ symmetric } \iff w(x) = 2m - x.$$  

$F$ is said to be skew (to the right) if $X - m$ is stochastically larger than $-X + m$, or equivalently, if

$$\text{pr}(X \geq m+t) \geq \text{pr}(X \leq m-t) \quad \text{for all } t > 0.$$  

In this case, $w(x)$ lies entirely on or above $2m - x$ and for the commonly used skew models, $w(x) + x$ will be U or bathtub shaped with its minimum at $x = m$. See Doksum (1975) and Gnanadesikan and Wilk (1968). Fig. 1 (a) essentially gives $w(x)$ and $2m - x$ when $F$ is the $\chi^2_{10}$ distribution. Note that in Fig. 1, the scale has been adjusted so that the graphs have the same range along both axis. In this
case, we have plotted
\[ w_1(x) = 13.57 + 0.55 w(x) \text{ and } m_1(x) = 13.57 + 0.55 (2m-x). \]

(B) The Tukey plot \( t_n(x) \).

As given by Gnanadesikan and Wilk (1968, p.4), this is a plot of
\[ (x(n+1-i), x(i) + x(n+1-i)) , i=1,...,[\frac{1}{2}(n+1)]. \]
It is the curve
\[ t_n(x) = (-G^{-1}_n(F_n(x)) - x, x - G^{-1}_n(F_n(x))) \]
evaluated at the \( x(i) \) order statistics, \( i \leq [\frac{1}{2}(n+1)] \). For fixed \( x \), it converges a.s. to
\[ t(x) = \{(t_1(x), t_2(x)) , x \leq m\} \]
where
\[ t_1(x) = -G^{-1}_n(F_n(x)) - x, t_2(x) = x - G^{-1}_n(F_n(x)) \]
\( t \) reduces to the line \( \{(2m-2x, 2m):x \leq m\} \) when \( F \) is symmetric and when \( F \) is skew, \( t_2(x) \) lies entirely on or above the line \( 2m \). Figure 1 (b) gives \( (2.16+t_1(x)/3.21 , t_2(x)) , x < m\), for the \( \chi^2_{10} \) distribution.

(C) The symmetry plot \( \theta_n(x) \).

The plot
\[ (x(i), \frac{1}{2}[x(i)+x(n+1-i)]) = (x(i), \theta_n(x(i))) , i=1,...,n \]
where
\[ \theta_n(x) = \frac{1}{2}\{x-G^{-1}_n(F_n(x))\} \]
was considered by Doksum (1975). \( \theta_n(x) \) converges for fixed \( x \)
a.s. to
\[ \theta(x) = \frac{1}{2}\{x-G^{-1}(F(x))\} \]
which is called the symmetry function. It reduces to the point of symmetry \( m \) when \( F \) is symmetric. In general it measures location as well as asymmetry. When \( F \) is skew to the right, it lies
entirely on or above $m$. See Doksum (1975) and Figure 1(c) which gives $2\theta(x)$ for the $\chi^2_{10}$ distribution.

3. THE CONFIDENCE BANDS

The bands for $h(x) = G^{-1}(F(x))$ will be based on statistics of the type

$$T_n = T(F_n, G_n) = \sup_{x : a \leq F_n(x) \leq 1 - a} \frac{|G_n(x) - F_n(x)|}{\Psi(F_n(x), G_n(x))}$$

where $0 \leq a \leq \frac{1}{2}$ is a given constant and $\Psi$ is a positive weight function satisfying $\Psi(u, v) = \Psi(1 - u, 1 - v)$ for $0 \leq u + v \leq 1$. It is known that $T(F_n, G_n)$ is distribution-free for testing symmetry about zero, thus we can find a critical value $K$ such that

$$\Pr_H\{T(F_n, G_n) \leq K\} = 1 - \alpha,$$  \hspace{1cm} (3.1)

where $H$ stands for the hypothesis: "$F(x) = 1 - F(-x)$ for all $x$". We will use the notation $G_n h$, $G^{-1} F$, etc., for the composite functions $G_n(h(x))$, $G^{-1}(F(x))$ etc.

The following result will allow us to use $T(F_n, G_n h)$ as a pivot for finding a confidence band for $h(x)$:

**Lemma 3.1.** If (3.1) holds, then

$$\Pr_F\{T(F_n, G_n h) \leq K\} = 1 - \alpha$$

for all continuous $F$.

**Proof:** We can write

$$T(F_n, G_n h) = \sup_{u : a \leq F_n^{-1}(u + \frac{1}{2}) \leq 1 - a} \frac{|F_n F^{-1}(u + \frac{1}{2}) - G_n h F^{-1}(u + \frac{1}{2})|}{\Psi(F_n F^{-1}(u + \frac{1}{2}), G_n h F^{-1}(u + \frac{1}{2}))}.$$

Let $U_n$ and $V_n$ denote the empirical distributions of the uniform samples $\{F(X_i) - \frac{1}{2}\}$ and $\{-F(X_i) - \frac{1}{2}\}$, respectively, and note that $F_n F^{-1}(u + \frac{1}{2}) = U_n(u)$, $-\frac{1}{2} \leq u \leq \frac{1}{2}$, while
Thus

\[ T(F_n, G_n h) = \sup_{\{u: a \leq U_n(u) \leq b\}} \frac{|U_n(u) - V_n(u)|}{\Psi(U_n(u), V_n(u))} \]

and the result follows.

The lemma can be used to obtain confidence bands for \( h \) if the solution of the inequality \( T(F_n, G_n h) \leq K \) for \( G_n h \) is an interval

\[ G_n F_n(x) \leq G_n h(x) \leq G_n F_n(x) \]

for all \( x \) in

\[ S_n = \{x : a \leq F_n(x) \leq 1 - a\} \]

Next we apply \( G_n^{-1} \) and \( G_n^{-1} \) to the left and right hand inequalities, respectively, and obtain

\[ G_n^{-1} G_n F_n(x) \leq h(x) \leq G_n^{-1} G_n F_n(x) \]

for all \( x \in S_n \).

Now this inequality holds for all \( x \in S_n \) with probability \( (1 - \alpha) \). Thus (3.3) defines a simultaneous confidence interval for \( h = G_n^{-1} F \) which is valid for all continuous \( F \). We consider two applications of (3.3).

(A) A band based on the Butler statistic.

Suppose \( a = 0 \) and \( \Psi = 1 \). Then \( T_n \) reduces to the Butler (1969) statistic

\[ B_n = \sup_{-\infty < x < \infty} |F_n(x) - G_n(x)| = \sup_{x < 0} |F_n(x) - G_n(x)| \]

Now \( G_n(u) = u - K \), \( G_n^*(u) = u + K \) and (3.3) becomes
\[ G_n^{-1}(F_n(x)-k) \leq h(x) \leq G_n^{-1}(F_n(x)+k) \]

We denote this band by \([B_*(x), B^*(x)]\) and call it the \textit{B-band}.

Let \(\langle t \rangle\) denote the least integer greater than or equal to \(t\).

For \(x\) in \([x(i), x(i+1))\), \(F_n(x) = \frac{i}{n}\) and thus
\[ [B_*(x), B^*(x)] = [-x(j), -x(l)] \]
where \(j = n + 1 - \langle i - nK \rangle\) and \(l = n - \lfloor i + nK \rfloor\).

Computing this band thus just involves finding the order statistics and \(K\).

The critical value \(K\) can be obtained from the distribution of \(B_n\) given by Butler. Tables for \(n \leq 16\) have been given by Chatterjee and Sen (1973). For other \(n\), we proceed as follows.

Let
\[ B_n^+ = \sup_x \{G_n(x) - F_n(x)\} \]
\[ B_n^- = \sup_x \{F_n(x) - G_n(x)\} \]

and consider the approximation
\[ p_{B_n}(nB_n \geq k) \approx p_{B_n}(nB_n^+ \geq k) + p_{B_n}(nB_n^- \geq k) \]  \(\text{(3.4)}\)

The error in this approximation is very small when \(pr(nB_n \geq k) \leq .5\) as indicated by Table 1.

<table>
<thead>
<tr>
<th>(n)</th>
<th>(k)</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td></td>
<td>.8</td>
<td>2.3</td>
<td>4.3</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>.6</td>
<td>1.3</td>
<td>2.3</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

The advantage of (3.4) is that, as shown by Butler, the right hand side is easy to compute. We give the following extension of
the tables of Chatterjee and Sen.

TABLE 2.

<table>
<thead>
<tr>
<th>k</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>.260</td>
<td>.189</td>
<td>.115</td>
<td>.078</td>
<td>.041</td>
<td>.027</td>
<td>.012</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>.266</td>
<td>.189</td>
<td>.130</td>
<td>.078</td>
<td>.052</td>
<td>.027</td>
<td>.017</td>
<td>.007</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>.266</td>
<td>.210</td>
<td>.130</td>
<td>.093</td>
<td>.052</td>
<td>.035</td>
<td>.017</td>
<td>.011</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>.307</td>
<td>.230</td>
<td>.150</td>
<td>.108</td>
<td>.064</td>
<td>.043</td>
<td>.023</td>
<td>.0146</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>.230</td>
<td>.169</td>
<td>.108</td>
<td>.075</td>
<td>.043</td>
<td>.028</td>
<td>.0146</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>.248</td>
<td>.169</td>
<td>.122</td>
<td>.075</td>
<td>.052</td>
<td>.028</td>
<td>.019</td>
<td>.009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>.248</td>
<td>.185</td>
<td>.122</td>
<td>.087</td>
<td>.052</td>
<td>.036</td>
<td>.019</td>
<td>.0125</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>.265</td>
<td>.185</td>
<td>.136</td>
<td>.087</td>
<td>.061</td>
<td>.036</td>
<td>.025</td>
<td>.013</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>.200</td>
<td>.136</td>
<td>.099</td>
<td>.061</td>
<td>.043</td>
<td>.025</td>
<td>.016</td>
<td>.008</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>.200</td>
<td>.150</td>
<td>.099</td>
<td>.071</td>
<td>.043</td>
<td>.029</td>
<td>.016</td>
<td>.010</td>
<td>.005</td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td></td>
<td>.175</td>
<td>.132</td>
<td>.090</td>
<td>.065</td>
<td>.041</td>
<td>.029</td>
<td>.0167</td>
<td>.011</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td></td>
<td>.150</td>
<td>.116</td>
<td>.080</td>
<td>.059</td>
<td>.038</td>
<td>.028</td>
<td>.017</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td></td>
<td>.135</td>
<td>.103</td>
<td>.072</td>
<td>.054</td>
<td>.036</td>
<td>.026</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td></td>
<td>.161</td>
<td>.120</td>
<td>.092</td>
<td>.065</td>
<td>.049</td>
<td>.033</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For larger n, we can use the approximation of the next section.

(B) A band based on an "equal weight" statistic.

We next consider choosing \( \gamma \) so as to give approximately equal weight to all \( x \). Thus we choose \( \gamma \) equal to an estimate of the standard deviation of \( \sqrt{n}|F_n(x) - G_n(x)| \). We find

\[
\text{var}_H \{F_n(x) - G_n(x)\} = \frac{2}{n}[1 - F(|x|)]
\]

Schuster (1975) has shown that, under \( H \), \( \frac{1}{2}|F_n(x) + G_n(x)| \) is a better estimate of \( F \) than \( F_n \). Thus we choose
\( \psi^2(u,v) = \psi_1^2(u,v) = u + v \quad \text{if} \quad 0 < u + v \leq 1 \\
= 2 - u - v \quad \text{if} \quad 1 < u + v < 2 \\
= 1 \quad \text{if} \quad u + v = 0 \text{ or } 2 \\
\)

and consider the statistic

\[
A_n = A(F_n, G_n) = \sup \left\{ a \leq F_n(x) \leq 1 - a \right\} \frac{|G_n(x) - F_n(x)|}{\psi_1^2(F_n(x), G_n(x))}
\]

where \( 0 \leq a < \frac{1}{2} \) is a given constant.

We will need the following continuous functions on \( [0,1] \)

\[
\lambda_1^+(u) = u + \frac{1}{2}K^2 + \frac{1}{2}K(8u + K^2)^{1/2} \\
\lambda_2^+(u) = u - \frac{1}{2}K^2 + \frac{1}{2}K(8(1-u) + K^2)^{1/2} = 1 - \lambda_1^-(1-u) \\
\lambda_1^-(u) = 0 , \quad u \leq K^2 \\
= \lambda_1^-(u) , \quad K^2 < u \leq \frac{1}{2} + \frac{1}{2}K \\
= \lambda_2^-(u) , \quad u > \frac{1}{2} + \frac{1}{2}K \\
\lambda_2^-(u) = \lambda_1^+(u) , \quad u \leq \frac{1}{2} - \frac{1}{2}K \\
= \lambda_2^+(u) , \quad \frac{1}{2} - \frac{1}{2}K < u \leq 1 - K^2 \\
= 1 , \quad u > 1 - K^2 
\]

Note that \( A_n \leq 1 \), which implies \( K \leq 1 \). The functions \( \lambda_* \) and \( \lambda^* \) will therefore be continuous and non-decreasing.

**Theorem 3.1.** The simultaneous confidence band for \( H(x) \) based on \( A_n \) is given by

\[
[A_*^-(x), A_*^+(x)] = [g_1^{-1} \lambda_*(F_n(x)), g_1^{-1} \lambda^*(F_n(x))] \\
\]

\( x \in \{ x : a \leq F_n(x) \leq 1 - a \} \).
Proof. As shown in (3.2), we first need to solve
\[ A(F_n, G_n h) \leq K \] for \( G_n h \). Set \( u = F_n(x), v = G_n h(x) \), then this is equivalent to solve
\[ (v-u)^2 \leq K^2 (u+v) , \quad 0 < u + v \leq 1 \] (3.5)
and
\[ (v-u)^2 \leq K^2 (2-u-v) , \quad 1 < u + v < 2 \] (3.6)
for \( v \). We find that (3.5) and (3.6) are equivalent to
\[ \lambda_1^-(u) \leq v \leq \lambda_1^+(u) , \quad 0 < u + v \leq 1 \]
and
\[ \lambda_2^-(u) \leq v \leq \lambda_2^+(u) , \quad 1 < u + v < 2 \]
respectively. Considering the graphs of \( \lambda_1^+(u) \) and \( \lambda_2^+(u) \) in Fig. 2, the result follows.

The \( A \)-band \( [A_*(x), A^*(x)] \) is
\[ -x(j) \leq h(x) \leq -x(k) , \quad x \in [x(i), x(i+1)] \]
where \( j = n + 1 - \lfloor n \lambda_1^*(\frac{1}{n}) \rfloor \) and \( k = n - \lfloor n \lambda_1^*(\frac{1}{n}) \rfloor \), defining
\( x(n+1) = +\infty \) and \( x(0) = -\infty \).

When considering the distribution of \( A_n \) it helps to note that, because of symmetry,
\[ A_n = \sup \left\{ x : x < 0, a \leq F_n(x) \right\} \frac{|F_n(x) - G_n(x)|}{\sqrt{F_n(x) + G_n(x)}} . \]
Moreover, \( \text{pr}_H(A_n > K) = \text{pr}_H(A_+^+ > K) + \text{pr}_H(A^- > K) = 2 \text{pr}_H(A_+^+ > K) \)
where
\[ A_+^+ = \sup \left\{ x : x < 0, a \leq F_n(x) \right\} \frac{+|G_n(x) - F_n(x)|}{\sqrt{F_n(x) + G_n(x)}} . \]

\( A_n \) takes on values between 0 and 1 while \( A_+^+ \) and \( A_-^- \) take on values between -1 and 1.
$A_n$ converges a.s. to

$$A(F, G) = \sup_{\{x : x < 0, a \leq \bar{F}(x)\}} \frac{|F(x) - G(x)|}{\sqrt{F(x) + G(x)}}$$

with the corresponding results for $A_n^+$ and $A_n^-$. It follows that the tests based on them are consistent. Asymptotic critical values are given in the next section.

**Remark 3.1.** From (3.2) we see that if we want upper and lower level $(1-a)$ confidence boundaries for $h$, then we can invert the statistics $T_n^+$ and $T_n^-$, respectively, where

$$T_n^+ = \sup_{\{x : a \leq F_n(x) \leq 1-a\}} \frac{\pm[G_n(x) - F_n(x)]}{\Psi(F_n(x), G_n(x))}.$$ 

The resulting boundaries differ from the $T_n$ band boundaries only in the critical constant $K$ and, because of the approximation

$$\Pr(T_n > K) \approx \Pr(T_n^+ > K) + \Pr(T_n^- > K),$$

are very close to the upper and lower boundaries of the level $(1-2a)$ $T_n$ band.

4. **ASYMPTOTIC DISTRIBUTIONS OF THE TEST STATISTICS**

We assume that $F$ is continuous and symmetric about zero and that $\Psi$ is continuous and bounded away from zero for $u, v \in [a, 1-a]$. Since $\Psi(u, v) = \Psi(1-u, 1-v)$, then

$$T_n = \sup_{\{x : x < 0, a \leq F_n(x) \geq a\}} \frac{|G_n(x) - F_n(x)|}{\Psi(F_n(x), G_n(x))} \quad (4.1)$$

We will also consider $T_n^+$ and $T_n^-$ obtained by replacing the numerator in (4.1) by $\{G_n(x) - F_n(x)\}$ and $\{F_n(x) - G_n(x)\}$, respectively. As in Section 3, it is often convenient to note that
$T_n = \max\{T_n^-, T_n^+\}$ and to use the approximation

$$\Pr(T_n > K) \approx \Pr(T_n^+ > K) + \Pr(T_n^- > K) = 2 \Pr(T_n^+ > K)$$

Let $W$ and $W_0$ denote the standard Wiener process and Brownian Bridge on $D[0,1]$, respectively. Since

$$G_n(t) = 1 - F_n(-t^-)$$

and

$$\sqrt{n}[F_n(t) - F(t)] \to W_0(F(t)),$$

we find that

$$\sqrt{n}[G_n(t) - F_n(t)] \to W_0(1-F(t)) - W_0(F(t))$$

where $\to$ in this section denotes convergence in distribution.

Since

$$\frac{1}{\sqrt{2n}} \{W_0(u) + W_0(1-u)\}$$

has the same distribution as $W(u)$ on $[0,\frac{1}{2}]$, we can conclude, as did Butler (1969), that

$$\sqrt{n}[G_n(t) - F_n(t)] \to \sqrt{2} W(F(t)) \ , \ F(t) \leq \frac{1}{2}.$$  \hspace{1cm} (4.2)

Thus

$$\sqrt{n} T_n \to \sup_{\{a < u \leq \frac{1}{2}\}} \frac{\sqrt{2} |W_0(u)|}{\psi(u)}$$

and

$$\sqrt{n} T_n^+ \to \sup_{\{a < u \leq \frac{1}{2}\}} \frac{\sqrt{2} |W_0(u)|}{\psi(u)}$$

where

$$\psi(u) = \psi(u,u).$$

Since $W(t)$ has the same distribution as $(1+t)W_0(t/(1+t))$, $t \geq 0$, we find that

$$\sqrt{n} T_n \to \sup_{\{a/(1+a) \leq u \leq \frac{1}{2}\}} \frac{\sqrt{2} |W_0(u)|}{(1-u)\psi\left(\frac{u}{1-u}\right)}$$

with the corresponding results holding for $T_n^+$ and $T_n^-$.  \hspace{1cm} (4.3)

**Lemma 4.1.** When $\psi(u,u) = 1$, $T_n$ becomes the Butler statistic $B_n(a)$ restricted to the set where $a \leq F_n(x) \leq 1 - a$.

**Example 4.1.** When $\psi(u,u) = 1$, $T_n$ becomes the Butler statistic $B_n(a)$ restricted to the set where $a \leq F_n(x) \leq 1 - a$.

Using Lemma 4.1 and the transformation $u \to 1 - u$, we find that the limiting distribution of $\sqrt{n} B_n^+(a)$ is the distribution of
Next we apply Theorem 7 of Rényi (1953) and we find that $B_n(a)$ has the same limiting distribution as the Rényi statistic
\[
\sqrt{n} \sup_{2 \leq x \leq 1/(1+a)} \frac{|F_n(x) - F(x)|}{F(x)}.
\]
Moreover,
\[
\lim_{n \to \infty} \Pr((\sqrt{n})^{\frac{1}{2}} B_n(a) \leq k) = \frac{k}{\sqrt{a}} \int_{-\infty}^{\frac{k}{\sqrt{a}}} e^{-\frac{1}{2}} d\Phi(\frac{2a}{1-2a})^{\frac{1}{2}} e^{-\frac{1}{2} t^2} dt dv.
\]
When $a = 0$, this reduces to
\[
\lim_{n \to \infty} \Pr(\sqrt{n} B_n^+ \geq t) = 2\{1 - \Phi(t)\},
\]
where $\Phi$ is the standard normal distribution. This approximation is good already when $n \geq 20$.

The error in the approximation
\[
\Pr(nB_n^+ \geq k) \approx 2\{1 - \Phi(k/\sqrt{n})\}
\]
can be ascertained by comparing Table 2 with Table 3 below.

<table>
<thead>
<tr>
<th>n</th>
<th>k</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>.264</td>
<td>.180</td>
<td>.118</td>
<td>.074</td>
<td>.044</td>
<td>.025</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>.230</td>
<td>.162</td>
<td>.110</td>
<td>.072</td>
<td>.045</td>
<td>.028</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>.200</td>
<td>.144</td>
<td>.100</td>
<td>.068</td>
<td>.045</td>
<td>.029</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>.176</td>
<td>.128</td>
<td>.091</td>
<td>.063</td>
<td>.043</td>
<td>.028</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>.154</td>
<td>.114</td>
<td>.082</td>
<td>.058</td>
<td>.040</td>
<td>.027</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>.136</td>
<td>.101</td>
<td>.074</td>
<td>.053</td>
<td>.037</td>
<td>.025</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>.120</td>
<td>.090</td>
<td>.066</td>
<td>.048</td>
<td>.034</td>
<td>.024</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
For $B_n$, we have now

$$\Pr(nB_n \geq k) \approx 4\{1 - \Phi(k/\sqrt{n})\}.$$ 

Next we turn to the statistics $A_n$, $A_n^+$ and $A_n^-$. 

Theorem 4.2:

$$\lim_{n \to \infty} \Pr(\sqrt{n}A_n^+ > t) = \lim_{n \to \infty} \Pr(\sqrt{n}A_n^- > t) = \frac{t \exp(-\frac{1}{2}t^2)}{2\sqrt{2\pi}} \log(1/2a).$$

Proof: (4.3) and Lemma 4.1 imply that 

$$\sqrt{n} A_n^+ \to \sup_{\{a/(1+a) \leq u \leq \frac{1}{2}\}} \frac{W_0(u)}{u(1-u)^{\frac{1}{2}}}.$$ 

The right hand side has been considered by Borokov and Sycheva (1968) and the result follows.

Note that $A_n$ has the same asymptotic distribution as statistics of the form 

$$C_n = \sup_{u \leq F(x) \leq v} \frac{|F_n(x) - F(x)|}{[F(x)(1-F(x))]^{\frac{1}{2}}}.$$ 

Such statistics were considered by Anderson and Darling (1952) while Borokov and Sycheva (1968) demonstrated the asymptotic optimality of the one-sided version $C_n^+$ of $C_n$.

Table 4 below gives some asymptotic critical values for $A_n^+$. To get approximate critical values for $A_n$, we again use 

$$\Pr(A_n \geq k) \approx 2\Pr(A_n^+ \geq k).$$

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>.01</th>
<th>.05</th>
<th>.10</th>
<th>.15</th>
<th>.20</th>
</tr>
</thead>
<tbody>
<tr>
<td>.100</td>
<td>2.425</td>
<td>2.137</td>
<td>1.902</td>
<td>1.665</td>
<td>1.337</td>
</tr>
<tr>
<td>.050</td>
<td>2.741</td>
<td>2.504</td>
<td>2.325</td>
<td>2.164</td>
<td>1.993</td>
</tr>
<tr>
<td>.025</td>
<td>3.015</td>
<td>2.808</td>
<td>2.657</td>
<td>2.525</td>
<td>2.392</td>
</tr>
<tr>
<td>.010</td>
<td>3.335</td>
<td>3.155</td>
<td>3.025</td>
<td>2.918</td>
<td>2.806</td>
</tr>
</tbody>
</table>
5. COMPARISON OF THE CONFIDENCE BANDS

The width of the band (3.3) at \( x \) is

\[
\begin{align*}
  w_n(x) &= G_n^{-1} G^* F_n(x) - G_n^{-1} G F_n(x) \\
                   \\
  \text{and the asymptotic width is} \\
  w(x) &= \lim_{n \to \infty} \sqrt{n} w_n(x)
\end{align*}
\]

where the limit is in probability. We assume that \( \sqrt{n} T_n \) has a limiting distribution. Thus when computing \( w(x) \) we can replace \( K \) by \( k \epsilon \) where \( k \) is the fixed asymptotic level critical value and \( \epsilon = n^{-\frac{1}{2}} \). To emphasize the dependence on \( \epsilon \), we write \( G_*(\epsilon,u) \) and \( G^*(\epsilon,u) \) for \( G_*(u) \) and \( G^*(u) \).

Combining the proof of Theorem 8.1 of Doksum (1975) with the proof of Theorem 2 of Doksum and Sievers (1976), we find

Theorem 5.1. Suppose that \( F \) has a continuous positive derivative on \([a,1-a]\) and that \( \text{g}_*(\epsilon,u) \) and \( \text{g}^*(\epsilon,u) \) have right hand derivatives \( \text{g}_*(u) \) and \( \text{g}^*(u) \) with respect to \( \epsilon \) at \( \epsilon = 0 \). Let \( u \in [a,1-a] \) be a point at which \( \text{g}_*(u) \) and \( \text{g}^*(u) \) are continuous. Then for \( x = F^{-1}(u) \)

\[
w(x) = \frac{\text{g}^*(F(x)) - \text{g}_*(F(x))}{f(-F^{-1}(1-F(x)))} \quad (5.1)
\]

If we have two bands \( B_1 \) and \( B_2 \) with asymptotic widths \( w_1(x) \) and \( w_2(x) \), we define the asymptotic relative efficiency of \( B_1 \) to \( B_2 \) as

\[
e(B_1,B_2 ; x) = \frac{w_2(x)}{w_1(x)}.
\]

From (5.1), we see that for each quantile \( x = F^{-1}(q) \), this efficiency is independent of the shape of \( F \).
In the case of the B band, \( g_{\ast}(u) = -k_B \) and \( g_{\ast}(u) = k_B \),
where \( k_B \) is the asymptotic critical value for \( \sqrt{n} B_n \). For the A-band, we set \( K = k \varepsilon \) in \( \lambda_1^* \) and \( \lambda_2^* \) and differentiate with respect to \( \varepsilon \). This yields

\[
\begin{align*}
g_{\ast}(u) &= -\frac{1}{2} k (8u)^{\frac{3}{2}}, \quad 0 < a \leq u \leq \frac{1}{2} \\
&= -\frac{1}{2} k \{8(1-u)\}^{\frac{1}{2}}, \quad \frac{1}{2} < u \leq 1 - a \\
g_{\ast}(u) &= \frac{1}{2} k (8u)^{\frac{3}{2}}, \quad 0 < a \leq u < \frac{1}{2} \\
&= \frac{1}{2} k \{8(1-u)\}^{\frac{1}{2}}, \quad \frac{1}{2} \leq u \leq 1 - a.
\end{align*}
\]

It follows that

\[
e(B,A ; x_q) = 2 \left( \frac{k_A}{k_B} \right)^q, \quad a \leq q \leq \frac{1}{2}
\]

\[
= 2 \left( \frac{k_A}{k_B} \right)^q (1-q), \quad \frac{1}{2} \leq q \leq 1 - a
\]

where \( k_A \) is the asymptotic critical value for \( \sqrt{n} A_n \).

For our statistics, \( \Pr(T_n \geq k) \) and \( 2 \Pr(T_n^+ \geq k) \) are very close. Table 5 below gives the efficiency using this approximation. Alternatively, it can be interpreted as the asymptotic efficiency of the boundaries given in Remark 3.1.

**TABLE 5.**

The asymptotic efficiency of the B band to the A band when \( a = .05 \).

<table>
<thead>
<tr>
<th>a</th>
<th>.01</th>
<th>.05</th>
<th>.10</th>
<th>.15</th>
<th>.20</th>
<th>.30</th>
<th>.40</th>
<th>.50</th>
</tr>
</thead>
<tbody>
<tr>
<td>.01</td>
<td>.04</td>
<td>.18</td>
<td>.36</td>
<td>.54</td>
<td>.72</td>
<td>1.09</td>
<td>1.45</td>
<td>1.81</td>
</tr>
<tr>
<td>.05</td>
<td>.16</td>
<td>.31</td>
<td>.47</td>
<td>.63</td>
<td>.94</td>
<td>1.26</td>
<td>1.57</td>
<td></td>
</tr>
<tr>
<td>.10</td>
<td>.28</td>
<td>.42</td>
<td>.56</td>
<td>.84</td>
<td>1.12</td>
<td>1.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.15</td>
<td>.38</td>
<td>.51</td>
<td>.76</td>
<td>1.02</td>
<td>1.27</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.20</td>
<td>.46</td>
<td>.68</td>
<td>.91</td>
<td>1.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The efficiencies for \( q > \frac{1}{2} \) are the same as for \( 1-q \). Other values of \( \alpha \) yield similar results.

The table shows that the B band is wider in the tails than the A band while it is narrower in the middle.

Note that the A band holds when \( F_n(x) \in [\min\{a,k_A^2/n\} , \max\{1-a,1-(k_A^2/n)\}] \) while the B band holds for \( F_n(x) \in [k_B/\sqrt{n}, 1-k_B/\sqrt{n}] \). If \( a \) is chosen so that these intervals are about the same, the A band would be preferable unless one wants to give more weight to the central quantiles.

6. TESTS FOR SYMMETRY

We consider the problem of testing the hypothesis \( H_s : \) 
"\( F \) is symmetric about some point \( s \)" against the alternative that \( F \) is skew in the sense of Section 2. We can use the confidence bands of Section 3 as follows:

Let \( [h^*(x), h(x)] \) be a level \( (1-\alpha) \) confidence band for \( h(x) = G^{-1}(F(x)) \), then

\[
[h^*(x), h(x)] = \left[ \frac{1}{2}[x-h(x)], \frac{1}{2}[x-h^*(x)] \right]
\]

is a level \( (1-\alpha) \) confidence band for the symmetry function \( \theta(x) = \frac{1}{2}[x-h(x)] \).

Suppose first that the point of symmetry \( s \) in \( H_s \) is specified. Since \( \theta(x) \) reduces to the point of symmetry when \( F \) is symmetric, the test which rejects \( H_s \) when the horizontal line \( y=s \) is not in the band \( \{[\theta^*(x), \theta(x)], x \in R\} \) has level \( \alpha \).

In terms of the \( T_n \) statistic, this amounts to computing \( T_n(s) \) for the sample \( \{X_i - s\} \) obtaining \( T_n(s) \), say, and then rejecting when \( T_n(s) \geq K \).

Next suppose that \( s \) in \( H_s \) is not specified. Then the test which rejects when no horizontal line fits in the confidence band
[θ*(x), θ*(x)] has level at most \( \alpha \). This test amounts to rejecting when
\[
\sup_x \theta_*(x) - \inf_x \theta_*(x) > 0.
\]
In terms of the \( T_n \) statistic it consists of rejecting when
\[
\inf_s T_n(s) > K.
\]
This test is conservative with level less than \( \alpha \). In the case of the \( B_n \) statistic, it is easy to check that the level is much less than \( \alpha \) and this is a very rough test.

Another possibility is to estimate the point of symmetry using \( \hat{s} \), say, and then reject \( H_0 \) if \( \hat{s} \) does not fall in the confidence band. This is the same as rejecting when \( T_n(\hat{s}) > t \) for some \( t \). \( T_n(\hat{s}) \) is not distribution-free under the hypothesis and the shape of the symmetric distribution must be specified in order to carry out the test. It could be performed by obtaining a Monte Carlo estimate of the critical value for the distribution and sample size needed. It would then be a computer implemented test in the sense of Birnbaum (1974). A natural specification of \( F \) would be the normal distribution, in which case \( T_n(\hat{s}) \) would be a test of normality against skew alternatives.

The choice of the location estimate \( \hat{s} \) is critical in the performance of the \( T_n(\hat{s}) \) test. This is illustrated in Figure 3 where \( \tilde{F} \) and \( \tilde{G} \) are the distributions of \( X_i - s \) and \( -(X_i - s) \). When \( s \) is a small location parameter such as the median \( m \), we get the situation of Figure 3(a), while if \( s \) is a large location parameter such as the mean \( \mu \), we get Figure 3(b). Here \( B = B(s) \) denotes the a.s. limit
\[
\lim_{n \to \infty} B_n(s) = \sup_x |\tilde{F}(x) - \tilde{G}(x)|
\]
and we denote \( B^+ = B^+(s) = \lim_{n \to \infty} B_n^+(s) \).
When $F$ is skew to the right as in the figure, $B(s)$ is maximized by choosing $s$ as small as possible or as large as possible. From Doksum (1975), it is known that for skew $F$, the smallest location parameter is the median $m$ while if $F$ is strongly skew to the right as defined in Doksum (1975), then

$$
\eta(F) = \lim_{\alpha \to 0^+} \frac{1}{2} [F^{-1}(\alpha) + F^{-1}(1-\alpha)]
$$

is the largest. Thus one would expect $T_n(s)$ to have good power properties when $s$ is the sample median $M$ or

$$
\eta = \frac{1}{2}(x(1) + x(n)).
$$

In fact, if the support $S(F) = \{x : 0 < F(x) < 1\}$ of $F$ is of the form $(b, \infty)$, which is typically the case for skew models such as the gamma, then $\eta \to \infty$ a.s. and $B_n(\eta)$ tends to its maximum value 1 a.s.

If one is reluctant to use extreme order statistics, one could modify the above by considering

$$
\eta(\beta, F) = \frac{1}{2} [F^{-1}(\beta) + F^{-1}(1-\beta)]
$$

and

$$
\hat{\eta}(\beta) = \frac{1}{2} \left\{ \begin{array}{l}
\frac{x}{[n\beta]+1} + \frac{x}{(n-[n\beta])}
\end{array} \right\}
$$

for some fixed $\beta \in (0, \frac{1}{2})$.

The last idea for a test statistic is as follows (see Doksum (1975)). Since $\theta(x)$ is a measure of skewness which is constant when $F$ is symmetric, we can use

$$
\sup_x \frac{\theta(x) - \inf_x \theta(x)}{\sigma}
$$
as an index of skewness, where \( \sigma \) is a scale parameter. From the results of Doksum (1975), it follows that when \( F \) is strongly skew, this is estimated by
\[
\frac{\hat{\eta} - M}{\hat{\sigma}}
\]
where \( \hat{\sigma} \) is a scale estimate.

Alternatively, we could use
\[
J_\beta = \frac{\hat{\eta}(\beta) - M}{\hat{\sigma}}
\]
as a test statistic. David and Johnson (1956) had a similar idea and studied this statistic whose excellent power properties when \( n=100 \) have been reported by Resek (1974). David and Johnsen found \( \beta = .0125 \) asymptotically optimal while Resek uses \( \beta = .02 \) and \( .03 \). He found \( \beta = .02 \) the better value when \( n=100 \).

\( J_\beta \) is approximately normal with asymptotic mean and variance given in the above references.

The Monte Carlo results of the next section support the remarks of this section.

7. MONTE CARLO RESULTS

Let
\[
\begin{align*}
B_n(s) &= \sup_x |F_n(x+s) - G_n(x-s)| \\
\bar{B}_n^+(s) &= \sup_x \{G_n(x-s) - F_n(x+s)\} \\
\bar{B}_n^-(s) &= \sup_x \{G_n(x-s) - F_n(x+s)\} .
\end{align*}
\]

For estimates of \( s \) we consider the median \( M \), mean \( \bar{x} \) and mid-means \( \hat{\eta}_1 = \frac{1}{2}(x_1 + x_n) \) and \( \hat{\eta}_2 = \frac{1}{2}(x_2 + x_{n-1}) \). From Figure 3, we expect \( B_n^+(M) \) to be better than \( B_n^-(M) \), while for the other estimates \( \hat{\eta} \), we expect \( B_n^+(\hat{\eta}) \) to be better than
This is indeed the case, and we omit the poorer test statistics below.

The last test statistic considered is

\[ J = \frac{\hat{\theta} - M}{\hat{\theta}} \]

where \( \hat{\theta} = x ([.75n]+1) - x ([.25n]) \).

Since the \( B_n(\hat{s}) \) statistics have discrete distributions, it is necessary to use randomized tests to get a fair power comparison. The first part of Table 6 below gives the critical constants \( k \) and \( \gamma \) such that the tests of the form

\[ \Psi(x) = 1 \quad \text{if} \quad nB_n(\hat{s}) > k \\
= \gamma \quad \text{if} \quad nB_n(\hat{s}) = k \\
= 0 \quad \text{if} \quad nB_n(\hat{s}) < k \]

has level \( \alpha \) for the normal distribution. \( k \) and \( \gamma \) are estimated on the basis of 10,000 Monte Carlo trials. By comparing with Table 2, we see that the critical values for \( B_n(M) \) and \( B_n(\hat{x}) \) are smaller than those for \( B_n(s) \), while \( B_n(\hat{\eta}) \) and \( B_n(\hat{\eta}_2) \) have larger critical values than \( B_n(s) \).

The next part of the table gives power estimates based on 2000 trials for the indicated distributions, where \( \log N \) stands for log normal with \( \mu=0, \sigma=1 \). We find that the \( J \) test is best among all the tests, while among the \( B_n(\hat{s}) \) tests, \( B_n(\hat{\eta}) \) is the best one. By comparing with Stephens (1974, Table 3), we find that \( B_n(\hat{\eta}) \) is better than the Kolmogorov test with mean and variance estimated by the sample mean and variance. More importantly, \( B_n(\hat{\eta}) \) and \( J \) both compare favorably with the Shapiro-Wilk statistic \( W \), whose power estimates (from Stephens) are given in the last row.
The last part of the table considers robustness of level, i.e., if we have a symmetric distribution other than the normal, can we use the normal critical values? The table gives estimates based on 2000 trials of rejection probabilities for uniform and Cauchy distributions and indicates that except for $B^+(M)$ and $B_n(M)$, the rejection probabilities are too large for heavy tailed distributions.

### TABLE 6

Power comparisons.
Test for normality versus skewness. $\alpha = .05$, $n = 20$.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Critical constants</th>
<th>Power</th>
<th>Rejection probabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$k$ $\gamma$</td>
<td>$\chi^2_{10}$ $\chi^2_4$ $\chi^2_2$ $\log N$ $\chi^2_1$</td>
<td>Uniform Cauchy</td>
</tr>
<tr>
<td>$B^+_n(M)$</td>
<td>5 0.29</td>
<td>0.13 0.24 0.46 0.53 0.77</td>
<td>0.11 0.03</td>
</tr>
<tr>
<td>$B^-_n(M)$</td>
<td>6 0.70</td>
<td>0.08 0.17 0.34 0.40 0.67</td>
<td>0.14 0.03</td>
</tr>
<tr>
<td>$B^-_n(\bar{x})$</td>
<td>5 0.46</td>
<td>0.26 0.45 0.69 0.86 0.90</td>
<td>0.04 0.37</td>
</tr>
<tr>
<td>$B^-_n(\bar{\eta})$</td>
<td>11 0.62</td>
<td>0.32 0.54 0.75 0.88 0.91</td>
<td>0.01 0.42</td>
</tr>
<tr>
<td>$B^-_n(\bar{\eta}_2)$</td>
<td>12 0.08</td>
<td>0.22 0.39 0.61 0.81 0.81</td>
<td>0.00 0.79</td>
</tr>
<tr>
<td>$B^-_n(\bar{\eta}_2)$</td>
<td>9 0.82</td>
<td>0.25 0.44 0.66 0.78 0.86</td>
<td>0.02 0.34</td>
</tr>
<tr>
<td>$B^-_n(\bar{\eta}_2)$</td>
<td>10 0.41</td>
<td>0.15 0.29 0.49 0.66 0.74</td>
<td>0.01 0.59</td>
</tr>
<tr>
<td>$J$</td>
<td>0.97 0</td>
<td>0.33 0.58 0.80 0.91 0.95</td>
<td>0.01 0.39</td>
</tr>
<tr>
<td>$W$</td>
<td></td>
<td>0.29 0.50</td>
<td>0.93 0.97</td>
</tr>
</tbody>
</table>
Acknowledgement.

The authors are grateful to Sven Øivind Wille who programmed the Monte Carlo experiments.

This paper was prepared with the partial support of the National Science Foundation and the Norwegian Research Council for Science and the Humanities. Doksum's work was done while visiting the Department of Mathematics, University of Oslo.
Figure 1(a). \( w(x) \) for the \( \chi^2_{10} \) distribution

\[
\begin{align*}
    w_1(x) &= 13.57 + 0.55 \cdot w(x) \\
    m_1(x) &= 13.57 + 0.55 \cdot (2m-x)
\end{align*}
\]
Figure 1 (b). $t_1$ for the $\chi^2_{10}$ distribution

$t_1 = 2.16 + t/3.21$
Figure 1 (c). $\theta(x)$ for the $\chi^2_{10}$ distribution
Figure 2. The functions $\lambda_1^+(u)$ and $\lambda_2^+(u)$ on $[0,1]$ for $K = 0.6$. 
Figure 3 (a). The median subtracted from the sample
Figure 3 (b). The mean subtracted from the sample
REFERENCES


