Abstract. The smooth tests for testing uniformity were introduced by Neyman [15]. The data driven method of selecting the number of components in a smooth test for uniformity is discussed, including the first-order asymptotic null distribution, consistency, empirical critical values and Monte Carlo powers. The first-order asymptotic null distribution is not sufficiently precise for approximation tools. A substantial improvement is made in this paper by deriving a second-order approximation of the null distribution, which turns out to be very accurate in numerical examples. The approximations are based on the second-order behaviour of Schwarz's selection rule $S$ under uniformity. The new results on $S$ are of independent interest.

1. Introduction. Let $X_1, X_2, \ldots, X_n$ be i.i.d. r.v.'s. Consider the goodness-of-fit problem of testing the simple null hypothesis $H_0$ that the $X_i$'s have distribution function $F_0$, where $F_0$ is a given continuous distribution function. Without loss of generality we assume that under $H_0$ the distribution of $X_i$ is the uniform distribution on $[0, 1]$.

Neyman's [15] paper of 1937, called 'Smooth test' for goodness of fit, is seen as the starting point of a subbranch of goodness-of-fit tests, the smooth tests of goodness-of-fit. In fact, Neyman's test is a particular example of the whole class of smooth goodness-of-fit tests, taking the Legendre polynomials as the orthonormal system. In [16], pp. 19–22, it is memorized that Neyman's test is intended as an "optimal" competitor to Pearson's chi-squared test. Optimal means here that the test is locally uniformly most powerful symmetric, unbiased for testing uniformity against alternatives of the form

$$c(\theta)\exp\left\{\sum_{j=1}^{k} \theta_j \pi_j(x)\right\}, \quad 0 \leq x \leq 1,$$

where $\{\pi_j\}$ are the orthonormal Legendre polynomials on $[0, 1]$ and $c(\theta)$ is a normalizing constant. Alternatives of the form (1.1) were vaguely described

* Research supported by Grant MEN 341114.
as smooth. The test statistic for testing \(H_0: \theta_j = 0, j = 1, \ldots, k\), corresponding to uniformity, is given by

\[
N_k = \sum_{j=1}^{k} \left\{n^{-1/2} \sum_{i=1}^{n} \pi_j(X_i) \right\}^2.
\]

Replacing \(\pi_j\) in (1.2) by

\[
p_j(x) = \left\{1_{(d_{j-1},d_j)}(x) - (d_j-d_{j-1})\right\}(d_j-d_{j-1})^{-1/2},
\]

where \(1_A(x)\) denotes the indicator function of the set \(A\), we get Pearson’s chi-square test

\[
P_k = \sum_{j=1}^{k} \left\{n^{-1/2} \sum_{i=1}^{n} p_j(X_i) \right\}^2 = \sum_{j=1}^{k} (O_j-E_j)^2/E_j
\]

with partition \(0 = d_0 < d_1 < \ldots < d_k = 1\), \(O_j\) and \(E_j\) denoting the observed and expected (under uniformity) number of observations in the \(j\)-th interval, respectively.

Hamdan [6, 8] implemented a suggestion of Lancaster to use different orthonormal systems, leading to what nowadays is called the class of smooth tests for goodness-of-fit, given by the test statistics

\[
T_k = \sum_{j=1}^{k} \left\{n^{-1/2} \sum_{i=1}^{n} \phi_j(X_i) \right\}^2,
\]

where \(\phi_0, \phi_1, \ldots\) is an orthonormal system in \(L_2([0, 1])\) with \(\phi_0(x) \equiv 1\). For more extensive and general information on the class of smooth tests we refer to Rayner and Best [18]. After the papers of Barton [2–4], Watson [21], and Hamdan [6–8] had appeared, there was no much progress in the field. A possible reason is that the required computations are quite heavy by hand. This is no longer a barrier with modern computers. Although smooth tests arose little interest for several years, nowadays Neyman’s paper [15] is considered to be ingenious (cf. Le Cam and Lehmann [12], p. ix). Recently, there was a renewed interest in smooth tests. For an overview we refer to Rayner and Best [19], who conclude in reviewing several tests of fit: “don’t use those other methods — use a smooth test!” The same conclusion is derived in Milbrodt and Strasser [14], p. 14. One of the important issues in applying \(T_k\) is the number \(k\) of components. We discuss three different approaches.

In the first one, \(k\) (and the orthonormal system) are selected such that the alternatives of particular interest are represented using only the first \(k\) components of the orthonormal system with \(k\) as small as possible. It requires that the user not automatically applies the test, but thinks why the goodness-of-fit test is executed and what types of alternatives are of particular interest. This imperative task is useful on itself, and therefore the method is more advantageous than disadvantageous. While especially focused on the selected (type of)
alternatives, the test keeps its omnibus character. A criterion of a simple structure to implement the idea is presented and extensively motivated in Inglot et al. [9]. This approach is closest to the classical application of smooth tests, since the test statistic is not modified and a formal test is executed.

The second approach is a more data-analytic approach. Obviously, it is not valid to apply a lot of significance tests, by varying $k$, to a data set and focus only on the most critical of them. Neyman was very clear about it as can be seen from the citation on p. 47 of [18]. Nevertheless, if the null hypothesis is rejected, the components may be used informally to suggest the nature of the departure from the null hypothesis. This second approach is described in detail in [18].

Returning to formal testing theory, the third approach concerns a data driven selection of $k$. This idea was introduced by Ledwina [13]. Roughly speaking, it works as follows. First, Schwarz’s [20] selection rule is applied to find a suitable dimension $S$, say, of an exponential family model for the data. Then Neyman’s test is applied within the fitted model, resulting in the test statistic $N_S$. So Schwarz’s rule serves as a kind of first selection, followed by a more precise instrument, being Neyman’s test in the “right” dimension. As with the original test of Neyman, the data driven version can also be applied using other orthonormal systems than the Legendre polynomials.

Observe that the “data driven approach” has also the extra advantage lying behind the idea of the second approach. Namely, when rejecting the null hypothesis, automatically a well-defined alternative model is provided for the data at hand.

The rest of this paper is mainly concerned with this third approach. In Section 2 the method is described and properties of the tests as obtained by Ledwina [13] and Kallenberg and Ledwina [11] are discussed. In Section 3 new results on Schwarz’s selection rule are presented. The results may be of independent interest. They are applied in Section 4, where an accurate new approximation of the null distribution of the data driven version of smooth goodness-of-fit tests is derived.

2. A data driven version of smooth goodness-of-fit tests. Consider the class of test statistics given by (1.5). Assume that the functions $\phi_1, \phi_2, \ldots$ are bounded, but not necessarily uniformly bounded in $j = 1, 2, \ldots$. Define for $k = 1, 2, \ldots$ exponential families by their densities $p_\theta(x)$ with respect to Lebesgue measure on $[0, 1]$ of the form

\[(2.1) \quad p_\theta(x) = \exp \{ \theta \circ \phi(x) - \psi_k(\theta) \}, \]

where

\[(2.2) \quad \theta = (\theta_1, \ldots, \theta_k), \quad \phi = (\phi_1, \ldots, \phi_k), \quad \psi_k(\theta) = \log \int_0^1 \exp \{ \theta \circ \phi(x) \} \, dx, \]
and $\circ$ stands for the inner product in $\mathbb{R}^k$. When there is no confusion, the dimension $k$ is sometimes suppressed in the notation.

It is well known that

\begin{equation}
\lambda(\theta) = \psi'(\theta) = \mathbb{E}_\theta \phi(X),
\end{equation}

where $'$ denotes derivative. Moreover, by orthonormality, we have

\begin{equation}
\lambda(0) = 0, \quad \psi''(0) = I \text{ (the identity matrix)}.
\end{equation}

Writing

\begin{equation}
Y_n = (\phi_1, \ldots, \phi_k), \quad \phi_j = n^{-1} \sum_{i=1}^{n} \phi_j(x_i)
\end{equation}

we obtain the expression for the density of $X_1, \ldots, X_n$, each $X_i$ having density (2.1), in the form

\begin{equation}
\exp \{n[Y_n \circ \theta - \psi_k(\theta)]\}.
\end{equation}

Schwarz's [20] Bayesian information criterion (BIC) for choosing sub-models corresponding to successive dimensions yields

\begin{equation}
S = \min \{k: 1 \leq k \leq d(n), \sup_{\theta \in \mathbb{R}^k} \{Y_n \circ \theta - \psi_k(\theta)\} - \frac{1}{2} k \log n \geq \sup_{t \in \mathbb{R}^j} \{Y_n \circ t - \psi_j(t)\} - \frac{1}{2} j \log n, \ j = 1, \ldots, d(n)\}.
\end{equation}

Although it is not mentioned in the notation, $S$ depends of course on the upper bound $d(n)$ of the exponential families under consideration. Schwarz's BIC is further discussed in Section 3.

The data driven smooth test statistic is defined by

\begin{equation}
T_S = \sum_{j=1}^{S} \left\{n^{-1/2} \sum_{i=1}^{n} \phi_j(X_i)\right\}^2
\end{equation}

with $S$ given by (2.7). The null hypothesis of uniformity is rejected for large values of $T_S$.

The idea behind (2.8) is that Schwarz's rule gives a first indication about the true density of the observations by fitting an exponential family model to the data and that the finishing touch comes from the smooth test in the selected exponential family.

Numerical results in, e.g., [9] show that Neyman's test performs very well provided a good choice of $k$ has been made. On the other hand, examples in Table 2 of [9] and in Tables 3–5 in [11] show that a considerable loss of power may occur when a wrong choice of $k$ is made. This illustrates that a good procedure for choosing $k$ based on the data is very welcome. Both by theoretical results and by the Monte Carlo simulation it will be argued that $T_S$ is such a procedure. To do this we start by considering $N_S$, the data driven
version of Neyman's test, i.e., (2.8) with \( \{\phi_j\} = \{\pi_j\} \), the orthonormal Legendre polynomials. The asymptotic null distribution of \( N_s \) is given by Corollary 3.5 of [11] and reads as follows. (By \( \chi^2 \) we denote an r.v. with a chi-square distribution with 1 degree of freedom.)

**THEOREM 2.1.** If

\[
\lim_{n \to \infty} d^3(n)n^{-1} \log n = 0,
\]

then

\[
N_s \xrightarrow{d} \chi^2_1 \text{ under } H_0.
\]

Using this limit theorem as an approximation theorem we obtain the asymptotic 0.05 and 0.01 critical values equal to 3.841 and 6.635, respectively. To see how this approximation performs compared to the finite sample case, simulated 0.05 and 0.01 critical values are presented in Table 1 (partially from Table 2 of [13] and Table 2 of [11]).

The accuracy of the simulation results are seen in Table 2, where exemplary empirical sizes of \( N_s \) are calculated for empirical critical values from

| Table 1. 5% and 1% critical values of \( N_s \) simulated for subsequent values of \( d(n) \), from 10,000 samples in each case |
|---|---|---|---|---|---|---|---|---|---|---|---|
| \( d(n) \) | \( n \) | 20 |  | 50 |  | 80 |  | 100 |  | 120 |  |
|   |   | 0.05 | 0.01 | 0.05 | 0.01 | 0.05 | 0.01 | 0.05 | 0.01 | 0.05 | 0.01 |
| 11 |  |  |  |  |  |  |  |  |  |  |
| 12 |  |  |  |  |  |  |  |  |  |  |

| Table 2. Empirical 5% sizes of \( N_s \) simulated for subsequent values of \( d(n) \), from 10,000 samples in each case |
|---|---|---|---|---|---|---|---|---|---|
| \( n \) | \( d(n) \) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 20 | 0.0540 | 0.0491 | 0.0471 | 0.0470 | 0.0452 | 0.0469 | -- | -- | -- |
| 50 | 0.0506 | 0.0489 | 0.0500 | 0.0499 | 0.0502 | 0.0504 | 0.0496 | 0.0497 | 0.0497 |
| 80 | 0.0480 | 0.0477 | 0.0498 | 0.0499 | 0.0500 | 0.0500 | 0.0500 | 0.0500 | 0.0500 |
Table 1 on pseudorandom sequences different from those used to prepare Table 1.

It is clearly seen from Table 1 that the difference between the asymptotic 0.05 (0.01) critical value equal to 3.841 (6.635) and the simulated ones for $d(n) \geq 2$ is substantial, even for $n = 120$. Therefore, we may conclude that the first-order limiting theorem (Theorem 2.1) is not sufficiently precise for approximation tools, and hence there is a need for a second-order limiting theorem to improve the accuracy of the approximation of $N_S$ under $H_0$. This is done, more generally, for $T_S$ in Section 4 on the basis of the results of Section 3.

To show that $N_S$ has good power properties under a wide class of alternatives we consider its consistency under the alternative distribution $P$ on $[0, 1]$. Suppose that

\begin{equation}
E_P \pi_1(X) = \ldots = E_P \pi_{K-1}(X) = 0, \quad E_P \pi_K(X) \neq 0
\end{equation}

for some $K = K(P)$, $X$ having distribution $P$. Essentially any alternative of interest satisfies (2.11). (Note that the orthonormal system $\{\pi_j\}$ is complete.) The consistency of $N_S$ is given by Corollary 4.5 of [11] and reads as follows.

**Theorem 2.2.** If (2.11) holds, and if

\begin{equation}
\liminf_{n \to \infty} d(n) \geq K(P), \quad \lim_{n \to \infty} d^3(n)n^{-1}\log n = 0,
\end{equation}

then the test based on $N_S$ is consistent at $P$.

Since $K(P)$ is fixed, (2.9) together with $\lim_{n \to \infty} d(n) = \infty$ implies that the test based on $N_S$ is consistent against any alternative of the form (2.11). Apart from the theoretical justification in Theorem 2.2, the good power properties of $N_S$ for a wide range of alternatives are shown by simulation results in Tables 3 and 4 of [13] and Tables 3–5 of [11]. For illustration here we present in one table simulated powers of $N_S$, the widely recommended tests of Watson (W), Anderson and Darling (AD), Neuhaus (N) and the recently introduced procedures of Bickel and Ritov (BR), and Eubank and LaRiccia (ELR). For definitions of the several test statistics we refer to [13] and [11]. The alternatives under consideration are given by the following densities:

\begin{align*}
g_1(x) &= \frac{1}{4}\{x^{-1/2} + (1-x)^{-1/2}\}, \quad g_2(x) = 1.5(1-x)^{1/2}, \\
g_3(x) &= (3\sqrt{2})\{x^{1/2}1_{[0,1/3]}(x) + (1-x)^{1/2}1_{(1/2,1]}(x)\}, \quad g_4(x) = 1 + \frac{1}{2}\cos(2\pi x), \\
g_5(x) &= \exp\{-0.5\pi_2(x) - 0.2\pi_4(x) - \psi(0, -0.5, -0.2)\} \text{ (cf. (2.1)).}
\end{align*}
Table 3 and other simulation results show that $N_S$ has a stable and relatively high power for a broad range of alternatives.

**TABLE 3. Estimated powers (in %) based on 10000 samples in each case; $\alpha = 0.05$, $n = 50$, $d(n) = 10$**

<table>
<thead>
<tr>
<th>Alternatives</th>
<th>$N_S$</th>
<th>W</th>
<th>AD</th>
<th>N</th>
<th>BR</th>
<th>ELR</th>
</tr>
</thead>
<tbody>
<tr>
<td>$g_1$</td>
<td>81</td>
<td>65</td>
<td>69</td>
<td>64</td>
<td>63</td>
<td>63</td>
</tr>
<tr>
<td>$g_2$</td>
<td>58</td>
<td>39</td>
<td>71</td>
<td>62</td>
<td>60</td>
<td>24</td>
</tr>
<tr>
<td>$g_3$</td>
<td>61</td>
<td>59</td>
<td>16</td>
<td>50</td>
<td>43</td>
<td>24</td>
</tr>
<tr>
<td>$g_4$</td>
<td>57</td>
<td>61</td>
<td>33</td>
<td>56</td>
<td>48</td>
<td>27</td>
</tr>
<tr>
<td>$g_5$</td>
<td>63</td>
<td>55</td>
<td>15</td>
<td>46</td>
<td>40</td>
<td>25</td>
</tr>
</tbody>
</table>

The theoretical results on the asymptotic null distribution of $N_S$ and its consistency may be extended to other orthonormal systems as well. To do this we make some assumptions. Define

$$V_k = \max_{1 \leq j \leq k} \sup_{x \in [0,1]} |\phi_j(x)|.$$  \hspace{1cm} (2.13)

**ASSUMPTION 1. The following condition holds:**

$$\lim_{n \to \infty} d(n) V_{d(n)}(n^{-1} \log n)^{1/2} = 0.$$ \hspace{1cm} (2.14)

For the orthonormal Legendre polynomials on $[0, 1]$ we get $V_k = (2k + 1)^{1/2}$. The asymptotic null distribution and the consistency of $T_S$ are given by Theorems 3.4 and 4.4 in [11] and read as follows:

**THEOREM 2.3. If Assumption 1 holds, then under $H_0$**

$$T_S \overset{d}{\to} \chi_1^2,$$ \hspace{1cm} (2.15)

**and the test based on $T_S$ is consistent against any alternative $P$ of the form (2.11), provided that $\lim \inf_{n \to \infty} d(n) \geq K(P)$.**

3. **Schwarz's selection rule.** It is seen in Section 2 that the first-order result $T_S \overset{d}{\to} \chi_1^2$ is not sharp enough to imply accurate approximations of the null distribution. To improve the approximations we need some second-order results. Both for technical reasons and for more insight in what is going on, we have to derive precise results on the behaviour of $S$ under $H_0$. The notation $P_0$ refers to the null hypothesis, where $X_1$ has the uniform distribution on $[0, 1]$. While at the first order only the event \{S = 1\} plays a role (cf. Theorems 3.2 and 3.4 in [11]), for the second-order results both \{S = 1\} and \{S = 2\} come in
the picture. The event \( \{ S \geq 3 \} \) may still be neglected under \( H_0 \) as is seen in the following theorem:

**Theorem 3.1.** If Assumption 1 holds, then

\[
P_0(S \geq 3) = O(n^{-1}(\log n)^{1/2}) \quad \text{as } n \to \infty.
\]

**Proof.** Let \( K \) be fixed with \( 3 \leq K \leq d(n) \). Write

\[
P_0(S \geq 3) = \sum_{k=3}^{K} P_0(S = k) + \sum_{k=K+1}^{d(n)} P_0(S = k).
\]

By the definition of \( S \) and the application of Lemma 3.2 in [10], there exists for each \( k \in \{3, \ldots, K\} \) a constant \( c_k^* \) and there exists a constant \( \tilde{c}_k \) such that

\[
P_0(S = k) \leq P_0(\max_{3 \leq j \leq d(n)} \{ Y_j \cap \varphi_k(j) \} \geq \frac{1}{2}(k-1)\log n)
\]

\[
\leq c_k^* (\log n)^{(1/2)(k-2)} \exp\{-\frac{1}{2}(k-1)\log n\} \leq \tilde{c}_k (\log n)^{1/2} n^{-1}.
\]

For \( k = K + 1, \ldots, d(n) \) it follows from Lemma 3.1 in [11] that for

\( \varepsilon \in (0, \min \{1, \frac{3}{2} k V_k^2\}) \),

uniformly for \( k \in \{1, \ldots, d(n)\} \),

\[
P_0(S = k) \leq P_0(n Y_n \cap Y_n \geq (1-\varepsilon)(k-1)\log n),
\]

provided that \( n \) is sufficiently large. Note that here we use Assumption 1. Next we apply formula (2) of [17] with

\[
\varrho = \{(1-\varepsilon)(k-1)\log n\}^{1/2}, \quad m = k, \quad \lambda = 1,
\]

\[
a = k^{1/2} V_k n^{-1/2} \{(1-\varepsilon)(k-1)\log n\}^{1/2},
\]

which yields

\[
P_0(n Y_n \cap Y_n \geq (1-\varepsilon)(k-1)\log n) \leq c_1 \left( \frac{\varrho^2}{2} \right)^{(1/2)(k-1)} \frac{1}{\Gamma(k/2)} \exp\{-\frac{1}{2} \varrho^2 (1-\eta(a))\},
\]

where \( c_1 \) is an absolute constant, \( \eta(a) \to 0 \) as \( a \to 0 \), provided that \( \varrho^2/2 \geq k \) and \( a \leq 1 \) (\( \eta(a) \) is explicitly given on p. 188 of [17] and satisfies \( \eta(a) \leq a \)). In view of Assumption 1, \( a \to 0 \) as \( n \to \infty \) (uniformly for \( k \in \{1, \ldots, d(n)\} \)). Moreover, \( \varrho^2/2 \geq k \) for sufficiently large \( n \) (independent of \( k \)). By taking \( \varepsilon \) small enough we therefore get for any \( \zeta > 0 \) and \( k \in \{1, \ldots, d(n)\} \)

\[
P_0(S = k) \leq n^{-(1/2)(1-\zeta(k-1))}
\]
Data driven Neyman's tests

417

for \( n = n(\zeta) \) large enough, independent of \( k \in \{1, \ldots, d(n)\} \). Taking, for instance,
\( K = 3 \) and \( \zeta = 0.2 \) we get for \( n \) sufficiently large

\[
P_0(S \geq 3) \leq \tilde{c}_3 (\log n)^{1/2} n^{-1} + \sum_{k=4}^{\infty} n^{-0.4(k-1)} = \tilde{c}_3 (\log n)^{1/2} n^{-1} + \frac{n^{-1.2}}{1 - n^{-0.4}},
\]

and hence (3.1) holds. \( \blacksquare \)

Remark 3.1. It is clear from the proof that the main contribution comes from \( S = 3 \), as expected. In other words, by virtually the same proof (cf. (3.3) and (3.4)), it follows that for each fixed \( K \)

\[
(3.5) \quad P_0(S \geq K) = O(n^{-(1/2)(K-1)}(\log n)^{(1/2)(K-2)}) \quad \text{as } n \to \infty. \quad \blacksquare
\]

In view of Theorem 3.1 we may concentrate on the events \( \{S = 1\} \) and \( \{S = 2\} \). The event \( \{S = 1\} \) means that dimension 1 is more important than dimension 2, dimension 3, dimension 4, etc. Under \( H_0 \) it turns out that the comparison of dimension 1 with dimension 2 is up to the second order the only comparison of interest. Similarly, on the event \( \{S = 2\} \) the only important contribution under \( H_0 \) comes from the comparison of dimension 2 with dimensions 1 and 3. This is shown in the next two theorems.

**Theorem 3.2.** There exists a positive constant \( c_2 \) such that for each event \( A \)

\[
(3.6) \quad P_0(A, S = 1) \leq P_0(A, n \phi_2^3 \leq \log n + c_2 n^{-1/2} (\log n)^{3/2}) + O(n^{-1} (\log n)^{-1/2})
\]

and

\[
(3.7) \quad P_0(A, S = 1) \geq P_0(A, n \phi_2^3 \leq \log n - c_2 n^{-1/2} (\log n)^{3/2})
+ O(n^{-1} (\log n)^{-1/2}) - P_0(S \geq 3)
\]

as \( n \to \infty \), uniformly in \( A \).

**Proof.** By the definition of \( S \) we get

\[
P_0(A, S = 1) \leq P_0(A, n \sup_{\theta \in \mathbb{R}^k} \{ Y_n g_j - \psi_j(\theta) \} - \frac{1}{2} \log n \geq n \sup_{t \in \mathbb{R}^2} \{ Y_n \circ t - \psi_2(t) \} - \log n). \]

For each fixed \( k \) we have

\[
\sup_{\theta \in \mathbb{R}^k} \{ y \circ \theta - \psi_k(\theta) \} = \frac{1}{2} \| y \|^2 + O(\| y \|^3) \quad \text{as } y \to 0,
\]

where \( \| \cdot \| \) denotes the Euclidean norm in \( \mathbb{R}^k \). Therefore, on the set

\[
\{ y: |y| \leq n^{-1/2}(2\log n)^{1/2}, j = 1, \ldots, k \}
\]

we get

\[
|\sup_{\theta \in \mathbb{R}^k} \{ y \circ \theta - \psi_k(\theta) \} - \frac{1}{2} \| y \|^2 | \leq \tilde{c}_k n^{-3/2} (\log n)^{3/2}
\]
for some constant \( \bar{c}_i \). Hence, uniformly in \( A \),

\[
P_0(A, S = 1) \leq P_0(A, n\{\frac{1}{2}\phi_1^2 + \bar{c}_1 n^{-3/2}(\log n)^{3/2}\} + \frac{1}{2}\log n
\geq n\{\frac{1}{2}(\phi_1^2 + \phi_2^2) - \bar{c}_2 n^{-3/2}(\log n)^{3/2}\}
+ P_0(n^{1/2}|\phi_1| > (2\log n)^{1/2}) + P_0(n^{1/2}|\phi_2| > (2\log n)^{1/2})
= P_0(A, n\phi_2^2 \leq \log n + c_2 n^{-1/2}(\log n)^{3/2}) + O(n^{-1}(\log n)^{-1/2})
\]

as \( n \to \infty \), with \( c_2 = 2(\bar{c}_1 + \bar{c}_2) \), where for \( j = 1, 2 \) we have used

\[
P_0(n^{1/2}|\phi_j| > (2\log n)^{1/2}) = O(n^{-1}(\log n)^{-1/2}),
\]

which follows by standard large deviation theory. This completes the proof of (3.6).

To prove (3.7) note that

\[
P_0(A, S = 1) \geq P_0(A, n \sup_{t \in R}|Y_n \circ t - \psi_2(t)| - \log n - P_0(S \geq 3).\]

The rest of the proof is similar to the proof of (3.6).

**Theorem 3.3.** Let \( d(n) \geq 2 \). There exists a positive constant \( c_3 \) such that for each event \( A \)

\[
P_0(A, S = 1) \leq P_0(A, n\phi_2^2 \geq \log n - c_3 n^{-1/2}(\log n)^{3/2},
\]

\[
n\phi_2^2 \leq \log n + c_3 n^{-1/2}(\log n)^{3/2}) + O(n^{-1}(\log n)^{-1/2})
\]

and

\[
P_0(A, S = 2) \geq P_0(A, n\phi_3^2 \geq \log n + c_3 n^{-1/2}(\log n)^{3/2},
\]

\[
n\phi_3^2 \leq \log n - c_3 n^{-1/2}(\log n)^{3/2}) + O(n^{-1}(\log n)^{-1/2}) - P_0(S \geq 3)
\]

as \( n \to \infty \), uniformly in \( A \).

**Proof.** By the definition of \( S \) we get (as in the proof of Theorem 3.2)

\[
P_0(A, S = 2)
\leq P_0(A, n \sup_{t \in R}|Y_n \circ t - \psi_2(t)| - \log n \geq n \sup_{t \in R}|Y_n \circ t - \psi_1(t)| - \frac{1}{2}\log n,
\]

\[
n \sup_{t \in R}|Y_n \circ t - \psi_2(t)| - \log n \geq n \sup_{\theta \in R}|Y_n \circ \theta - \psi_3(\theta)| - \frac{1}{2}\log n
\]

\[
\leq P_0(A, n\{\frac{1}{2}(\phi_1^2 + \phi_2^2) + \bar{c}_2 n^{-3/2}(\log n)^{3/2}\}
- \log n \geq n\{\frac{1}{2}(\phi_1^2 + \phi_2^2) + \bar{c}_3 n^{-3/2}(\log n)^{3/2}\}
\]

\[
- \log n \geq n\{\frac{1}{2}(\phi_1^2 + \phi_2^2) + \phi_3^2 - \bar{c}_3 n^{-3/2}(\log n)^{3/2}\} - \frac{1}{2}\log n
\]
By taking 
\[ c_3 = \max\{2(\tilde{c}_1 + \tilde{c}_2), 2(\tilde{c}_2 + \tilde{c}_3)\} \]
the proof of (3.8) is completed. Inequality (3.9) is proved in a similar way.

Remark 3.2. The \( O(n^{-1}(\log n)^{1/2}) \)-terms in (3.6)–(3.9) may be replaced by \( O(n^{-r}(\log n)^{-1/2}) \) for arbitrary \( r \). This is seen by replacing the set \( \{ y: |y_j| \leq n^{-1/2}(2\log n)^{1/2}, j = 1, \ldots, k \} \) by the set \( \{ y: |y_j| \leq n^{-1/2}(2r\log n)^{1/2}, j = 1, \ldots, k \} \) and noting that

\[
P_0(n^{1/2}|\tilde{\phi}| > (2r\log n)^{1/2}) = O(n^{-r}(\log n)^{-1/2}) \quad \text{as } n \to \infty.
\]

4. Second-order approximation to the null distribution of \( T_S \). As was seen from Theorem 2.1 and Table 1, the first-order \( \chi^2 \)-approximation of the null distribution of \( T_S \) is not very precise. Therefore, we consider the second-order approximation, based on the second-order limiting theorems for \( S \) under \( H_0 \) in the previous section. Denoting by \( U_1, U_2 \) two independent r.v.'s, each with a standard normal distribution, we shall prove the following result:

**Theorem 4.1.** Let \( d(n) \geq 2 \). If Assumption 1 holds, and Cramér’s condition is satisfied, i.e.,

\[
(4.1) \quad \limsup_{|t| \to \infty} |E_{P_0} \exp \{ it \phi_1(X_1) \}| < 1,
\]

then

\[
(4.2) \quad P_0(T_S \leq x) = \Pr(U_1^2 \leq x)\Pr(U_2^2 \leq \log n) + \Pr(U_1^2 + U_2^2 \leq x, U_2^2 \geq \log n) + O(n^{-1}\log n)
\]

as \( n \to \infty \), uniformly in \( x \).

It is easily seen that the approximation on the right-hand side of (4.2) may also be written as

\[
(4.3) \quad \Pr(U_1^2 \leq x) - \Pr(U_1^2 \leq x, U_1^2 + U_2^2 \geq x, U_2^2 \geq \log n).
\]

In this formulation the correction term in comparison to the first-order approximation \( \Pr(U_1^2 \leq x) \) is clearly seen. Corresponding to the simulation results in Table 1, indeed, \( P_0(T_S \leq x) \) is overestimated by its first-order approximation. The two terms on the right-hand side of (4.2) are the leading terms of \( P_0(T_S \leq x, S = 1) \) and \( P_0(T_S \leq x, S = 2) \), respectively. For \( x \leq \log n \)}
the second term disappears and the second-order approximation in that case is simply obtained by multiplying the first-order approximation by $\Pr(U_2^2 \leq \log n)$. Condition (4.1) is satisfied if $\phi_i(X)$ has an absolutely continuous component, and therefore it is fulfilled in many cases, e.g., for Neyman's test with the orthonormal Legendre polynomials. To prove Theorem 4.1 we state and prove first several lemmas.

**Lemma 4.1.** For each fixed $k$ and for each constant $c_4 \in \mathbb{R}$ we have

\[ P_0(n\phi_k^2 \leq \log n + c_4 n^{-1/2}(\log n)^{3/2}) - P_0(n\phi_k^2 \leq \log n) = O(n^{-1}\log n). \]

**Proof.** By standard large deviation theory (cf., e.g., [5], p. 553) we have

\[ \Pr(U_i \leq \log n) = \Pr(U_i \leq \log n + c_4 n^{-1/2}(\log n)^{3/2}) \]

\[ = 2\Pr(U_1 > (\log n)^{1/2})\{1 + O(n^{-1/2}(\log n)^{3/2})\} \]

\[ - 2\Pr(U_1 > \{\log n + c_4 n^{-1/2}(\log n)^{3/2}\})\{1 + O(n^{-1/2}(\log n)^{3/2})\} \]

\[ = O(n^{-1}\log n) \quad \text{as } n \to \infty. \]

**Lemma 4.2.** If Assumption 1 holds, then

\[ P_0(T_6 \leq x, S = 1) = P_0(n\phi_1^2 \leq \log n) + O(n^{-1}\log n) \]

as $n \to \infty$, uniformly in $x$.

**Proof.** Taking $A = \{T_1 \leq x\}$ we see by Theorem 3.2 that there exists a positive constant $c_2$ such that

\[ P_0(T_6 \leq x, S = 1) = P_0(T_1 \leq x, S = 1) \]

\[ \leq P_0(T_1 \leq x, n\phi_2^2 \leq \log n + c_2 n^{-1/2}(\log n)^{3/2}) + O(n^{-1}(\log n)^{-1/2}) \]

\[ \leq P_0(T_1 \leq x, n\phi_2^2 \leq \log n) + P_0(n\phi_2^2 \leq \log n + c_2 n^{-1/2}(\log n)^{3/2}) \]

\[ - P_0(n\phi_2^2 \leq \log n) + O(n^{-1}(\log n)^{-1/2}) \]

as $n \to \infty$, uniformly in $x$. The application of Lemma 4.1 and noting that $T_1 = n\phi_1^2$ yield

\[ P_0(T_6 \leq x, S = 1) \leq P_0(n\phi_1^2 \leq \log n, n\phi_2^2 \leq \log n) + O(n^{-1}\log n) \]

as $n \to \infty$, uniformly in $x$. Similarly, the converse inequality is obtained by using (3.7) in combination with Theorem 3.1. ■

**Lemma 4.3.** For each constant $c_5 \in \mathbb{R}$ we have

\[ P_0(n\phi_2^2 \geq \log n, n\phi_3^2 > \log n - c_5 n^{-1/2}(\log n)^{3/2}) = O(n^{-1}(\log n)^{-1}) \]

as $n \to \infty$. 

W. C. M. Kallenberg and T. Ledwina
Proof. Write
\[ D_n = \{ (y_1, y_2) : y_1 \geq (\log n)^{1/2}, y_2 \geq (\log n - c_5 n^{-1/2} (\log n)^{3/2})^{1/2} \}. \]
The set \( D_n \) is a convex Borel set and the point
\[ a = ( (\log n)^{1/2}, (\log n - c_5 n^{-1/2} (\log n)^{3/2})^{1/2} ) \]
is the nearest point to the origin. The application of Assertion 2 in Remark 1 on p. 67 of [1] yields
\[
P_0( (n^{1/2} \bar{\phi}_2, n^{1/2} \bar{\phi}_3) \in D_n ) = \Pr((U_1, U_2) \in D_n) \{ 1 + O(n^{-1/2} (\log n)^{3/2}) \}
\]
\[ = O(n^{-1} (\log n)^{-1}) \quad \text{as } n \to \infty. \]

By the same argument for the other three regions of similar form, which together with \( D_n \) constitute the event in (4.7), the proof is completed. \( \blacksquare \)

**Lemma 4.4.** If Assumption 1 holds, then
\[
P_0(T_S \leq x, S = 2) = P_0(n(\bar{\phi}_1^2 + \bar{\phi}_2^3) \leq x, n\bar{\phi}_2^2 \geq \log n) + O(n^{-1} \log n) \quad (4.8)
\]
as \( n \to \infty \), uniformly in \( x \).

**Proof.** Taking \( A = \{ T_2 \leq x \} \) we see by Theorem 3.3 that there exists a positive constant \( c_3 \) such that
\[
P_0(T_S \leq x, S = 2) \\
\leq P_0(T_2 \leq x, n\bar{\phi}_2^2 \geq \log n - c_3 n^{-1/2} (\log n)^{3/2}) + O(n^{-1} (\log n)^{-1/2}) \\
\leq P_0(T_2 \leq x, n\bar{\phi}_2^3 \geq \log n) + P_0(\log n - c_3 n^{-1/2} (\log n)^{3/2} \leq n\bar{\phi}_2^2 < \log n) \\
+ O(n^{-1} (\log n)^{-1/2})
\]
as \( n \to \infty \), uniformly in \( x \). The application of Lemma 4.1 and noting that 
\( T_2 = n(\bar{\phi}_1^2 + \bar{\phi}_2^3) \) yield
\[
P_0(T_S \leq x, S = 2) \leq P_0(n(\bar{\phi}_1^2 + \bar{\phi}_2^3) \leq x, n\bar{\phi}_2^2 \geq \log n) + O(n^{-1} \log n)
\]
as \( n \to \infty \), uniformly in \( x \).

For the converse inequality we consider (3.9) in combination with Theorem 3.1 and we obtain
\[
P_0(T_S \leq x, S = 2) \\
\geq P_0(T_2 \leq x, n\bar{\phi}_2^2 \geq \log n) - P_0(\log n \leq n\bar{\phi}_2^2 < \log n + c_3 n^{-1/2} (\log n)^{3/2}) \\
- P_0(n\bar{\phi}_2^3 \geq \log n, n\bar{\phi}_3^2 > \log n - c_3 n^{-1/2} (\log n)^{3/2}) + O(n^{-1} (\log n)^{1/2}) \\
\geq P_0(T_2 \leq x, n\bar{\phi}_2^2 \geq \log n) + O(n^{-1} \log n)
\]
by Lemmas 4.1 and 4.3. This completes the proof of the lemma. \( \blacksquare \)
LEMMA 4.5. If (4.1) holds, then, uniformly for \( x \), as \( n \to \infty \)

\[
P_0(n\bar{\phi}_1^2 \leq x) = \Pr(U_1^2 \leq x) + O(n^{-1}).
\]

Proof. By the standard Edgeworth expansion (cf., e.g., [5], Theorem 3 on p. 541) it follows that, uniformly in \( x \),

\[
P_0(n\bar{\phi}_1^2 \leq x) = P_0(n^{1/2}\bar{\phi}_1 \leq x^{1/2}) - P_0(n^{1/2}\bar{\phi}_1 < -x^{1/2})
\]

\[
= \Pr(U_1 \leq x^{1/2}) - \left\{ \frac{E\phi_1^3}{6n^{1/2}} \phi(x^{1/2})(x-1) - \Pr(U_1 \leq -x^{1/2}) \right\}
\]

\[
+ \left\{ \frac{E\phi_1^3}{6n^{1/2}} \phi(-x^{1/2})(x-1) + O(n^{-1}) \right\}
\]

\[
= \Pr(U_1^2 \leq x) + O(n^{-1}),
\]

where \( \phi \) denotes the standard normal density. \( \blacksquare \)

LEMMA 4.6. Uniformly for \( x \) as \( n \to \infty \) we have

\[
P_0(n\bar{\phi}_1^2 \leq x, n\bar{\phi}_2^2 > \log n) = \Pr(U_1^2 \leq x) \Pr(U_2^2 > \log n) + O(n^{-1} \log n)
\]

and

\[
P_0(n(\bar{\phi}_1^2 + \bar{\phi}_2^2) \leq x, n\bar{\phi}_2^2 \geq \log n)
\]

\[
= \Pr(U_1^2 + U_2^2 \leq x, U_2^2 \geq \log n) + O(n^{-1} \log n).
\]

Proof. We give the proof of (4.11); the relation (4.10) can be shown in a quite similar way. Without loss of generality assume that \( x \geq \log n \). Write

\[
D_{n1} = \{(y_1, y_2); y_2 \geq (\log n)^{1/2}, y_1^2 + y_2^2 \leq x\},
\]

\[
D_{n2} = \{(y_1, y_2); y_2 \leq -(\log n)^{1/2}, y_1^2 + y_2^2 \leq x\}.
\]

The sets \( D_{n1} \) and \( D_{n2} \) are convex. It follows by Assertion 2 in Remark 1 on p. 67 of [1] that for \( i = 1, 2 \), with \( a = (0, (\log n)^{1/2}) \) if \( i = 1 \) and \( a = (0, -(\log n)^{1/2}) \) if \( i = 2 \), we have

\[
P_0((n^{1/2}\bar{\phi}_1, n^{1/2}\bar{\phi}_2) \in D_{n1}) = \Pr((U_1, U_2) \in D_{n1})\{1 + O(n^{-1/2}(\log n)^{3/2})\}
\]

\[
= \Pr((U_1, U_2) \in D_{n1}) + O(n^{-1} \log n),
\]

which implies (4.11). \( \blacksquare \)

Proof of Theorem 4.1. By (4.6), (4.8)–(4.11) and (3.1) we have, uniformly in \( x \), as \( n \to \infty \)

\[
P_0(T_S \leq x) = P_0(T_S \leq x, S = 1) + P_0(T_S \leq x, S = 2) + P_0(T_S \leq x, S \geq 3)
\]

\[
= P_0(n\bar{\phi}_1^2 \leq x) - P_0(n\bar{\phi}_1^2 \leq x, n\bar{\phi}_2^2 > \log n)
\]

\[
+ P_0(n(\bar{\phi}_1^2 + \bar{\phi}_2^2) \leq x, n\bar{\phi}_2^2 \geq \log n) + O(n^{-1} \log n)
\]
\[
= \Pr(U_1^2 \leq x) - \Pr(U_1^2 \leq x, U_2^2 > \log n) + \Pr(U_1^2 + U_2^2 \leq x, U_2^2 \geq \log n) + O(n^{-1}\log n),
\]
and hence (4.2) holds.

The approximation on the right-hand side of (4.2) is surprisingly simple and very accurate, especially when compared to the first-order approximation. The structure of the approximation clarifies that under \(H_0\) up to the second order only \(S = 1\) and \(S = 2\) are important. Moreover, \(S = 1\) and \(S = 2\) may be further restricted to dimension 1 beats dimension 2, corresponding to \(U_2^2 \leq \log n\), and dimension 2 beats dimension 1, corresponding to \(U_2^2 > \log n\), respectively. (As a side-note, remark that here also dimension 2 beats dimension 3 is cancelled out.)

Both to investigate the accuracy of the approximation and to facilitate its evaluation we derive an upper and lower bound of \(\Pr(U_1^2 + U_2^2 \leq x, U_2^2 \geq \log n)\). Let for \(x \geq \log n\)

(4.12) \[
D^* = \{(y_1, y_2) : y_1^2 + y_2^2 \geq x, 0 \leq y_1 \leq (x - \log n)^{1/2}, (\log n)^{1/2} \leq y_2 \leq x^{1/2}\}.
\]

The area \(a(x)\) of this set equals

(4.13) \[
a(x) = \left\{x^{1/2} - (\log n)^{1/2}\right\}(x - \log n)^{1/2} - \frac{1}{2}x \arccos \left(\frac{\log n}{x}\right)^{1/2}
+ \frac{1}{2}(\log n)^{1/2}(x - \log n)^{1/2},
\]
while for \((y_1, y_2) \in D^*\) we have

(4.14) \[
n^{1/2}\exp(-x) \leq \exp\left\{-\frac{1}{2}(y_1^2 + y_2^2)\right\} \leq \exp\left(-\frac{1}{2}x\right).
\]

Using symmetry we therefore get

(4.15) \[
\Pr(U_1^2 + U_2^2 \leq x, U_2^2 \geq \log n)
= 4\Pr(U_1^2 + U_2^2 \leq x, 0 \leq U_1 \leq (x - \log n)^{1/2}, (\log n)^{1/2} \leq U_2 \leq x^{1/2})
= 4\{\Pr(0 \leq U_1 \leq (x - \log n)^{1/2})\Pr((\log n)^{1/2} \leq U_2 \leq x^{1/2}) - (2n)^{-1}a(x)b(x)\}
\]
with

(4.16) \[
n^{1/2}\exp(-x) \leq b(x) \leq \exp\left(-\frac{1}{2}x\right).
\]

Replacing \(\Pr(U_1^2 + U_2^2 \leq x, U_2^2 \geq \log n)\) on the right-hand side of (4.2) by its lower and upper bounds obtained from (4.15) and (4.16) and inserting for \(x\) the empirical critical values of Table 1 we get the results compiled in Table 4. It is seen that for \(d(n) = 2\) the approximation is excellent, even for small values of \(n\). For \(d(n) \geq 3\) we have a slight overestimation, but the approximation is
also in that case rather accurate, and far more accurate than the first-order approximation.

As was noted before, if \( x \leq \log n \), the second term in the approximation (4.2) disappears. On the other hand, if \( x \geq 2 \log n \), then

\[
0 \leq \Pr(U_2^2 \geq \log n) - \Pr(U_1^2 + U_2^2 \leq x, U_2^2 \geq \log n) = \Pr(U_1^2 + U_2^2 > x) = \exp(-\frac{1}{2}x) \leq n^{-1},
\]

which implies for \( x \geq 2 \log n \)

\[
P_0(T_S \leq x) = \Pr(U_1^2 \leq x) \Pr(U_2^2 \leq \log n) + \Pr(U_2^2 > \log n) + O(n^{-1} \log n).
\]

### Table 4. Lower and upper bounds of the second-order approximation of \( P_0(T_S \leq x) \) for empirical critical values \( x \) of Table 1 (probabilities \( \times 10^3 \))

<table>
<thead>
<tr>
<th>( n )</th>
<th>( x )</th>
<th>Lower bounds</th>
<th>Upper bounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>5.367</td>
<td>94055</td>
<td>94912</td>
</tr>
<tr>
<td></td>
<td>6.537</td>
<td>96577</td>
<td>97341</td>
</tr>
<tr>
<td></td>
<td>7.213</td>
<td>97273</td>
<td>98182</td>
</tr>
<tr>
<td></td>
<td>7.617</td>
<td>97698</td>
<td>98552</td>
</tr>
<tr>
<td></td>
<td>7.854</td>
<td>97915</td>
<td>98733</td>
</tr>
<tr>
<td>50</td>
<td>5.350</td>
<td>94758</td>
<td>95028</td>
</tr>
<tr>
<td></td>
<td>5.865</td>
<td>95784</td>
<td>96204</td>
</tr>
<tr>
<td></td>
<td>6.015</td>
<td>96043</td>
<td>96493</td>
</tr>
<tr>
<td></td>
<td>6.094</td>
<td>96173</td>
<td>96637</td>
</tr>
<tr>
<td></td>
<td>6.096</td>
<td>96176</td>
<td>96641</td>
</tr>
<tr>
<td></td>
<td>6.117</td>
<td>96210</td>
<td>96678</td>
</tr>
<tr>
<td></td>
<td>6.121</td>
<td>96217</td>
<td>96685</td>
</tr>
<tr>
<td>80</td>
<td>5.260</td>
<td>94928</td>
<td>95028</td>
</tr>
<tr>
<td></td>
<td>5.508</td>
<td>95426</td>
<td>95579</td>
</tr>
<tr>
<td></td>
<td>5.581</td>
<td>95563</td>
<td>95732</td>
</tr>
<tr>
<td></td>
<td>5.592</td>
<td>95584</td>
<td>95755</td>
</tr>
<tr>
<td>100</td>
<td>5.269</td>
<td>95122</td>
<td>95174</td>
</tr>
<tr>
<td></td>
<td>5.499</td>
<td>95562</td>
<td>95652</td>
</tr>
<tr>
<td></td>
<td>5.557</td>
<td>95667</td>
<td>95768</td>
</tr>
<tr>
<td></td>
<td>5.571</td>
<td>95692</td>
<td>95796</td>
</tr>
<tr>
<td></td>
<td>5.581</td>
<td>95710</td>
<td>95815</td>
</tr>
<tr>
<td></td>
<td>5.586</td>
<td>95719</td>
<td>95825</td>
</tr>
<tr>
<td>120</td>
<td>5.032</td>
<td>94813</td>
<td>94819</td>
</tr>
<tr>
<td></td>
<td>5.240</td>
<td>95214</td>
<td>95235</td>
</tr>
<tr>
<td></td>
<td>5.314</td>
<td>95353</td>
<td>95382</td>
</tr>
<tr>
<td></td>
<td>5.321</td>
<td>95366</td>
<td>95396</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( n )</th>
<th>( x )</th>
<th>Lower bounds</th>
<th>Upper bounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>8.451</td>
<td>98376</td>
<td>99095</td>
</tr>
<tr>
<td></td>
<td>10.216</td>
<td>99226</td>
<td>99662</td>
</tr>
<tr>
<td></td>
<td>11.332</td>
<td>99518</td>
<td>99817</td>
</tr>
<tr>
<td></td>
<td>12.183</td>
<td>99664</td>
<td>99885</td>
</tr>
<tr>
<td></td>
<td>12.997</td>
<td>99763</td>
<td>99926</td>
</tr>
<tr>
<td>50</td>
<td>8.592</td>
<td>98654</td>
<td>99152</td>
</tr>
<tr>
<td></td>
<td>10.447</td>
<td>99378</td>
<td>99698</td>
</tr>
<tr>
<td></td>
<td>11.191</td>
<td>99545</td>
<td>99800</td>
</tr>
<tr>
<td></td>
<td>11.827</td>
<td>99652</td>
<td>99859</td>
</tr>
<tr>
<td></td>
<td>12.128</td>
<td>99693</td>
<td>99880</td>
</tr>
<tr>
<td></td>
<td>12.217</td>
<td>99704</td>
<td>99886</td>
</tr>
<tr>
<td></td>
<td>12.465</td>
<td>99734</td>
<td>99900</td>
</tr>
<tr>
<td>80</td>
<td>8.511</td>
<td>98697</td>
<td>99109</td>
</tr>
<tr>
<td></td>
<td>9.853</td>
<td>99253</td>
<td>99578</td>
</tr>
<tr>
<td></td>
<td>10.239</td>
<td>99364</td>
<td>99659</td>
</tr>
<tr>
<td></td>
<td>10.386</td>
<td>99401</td>
<td>99686</td>
</tr>
<tr>
<td></td>
<td>10.388</td>
<td>99402</td>
<td>99686</td>
</tr>
<tr>
<td>100</td>
<td>8.449</td>
<td>98704</td>
<td>99076</td>
</tr>
<tr>
<td></td>
<td>9.941</td>
<td>99301</td>
<td>99597</td>
</tr>
<tr>
<td></td>
<td>10.227</td>
<td>99380</td>
<td>99656</td>
</tr>
<tr>
<td></td>
<td>10.352</td>
<td>99411</td>
<td>99679</td>
</tr>
<tr>
<td>120</td>
<td>8.967</td>
<td>98981</td>
<td>99306</td>
</tr>
<tr>
<td></td>
<td>10.272</td>
<td>99406</td>
<td>99664</td>
</tr>
<tr>
<td></td>
<td>10.522</td>
<td>99464</td>
<td>99707</td>
</tr>
<tr>
<td></td>
<td>10.603</td>
<td>99482</td>
<td>99720</td>
</tr>
<tr>
<td></td>
<td>10.615</td>
<td>99485</td>
<td>99722</td>
</tr>
</tbody>
</table>
Therefore we may take the following very simple approximation of $P_0(T_s \leq x)$:

\begin{align*}
(4.18) \quad \Pr(U^2 \leq x) & \Pr(U_2^2 \leq \log n) \quad \text{if } x \leq \log n, \\
\Pr(U^2 \leq x) \Pr(U_2^2 \leq \log n) + \Pr(U_2^2 \geq \log n) & \quad \text{if } x \geq 2\log n, \\
\text{linearize} & \quad \text{if } \log n \leq x \leq 2\log n.
\end{align*}

To illustrate the approximation (4.18) consider $n = 50$ and the empirical 5% critical values $x$ of Table 1. We obtain the following:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>[approximation (4.18)]</td>
<td>93901</td>
<td>95068</td>
<td>95408</td>
<td>95587</td>
<td>95592</td>
<td>95639</td>
<td>95649</td>
</tr>
</tbody>
</table>

$x \times 10^5$

It is seen that the approximation works quite well in this case. As is seen from (4.17) the approximation (4.18) is larger than the right-hand side of (4.2) if $x \geq 2\log n$. It was seen in Table 4 that for the 1% critical values the right-hand side of (4.2) gives already a (slight) overestimation. Since as a rule in Table 4 the 1% critical values are larger than $2\log n$, the approximation (4.18) gives a (higher)-overestimation of $P_0(T_s \leq x)$. A possible remedy is to replace the interval $[\log n, 2\log n]$ in (4.18) by, for instance, $[\log n, 3\log n]$ if small levels are concerned.

REFERENCES


Faculty of Applied Mathematics
University of Twente
P.O. Box 217
7500 AE Enschede, The Netherlands

Institute of Mathematics
Technical University of Wroclaw
Wybrzeże Wyspiańskiego 27
50-370 Wroclaw, Poland

Received on 14.12.1993