ON THE POWER FUNCTION OF THE
CHI-SQUARE GOODNESS OF FIT TEST

(Submitted for publication to
Mathematische Operationsforschung und Statistik)
ON THE POWER FUNCTION OF THE
CHI-SQUARE GOODNESS OF FIT TEST

B. Schorr*)

1. INTRODUCTION

In this paper we investigate the power function of the chi-square
goodness of fit test in the case of a simple hypothesis. We are particu-
larly interested in the question of the choice of the number of class
intervals of the $\chi^2$-test where the null hypothesis consists of a one-
dimensional continuous distribution function. In this case the test
problem can always be reduced by the well-known probability transformation
to the specific case of testing the null hypothesis:

$$H_0: F(x) = F_0(x) = x, \quad 0 \leq x \leq 1.$$  (1.1)

To apply the $\chi^2$-test, the interval $[0,1]$ is divided into $k$ different
class intervals by choosing the corresponding end points as

$$0 = x_0 < x_1 < \ldots < x_{k-1} < x_k = 1.$$  (1.2)

Suppose we have a sample of size $n$ of the corresponding random vari-
able $X$, and let $N_i$, $i = 1, 2, \ldots, k$, be the number of sample elements in
the $i^{th}$ class interval, then the test statistic of the $\chi^2$-test for $H_0$ is
given by

$$\chi^2: = \sum_{i=1}^{k} \frac{(N_i - n\pi_i)^2}{n\pi_i},$$  (1.3)

where

$$\sum_{i=1}^{k} N_i = n.$$  (1.4)

*) Dr. Benno Schorr, CERN, 1211 Geneva 23, Switzerland.
\[ \pi_i = x_i - x_{i-1}, \quad i = 1, \ldots, k. \quad (1.5) \]

[On the left side in (1.3) we used the common symbol for the statistic of the \( \chi^2 \)-test introduced by Cochran.]

Assuming that \( H_0 \) is true, \( X^2 \) has for large \( n \) a \( \chi^2 \)-distribution with \( k - 1 \) degrees of freedom. Therefore, one defines the \( \chi^2 \)-test by the rejection region

\[
R = \{ (n_1, \ldots, n_k) : x^2 = \sum_{i=1}^{k} \frac{(n_i - n_{\pi_i})^2}{n_{\pi_i}} > \chi^2_{\alpha, k-1} \}, \quad (1.6)
\]

where \( \chi^2_{\alpha, k-1} \) is the upper \( \alpha \)-point of the \( \chi^2 \)-distribution with \( k - 1 \) degrees of freedom.

From a practical point of view an important question is, When can \( n \) be considered large enough for one to use the \( \chi^2 \)-distribution for \( X^2 \)? For this question we refer, for example, to Vesserau's [1] and Slakter's [2] papers. We are more interested in the choice of \( k \) and the \( \pi_i \).

In 1942 Mann and Wald found that in the statistical literature there only existed rules of thumb on the choice of \( k \) and \( \pi_i \). Therefore they tried in their paper [3] to formulate exact principles for this choice. They first proved that the \( \chi^2 \)-test is locally unbiased in the special case of

\[
\pi_i = \frac{1}{k}, \quad i = 1, 2, \ldots, k. \quad (1.7)
\]

To find an optimal choice of \( k \) under the condition (1.7), they considered an alternative hypothesis

\[
H_1 : F(x) = F_1(x) \quad (1.8)
\]
and introduced the following distance

$$d(F_0, F_1) = \sup_{x \in (0,1)} |F_1(x) - F_0(x)|.$$  \hspace{1cm} (1.9)

Let $C(\Delta)$ for $\Delta > 0$ be the class of alternative distributions with $d(F_0, F_1) \geq \Delta$. Let $f(n, k, F_1)$ be the power of the $\chi^2$-test for fixed $n$, $k$, and $F_1$ under the condition (1.7). Then we can summarize their results in the following theorem:

**Theorem 1**

Let

$$f_0(n, k, \Delta) = \inf_{F \in C(\Delta)} f(n, k, F)$$  \hspace{1cm} (1.10)

and let $k_n$ be that $k$ which maximizes $f_0(n, k, \Delta)$. Then, for

$$\Delta_n = \frac{5}{k_n} - \frac{4}{k_n^2}$$  \hspace{1cm} (1.11)

and

$$k_n = \left[ \frac{4 \sqrt{2(n-1) \alpha}}{c_\alpha^2} \right],$$  \hspace{1cm} (1.12)

we have

$$\lim_{n \to \infty} f_0(n, k_n, \Delta_n) = \frac{1}{2},$$  \hspace{1cm} (1.13)

where $c_\alpha$ is the upper $\alpha$-point of the standardized normal distribution.

Theorem 1 says that for large $n$ one can reach a power of about a half or more against alternatives $F_1$ with a distance of at least $\Delta_n$ from $F_0$ by taking $k_n$ as the number of class intervals. The choice (1.12) for the "optimal" $k$ leads to a rather high number of class intervals. For example,
for \( n = 200 \) and \( \alpha = 0.05 \) the formula (1.12) leads to \( k_n = 31 \). Williams [4] investigated formula (1.12), and he found that taking \([0.5 \, k_n]\) as the "optimal" number of class intervals for \( n \geq 200 \) would not significantly deteriorate the power of the test, in practice.

In 1970 Beier-Küchler and Neumann [5] again tried to improve Mann and Wald's results. The proof of theorem 1 is based on the assumption that, for large \( n \), \( X^2 \) has a normal distribution. In Ref. [5] a better approximation of the distribution of \( X^2 \) given by Patnaik [6] was used. Beier-Küchler and Neumann came to the following Rule of thumb.

If \( \alpha = 0.05 \) and the \( \pi_i \) are all equal, the choice of \( k = 16 \) leads to as few as possible wrong decisions. The choice of \( k \) too large is disadvantageous.

Since in Ref. 5 it is not clearly defined how small \( n \) may be in order to apply the rule of thumb, we refer for small \( n \) (\( \leq 50 \)) to Slakter's paper [7] in which the power of the \( \chi^2 \)-test for small \( n \) and small ratios \( n/k \) is investigated by Monte Carlo simulation.

In this paper we will study the choice of the number of class intervals by using the distance

\[
\rho(F_0, F_1) = \int_0^1 |F_1(x) - F_0(x)| \, dx \tag{1.15}
\]

instead of Mann and Wald's notion (1.9).

2. APPROXIMATION OF THE DISTRIBUTION OF \( X^2 \)

We assume that the \( N_i \) have a multinomial distribution with the probabilities \( p_i, i = 1, 2, \ldots, k \). Especially if \( H_0 \) is true we have \( p_i = \pi_i \) for \( i = 1, 2, \ldots, k \).
A cumbersome but not difficult calculation leads to the following expressions for the expectation and the variance of $X^2$:

\[
E(X^2) = k - 1 + n \cdot \delta^2 - \delta^2 + \sum_{i=1}^{k} \frac{Y_i}{\pi_i} , \tag{2.1}
\]

\[
\text{Var}(X^2) = 2(k - 1) + 4n \cdot \delta^2 - \frac{2(k - 1)}{n} + \frac{R_1}{n} + \frac{R_2}{n} \tag{2.2}
\]

with

\[
Y_i = p_i - \pi_i , \quad i = 1, 2, \ldots, k , \tag{2.3}
\]

\[
\delta^2 = \sum_{i=1}^{k} \frac{Y_i^2}{\pi_i} , \tag{2.4}
\]

\[
R_1 = 6(n - 1) \sum_{i=1}^{k} \frac{Y_i^2}{\pi_i} - 4[(4 + k)(n - 1) + 1] \cdot \delta^2 +
\]

\[
+ 4(n - 1)(n - 2) \sum_{i=1}^{k} \frac{Y_i^3}{\pi_i^2} - 2(n - 1)(2n - 3) \cdot \delta^4 , \tag{2.5}
\]

\[
R_2 = \left[4(n - 1)(2 - \delta^2) - 2k\right] \left[1 - \sum_{i=1}^{k} \frac{Y_i}{\pi_i}\right] \sum_{i=1}^{k} \frac{Y_i}{\pi_i} +
\]

\[
+ \sum_{i=1}^{k} \frac{Y_i}{\pi_i^2} + \sum_{i=1}^{k} \frac{1 - k \cdot \pi_i}{\pi_i} . \tag{2.6}
\]

Under the condition that all $\pi_i = 1/k$ the expressions (2.1) and (2.2) are simplified to:

\[
E(X^2) = k - 1 + n\delta^2 - \delta^2 , \tag{2.7}
\]

\[
\text{Var}(X^2) = 2(k - 1) + 4n\delta^2 - \frac{2(k - 1)}{n} + \frac{R_4}{n} \cdot \delta^2 + \frac{R_2}{n} , \tag{2.8}
\]
with

\[ R_3 = 2 \left[ (k - 8)(n - 1) - 2 \right] - 2(n - 1)(2n - 3) \delta^2 , \quad (2.9) \]

\[ R_4 = 4k^2(n - 1)(n - 2) \sum_{i=1}^{k} \gamma_i^3 . \quad (2.10) \]

Let us now consider the non-central \( \chi^2 \)-distribution with \( k - 1 \) degrees of freedom and the non-central parameter \( \lambda^2 \). Its characteristic function is given by

\[ \phi(t) = \frac{e^{\lambda^2 it/(1-2it)}}{(1-2it)^{(k-1)/2}} \quad (2.11) \]

(see Ref. 6). If we denote by \( \chi'^2 \) the random variable which corresponds to \( \phi(t) \), it is easy to derive from (2.11) that

\[ E(\chi'^2) = k - 1 + \lambda^2 , \quad (2.12) \]

\[ \text{Var}(\chi'^2) = 2(k - 1) + 4\lambda^2 . \quad (2.13) \]

setting

\[ \lambda^2 = n \cdot \delta^2 \quad (2.14) \]

and comparing (2.1), (2.2), or (2.7), (2.8) with (2.12) and (2.13), respectively, we see that the first two terms coincide. Generally it is possible to show (see Eisenhart [8]) that for

\[ p_i = \pi_i + \frac{\Delta_i}{\sqrt{n}}, \quad \Delta_i \geq 0 , \quad i = 1, 2, \ldots, k , \]

and for \( n \to \infty \) the random variable \( X^2 \) has a non-central \( \chi^2 \)-distribution with \( k - 1 \) degrees of freedom and the non-central parameter

\[ \lambda^2 = \sum_{i=1}^{k} \frac{\Delta_i^2}{\pi_i} . \]
In the special case where all $\pi_i = 1/k$, we can see by comparing (2.1), (2.2) with (2.7), (2.8) that the approximation of the distribution of $X^2$ by the non-central $\chi^2$-distribution will be better, in general, than for arbitrary $\pi_i$.

Indeed we know the characteristic function $\phi(t)$ of $\chi^2$, but the corresponding distribution function can only be represented as a complicated infinite series. Therefore it is more convenient for analytic investigations to approximate the distribution function of $\chi^2$ by a simple expression. There are several approximations in the literature. We use the one given by Patnaik [6] which is also used in Ref. 5.

If

$$Y = \sqrt{2\chi^2(k - 1 + \lambda^2)} \frac{k - 1 + 2\lambda^2}{k - 1 + 2\lambda^2},$$

(2.15)

then $Y$ is asymptotically normally distributed with the mean value

$$m = \sqrt{\frac{2(k - 1 + \lambda^2)^2}{k - 1 + 2\lambda^2}} - 1$$

(2.16)

and variance 1. The examples calculated by Patnaik show that the approximation may be accurate up to two digits if

$$n + \lambda^2 \geq 50.$$  

(2.17)

Taking into account that Slakter [7] has covered the case $n \leq 50$, we shall confine our investigations to $n \geq 50$. Therefore formula (2.17) will always be satisfied.

Summarizing what we have found in this paragraph and using the notation from paragraph 1, we may say that the power function

$$\beta(H_1) = \Pr(X^2 > \chi^2_{\alpha, k-1} | H_1)$$

(2.18)
can be approximated by

\[ \beta_0(H_1) = \Pr(\chi^2 > \chi^2_{\alpha, k-1} | H_1) \] (2.19)

for small \( \lambda^2 \) and large \( n \). In addition, \( \beta_0(H_1) \) may be approximated under the condition (2.17) by

\[ \beta_1(H_1) = 1 - \Phi(z) , \] (2.20)

where

\[ z = \sqrt{\frac{2\chi^2_{\alpha, k-1} (k - 1 + \lambda^2)}{k - 1 + 2\lambda^2}} - m \] (2.21)

and \( \Phi \) is the distribution function of the standardized normal distribution.

Finally, we mention a lemma which is given in Ref. 5 and which we shall use later.

**Lemma**

For fixed \( k \), \( \beta_0(H_1) \) is monotonically increased with \( \lambda^2 \).

3. **MINIMIZATION OF \( \beta_0(H_1) \)**

Let \( K \) be the class of distribution functions \( F \) on (0,1) with continuous probability density function \( f \), and for \( 0 \leq \rho < 1/2 \) let

\[ S(\rho) = \{ F : F \in K, \rho(x, F) \geq \rho \} , \] (3.1)

where \( \rho(x, F) \) is defined by (1.15).

We want to determine

\[ \inf_{F_1 \in S(\rho)} \beta_0(H_1) \] (3.2)
for arbitrary and fixed \( \pi_i \) (\( i = 1, 2, \ldots, k \)), \( n, k \) and \( \rho > 0 \). To solve this problem it is sufficient, because of the lemma of paragraph 2, to determine

\[
I = \inf_{F_1 \in S(\rho)} \delta^2 = \inf_{F_1 \in S(\rho)} \sum_{i=1}^{k} \frac{(p_i - \pi_i)^2}{\pi_i}
\]  

(3.3)

and to insert it in (2.19), taking into account (2.14). To determine (3.3) it is not hard to see that, analogously to the case of the distance (1.9) used by Mann and Wald, it is sufficient to confine \( F \in S(\rho) \) -- we now drop the index "1" -- to the case where, for example,

\[
F(x) \geq x, \quad x \in (0,1).
\]  

(3.4)

Since we are only interested in \( I > 0 \), we have another condition for \( \rho \):

\[
\frac{1}{2} \sum_{i=1}^{k} \pi_i^2 < \rho ,
\]  

(3.5)

which is easy to see from Fig. 1, where the shadowed area is the value of the left side of (3.5) for \( k = 4 \).

![Fig. 1](image-url)
Under the hypothesis (3.4) we have for \( F \in S(\rho) \):

\[
\int_0^1 [F(x) - x] \, dx \geq \rho .
\] (3.6)

By partial integration and because of (1.5) we can write (3.6) as

\[
\sum_{i=1}^{k} \int_{x_{i-1}}^{x_i} x \cdot f(x) \, dx \leq \frac{1}{2} - \rho .
\] (3.7)

Since \( f \) is continuous and non-negative, (3.7) can be written as

\[
\sum_{i=1}^{k} \xi_i \cdot p_i \leq \frac{1}{2} - \rho ,
\] (3.8)

where

\[
x_{i-1} \leq \xi_i \leq x_i ,
\] (3.9)

and

\[
p_i = \int_{x_{i-1}}^{x_i} f(x) \, dx
\]

for \( i = 1, 2, \ldots, k \).

For the determination of \( I \) in (3.3) it is sufficient to take the equality sign in (3.8). Therefore we can now formulate the problem of searching the infimum of \( \delta^2 \) as follows:

**Problem A**

Find \( p_i \) for \( i = 1, 2, \ldots, k \) such that

\[
\delta^2 = \sum_{i=1}^{k} \frac{(p_i - \pi_i)^2}{\pi_i}
\] (3.10)
takes its minimum under the conditions

\[ p_i \geq 0, \quad i = 1, 2, \ldots, k, \quad (3.11) \]

\[ \sum_{i=1}^{k} p_i = 1, \quad (3.12) \]

\[ \sum_{i=1}^{r} p_i \geq \sum_{i=1}^{r} \pi_i, \quad r = 1, 2, \ldots, k - 1 \quad (3.13) \]

\[ \sum_{i=1}^{k} \xi_i p_i = \frac{1}{2} - \rho, \quad (3.14) \]

where the \( \xi_i \) satisfy (3.9).

The conditions (3.11) and (3.12) follow from the fact that the \( p_i \) are probabilities. Formula (3.13) follows from (3.4), and (3.14) from (3.6) and (3.8), respectively.

Suppose that in (3.14) the \( \xi_i \) are known constants, then problem A is a problem of mathematical programming with the quadratic objective function \( \delta^2 \) and the linear restrictions (3.11) through (3.14), and therefore especially a problem of convex programming for which the Kuhn-Tucker theorem (see Künzi et al. [9]) gives necessary and sufficient conditions for the solution.

Therefore we first determine the \( \xi_i \). Suppose we have a step function, such as, for example, that shown in Fig. 2, which with \( p_1^*, p_2^*, \ldots, p_k^* \) yields the minimum of \( \delta^2 \) where in (3.6) the equality sign is valid. Such \( p_i^* \) always exist for every \( \rho \) with \( 0 < \rho < 1/2 \). Then we can show that this step function is the only distribution function for which \( \delta^2 \) takes its minimum under the given conditions. For, a distribution function \( F \) with
\[ p_i^* = \int_{x_{i-1}}^{x_i} dF, \quad i = 1, 2, \ldots, k, \]  \hfill (3.15)

which lies under the step function as in Fig. 2 has \( p(x,F) < p \). The function \( F \) cannot lie above the step function, because of the monotonicity of a distribution function.

![Fig. 2](image-url)

Therefore we can see that the density functions \( f \) of the distribution functions \( F \in S(\rho) \) which lead to the infimum of \( \delta^2 \) must be concentrated in the limit into the left end points of the intervals \( (x_{i-1}, x_i) \).

Consequently in the limit we must have

\[ \xi_i = x_{i-1}, \quad i = 1, 2, \ldots, k. \]  \hfill (3.16)

With (3.16) we now have in problem A the programming problem already mentioned. We shall denote by problem \( A_0 \) the special case of problem A where all the \( \pi_i = 1/k \). The solution of problem \( A_0 \) will be given in theorem 2. Theorem 3 gives the solution of problem A, but only for the case where the \( \pi_i \) are all in a certain neighbourhood of \( 1/k \). We conjecture that the given solution is even valid for arbitrary \( \pi_i \) -- for \( k = 2 \) this can easily be shown -- but the exact proof is still lacking.
Theorem 2

The solution of problem \( A_0 \) is given by

\[
p_{i} = \frac{1}{k} \left\{ 1 + \frac{2(k - r)}{r(r + 1)} \left[ 2(r + 1) - 3i \right] + \right. \\
+ \frac{6k^2}{r(r^2 - 1)} \left[ \rho - \frac{k^2 - r(r - 1)}{2k^2} \right] (r + 1 - 2i) \left\} 
\]

for \( i = 1, 2, \ldots, r \); \hspace{1cm} (3.17)

\[
p_{i} = 0 \quad \text{for} \quad i = r + 1, \ldots, k , \hspace{1cm} (3.18)
\]

where \( r = r(\rho) \) is a positive integer, determined by

\[
\begin{aligned}
r &= k \quad \text{for} \quad \frac{1}{2k} < \rho < L_0(k) , \\
2 \leq r &\leq k - 1 \quad \text{for} \quad L_0(r + 1) \leq \rho < L_0(r) 
\end{aligned} \hspace{1cm} (3.19) \hspace{1cm} (3.20)
\]

with

\[
L_0(r) = \frac{3k + 4 - 2r}{6k} . \hspace{1cm} (3.21)
\]

The minimum of \( \delta^2 \) is given by

\[
\begin{aligned}
\delta^2_{r,k} &= \frac{r}{k} \left[ \frac{\rho - \frac{k^2 - r(r - 1)}{2k^2}}{r(r + 1)} \right]^2 \\
+ \frac{k - r}{k} \left\{ 1 + \frac{4k^2}{r(r + 1)} \left[ \rho - \frac{(k - r)^2 + k}{2k^2} \right] \right\} 
\end{aligned} \hspace{1cm} (3.22)
\]

with

\[
D_{r,k} = \frac{r^2(r^2 - 1)}{12k^4} , \quad r = 2, 3, \ldots, k . \hspace{1cm} (3.23)
\]
Proving theorem 2 we shall formally do the steps for arbitrary $\pi_i$. In this way we can immediately give the proof for theorem 3 which is to be formulated later. The proof is given in three stages:

I) we solve problems $A$ and $A_0$, respectively, without considering the conditions (3.11) and (3.13);

II) formula (3.13) is taken into account;

III) formula (3.11) is taken into account.

The proof of stage (I) leads to a problem which can be solved by the method of Lagrange multipliers. Let $\lambda_1$, $\lambda_2$ be the multipliers. Then we consider the Lagrange function

$$
\psi(p_1, \ldots, p_k; \lambda_1, \lambda_2) = \delta^2 + 2\lambda_1 \left( \sum_{i=1}^{k} p_i - 1 \right) + \\
+ 2\lambda_2 \left( \sum_{i=1}^{k} x_{i-1} - p_i + \rho + \frac{1}{2} \right). \quad (3.24)
$$

The application of the well-known Lagrange method leads, after some calculation, to the equations

$$
p_i = \pi_i (1 + \lambda_1 + x_{i-1} \lambda_2), \quad i = 1, 2, \ldots, k, \quad (3.25)
$$

$$
\lambda_1 + \left( \sum_{i=1}^{k} x_{i-1} \pi_i \right) \lambda_2 = 0, \quad (3.26)
$$

$$
\left( \sum_{i=1}^{k} x_{i-1} \pi_i \right) \lambda_1 + \left( \sum_{i=1}^{k} x_{i-1}^2 \pi_i \right) \lambda_2 = \frac{1}{2} - \rho - \sum_{i=1}^{k} x_{i-1} \pi_i. \quad (3.27)
$$

The determinant of the system of equations (3.26), (3.27) is given by

$$
D_k = \sum_{i=1}^{k} x_{i-1}^2 \pi_i - \left( \sum_{i=1}^{k} x_{i-1} \pi_i \right)^2. \quad (3.28)
$$
is always positive, for it can be written as

\[ D_k = \sum_{i=1}^{k} x_{i-1} \, \pi_i \left( x_{i-1} - \sum_{j=1}^{k} x_{j-1} \, \pi_j \right)^2. \]  

(3.29)

Simple geometric arguments lead to

\[ \frac{1}{2} \sum_{j=1}^{k} \pi_j^2 = \frac{1}{2} - \sum_{j=1}^{k} x_{j-1} \, \pi_j. \]  

(3.30)

Using (3.30) in (3.29) it is easily seen that \( D_k > 0 \). Especially for problem A\( \delta \), \( D_k \) is given by

\[ D_k = D_{k,k} = \frac{k^2 - 1}{12k^2}, \quad k \geq 2. \]  

(3.31)

Solving the system (3.26), (3.27) and inserting the solution into (3.25) yields the solution of stage (I) of problem A as

\[ p_i = \pi_i \left( 1 + \frac{1}{D_k} \left[ \sum_{j=1}^{k} x_{j-1} \, \pi_j - x_{i-1} \right] \right) \left[ \rho - \frac{1}{2} + \sum_{j=1}^{k} x_{j-1} \, \pi_j \right], \]

(3.32)

for \( i = 1, 2, \ldots, k \).

Since \( \delta^2 \) is a positive definite quadratic form in the \( p_i \), (3.32) furnishes the absolute minimum of \( \delta^2 \) for stage (I) of problem A.

To treat stage (II) of problem A we must find out whether (3.32) satisfies condition (3.13). We insert (3.32) into (3.13) and, taking into account that \( D_k > 0 \) and because of (3.5), it can easily be seen that (3.13) in this case is equivalent to

\[ \sum_{i=1}^{r} \pi_i \left[ \frac{1}{2} - \sum_{j=1}^{k} \pi_j^2 - x_{i-1} \right] \geq 0 \]

(3.33)

for \( r = 1, 2, \ldots, k - 1 \). Let us first take \( \pi_i = 1/k \), i.e., \( x_i = i/k \).
Then a rather simple calculation shows that (3.33) is equivalent to \( k \geq r \)
for \( r = 1, 2, \ldots, k - 1 \), which is always true. Since the left side of
(3.33) is continuous in the \( \pi_j \), (3.33) must also be true for all \( \pi_j \) which
are in a certain neighbourhood of \( 1/k \). Thus stage (II) of problem A_0 is
treated.

We now come to stage (III) of problem A. We have to check whether
the \( p_i \) in (3.32) are non-negative for all \( \rho \) which are less than \( 1/2 \) and
which satisfy the condition (3.5).

It is immediately seen that the \( p_i \) in (3.32) are monotonically de-
creasing in \( i \). Especially we have the fact that if \( p_k \geq 0 \), then \( p_i > 0 \)
for \( i = 1, 2, \ldots, k - 1 \). Taking in (3.32) \( p_k \geq 0 \) leads, after some cal-
culation, to the condition

\[
\rho \leq \frac{D_k}{x_k} + \frac{1}{2} - \sum_{j=1}^{k} x_{j-1} \pi_j .
\]  

(3.34)

Setting in (3.34) \( \pi_j = 1/k \) for all \( i = 1, 2, \ldots, k \) yields

\[
\rho \leq \frac{k + 4}{6k} = L_0(k) .
\]  

(3.35)

Summarizing what we have proved until now, we can say that for

\[
\frac{1}{2} - \sum_{j=1}^{k} x_{j-1} \pi_j < \rho < \frac{D_k}{x_k} + \frac{1}{2} - \sum_{j=1}^{k} x_{j-1} \pi_j ,
\]  

(3.36)

and for

\[
\frac{1}{2k} < \rho < L_0(k)
\]  

(3.37)

in the special case, (3.32) yields the solution of problem A for \( \pi_j \) in the
neighbourhood of 1/k, and especially the solution of problem $A_0$. In addition, we can say that if (3.36) is valid, all $p_i$ in (3.32) are positive. Therefore taking $\pi_i = 1/k$ in (3.32) we obtain (3.17) in theorem 2 for $r = k$.

To prove the remainder of theorem 2 we first make a guess at how the solution of problem $A$ could behave if $\rho$ is larger than in condition (3.36). We observe that $p_k = 0$ if $\rho$ is equal to the right side of (3.36). In addition, geometric arguments give us a hint that step by step, with $\rho$ increasing, all the $p_i$ should vanish for $i = k - 1, k - 2, \ldots, 3$. Therefore, we now make the assumption that the solution is of the form:

\begin{align*}
& p_i > 0 \quad \text{for } i = 1, \ldots, r \\
& p_i = 0 \quad \text{for } i = r, r + 1, \ldots, k ,
\end{align*}

(3.38) (3.39)

where $r$ is dependent on $\rho$. We will solve problem $A$ under this assumption, and later we shall justify the assumption.

Again we have to go through all the three stages taking into account that the solution is of the form (3.38) and (3.39). Stage (I) leads to the problem:

Find $r = r(\rho)$ and $p_1, \ldots, p_r$ such that

\begin{equation}
\sum_{i=1}^{r} \frac{(p_i - \pi_i)^2}{\pi_i}
\end{equation}

(3.40)

takes its minimum under the conditions

\begin{equation}
\sum_{i=1}^{r} p_i = 1 ,
\end{equation}

(3.41)
\[
\sum_{i=1}^{r} x_{i-1} p_i = \frac{1}{2} - \rho .
\] (3.42)

The solution of this problem is again obtained by the method of Lagrange multipliers. After cumbersome calculations we have:

\[
p_i = \pi_i \left[ 1 + \frac{1}{D_r} \left( \left( \sum_{j=1}^{r} x_{j-1} \bar{\pi}_j - x_r x_{i-1} \right) \left( \rho - \frac{1}{2} + \sum_{j=1}^{r} x_{j-1} \pi_j \right) + \left( 1 - x_r \right) \left( \sum_{j=1}^{r} x_{j-1}^2 \bar{\pi}_j - \left( \sum_{j=1}^{r} x_{j-1} \pi_j \right) x_{i-1} \right) \right) \right]
\] (3.43)

for \( i = 1, 2, \ldots, r \) with

\[
D_r = x_r \sum_{i=1}^{r} x_{i-1}^2 \bar{\pi}_i - \left( \sum_{i=1}^{r} x_{i-1} \bar{\pi}_i \right)^2 .
\] (3.44)

For problem \( A_0 \) we have

\[
D_r = D_{r,k} = \frac{r^2 (r^2 - 1)}{12k^3} , \quad r = 2, 3, \ldots, k .
\] (3.45)

Again because of (3.45) and since \( D_r \) is continuous in the \( \bar{\pi}_i \), we have
\( D_r > 0 \) at least for all \( \pi_i \) in a certain neighbourhood of \( 1/k \).

Thus stage (I) is treated.

Stage (II) is expressed by the condition

\[
\sum_{i=1}^{s} p_i \geq \sum_{i=1}^{S} \pi_i \quad \text{for } s = 1, 2, \ldots, r - 1 .
\] (3.46)

Here we first take only the special case \( \pi_i = 1/k \) for \( i = 1, 2, \ldots, k \). Inserting (3.43) in this case into (3.46) leads, after some calculation, to the following condition for \( \rho \) for given \( r \):

\[ \rho \geq \frac{k^2 - r(r - 1)}{2k^2} - \frac{(k - r)(2r - 1)}{3k^2}. \] (3.47)

Stage (III) says that in (3.43) all the \( p_i \) must be positive. Again we first take \( \pi_i = 1/k \) and insert them into (3.43), which leads to (3.17) in theorem 2. For this formula we first look for a condition for \( \rho \) such that the \( p_i \) are monotonically decreasing with \( i \). Formal differentiation with respect to \( i \) and setting the derivative less than zero leads to the condition
\[ \rho > \frac{k - r + 1}{2k}. \] (3.48)

Assume that \( \rho \) satisfies (3.48) for given \( r \). Then, of \( p_r > 0 \), all the \( p_i \) are greater than zero. But \( p_r > 0 \) yields the condition
\[ \rho < \frac{3k + 4 - 2r}{6k} = L_0(r) \] (3.49)
after some calculation.

We now have to show the compatibility of the three conditions (3.47), (3.48), and (3.49). We first see that for \( r = k \) (3.47) and (3.49) yield (3.19), our previous result. In addition, for continuity we have to require that
\[ L_0(r + 1) \leq \rho < L_0(r), \] (3.50)
i.e. for given \( r \) the distance \( \rho \) must satisfy (3.50); or, we can interpret it in the opposite direction, i.e. if \( \rho \) is given then \( r \) must be such that (3.50) is satisfied.

It now remains to show that the validity of the left inequality of (3.50) implies both inequalities (3.47) and (3.48). But this can easily be done by simple calculation.
Summarizing we can say that (3.17) through (3.20) in theorem 2 are proved under the condition that our assumption that the solution is of the form (3.38), (3.39) is true.

To show that this assumption holds, we use the earlier mentioned Kuhn-Tucker theorem (see Ref. 9) for convex programming. A rather formal but cumbersome calculation shows that the expressions given in theorem 2 satisfy the necessary and sufficient conditions of the Kuhn-Tucker theorem for the solution.

Finally, when we insert (3.17) and (3.18) into \( \delta^2 \) we obtain (3.22) and the proof of theorem 2 is completed. But we have shown more; namely, that for all \( \pi_i \) in a certain neighbourhood of 1/k, (3.43) together with \( p_i = 0 \) for \( i = r + 1, \ldots, k \) solve problem A. However, for problem A we still have to formulate the condition equivalent to (3.50) which determines the relations between \( r \) and \( p \). Analogous arguments to those which gave us (3.50) lead to the condition

\[
L(r + 1) \leq p < L(r) \quad \text{for } r = 2, 3, \ldots, k - 1
\]

with

\[
L(r) = \frac{1}{x_{r-1} x_r - \sum_{j=1}^{r} x_{j-1} \pi_j} \left\{ D_r + (1 - x_r) \right\} + \frac{1}{2} - \sum_{j=1}^{r} x_{j-1} \pi_j.
\]

Inserting (3.43) and \( p_i = 0 \) for \( i = r + 1, \ldots, k \) into \( \delta^2 \) and denoting the result by \( \frac{\delta^2}{r} \), we finally can formulate the following theorem.
Theorem 3

Under the condition that all the \( \pi_i \) are in a certain neighbourhood of \( 1/k \), (3.43) together with \( p_i = 0 \) for \( i = r + 1, \ldots, k \) yields the solution of problem A, where \( r = k \) if (3.36) is valid and \( 2 \leq r \leq k - 1 \) if (3.51) is satisfied. The minimum of \( \delta^2 \) is given by

\[
\delta^2_r = \frac{1}{n^2} \sum_{i=r}^{r} \left( \sum_{j=1}^{r} x_{j-1} \pi_j - x_r x_{i-1} \right) \left( \rho - \frac{1}{2} + \sum_{j=1}^{r} x_{j-1} \pi_j \right) + (1 - x_r) \left[ \sum_{j=1}^{r} x_{j-1} \pi_j - \left( \sum_{j=1}^{r} x_{j-1} \pi_j \right) x_{i-1} \right] \pi_i^2 + \sum_{i=r+1}^{k} \pi_i^2.
\]

(3.53)

From theorem 2 we can derive the following corollary for the approximation of the distribution of \( X^2 \) by a non-central \( \chi^2 \)-distribution.

Corollary

For all \( \pi_i = 1/k \) and for

\[
\frac{1}{2k} < \rho < \frac{k + 4}{6k}
\]

(3.54)

we have \( R_n = 0 \) in (2.8) if the \( p_i \) are given by (3.17).

The proof of this corollary is easy and may be omitted. Our corollary indicates that the approximation of the distribution of \( X^2 \) by a non-central \( \chi^2 \)-distribution in this particular case may be even better than for arbitrary \( p_i \).

Since, according to Eisenhart [8], \( X^2 \) follows a non-central \( \chi^2 \)-distribution only for small \( \delta^2 \) and for large \( n \), we shall confine the rest
of the paper to the case where \( \rho \) satisfies the inequality (3.36) and its special case (3.37), respectively, for \( \pi_i = 1/k \). This means that we are dealing in theorems 2 and 3 only with the case \( r = k \). This implies a fairly simple form of \( \delta^2_r = \delta^2_k \). It is easily seen that we then obtain

\[
\delta^2_k = \left( \rho - \frac{1}{2} \sum_{j=1}^{k} \pi_j^2 \right)^2
\]

(3.55)

for arbitrary \( \pi_i \), and

\[
\delta^2_k = \delta^2_{k,k} = \frac{12k^2}{k^2 - 1} \left( \rho - \frac{1}{2k} \right)^2
\]

(3.56)

in the special case of \( \pi_i = 1/k \).

For comparison with the results of Mann and Wald we cite the corresponding expression for \( \delta^2_{k,k} \) if \( \delta^2 \) is minimized with respect to the class \( C(\Delta) \) of distribution functions, which is defined in Section 1 immediately after formula (1.9). It turns out that in this case we have

\[
\delta^2_{k,k} = \delta^2_{\text{MW}} = 4 \left( \Delta - \frac{1}{k} \right)^2.
\]

(3.57)

Formulae (3.56) and (3.57) are of the same type (at least for large \( k \)) with respect to the dependence of the different distances \( \rho \) and \( \Delta \), respectively.

4. THE CHOICE OF \( k \) IN PRACTICE

It is our goal to find out how to choose in practice the number of class intervals for given \( \alpha \) and \( n \) in order to reach a certain power and smallest distance \( \rho \). A numerical investigation of approximate power function \( \beta_1(H_1) \) defined by (2.20) will bring us to our goal.
The power function \( \beta_1(H_1) \) is dependent on the four parameters \( (\alpha, n, k, \rho) \). We assume \( \alpha \) and \( n \) to be given. In addition, let \( \beta \) be the probability of the error of the second kind. Since

\[
\beta = 1 - \beta_1(H_1)
\]  

(4.1)

under the condition that \( H_1 \) is true, we obtain a functional relationship \( \rho = \rho(k) \), when we fix the three parameters \( (\alpha, \beta, n) \).

In a numerical calculation the following different combinations have been selected:

\[
\alpha = 0.01, 0.05, 0.10, \quad (4.2)
\]

\[
\beta = 0.50, 0.40, 0.30, 0.20, 0.10, 0.05, \quad (4.3)
\]

\[
n = 50 \ (50) \ 1000; 1100 \ (100) \ 1500; \ 2000. \quad (4.4)
\]

For each combination \( (\alpha, \beta, n) \) of values from (4.2) through (4.4) the curves \( \rho = \rho(k) \) for \( k = 10, 11, \ldots, 90 \) have been calculated, and for each of the 468 different curves the \( k \)-value which yields the minimum \( \rho_{min} = \rho_{min}(\alpha, \beta, n) \) have been determined. Table 1 shows the different optimal \( k \)-values. In Table 2 one can find the corresponding values of \( \rho_{min} \). Table 3 shows the optimal \( k \)-values which follow from Mann and Wald's formula given in (1.12). A comparison of these values with the corresponding values of Table 1 for \( \beta = 0.5 \) shows that Mann and Wald's values are higher than our values but not drastically so.

Now the 468 curves all show a rather flat behaviour in the neighbourhood of the \( k \)-values given in Table 1. This indicates that one may reduce the \( k \)-values without changing the corresponding \( \rho \)-values too much. Keeping this in mind and in addition being willing to have simple rules for the choice of \( k \), the class of 468 curves was subdivided into five different groups, dependent only on \( n \). A fixed \( k \)-value was assigned to each group,
to be considered as the optimal \( k \) in practice. The criterion used to select the groups and the corresponding \( k \) was that the relative error \( e_r \) satisfied the inequality

\[
e_r = \frac{\rho(k; \alpha, \beta, n) - \rho_{\text{min}}}{\rho_{\text{min}}} \leq 0.05 .
\]  (4.5)

This procedure led to the following Rule:

In order to satisfy (4.5) one may choose

<table>
<thead>
<tr>
<th>( k )</th>
<th>for ( n ) between</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>50-150</td>
</tr>
<tr>
<td>20</td>
<td>200-350</td>
</tr>
<tr>
<td>25</td>
<td>400-600</td>
</tr>
<tr>
<td>30</td>
<td>550-800</td>
</tr>
<tr>
<td>33</td>
<td>850-1000</td>
</tr>
<tr>
<td>36</td>
<td>1100-1300</td>
</tr>
<tr>
<td>40</td>
<td>1400-2000</td>
</tr>
</tbody>
</table>

In Table 4 one can find the relative errors \( e_r \) corresponding to this rule. We see that only for \( \alpha = 0.01 \) and \( \beta = 0.50, 0.40 \), \( e_r \) is higher than 0.05. It can also be seen that \( e_r \) is increasing with increasing \( \alpha \). In other words, for \( \alpha = 0.01 \) or 0.05 we may even slightly reduce \( k \) in the different groups without violating (4.5) very much.

For \( n \) up to 200 our rule is almost the same as the rule of thumb given by Beier-Küchter and Neumann in [5]. In addition our rule is also in good agreement with Williams [4] results which we mentioned in Section 1.

To have an impression of the order of magnitude of the distances \( \rho \), an example may be useful. Let

\[
F_0(x) = \frac{x}{\sigma}, \quad \sigma > 0 ,
\]  (4.6)
\[ F_1(x) = \phi\left(\frac{x + m}{\sigma}\right), \quad m > 0, \quad (4.7) \]

be two distribution functions of the normal type. Then

\[ \rho = \rho(F_0, F_1) = \int_{-\infty}^{\infty} \left[ F_1(x) - F_0(x) \right] dF_0(x) \quad (4.8) \]

can be evaluated by first differentiating \( \rho = \rho(u) \) with respect to \( u = m/\sigma \) and then integrating from 0 to \( u \), using that \( \rho(0) = 0 \). The result is

\[ \rho = \frac{1}{2} \text{erf}\left(\frac{u}{2}\right) \quad (4.9) \]

with

\[ \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt. \quad (4.10) \]

Table 5 shows some values of \( \rho \) for different \( u \).

**FINAL REMARK**

Analogously to Beier-Küchler and Neumann [5] one can try to maximize \( \beta_1(H_1) \) with respect to \( \pi_1 \) for fixed \( k \) and given \( \rho \). However, this can only be done numerically since the corresponding optimization problem is too complicated to be solved analytically.
REFERENCES


Table 1
Optimal k for $\alpha = 0.01, 0.05, 0.1; \beta = 0.5, \ldots, 0.05$, and $n = 50, \ldots, 2000$ (p-distance)

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>50</th>
<th>40</th>
<th>30</th>
<th>20</th>
<th>10</th>
<th>0.5</th>
<th>50</th>
<th>40</th>
<th>30</th>
<th>20</th>
<th>10</th>
<th>0.5</th>
<th>50</th>
<th>40</th>
<th>30</th>
<th>20</th>
<th>10</th>
<th>0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></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>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></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>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Table 2

Minimal $\rho$ corresponding to Table 1

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$\gamma$</th>
<th>0.01</th>
<th>0.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>0.196</td>
<td>0.208</td>
<td>0.214</td>
</tr>
<tr>
<td>100</td>
<td>0.147</td>
<td>0.156</td>
<td>0.163</td>
</tr>
<tr>
<td>150</td>
<td>0.124</td>
<td>0.129</td>
<td>0.140</td>
</tr>
<tr>
<td>200</td>
<td>0.111</td>
<td>0.117</td>
<td>0.124</td>
</tr>
<tr>
<td>250</td>
<td>0.101</td>
<td>0.107</td>
<td>0.113</td>
</tr>
<tr>
<td>300</td>
<td>0.094</td>
<td>0.099</td>
<td>0.105</td>
</tr>
<tr>
<td>350</td>
<td>0.088</td>
<td>0.093</td>
<td>0.104</td>
</tr>
<tr>
<td>400</td>
<td>0.083</td>
<td>0.088</td>
<td>0.098</td>
</tr>
<tr>
<td>450</td>
<td>0.079</td>
<td>0.084</td>
<td>0.094</td>
</tr>
<tr>
<td>500</td>
<td>0.075</td>
<td>0.080</td>
<td>0.090</td>
</tr>
<tr>
<td>550</td>
<td>0.073</td>
<td>0.078</td>
<td>0.086</td>
</tr>
<tr>
<td>600</td>
<td>0.070</td>
<td>0.074</td>
<td>0.083</td>
</tr>
<tr>
<td>650</td>
<td>0.068</td>
<td>0.072</td>
<td>0.081</td>
</tr>
<tr>
<td>700</td>
<td>0.066</td>
<td>0.070</td>
<td>0.078</td>
</tr>
<tr>
<td>750</td>
<td>0.064</td>
<td>0.068</td>
<td>0.072</td>
</tr>
<tr>
<td>800</td>
<td>0.063</td>
<td>0.066</td>
<td>0.070</td>
</tr>
<tr>
<td>850</td>
<td>0.061</td>
<td>0.064</td>
<td>0.069</td>
</tr>
<tr>
<td>900</td>
<td>0.060</td>
<td>0.063</td>
<td>0.066</td>
</tr>
<tr>
<td>950</td>
<td>0.058</td>
<td>0.061</td>
<td>0.065</td>
</tr>
<tr>
<td>1000</td>
<td>0.057</td>
<td>0.060</td>
<td>0.064</td>
</tr>
<tr>
<td>1100</td>
<td>0.055</td>
<td>0.058</td>
<td>0.061</td>
</tr>
<tr>
<td>1200</td>
<td>0.053</td>
<td>0.056</td>
<td>0.059</td>
</tr>
<tr>
<td>1300</td>
<td>0.051</td>
<td>0.054</td>
<td>0.057</td>
</tr>
<tr>
<td>1400</td>
<td>0.050</td>
<td>0.052</td>
<td>0.055</td>
</tr>
<tr>
<td>1500</td>
<td>0.048</td>
<td>0.051</td>
<td>0.054</td>
</tr>
<tr>
<td>2000</td>
<td>0.042</td>
<td>0.045</td>
<td>0.048</td>
</tr>
</tbody>
</table>

*Values represent minimal $\rho$ for various $\alpha$ and $\gamma$ combinations.*
Table 3

Optimal k for $\alpha = 0.01, 0.05, 0.1$; $\beta = 0.5$ and $n = 50, \ldots, 2000$ (Mann-Wald formula)

<table>
<thead>
<tr>
<th>ALPHA</th>
<th>0.01</th>
<th>0.05</th>
<th>0.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>BETA</td>
<td>50</td>
<td>15</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>20</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>24</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>27</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>250</td>
<td>29</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>32</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>350</td>
<td>34</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>35</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>450</td>
<td>37</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>39</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>550</td>
<td>40</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>600</td>
<td>42</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>650</td>
<td>43</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>700</td>
<td>45</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>750</td>
<td>46</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td>800</td>
<td>47</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>850</td>
<td>48</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>900</td>
<td>49</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>950</td>
<td>50</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>51</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>1100</td>
<td>53</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>1200</td>
<td>55</td>
<td>64</td>
</tr>
<tr>
<td></td>
<td>1300</td>
<td>57</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>1400</td>
<td>59</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>1500</td>
<td>61</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>68</td>
<td>78</td>
</tr>
</tbody>
</table>
Table 4:
Relative errors corresponding to the "rule".

<table>
<thead>
<tr>
<th>ALPHA</th>
<th>0.00</th>
<th>0.01</th>
<th>0.02</th>
<th>0.03</th>
<th>0.04</th>
<th>0.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 5

Distances between two normal distributions

<table>
<thead>
<tr>
<th>$u$</th>
<th>.1</th>
<th>.2</th>
<th>.3</th>
<th>.4</th>
<th>.5</th>
<th>.6</th>
<th>.7</th>
<th>.8</th>
<th>.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td>.023</td>
<td>.056</td>
<td>.084</td>
<td>.112</td>
<td>.138</td>
<td>.164</td>
<td>.189</td>
<td>.214</td>
<td>.238</td>
<td>.251</td>
</tr>
</tbody>
</table>