CERN - DATA HANDLING DIVISION
DD/73/2
B. Schorr
January 1973

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 particularly interested in the question of the choice of the number of class intervals of the χ^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:

$$H_0: F(x) = F_0(x) = x, \qquad 0 \le x \le 1 . \tag{1.1}$$

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

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

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

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

where

$$\sum_{i=1}^{k} N_i = n , \tag{1.4}$$

*) Dr. Benno Schorr, CERN, 1211 Geneva 23, Switzerland.

$$\pi_i = x_i - x_{i-1}, \qquad i = 1, \ldots, k . \tag{1.5}$$

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

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

$$R = \left\{(n_1, \ldots, n_k): \quad x^2 = \sum_{i=1}^{k} \frac{(n_i - n\pi_i)^2}{n\pi_i} > \chi^2_{\alpha,k-1}\right\}, \tag{1.6}$$

where $\chi^2_{\alpha,k-1}$ is the upper α-point of the χ^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 χ^2-distribution for X^2? For this question we refer, for example, to Vessereau's [1] and Slakter's [2] papers. We are more interested in the choice of k and the π_i.

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

$$\pi_i = \frac{1}{k}, \qquad i = 1, 2, \ldots, k . \tag{1.7}$$

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

$$H_1: F(x) = F_1(x) \tag{1.8}$$

and introduced the following distance

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

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 χ^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

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

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

$$\Delta = \Delta_n = \frac{5}{k_n} - \frac{4}{k_n^2} \tag{1.11}$$

and

$$k_n = \left[4\sqrt[5]{\frac{2(n-1)^2}{c_\alpha^2}} \right] , \tag{1.12}$$

we have

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

where c_α is the upper α-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 Δ_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 π_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 (≤ 50) to Slakter's paper [7] in which the power of the χ^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

$$\rho(F_0, F_1) = \int_0^1 |F_1(x) - F_0(x)|\, dx \qquad (1.15)$$

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:

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

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

with

$$\gamma_i = p_i - \pi_i , \qquad i = 1, 2, \ldots, k , \tag{2.3}$$

$$\delta^2 = \sum_{i=1}^{k} \frac{\gamma_i^2}{\pi_i} , \tag{2.4}$$

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

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

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

$$+ \sum_{i=1}^{k} \frac{\gamma_i}{\pi_i^2} + \sum_{i=1}^{k} + \frac{1 - k \cdot \pi_i}{\pi_i} . \tag{2.6}$$

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

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

$$Var(X^2) = 2(k - 1) + 4n\delta^2 - \frac{2(k - 1)}{n} + \frac{R_3}{n} \delta^2 + \frac{R_4}{n} , \tag{2.8}$$

with

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

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

Let us now consider the non-central χ^2-distribution with $k - 1$ degrees of freedom and the non-central parameter λ^2. Its characteristic function is given by

$$\phi(t) = \frac{e^{\lambda^2 it/(1-2it)}}{(1 - 2it)^{(k-1)/2}} \qquad (2.11)$$

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

$$E(\chi'^2) = k - 1 + \lambda^2 , \qquad (2.12)$$

$$\mathrm{Var}(\chi'^2) = 2(k - 1) + 4\lambda^2 . \qquad (2.13)$$

setting

$$\lambda^2 = n \cdot \delta^2 \qquad (2.14)$$

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

$$p_i = \pi_i + \frac{\Delta_i}{\sqrt{n}} , \qquad \Delta_i \geq 0 , \qquad i = 1, 2, \ldots, k ,$$

and for $n \to \infty$ the random variable X^2 has a non-central χ^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 χ^2-distribution will be better, in general, than for arbitrary π_i.

Indeed we know the characteristic function $\phi(t)$ of χ'^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 χ'^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

$$Y = \sqrt{\frac{2\chi'^2(k - 1 + \lambda^2)}{k - 1 + 2\lambda^2}} , \tag{2.15}$$

then Y is asymptotically normally distributed with the mean value

$$m = \sqrt{\frac{2(k - 1 + \lambda^2)^2}{k - 1 + 2\lambda^2} - 1} \tag{2.16}$$

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

$$n + \lambda^2 \geq 50 . \tag{2.17}$$

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

$$\beta(H_1) = \Pr(X^2 > \chi^2_{\alpha,k-1} \mid H_1) \tag{2.18}$$

can be approximated by

$$\beta_0(H_1) = \Pr(\chi'^2 > \chi^2_{\alpha,k-1} \mid H_1) \tag{2.19}$$

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

$$\beta_1(H_1) = 1 - \Phi(z) , \tag{2.20}$$

where

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

and Φ 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 λ^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 $0 \le \rho < 1/2$ let

$$S(\rho) = \{F: F \in K,\ \rho(x,F) \ge \rho\} , \tag{3.1}$$

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

We want to determine

$$\inf_{F_1 \in S(\rho)} \beta_0(H_1) \tag{3.2}$$

for arbitrary and fixed π_i (i = 1, 2, ..., k), n, k and $\rho > 0$. To solve this problem it is sufficient, because of the lemma of paragraph 2, to determine

$$I = \inf_{F_1 \in S(\rho)} \delta^2 = \inf_{F_1 \in S(\rho)} \sum_{i=1}^{k} \frac{(p_i - \pi_i)^2}{\pi_i} \tag{3.3}$$

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,

$$F(x) \geq x, \qquad x \in (0,1) \ . \tag{3.4}$$

Since we are only interested in $I > 0$, we have another condition for ρ:

$$\frac{1}{2} \sum_{i=1}^{k} \pi_i^2 < \rho \ , \tag{3.5}$$

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

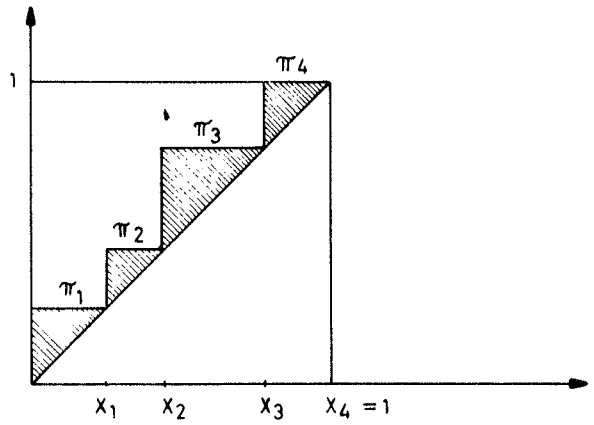

Fig. 1

Under the hypothesis (3.4) we have for $F \in S(\rho)$:

$$\int_0^1 \left[F(x) - x\right] dx \geq \rho \ . \tag{3.6}$$

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

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

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

$$\sum_{i=1}^{k} \xi_i \cdot p_i \leq \frac{1}{2} - \rho \ , \tag{3.8}$$

where

$$x_{i-1} \leq \xi_i \leq x_i \ , \tag{3.9}$$

and

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

for i = 1, 2, ..., 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 δ^2 as follows:

Problem A

Find p_i for i = 1, 2, ..., k such that

$$\delta^2 = \sum_{i=1}^{k} \frac{(p_i - \pi_i)^2}{\pi_i} \tag{3.10}$$

takes its minimum under the conditions

$$p_i \geq 0, \qquad i = 1, 2, \ldots, k , \tag{3.11}$$

$$\sum_{i=1}^{k} p_i = 1, \tag{3.12}$$

$$\sum_{i=1}^{r} p_i \geq \sum_{i=1}^{r} \pi_i , \qquad r = 1, 2, \ldots, k - 1 \tag{3.13}$$

$$\sum_{i=1}^{k} \xi_i p_i = \frac{1}{2} - \rho , \tag{3.14}$$

where the ξ_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 ξ_i are known constants, then problem A is a problem of mathematical programming with the quadratic objective function δ^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 ξ_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 δ^2 where in (3.6) the equality sign is valid. Such p_i^* always exist for every ρ with $0 < \rho < 1/2$. Then we can show that this step function is the only distribution function for which δ^2 takes its minimum under the given conditions. For, a distribution function F with

$$p_i^* = \int_{x_{i-1}}^{x_i} dF , \qquad i = 1, 2, \ldots, k , \tag{3.15}$$

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

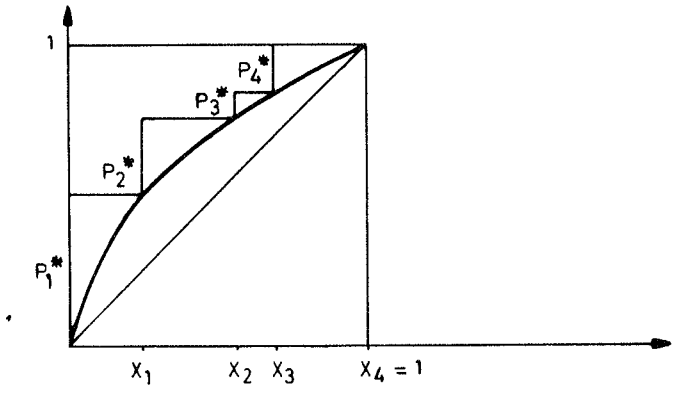

Fig. 2

Therefore we can see that the density functions f of the distribution functions $F \in S(\rho)$ which lead to the infimum of δ^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

$$\xi_i = x_{i-1} , \qquad i = 1, 2, \ldots, k . \tag{3.16}$$

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 π_i are all in a certain neighbourhood of 1/k. We conjecture that the given solution is even valid for arbitrary π_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

$$p_i = \frac{1}{k}\left\{1 + \frac{2(k - r)}{r(r + 1)}\left[2(r + 1) - 3i\right] + \right.$$

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

$$\text{for } i = 1, 2, \ldots, r \; ; \tag{3.17}$$

$$p_i = 0 \quad \text{for } i = r + 1, \ldots, k \; , \tag{3.18}$$

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

$$r = k \qquad \text{for } \frac{1}{2k} < \rho < L_0(k) \; , \tag{3.19}$$

$$2 \le r \le k - 1 \quad \text{for } L_0(r + 1) \le \rho < L_0(r) \tag{3.20}$$

with

$$L_0(r) = \frac{3k + 4 - 2r}{6k} \; . \tag{3.21}$$

The minimum of δ^2 is given by

$$\delta^2_{r,k} = \frac{r}{k}\frac{\left[\rho - \frac{k^2 - r(r - 1)}{2k^2}\right]^2}{D_{r,k}} + \tag{3.22}$$

$$+ \frac{k - r}{k}\left\{1 + \frac{4k^2}{r(r + 1)}\left[\rho - \frac{(k - r)^2 + k}{2k^2}\right]\right\}$$

with

$$D_{r,k} = \frac{r^2(r^2 - 1)}{12k^4} \; , \qquad r = 2, 3, \ldots, k \; . \tag{3.23}$$

Proving theorem 2 we shall formally do the steps for arbitrary π_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 λ_1, λ_2 be the multipliers. Then we consider the Lagrange function

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

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

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

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

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

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

$$D_k = \sum_{i=1}^{k} x_{i-1}^2\,\pi_i - \left(\sum_{i=1}^{k} x_{i-1}\,\pi_i\right)^2\ . \qquad (3.28)$$

D_k is always positive, for it can be written as

$$D_k = \sum_{i=1}^{k} x_{i-1}\,\pi_i \left(x_{i-1} - \sum_{j=1}^{k} x_{j-1}\,\pi_j \right)^2 . \tag{3.29}$$

Simple geometric arguments lead to

$$\frac{1}{2} \sum_{j=1}^{k} \pi_j^2 = \frac{1}{2} - \sum_{j=1}^{k} x_{j-1}\,\pi_j . \tag{3.30}$$

Using (3.30) in (3.29) it is easily seen that $D_k > 0$. Especially for problem A_0, D_k is given by

$$D_k = D_{k,k} = \frac{k^2 - 1}{12k^2} , \qquad k \geq 2 . \tag{3.31}$$

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

$$p_i = \pi_i \left(1 + \frac{1}{D_k} \left[\sum_{j=1}^{k} x_{j-1}\,\pi_j - x_{i-1} \right] \left[\rho - \frac{1}{2} + \sum_{j=1}^{k} x_{j-1}\,\pi_j \right] \right) \tag{3.32}$$

for $i = 1, 2, \ldots, k$.

Since δ^2 is a positive definite quadratic form in the p_i, (3.32) furnishes the absolute minimum of δ^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

$$\sum_{i=1}^{r} \pi_i \left[\frac{1}{2} - \frac{1}{2} \sum_{j=1}^{k} \pi_j^2 - x_{i-1} \right] \geq 0 \tag{3.33}$$

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 π_i, (3.33) must also be true for all π_i 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 ρ 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 decreasing 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 calculation, to the condition

$$\rho \leq \frac{D_k}{x_{k-1} - \sum_{j=1}^{k} x_{j-1}\,\pi_j} + \frac{1}{2} - \sum_{j=1}^{k} x_{j-1}\,\pi_j \,. \tag{3.34}$$

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

$$\rho \leq \frac{k + 4}{6k} = L_0(k) \,. \tag{3.35}$$

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

$$\frac{1}{2} - \sum_{j=1}^{k} x_{j-1}\,\pi_j < \rho < \frac{D_k}{x_{k-1} - \sum_{j=1}^{k} x_{j-1}\,\pi_j} + \frac{1}{2} - \sum_{j=1}^{k} x_{j-1}\,\pi_j \,, \tag{3.36}$$

and for

$$\frac{1}{2k} < \rho < L_0(k) \tag{3.37}$$

in the special case, (3.32) yields the solution of problem A for π_i 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 ρ is larger than in condition (3.36). We observe that $p_k = 0$ if ρ is equal to the right side of (3.36). In addition, geometric arguments give us a hint that step by step, with ρ 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:

$$p_i > 0 \quad \text{for } i = 1, \ldots, r \tag{3.38}$$

$$p_i = 0 \quad \text{for } i = r, r + 1, \ldots, k\ , \tag{3.39}$$

where r is dependent on ρ. 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

$$\sum_{i=1}^{r} \frac{(p_i - \pi_i)^2}{\pi_i} \tag{3.40}$$

takes its minimum under the conditions

$$\sum_{i=1}^{r} p_i = 1\ , \tag{3.41}$$

$$\sum_{i=1}^{r} x_{i-1}\, p_i = \frac{1}{2} - \rho \ . \tag{3.42}$$

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

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

for $i = 1, 2, \ldots, r$ with

$$D_r = x_r \sum_{i=1}^{r} x_{i-1}^2\, \pi_i - \left(\sum_{i=1}^{r} x_{i-1}\, \pi_i \right)^2 \ . \tag{3.44}$$

For problem A_0 we have

$$D_r = D_{r,k} = \frac{r^2(r^2 - 1)}{12k^4} \ , \qquad r = 2, 3, \ldots, k \ . \tag{3.45}$$

Again because of (3.45) and since D_r is continuous in the π_i, we have $D_r > 0$ at least for all π_i in a certain neighbourhood of $1/k$.

Thus stage (I) is treated.

Stage (II) is expressed by the condition

$$\sum_{i=1}^{s} p_i \geq \sum_{i=1}^{s} \pi_i \quad \text{for } s = 1, 2, \ldots, r - 1 \ . \tag{3.46}$$

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 ρ for given r:

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

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 ρ 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

$$\rho > \frac{k - r + 1}{2k} . \tag{3.48}$$

Assume that ρ 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

$$\rho < \frac{3k + 4 - 2r}{6k} = L_0(r) \tag{3.49}$$

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

$$L_0(r+1) \leq \rho < L_0(r) , \tag{3.50}$$

i.e. for given r the distance ρ must satisfy (3.50); or, we can interpret it in the opposite direction, i.e. if ρ 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 δ^2 we obtain (3.22) and the proof of theorem 2 is completed. But we have shown more; namely, that for all π_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 ρ. Analogous arguments to those which gave us (3.50) lead to the condition

$$L(r + 1) \leq \rho < L(r) \quad \text{for } r = 2, 3, \ldots, k - 1 \tag{3.51}$$

with

$$L(r) = \frac{1}{x_{r-1}\, x_r - \sum_{j=1}^{r} x_{j-1}\, \pi_j} \left\{ D_r + (1 - x_r) \left(\sum_{j=1}^{r} x_{j-1}^2\, \pi_j - x_r \sum_{j=1}^{r} x_{j-1}\, \pi_j \right) \right\} + \frac{1}{2} - \sum_{j=1}^{r} x_{j-1}\, \pi_j \,. \tag{3.52}$$

Inserting (3.43) and $p_i = 0$ for $i = r + 1, \ldots, k$ into δ^2 and denoting the result by δ_r^2, we finally can formulate the following theorem.

Theorem 3

Under the condition that all the π_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 δ^2 is given by

$$\delta_r^2 = \frac{1}{D_r^2} \sum_{i=1}^{r} \left\{ \left(\sum_{j=1}^{r} x_{j-1}\, \pi_j - x_r\, x_{i-1} \right) \left(\rho - \frac{1}{2} + \sum_{j=1}^{r} x_{j-1}\, \pi_j \right) + \right.$$

$$\left. + (1 - x_r) \left[\sum_{j=1}^{r} x_{j-1}^2\, \pi_j - \left(\sum_{j=1}^{r} x_{j-1}\, \pi_j \right) x_{i-1} \right] \right\}^2 \pi_i + \qquad (3.53)$$

$$+ \sum_{i=r+1}^{k} \pi_i \, .$$

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

Corollary

For all $\pi_i = 1/k$ and for

$$\frac{1}{2k} < \rho < \frac{k + 4}{6k} \qquad (3.54)$$

we have $R_4 = 0$ 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 χ^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 χ^2-distribution only for small δ^2 and for large n, we shall confine the rest

of the paper to the case where ρ 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 $\delta_r^2 = \delta_k^2$. It is easily seen that we then obtain

$$\delta_k^2 = \frac{\left(\rho - \frac{1}{2}\sum_{j=1}^{k} \pi_j^2\right)^2}{D_k} \tag{3.55}$$

for arbitrary π_i, and

$$\delta_k^2 = \delta_{k,k}^2 = \frac{12k^2}{k^2 - 1}\left(\rho - \frac{1}{2k}\right)^2 \tag{3.56}$$

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

For comparison with the results of Mann and Wald we cite the corresponding expression for $\delta_{k,k}^2$ if δ^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

$$\delta_{k,k}^2 = \delta_{MW}^2 = 4\left(\Delta - \frac{1}{k}\right)^2 . \tag{3.57}$$

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 ρ and Δ, 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 α and n in order to reach a certain power and smallest distance ρ. 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 (α, n, k, ρ). We assume α and n to be given. In addition, let β be the probability of the error of the second kind. Since

$$\beta = 1 - \beta_1(H_1) \tag{4.1}$$

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

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

$$\alpha = 0.01,\ 0.05,\ 0.10, \tag{4.2}$$

$$\beta = 0.50,\ 0.40,\ 0.30,\ 0.20,\ 0.10,\ 0.05, \tag{4.3}$$

$$n = 50\ (50)\ 1000;\ 1100\ (100)\ 1500;\ 2000. \tag{4.4}$$

For each combination (α, β, 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 ρ_{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 ρ-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

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

This procedure led to the following Rule:

In order to satisfy (4.5) one may choose

k	for n between
15	50-150
20	200-350
25	400-600
30	550-800
33	850-1000
36	1100-1300
40	1400-2000

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 α. 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-Küchler 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 ρ, an example may be useful. Let

$$F_0(x) = \Phi\left(\frac{x}{\sigma}\right) , \qquad \sigma > 0 , \qquad (4.6)$$

$$F_1(x) = \Phi\left(\frac{x + m}{\sigma}\right), \qquad m > 0 , \tag{4.7}$$

be two distribution functions of the normal type. Then

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

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

$$\rho = \frac{1}{2} \operatorname{erf}\left(\frac{u}{2}\right) \tag{4.9}$$

with

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt . \tag{4.10}$$

Table 5 shows some values of ρ for different u.

FINAL REMARK

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

REFERENCES

[1] A. Vessereau, Sur les conditions d'application du critérium χ^2 de Pearson, Bull de l'Inst. Int. Stat. 36 (1958) 87-101.

[2] M.J. Slakter, Comparative validity of the chi-square and two modified goodness of fit tests for small but equal frequencies, Biometrika 53 (1966) 619-22.

[3] H.B. Mann and A. Wald, On the choice of the number of class intervals in the application of the chi-square test, Ann. Math. Stat. 13 (1942) 306-317.

[4] C.A. Williams, On the choice of the number and width of classes for the chi-square test of goodness of fit, Ann. Math. Stat. 45 (1950) 77-86.

[5] I. Beier-Küchler und P. Neumann, Über die Klassenwahl beim χ^2-Anpassungstest, Math. OF Stat. 1 (1970) 55-68.

[6] P.B. Patnaik, The non-central χ^2- and F-distributions and their applications, Biometrika 36 (1949) 202-232.

[7] M.J. Slakter, Accuracy of an approximation to the power of the chi-square goodness of fit test with small but equal expected frequencies, J. Amer. Stat. Assoc. 63 (1968) 912-918.

[8] C. Eisenhart, The power function of the χ^2-test, Bull. Amer. Math. Soc. 44 (1938) 32.

[9] H.P. Künzi, H.G. Tzschach and C.A. Zehnder, Numerical methods of mathematical optimization (Academic Press, New York and London, 1968).

Table 1

Optimal k for α = 0.01, 0.05, 0.1; β = 0.5, ..., 0.05, and n = 50, ..., 2000 (ρ-distance)

ALPHA	.01						.05						.10					
BETA	.50	.40	.30	.20	.10	.05	.50	.40	.30	.20	.10	.05	.50	.40	.30	.20	.10	.05
50	15	14	14	13	13	12	16	16	15	14	14	13	18	17	16	15	14	14
100	19	19	18	18	17	17	22	21	20	19	18	17	24	22	21	20	19	18
150	23	22	22	21	20	20	26	24	23	23	21	21	28	26	25	24	22	21
200	26	25	24	23	23	22	29	27	26	25	24	23	32	30	28	27	25	24
250	28	27	26	26	25	24	32	30	29	28	26	25	34	32	31	29	27	26
300	30	29	28	28	26	26	34	32	31	30	28	27	37	35	33	31	30	28
350	32	31	30	29	28	27	36	34	33	32	30	29	40	37	35	33	32	30
400	34	33	32	31	30	29	38	36	35	33	32	31	42	39	37	35	33	32
450	36	34	33	32	31	30	40	38	36	35	33	32	43	41	39	37	35	33
500	37	36	35	34	32	32	42	40	38	36	34	33	45	43	40	38	36	35
550	38	37	36	35	34	33	43	41	39	38	36	35	47	44	42	40	38	36
600	40	38	37	36	35	34	44	42	41	39	37	36	49	46	43	41	39	37
650	41	40	38	37	36	35	46	44	42	40	38	37	50	47	45	43	40	38
700	42	41	40	38	37	36	48	45	43	42	40	38	52	49	46	44	41	40
750	43	42	41	40	38	37	49	46	44	43	41	39	53	50	48	45	42	41
800	44	43	42	41	39	38	50	48	46	44	42	40	55	51	49	46	44	42
850	46	44	43	42	40	39	51	49	47	45	43	41	56	53	50	47	45	43
900	46	45	44	42	41	40	52	50	48	46	44	42	57	54	51	48	46	44
950	48	46	44	43	42	40	54	51	49	47	44	43	59	55	52	49	47	45
1000	48	47	46	44	42	41	55	52	50	48	46	44	60	56	53	50	48	46
1100	50	49	47	46	44	43	57	54	52	50	47	46	62	58	55	52	50	48
1200	52	50	49	48	46	44	58	56	54	51	49	47	64	60	57	54	51	49
1300	54	52	50	49	47	46	60	58	55	53	50	49	66	61	59	56	53	51
1400	56	54	52	50	48	47	62	60	57	55	52	50	68	64	62	58	54	52
1500	57	55	54	52	50	48	64	62	58	56	53	52	70	66	65	59	56	54
2000	63	62	60	58	56	54	72	68	65	63	60	58	79	74	70	66	62	60

Table 2

Minimal ρ corresponding to Table 1

ALPHA	.01						.05						.10					
BETA	.50	.40	.30	.20	.10	.05	.50	.40	.30	.20	.10	.05	.50	.40	.30	.20	.10	.05
50	.196	.208	.221	.236	.256	.273	.165	.178	.192	.207	.229	.246	.146	.160	.175	.192	.214	.231
100	.147	.156	.165	.176	.191	.203	.124	.134	.144	.155	.171	.184	.110	.121	.132	.144	.160	.173
150	.124	.132	.140	.149	.161	.171	.105	.113	.122	.131	.144	.155	.094	.103	.112	.122	.135	.146
200	.111	.117	.124	.132	.142	.151	.094	.101	.108	.117	.128	.137	.083	.091	.099	.108	.120	.129
250	.101	.107	.113	.120	.130	.137	.085	.092	.099	.106	.117	.125	.076	.083	.090	.099	.109	.118
300	.094	.099	.105	.111	.120	.127	.079	.085	.092	.099	.108	.116	.071	.077	.084	.091	.101	.109
350	.088	.093	.098	.104	.113	.119	.075	.080	.086	.092	.101	.108	.067	.073	.079	.086	.095	.102
400	.083	.088	.093	.099	.106	.113	.071	.076	.081	.087	.096	.102	.063	.069	.075	.081	.090	.097
450	.079	.084	.088	.094	.101	.107	.067	.072	.077	.083	.091	.097	.060	.066	.071	.077	.086	.092
500	.076	.080	.085	.090	.097	.103	.064	.069	.074	.080	.087	.093	.058	.063	.068	.074	.082	.088
550	.073	.077	.081	.086	.093	.098	.062	.067	.071	.077	.084	.090	.055	.060	.066	.071	.079	.085
600	.070	.074	.078	.083	.090	.095	.060	.064	.069	.074	.081	.086	.054	.058	.063	.069	.076	.082
650	.068	.072	.076	.081	.087	.092	.058	.062	.067	.072	.078	.084	.052	.056	.061	.066	.073	.079
700	.066	.070	.074	.078	.084	.089	.056	.060	.065	.069	.076	.081	.050	.055	.059	.064	.071	.077
750	.064	.068	.072	.076	.082	.086	.055	.059	.063	.067	.074	.079	.049	.053	.058	.063	.069	.074
800	.063	.066	.070	.074	.080	.084	.053	.057	.061	.066	.072	.077	.048	.052	.056	.061	.067	.072
850	.061	.064	.068	.072	.078	.082	.052	.056	.060	.064	.070	.075	.046	.051	.055	.060	.066	.071
900	.060	.063	.066	.070	.076	.080	.051	.054	.058	.063	.068	.073	.045	.049	.054	.058	.064	.069
950	.058	.061	.065	.069	.074	.078	.050	.053	.057	.061	.067	.071	.044	.048	.052	.057	.063	.067
1000	.057	.060	.064	.067	.072	.077	.049	.052	.056	.060	.065	.070	.044	.047	.051	.056	.061	.066
1100	.055	.058	.061	.065	.070	.074	.047	.050	.054	.058	.063	.067	.042	.046	.049	.054	.059	.063
1200	.053	.056	.059	.062	.067	.071	.045	.048	.052	.056	.061	.065	.040	.044	.048	.052	.057	.061
1300	.051	.054	.057	.060	.065	.069	.044	.047	.050	.054	.059	.062	.039	.043	.046	.050	.055	.059
1400	.050	.052	.055	.059	.063	.066	.042	.045	.049	.052	.057	.061	.038	.041	.045	.048	.053	.057
1500	.048	.051	.054	.057	.061	.065	.041	.044	.047	.051	.055	.059	.037	.040	.043	.047	.052	.056
2000	.042	.045	.048	.050	.054	.057	.037	.039	.042	.045	.049	.052	.033	.036	.039	.042	.046	.049

Table 3

Optimal k for α = 0.01, 0.05, 0.1;
β = 0.5 and n = 50, ..., 2000 (Mann-Wald formula)

ALPHA	.01	.05	.10
BETA	.50	.50	.50
50	15	17	19
100	20	23	26
150	24	27	30
200	27	31	34
250	29	34	37
300	32	36	40
350	34	39	43
400	35	41	45
450	37	43	47
500	39	45	49
550	40	46	51
600	42	48	53
650	43	50	55
700	45	51	57
750	46	53	58
800	47	54	60
850	48	55	61
900	49	57	63
950	50	58	64
1000	51	59	65
1100	53	61	68
1200	55	64	70
1300	57	66	73
1400	59	68	75
1500	61	70	77
2000	68	78	86

Table 4

Relative errors e_r corresponding to the "rule"

ALPHA	.01						.05						.10					
BETA	.50	.40	.30	.20	.10	.05	.50	.40	.30	.20	.10	.05	.50	.40	.30	.20	.10	.05
50	0.00	.00	.00	.00	.00	.00	.00	.00	0.00	.00	.00	.00	.00	.00	.00	0.00	.00	.00
100	.01	.01	.00	.00	.00	.00	.02	.01	.01	.01	.00	.00	.03	.02	.01	.01	.01	.00
150	.02	.02	.01	.01	.01	.01	.04	.03	.02	.02	.01	.01	.05	.04	.03	.02	.02	.01
200	.01	.01	.00	.00	.00	.00	.02	.01	.01	.01	.00	.00	.03	.02	.01	.01	.01	.00
250	.01	.01	.01	.01	.00	.00	.03	.02	.02	.01	.01	.01	.04	.03	.02	.02	.01	.01
300	.02	.02	.01	.01	.01	.01	.04	.03	.02	.02	.01	.01	.05	.04	.03	.02	.02	.01
350	.01	.01	.00	.00	.00	.00	.02	.01	.01	.01	.00	.00	.03	.02	.01	.01	.01	.00
400	.01	.01	.01	.00	.00	.00	.02	.02	.01	.01	.01	.00	.03	.03	.02	.01	.01	.01
450	.01	.01	.01	.01	.01	.00	.03	.02	.02	.01	.01	.01	.04	.03	.02	.02	.01	.01
500	.02	.02	.01	.01	.01	.01	.03	.03	.02	.02	.01	.01	.05	.04	.03	.02	.02	.01
550	.01	.01	.00	.00	.00	.00	.02	.01	.01	.01	.00	.00	.03	.02	.01	.01	.01	.00
600	.01	.01	.01	.00	.00	.00	.02	.02	.01	.01	.01	.00	.03	.02	.02	.01	.01	.01
650	.01	.01	.01	.01	.00	.00	.02	.02	.01	.01	.01	.00	.04	.03	.02	.01	.01	.01
700	.01	.01	.01	.01	.00	.00	.03	.02	.02	.01	.01	.01	.04	.03	.02	.02	.01	.01
750	.02	.01	.01	.01	.01	.00	.03	.02	.02	.01	.01	.01	.05	.04	.03	.02	.01	.01
800	.02	.02	.01	.01	.01	.01	.04	.03	.02	.02	.01	.01	.05	.04	.03	.02	.02	.01
850	.01	.01	.01	.01	.00	.00	.03	.02	.02	.01	.01	.01	.04	.03	.02	.02	.01	.01
900	.01	.01	.01	.01	.00	.00	.03	.02	.02	.01	.01	.01	.04	.03	.02	.02	.01	.01
950	.02	.01	.01	.01	.01	.00	.03	.02	.02	.01	.01	.01	.05	.04	.03	.02	.01	.01
1000	.02	.02	.01	.01	.01	.01	.03	.03	.02	.02	.01	.01	.05	.04	.03	.02	.02	.01
1100	.01	.01	.01	.01	.00	.00	.03	.02	.02	.01	.01	.01	.04	.03	.02	.02	.01	.01
1200	.02	.01	.01	.01	.01	.00	.03	.03	.02	.02	.01	.01	.05	.04	.03	.02	.01	.01
1300	.02	.02	.01	.01	.01	.01	.04	.03	.02	.02	.01	.01	.06	.04	.03	.02	.02	.01
1400	.01	.01	.01	.01	.00	.00	.03	.02	.02	.01	.01	.01	.04	.03	.02	.02	.01	.01
1500	.02	.01	.01	.01	.01	.00	.04	.03	.02	.01	.01	.01	.05	.04	.03	.02	.01	.01
2000	.04	.03	.02	.02	.01	.01	.05	.04	.04	.03	.02	.02	.07	.06	.04	.03	.03	.02

Table 5

Distances between two normal distributions

u	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
ρ	.023	.056	.084	.112	.138	.164	.189	.214	.238	.251