THE "MEDIAN-QUARTILE TEST"

JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

NAS 7-100

108871
2

THE "MEDIAN-QUARTILE TEST"

by

R.K. Bauer

Translation by
Marjorie K. Feitis

The "Median-Quartile-Test": A Procedure for the Non-Parametric Testing of Two Independent Samples as to Unspecified Distribution Differences

by

R.K. Bauer, Dusseldorf[1]

Summary: The test procedure proposed is theoretically proved, methodically presented, and represented by a numerical example from research work in melting shops. The example result is compared with the results obtained from related test procedures.

Problem

We are given two, independent samples, which were chosen at random, of observed values for a chance variable of a numerical scale having an unknown distribution type. The zero hypothesis H_0 is to be investigated; according to this, the two samples are chosen from either a common, higher, fundamental totality or from two fundamental totalities of identical distribution type. The opposing hypothesis H_1 consists of an unspecified hypothesis totality, according to which the differences in distribution are not only limited to localization, but differences of an arbitrary nature are allowed - such as, dispersion, asymmetry, excess or modality.

The following testing procedures are possible (see

(1) Dipl.Math. Rainald K. Bauer, Verein Deutscher Eisenhuttenleute, Energie- und Betriebswirtschaftsstelle, Dusseldorf, Breite Str. 27 (author's address).

Siegel, 1956, pg 157 f):

1. The classical X^2-test of K. Pearson for the case of the 2xm-field.[1] It can be also replaced by the exact probability test of R.A. Fisher (see W.G. Cochran 1952, 1954), which can be extended to the 2xm-field case (for instance, according to S.N. Roy 1957: formula 15.2.3.1). It can be replaced by the latter depending on whether or not the modification of K.D. Tocher is followed by considering the sample nature of the border sums;

2. The (two-sided) 2-sample test of A. Kolmogoroff and N.V. Smirnoff; and finally,

3. The iteration test of A. Wald and J. Wolfowitz.

Let us first disregard the problem of determining the accuracy with regard to the three procedures mentioned above (in any case, they have never been adequately studied, up to the present time). At least three objections then remain, making the method proposed by Siegel somewhat unsatisfactory, in a theoretical sense as well as in a practical sense:

a) The X^2 and the Fisher test and - depending on the nature of the data - the Kolmogoroff-Smirnoff test, make a value classification, which is basically arbitrary, of the numerically-scaled random variables (such as we assumed at the beginning of this article). Without question, this introduces a completely arbitrary element into the testing process.

b) The Wald-Wolfowitz test is quite involved in terms of computation (this is also true of the Kolmogoroff-Smirnoff

(1) In this paper, X is used to designate the Greek letter chi (χ).

test under certain conditions), and, in addition, it is completely useless if identical observational values simultaneously occur many times in the two samples. Also, there is some doubt as to its asymptotic normality, a basic condition for its application to larger and/or unequal sample sizes. (See E.L. Lehmann 1953).

c) None of the three procedures permits an interpretation of the nature of the determined distribution difference, as one of its significant results.

Therefore, it is desirable that a procedure be found which:

A) objectifies the reduction in data scaling which may be necessary,

B) holds the computational work within bounds, and

C) interprets the hypothesis H_1.

The modification of the so-called median test (originally described by A.M. Mood 1950) in the form of a generalized testing hypothesis has proven to be suitable in a few problems, such as the calculation example presented below. The procedure in no way fulfills the conditions (A), (B), (C) in an ideal way. It is called median-quartile test in this text. The next section presents a description of the test method, followed by an example, presented in full detail, which is encountered in statistical work at a steel mill.

Method

1) Let the N observational values:

$$x = x_{ik} \qquad (i=1,2;\ k=1,2,\ldots\ldots,\ N_i)$$

of the two samples 1 and 2, having sizes N_1 and $N_2 (N_1 + N_2 = N)$ respectively, be ordered in a common sequence:

$$\{x^{(r)}\} \qquad (r=1,2,\ldots.N)$$

in such a way that always

$$x^{(r)} < x^{(r+1)} \quad .$$

If it is impossible to meet this requirement, due to identical observational values x, then a mean rank number $\bar{r}$ is assigned to these values, according to the formula:

$$\bar{r} = r_{min} + \frac{r_{max} - r_{min}}{2}$$

in which r_{min} is the rank number following the observational value which is next smaller in size to the identical x, $u = x - \epsilon$ ($\epsilon \leqslant 1$), and $r_{max} = r_{min} + t - 1$, if t identical x are present. If u occurs many times, then $r_{min}(x)$ is the rank number after $r_{max}(u)$.

2) Let N=4q-1 and q be a whole number. Find the lower quartile in($x^{(r)}$)

$$Q_1 = x^{(q)}$$

the median

$$Z = Q_2 = x^{(2q)}$$

and the upper quartile

$$Q_3 = x^{(3q)} \quad .$$

If $q = \frac{N+1}{4}$ is not a whole number and/or the Q's cannot be uniquely defined because of identical observational values (having mean rank numbers), then they are determined in such a way that, at least in an approximate manner, the following sample probabilities are valid, as exactly as is possible:

$$p(x \leqslant Q_1) = \frac{1}{4}$$

$$p(x \leqslant Q_2) = \frac{1}{2} \quad \text{and}$$

$$p(x \leqslant Q_3) = \frac{3}{4} \quad .$$

3) Let the N observational values x be reduced to the following 2x4 - field table:

	Sample 1	Sample 2	
$x \leqslant Q_1$			
$Q_1 < x \leqslant Q_2$			
$Q_2 < x \leqslant Q_3$			
$x > Q_3$			

4) Let $n_{ij} = n_{ij}(x)$ $(j = 1,2,3,4)$

be the abundance of the observational value x of the i-sample in the j-quarter of the rank sequence $(x^{(r)})$ (which is common to both samples), that is, the abundance of

$$Q_{j-1} < x_{ik} \leqslant Q_j .$$

Furthermore, let

$$n_j = \sum_{i=1}^{2} n_{ij} \text{ (or at least } \approx \frac{N}{4} \text{)}$$

be the right border sums of the 2x4-field table, shown in (3). The lower border sums

$$N_i = \sum_{j=1}^{4} n_{ij}$$

are the given sample sizes N_1 and N_2 by definition.

Then the sample probability for the occupation of the 2x4- field table, given by a concrete n_{ij} and the border sums N_i and n_j, is:

$$p(n_{ij}|N_i;n_j) = \frac{\prod_{i=1}^{2} N_i! \prod_{j=1}^{4} n_j!}{N! \prod_{i=1}^{2} \prod_{j=1}^{4} n_{ij}!} = P_{beob} \qquad (P_{beob} = P_{obs})$$

or, if $n_j = \frac{N}{4}$, it falls exactly:

$$p(n_{ij}|N_i;n_j) = \frac{\prod_{i=1}^{2} N_i! \, 4\left(\frac{N}{4}\right)!}{N! \prod_{i=1}^{2} \prod_{j=1}^{4} n_{ij}!} = P_{beob}$$

and the expected value of the sample of n_{ij} is

$$E(n_{ij}|N_i;n_j) = \frac{N_i}{4} = En_{ij} \quad ,$$

or, again, if $n_j = \frac{N}{4}$ holds exactly,

$$E(n_{ij}|N_i;n_j) = \frac{N_i}{4} = En_{ij}.$$

5) The hypothesis H_0 (see "problem") can be tested as follows:

a) without any condition with

$$P_{obs} = \frac{N_1!N_2!n_1!n_2!n_3!n_4}{N!n_{11}!n_{12}!n_{13}!\ n_{14}!n_{21}!n_{22}!n_{23}!n_{24}}$$

or, when $P_{obs} < a$ (a is the given probability of air), with

$$P = P_{obs} + \sum P_{extr}$$

(where $\sum P_{extr}$ is the sum of all the p's, calculated according to the defining formulas for P_{obs}, which result from changing the n_{ij} , retaining the border sums N_i and n_j . These p's also show that H_0 is wrong, to a greater extent than does the 2x4-field table, through P_{obs} in the manner already indicated. Also, one always has $P_{extr} < P_{obs}$). Or, when $P > a$, with

$$T = \frac{a - \sum P_{extr}}{P_{obs}} \quad .$$

According to this three-stage test procedure (= Tocher's modification of the exact probability test of Fisher, applied to a 2x4-field table), H_0 is to be rejected if $p < a$ or, if at least T is not smaller than a given coincidental number Z between 0 and 1. On the other hand, H_0 has still not been disproved if $P_{obs} > a$, or if $T < Z$ if $P_{obs} < a$ (and, therefore, $P > a$).

b) Under the condition that, at the most, one $En_{ij} < 5$ and that all $En_{ij} > 1$, with

$$x^2 = \sum_{i=1}^{2} \sum_{j=1}^{4} \frac{(n_{ij} - En_{ij})^2}{En_{ij}}$$

(= X^2-test in the 2x4-field case). Therefore, H_0 is to be rejected if X_0^2 is the table value having the probability a, for 3 degrees of freedom and $X^2 \geqslant X_0^2$. If $X^2 < X_0^2$, then H_0 can be retained.

6) Up until the present, the proposed testing method has fulfilled the first two of the three conditions mentioned at the beginning of this paper: the classification of the values was objectified by introducing the quartiles, in contrast to the usual X^2- or Fisher test and the Kolmogoroff-Smirnoff test (condition A). The computational effort has remained the same as for the X^2- or Fisher test. It has been reduced from that required for the Wald-Wolfowitz test, and it has been made smaller, rather than enlarged, with respect to the Kolmogoroff-Smirnoff test (condition B).

In contrast to this, it is only possible to meet condition C with the test procedure 5b (possibility of interpretation of the opposing hypothesis H_1 for a significant testing result, see "Problem"). An analysis of the X^2-partial values

$$X_j^2 = \sum_{i=1}^{2} \frac{(n_{ij} - En_{ij})^2}{En_{ij}} \qquad \left(\sum_{j=1}^{4} X_j^2 = X^2 \right)$$

of the median-quartile test for independent X^2-values, having one degree of freedom, determines in which quarter, or which quarters, of the rank series $(x^{(r)})$ (which is common to both samples) the significant distribution differences lie. A separate analysis of X^2-partial values of this type , or similar type, is already well known from genostatic problems (see, for instance, R.A. Fisher 1956, pg. 103 ff). However, it is necessary in the present case to keep the fact in mind that the individual X_j^2 are not independent. Instead, it is necessary to deal with one additional degree of freedom, as compared to the analysis of the total X^2, which , so to speak, has the effect of an assurance bonus of the testing result from the X_j^2.

If the observed values x are given the following notation:

a) "normal" within the semi-region of the rank series ($Q_1 < x \leq Q_3$),

b) "abnormally low" below the semi-region ($x \leq Q_1$), and

c) "abnormally high" above the semi-region ($x > Q_3$),

so that then, for instance, the result

$$\begin{array}{ll} X_1^2 & \text{significant} \\ \left.\begin{array}{l} X_2^2 \\ X_3^2 \\ X_4^2 \end{array}\right\} & \text{not significant} \end{array}$$

can be interpreted in the sense that in one of the two fundamental ensembles, governing one of the two samples, there is a more frequent occurrence of abnormally low observational values than there is in the other. It is evident that this result cannot be based on a difference in localization of the two distributions alone, but also on the fact that there must also be differences in the dispersion and/or in the symmetry properties of the distributions.

On the other hand, for instance, the result

$$\begin{array}{ll} \chi_1^2 & \text{significant} \\ \left.\begin{array}{l}\chi_2^2 \\ \chi_3^2\end{array}\right\} & \text{not significant} \\ \chi_4^2 & \text{significant} \end{array}$$

can occur for distributions which differ only in their localization.

Even though such an analysis does not give an unique determination of H_1, at least it presents an idea of the manner in which the determined difference in distribution came about. This indication is often sufficient for practical application, such as is indicated by the following example.

Example

A certain type of steel, from which clubs of the same dimensions and manufactured by the same process are made, is melted in two different steel factories. We want to investigate the question of whether there is a difference,in the cleaning effort necessary to improve the clubs' surface,between the shipments of the two manufacturers. The magnitude of this difference is to be also investigated.

The investigated quantities are the specific club cleaning times (measured in h/t) for 30 melts from steel mill 1, and for 38 melts from steel mill 2. These data are presented in Table 1 - from left to right, in the order of their occurrence.

Table 1

Steel Mill i	Specific Club Cleaning Time per Melt (h/t) x_{ik} $(i=1,2; k=1,2....,N_i)$	No. of Melts N_i
1	2.6 2.7 2.9 2.5 2.7 4.4 4.6 1.6 1.1 1.4	
	2.1 2.4 5.2 1.1 1.4 5.0 2.0 0.9 2.3 2.7	
	1.5 5.8 1.6 1.5 2.7 2.0 3.8 2.0 4.1 2.8	30
2	1.3 2.0 4.7 1.4 1.0 1.0 2.5 1.3 2.4 5.5	
	2.2 1.8 1.5 2.1 1.7 1.7 2.0 2.1 2.1 1.8	
	2.4 1.4 2.7 2.4 0.7 2.9 1.4 1.7 1.3 2.0	
	1.7 2.8 2.1 1.9 2.2 0.9 1.9 2.3	38

Table 2

$x^{(r)}_{1k}$	r_{1k}	$[r_{min}, r_{max}]$		$[r_{min}, r_{max}]$	r_{2k}	$x^{(r)}_{2k}$
0.9	2.5	[2,3]			1	0.7
1.1	6.5	[6,7]		[2,3]	2.5	0.9
1.1	6.5	"		[4,5]	4.5	1.0
1.4	13	[11,15]	(1)	"	4.5	1.0
1.4	13	"		[8,10]	9	1.3
1.5	17	[16,18]		"	9	1.3
1.5	17	"		"	9	1.3
1.6	19.5	[19,20]		[11,15]	13	1.4
1.6	19.5	"		"	13	1.4
2.0	31.5	[29,34]	(2)	"	13	1.4
2.0	31.5	"		[16,18]	17	1.5
2.0	31.5	"		[21,24]	22.5	1.7
2.1	37	[35,39]		"	22.5	1.7
2.3	42.5	[42,43]		"	22.5	1.7
2.4	45.5	[44,47]	(3)	"	22.5	1.7
2.5	48.5	[48,49]		[25,26]	25.5	1.8
2.6	50			"	25.5	1.8
2.7	53	[51,55]		[27,28]	27.5	1.9
2.7	53	"		"	27.5	1.9
2.7	53	"		[29,34]	31.5	2.0
2.7	53	"		"	31.5	2.0
2.8	56.5	[56,57]		"	31.5	2.0
2.9	58.5	[58,59]		[35,39]	37	2.1
3.8	60			"	37	2.1
4.1	61			"	37	2.1
4.4	62			"	37	2.1
4.6	63		(4)	[40,41]	40.5	2.2
5.0	65			"	40.5	2.2
5.2	66			[42,43]	42.5	2.3
5.8	68			[44,47]	45.5	2.4
				"	45.5	2.4
				"	45.5	2.4
				[48,49]	48.5	2.5
				[51,55]	53	2.7
				[56,57]	56.5	2.8
				[58,59]	58.5	2.9
					64	4.7
					67	5.5

(1) $x \leqslant Q_1$

(2) $Q_1 < x \leqslant Q_2$

(3) $Q_2 < x \leqslant Q_3$

(4) $x > Q_3$

Let us first investigate the hypothesis that the two samples of specific club cleaning times have the same "central tendency", i.e., the fundamental ensembles superimposed on them are located at the same position along the numerical scale. (It is true that the arithmetic mean of the specific club cleaning times for the 30 melts from steel mill 1 equals 2.65 h/t. For the 38 melts from steel mill 2, it equals 2.02 h/t). The observational values x_{ik} ($i=1,2$; $k=1,2....N_i$; $N_1=30$; $N_2= 38$) are also brought into a common rank order series ($x_{ik}^{(r)}$) ($k=1,2.....N_i$). Table 2 for $x_{ik}^{(r)}$, having the same absolute value, shows the common (mean) rank number $r_{ik}=\bar{r}$ and the value range of the effective rank numbers (r_{min}, r_{max}).

If we choose the common sample-median of the $x = x_{ik}=x_{ik}^{(r)}$, as:

$$Q_2 = 2.0 \text{ h/t},$$

then the condition $p(x \leqslant Q_2) = ½$ is exactly fulfilled, and the median test can be applied as follows:

$n_{ij}(x)$	$i = 1$	$i = 2$	total
$j = 1,2$	12	22	34
$j = 3,4$	18	16	34
total	30	38	68

The X^2-test can serve as a testing procedure for the 2x2-field case (with the continuity correction of F.Yates). We obtain $X^2 = 1.49$.
This value is not significant, with a presupposed value of $\alpha = 0.05$ for the error probability ($X_0^2 = 3.84$ for 1 degree of freedom).

A procedure for testing the hypothesis of a common central tendency of two, independent samples, which has better resolution than the median test for the size of the present investigational data (N = 68), is, for example, the corresponding modification of the Wilcoxon test. We shall now apply this procedure in order to control the result of the median test, which is easier to compute.

The test number of the Wilcoxon test, the sum of the rank number in the smaller of the two samples (sample 1), is:

$$R = \sum_{k'=1}^{N_1} r_{1k'} = 1204.5 \quad .$$

Its sample expectation value, assuming that the testing hypothesis holds, is:

$$E(R) = \frac{N_1(N+1)}{2} = 1035 \quad ,$$

and the mean quadratic deviation, under the same condition, is:

$$MQA(R) = \sqrt{\frac{N_1 N_2 (N+1)}{12}} = 81 \quad .$$

For the size of the given N, R can be assumed to be normally distributed in the case of the testing hypothesis, and therefore

$$u = \frac{R - E(R)}{MQA(R)}$$

can be assumed to be normally distributed with the expectation value 0 and the scatter 1. Corresponding to the caculated $u = 2.09$, there is a table value $u_0 = 1.96$ across from $\alpha = 0.05$. The hypothesis of a common, central tendency in the two samples can be rejected, with the probability of error of 5%, whereas the non-significant result of the median test should be due to its small resolution. (In the case of the Wilcoxon test, the necessary continuity correction on the right side, for taking into account the observational values having common rank numbers, can be disregarded, because $u = u_{uncorr.}$ is already significant and we always have $u_{corr} > u_{uncorr.}$)

If Table 2 is examined in greater detail, it would be expected that the difference between the two sample distributions is not so much in the central tendency, as it is in the dispersion and/or in the asymmetry of the distributions. In any case, it should be noted that the observational values, having the order of magnitude $2 \cdot Q_2$ and more in sample 1, appear more frequently than in sample 2. Therefore, in the rest of the investigation of our data, we shall no longer restrict the testing hypothesis to a special characteristic of the distribution, but also to non-specific distribution differences, with the aid of the median-quartile test.

For this purpose, we shall choose the quartiles in such a way that each quarter of the common distribution contains

$$\frac{N}{4} = 17$$

observational values, as close as possible. Therefore, we obtain, for the lower quartile, $Q_1 = 1.5$ h/t, and for the upper quartile $Q_3 = 2.6$ h/t. (The deviation from the desired relationship $p(x \leqslant Q_3) = 3/4$ is smaller for

Q_3 = 2.6 h/t with p= 0.74, than for Q_3 = 2.7 h/t with p=0.81).
The median-quartile test must be set up as follows:

$n_{ij}(x)$	i = 1	i = 2	total
j = 1	7	11	18
j = 2	5	11	16
j = 3	5	11	16
j = 4	13	5	18
total	30	38	68

Let us again test with X^2 and we obtain

$$X^2 = 8.28 ,$$

a value which exceeds the table value $X_0^2 = 7.81$ for $\alpha = 0.05$ and three degrees of freedom.

Thus, the hypothesis of the existence of a general distribution identity of the fundamental ensembles, which are superimposed on the two samples, must be rejected. The remaining question is, if, considering the results derived from the Wilcoxon test, there is only one difference in the central tendency or if other differences play a decisive role in the result of the median-quartile test.

In order to answer this question, the X^2-partial values of the median-quartile test

$$X_1^2 = 0.18$$

$$X_2^2 = 1.12$$

$$X_3^2 = 1.12 \text{ and}$$

$$X_4^2 = 5.86$$

are analyzed as independent X^2-values having one degree of freedom each. The result is that only X_4^2 is statistically assured for $\alpha = 0.05$ ($X_0^2 = 3.84$). It follows from this: the differences in distribution, detected by the median-quartile test, are based essentially - at least in part - on the fact that the occurrence probability of observational values above the (common) semi-region, that is, $p(x > Q_3)$, is different in the two ensembles, and therefore the dispersion and/or the symmetry behaviour is different.

If the results of the Wilcoxon and the median-quartile tests are summarized, the concrete question of the first paragraph of this section "Example" can be answered as follows.

The melts from steel mill 1 require a higher club cleaning effort, on the average, than do the melts from steel mill 2. This effect is,in particular, due to the fact that there are many more melts with abnormally high cleaning requirements from steel mill 1 than there are from steel mill 2.

Let us compare the median-quartile test finally with the pure X^2-procedure and the Kolmogoroff-Smirnoff test, even though we know in advance that their testing results cannot be interpreted in the form described above.

Let us choose a four-stage value classification of the investigated variables for the pure X^2 test. For this purpose, the observational values, calculated to 0.1 h/t, are rounded off to an accuracy of 0.5 h/t (so that, for example, the rounded value 2 h/t corresponds to the observational values 1.8 - 2.2 h/t, and the rounded value 2.5 h/t corresponds to the observational values 2.3 - 2.7 h/t). The X^2 test is done as follows:

$n(x_{ik})$	i = 1	i = 2	total
$x_{ik} \leq 1.5$	9	15	24
$x_{ik} = 2$	4	13	17
$x_{ik} = 2.5$	8	6	14
$x_{ik} \geq 3$	9	4	13
total	30	38	68

We then obtain $X^2 = 7.70$, which is not secured for $\alpha = 0.05$ and three degrees of freedom ($X_0^2 = 7.81$). From this, it can be seen that a value classification, which is not pulled out of thin air, of numerically-scaled data can lead to an entirely different result in the case of the X^2 test, than another classification (also "reasonable") such as that presupposed by the median-quartile test. This is another reason for giving preference to the classification procedure of the median-quartile test, which is objective,i.e., fixed, once and for all.

In order to exclude any arbitrary element from the value classification of the Kolmogoroff-Smirnoff test in advance, we shall retain the original, observational values in this test procedure. Table 3 shows the sum, percentile abundancies $F_i(x)$ of the two samples of $F_i(x_{min})$ to $F_i(x_{max})$, and the absolute values of their differences, $d(x) = |F_1(x) - F_2(x)|$.

Table 3

x	$F_1(x)$	$F_2(x)$	$d(x)$	x	$F_1(x)$	$F_2(x)$	$d(x)$
0.7	0.000	0.026	0.026	2.7	0.700	0.895	0.195
0.8	.000	.026	.026	2.8	.733	.921	.188
0.9	.033	.053	.020	2.9	.767	.947	.180
1.0	.033	.105	.072	...	...	...	...
1.1	.100	.105	.005	3.8	.800	.947	.147
1.2	.100	.105	.005	...	...	...	...
1.3	.100	.184	.084	4.1	.833	.947	.114
1.4	.167	.263	.096	...	...	...	...
1.5	.233	.289	.056	4.4	.867	.947	.080
1.6	.300	.289	.011	4.5	.867	.947	.080
1.7	.300	.395	.095	4.6	.900	.947	.047
1.8	.300	.447	.147	4.7	.900	.973	.073
1.9	.300	.500	.200	...	...	...	...
2.0	.400	.579	.179	5.0	.933	.973	.040
2.1	.433	.684	.251	5.1	.933	.973	.040
2.2	.433	.737	.304	5.2	.967	.973	.006
2.3	.467	.763	.296	...	...	...	...
2.4	.500	.842	.342	5.5	.967	1.000	.033
2.5	.533	.868	.335	...	...	...	...
2.6	.567	.868	.301	5.8	1.000	1.000	.000

The calculated value of the tested quantity of the Kolmogoroff-Smirnoff test, for x = 2.4 h/t , is

$$D = \max\,(d(x)) = 0.342,$$

and the comparison value - for $N_1 \neq N_2$ and $N_1, N_2 \geq 30$ for $\alpha = 0.05$ (according to Smirnoff 1948) - is

$$D_0 = 1.36 \sqrt{\frac{N_1 + N_2}{N_1 \cdot N_2}} = 0.244.$$

Thus, the Kolmogoroff-Smirnoff test leads to a rejection of H_0, as was the case for the median-quartile test.

It might still be of interest to examine the question of the influence of the value classification on the testing result of the Kolmogoroff-Smirnoff test. As Table 4 shows, (Table 4 is an excerpt of Table 3, in which the observational values are rounded off to an accuracy of 0.5 h/t) one would have found significance in the present case, if the value classification of the pure X^2 test were used:

Table 4

x	$F_1(x)$	$F_2(x)$	$d(x)$	x	$F_1(x)$	$F_2(x)$	$d(x)$
0.5	0.000	0.026	0.026	3.5	0.767	0.947	0.180
1.0	.100	.105	.005	4.0	.833	.947	.114
1.5	.300	.395	.095	4.5	.900	.973	.073
2.0	.433	.737	.304	5.0	.967	.973	.006
2.5	.700	.895	.195	5.5	.967	1.000	.033
3.0	.767	.947	.180	6.0	1.000	1.000	.000

$$D = 0.304$$

Also it becomes more than the comparison value D_0 for $x = 2.0$ h/t.

REFERENCES

(1) Cochran, W.G. (1952): The X^2 test of goodness of fit. Ann. Math. Statist. 23, 315-345.

(2) Cochran, W.G. (1954): Some methods for strengthening the common X^2 tests. Biometrics 10, 417-451.

(3) Fisher, R.A. (1956): Statistische Methoden fur die Wissenschaft (Statistical Methods for Science. 12. Oliver and Boyd, Edinburgh-London.

(4) Lehmann, E.L. (1953): The power of rank tests. Ann. Math. Statist., 24, 23-43.

(5) Mood, A.M. (1950): Introduction to the theory of statistics. McGraw-Hill, New York.

(6) Roy, S.N. (1957): Some aspects of multivariate analysis. Wiley, New York. Indian Statistical Institute, Calcutta.

(7) Siegel, S. (1956): Nonparametric statistics for the behavioral sciences. McGraw-Hill, New York.

(8) Smirnoff, N.V. (1948): Tables for estimating the goodness of fit of empirical distributions. Ann. Math. Statist. 19, 280-281.