5 ex.
niet stencilen

MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM
STATISTISCHE AFDELING

Leiding: Prof. Dr D. van Dantzig
Chef van de Statistische Consultatie: Prof. Dr J. Hemelrijk

SP 20

Some remarks on the combination of independent tests,

by

J. Hemelrijk.

1953.

MATHEMATISCH CENTRUM
AMSTERDAM

Some remarks on the combination of independent tests.

by J. Hemelrijk.

1. ~~Introduction~~. The use of the integral transformation for the combination of independent tests of ~~ty~~ statistical hypotheses has first been introduced by R.A. Fisher (1932). The present note is an extension of a paper of E.S. Pearson (1938) on this subject and ~~is mainly concer~~ contains some results about the combination of twosided tests, which are not given in that paper ~~In order to give a clear picture~~ and which ~~are of some~~ may be of practical importance. The methods used are the same as Pearson's and ~~by~~ in order to give a clear picture of the situation, some of his results are summarized in the following section.

~~2.~~ We consider n independent tests of n hypotheses $H_1, H_2, \dots, H_n$, which may, but need not, be identical. Let the test of H_i $(i=1,\dots,n)$ be based on ~~the~~ a statistic $\underline{t}_i$ [1]). ~~Under the hypotheses $H_1,\dots,H_n$~~ The statistics $\underline{t}_i$ ~~must have to be~~ are ~~supposed to be~~ stochastically independent and we denote their distribution functions, under H_i, by

$$(1) \qquad F_i(x \mid H_i) \overset{\text{def}}{=} P[\underline{t}_i \leq x \mid H_i]$$

and, on the other hand

$$(2) \qquad R_i(x \mid H_i) \overset{\text{def}}{=} P[\underline{t}_i \geq x \mid H_i].$$

~~Then~~ Defining

$$(3) \qquad \begin{cases} \underline{u}_i \overset{\text{def}}{=} F_i(\underline{t}_i \mid H_i) \\ \underline{v}_i \overset{\text{def}}{=} R_i(\underline{t}_i \mid H_i) \\ \underline{w}_i \overset{\text{def}}{=} 2\,\mathrm{Min}(\underline{u}_i, \underline{v}_i) \end{cases} \qquad (i=1,\dots,n)$$

---

1) Random variables ~~are~~ have been underlined in order to ~~give~~ distinguish them from ordinary variables ~~as~~ and ~~from values assumed by random~~ ordinary numbers (e.g. values assumed by random variables, which ~~as~~ are sometimes denoted by the same letter, ~~as the random variables, but~~ not underlined).

2. Consider m hypotheses $H_1, H_2, \dots, H_m$, which may, but need not, be identical. ~~Suppose, that, in n independent experiments have been performed, to test these order to test these hypotheses, an experiment has been perfor~~ Suppose that these hypotheses are to be tested on the evidence of m independent experiments, one experiment for each ~~of the~~ hypothesis. Let $\underline{t}_i$ $(i=1,\dots,m)$ [1] be the test criterion ~~for~~ used for testing H_i. The experiments being independent, the statistics $\underline{t}_i$ are stochastically independent. The problem we wish to consider is, how to test $H_1, \dots, H_m$ simultaneously on the available data, i.e. how to combine the tests based on $\underline{t}_1, \dots, \underline{t}_m$ into one test of the hypothesis H_0, which is the conjunction of $H_1, \dots, H_m$:

$$(1) \qquad H_0 \equiv H_1 \text{ and } H_2 \text{ and} \dots \text{and } H_m.$$

Pearson derived a solution for this problem by means of the likelihood ratio principle [2] for several sets of admissable hypotheses.

If we denote the (cummulative) ~~probability~~ distribution function of $\underline{t}_i$, under hypothesis H_i, by

$$(2) \qquad F_i(x \mid H_i) \overset{\text{def}}{=} P[\underline{t}_i \leqq x \mid H_i]$$

and if we define

$$(3) \qquad R_i(x \mid H_i) \overset{\text{def}}{=} P[\underline{t}_i \geqq x \mid H_i],$$

then the variables

$$(4) \qquad \left\{ \begin{aligned} \underline{u}_i &\overset{\text{def}}{=} F_i(\underline{t}_i \mid H_i) \\ \underline{v}_i &\overset{\text{def}}{=} R_i(\underline{t}_i \mid H_i) \\ \underline{w}_i &\overset{\text{def}}{=} 2 \operatorname{Min}(\underline{u}_i, \underline{v}_i) \end{aligned} \right. \qquad (i=1,\dots,m)$$

are random variables, which satisfy the following inequalities

$$(5) \qquad \left\{ \begin{aligned} P[\underline{u}_i \leqq y] &\leqq y \\ P[\underline{v}_i \leqq y] &\leqq y \\ P[\underline{w}_i \leqq y] &\leqq y. \end{aligned} \right. \qquad (0 \leqq y \leqq 1;\ i=1,\dots,m).$$

The equality sign holds for every y between 0 and 1 ~~when~~ if the corresponding $\underline{t}_i$ is continuously distributed;

---

1) Random variables ~~have been underlined~~ are denoted by underlined symbols, to distinguish them from ordinary variables or numbers. Values ~~of the~~ assumed by random variables are sometimes denoted by the same symbol, not underlined.

2) Cf. J. Neyman and E. S. Pearson (1933).

(4) denotes the above mentioned integral transformation in three slightly different forms and $\underline{u}_i$, e.g., is the random probability, that a second observation of $\underline{t}_i$, taken independently of a first observation of this statistic, will not exceed this first one.

We shall mainly consider the case, ~~with~~ that $\underline{t}_1, \dots, \underline{t}_n$ are continuously distributed (some remarks about the discontinuous case are given in section 6); ~~then~~ $\underline{u}_i$, $\underline{v}_i$ and $\underline{w}_i$ will, under H_0, have a ~~to~~ rectangular distribution.

Pearson considered i.a. the following sets of Ω_1, Ω_2 and Ω_3 of admissable hypotheses.

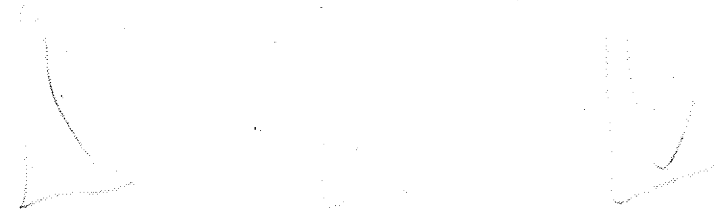
$$(6)\quad \Omega_1: \quad f_i(u \mid H') = (a+1)\, u^a \quad (-1 < a \leq 0;\ i = 1, \dots, n),$$

$$(7)\quad \Omega_2: \quad g_i(v \mid H'') = (b+1)\, v^b \quad (-1 < b \leq 0;\ i = 1, \dots, n),$$

$$(8)\quad \Omega_3: \quad h_i(u \mid H''') = \frac{\Gamma(2c+2)}{\{\Gamma(c+1)\}^2}\, u^c (1-u)^c \quad (-1 < c \leq 0;\ i = 1, \dots, n).$$

Here f_i ~~..~~ and h_i denote the probability density of $\underline{u}_i$, ~~under~~ g_i the probability density of $\underline{v}_i$; a, b and c respectively characterize simple admissable hypotheses H', H'' and H''' respectively, which coincide with H_0 if $a=0$, $b=0$ and $c=0$ respectively. The statistics $\underline{u}_1, \dots, \underline{u}_n$ are again independently distributed, corresponding to the independence of the n original tests of $H_1, \dots, H_n$. The same applies to $\underline{v}_1, \dots, \underline{v}_n$.

The type of the hypotheses of these three classes of admissable hypotheses has been sketched in figure 1.



a. b. c.

Fig. 1. Classes Ω_1, Ω_2 and Ω_3 of admissable hypotheses.

Using ~~Using~~ the likelihood ratio principle, Pearson found, that the statistics

(9) $\quad \underline{U} \overset{\text{def}}{=} -2 \ln \prod_{i=1}^{m} \underline{u}_i ,$

(10) $\quad \underline{V} \overset{\text{def}}{=} -2 \ln \prod_{i=1}^{m} \underline{v}_i$

and

(11) $\quad \underline{W} \overset{\text{def}}{=} -2 \ln \prod_{i=1}^{m} \underline{w}_i ,$

where $\underline{w}_i$ is defined by the last equation (4), may be used as test criteria for H_0 with Ω_1, Ω_2 and Ω_3 respectively as the set of admissable hypotheses, large values of $\underline{U}, \underline{V}$ and $\underline{W}$ respectively being taken as critical values. It is not difficult to prove, that under H_0 each of these statistics has a χ^2-distribution with $2m$ degrees of freedom (not approximately ~~this time~~ in this case, but exactly). ~~Large values~~ These tests, being λ-tests, may be expected to ~~b~~ have a satisfactory power function.

3. The classes Ω_1 and Ω_2 may be considered as a satisfactory model of situations, often encountered in the practice of statistical analysis, where onesided tests are to be combined. Ω_3 was meant to represent the twosided case. For most twosided tests, however, — in fact for all twosided tests, which the present author can call to mind — the alternative hypotheses would never take the form ~~G~~ of ~~H~~ a hypothesis H^m of Ω_3. ~~Take e.g. the~~ Let e.g. t_i be Students $\underline{t}$ ~~test~~ for testing the hypothesis, that the mean of a normally distributed variable is equal to zero. If this is not the case, the mean will either be negative — and then the probability distribution of ~~$\underline{u}_i$~~ $\underline{u}_i$ will ~~resemble have a~~ resemble fig. 1,a — or it will be positive — and then the distribution of $\underline{u}_i$ will be something like fig. 1,b — but in no case will the distribution of $\underline{u}_i$ have the form of figure 1,c.

The relevant class of admissable hypotheses for the ~~two side~~ combination of two sided tests seems to be

(12) Ω_4: for every $i = 1, \ldots, m$ separately either
$f_i(u \mid H^*) = (a+1)\, u^a$, or $g_i(v \mid H^*) = (a+1)\, v^a$;
with $-1 < a \leq 0$,

~~a again characterizing the simple hypo admissable hypothesis H^* of Ω_4, which coincides with H_0~~

f_i and g_i again denoting the probability densities of $\underline{u}_i$ and $\underline{v}_i$ under the simple hypothesis H^* of Ω_4, H^* being characterized by the value of a and coinciding with H_0 if $a=0$; $\underline{u}_1, \dots, \underline{u}_m$ (and also $\underline{v}_1, \dots, \underline{v}_m$) are again stochastically independent.

It is to be expected, that ~~the difference between the classes Ω_3 and Ω_4 will not influence the result of~~ the application of the likelihood ratio principle with Ω_4 as class of admissable hypotheses will yield the same result as found for Ω_3. This is indeed the case, as will be proved in section 4.

There is, however, another class of admissable hypotheses, which is worth considering and for which the likelihood ratio test does not coincide with ~~the one case~~ one of the above mentioned tests. When using Ω_4, we admit the possibility, that some of the probability distributions of $\underline{u}_1, \dots, \underline{u}_m$ will ~~have a distribution~~ be of the type of fig. 1,a and some others of the type of fig. 1,b. It may be, and as a matter of fact it is far from rare, that either all of these ~~its~~ probability distributions are of the first or all of them are of the second type mentioned. This will e.g. be true, if the m original hypotheses to be tested, $H_1, \dots, H_m$, are identical, with identical ~~or~~ or nearly identical alternatives hypotheses. We may, as an example, think of $\underline{t}_i$ as Student's $\underline{t}$, ~~for every i with~~ for testing the mean of normal distributions, which have the same mean for $i=1, \dots, m$ but ~~possibly~~ different standard deviations. In that case, the hypothesis, that this mean is equal to 0, cannot be tested by pooling the m sets of data, because of the different standard deviations, ~~and thus a method of combining the~~ ~~and as the mean is the same for every i the class Ω_4 contains hypotheses, which should not~~ and thus another method of combining the data is necessary. ~~The class Ω_4, for is~~ The ~~means~~ (unknown) mean being the same for every i, the class Ω_4 contains hypotheses which should be excluded and we must look for another class of admissable hypotheses, which is relevant for two sided tests of this type. ~~This class seems~~ We may then consider the class

~~(13) Ω_5: for every $i=1, \dots, m$ we have either~~

(13) Ω_5: either $f_i(u \mid H^{**}) = (a+1)u^a$ ~~(*) (...)~~
for every $i=1, \dots, m$, or $g_i(v \mid H^{**}) = (a+1)v^a$
for every $i=1, \dots, m$, with $-1 < a \leq 0$,

with f_i, g_i and a ~~as the~~ having the same meaning as before, ~~again~~ $\underline{u}_1, \dots, \underline{u}_m$ (and also $\underline{v}_1, \dots, \underline{v}_m$) being stochastically independent. ~~We may~~ Ω_5 may also be defined as the union of Ω_1 and Ω_2.

We shall prove, in section 4, that within this case the likelihood ratio principle yields a test based on the statistic

$$(14) \qquad \underline{X} \overset{\text{def}}{=} \operatorname{Max}(\underline{U}, \underline{V}),$$

with $\underline{U}$ and $\underline{V}$ defined by (9) and (10), large values of $\underline{X}$ being critical. This criterion has been proposed by K. Pearson (1934), who, however, did not take into account, that $\underline{X}$ does not have a χ^2-distribution with $2m$ degrees of freedom. Consequently he underrated the probability of an error of the first kind [illegible] value ($\frac{1}{2}$). We shall show, that under H_0 we have

$$(15) \qquad P[\underline{X} \geq d(\alpha) \mid H_0] < 2\alpha,$$

if $d(\alpha)$ is the $(1-\alpha)$-quantile of a χ^2-distribution with $2m$ degrees of freedom. ~~By~~ Using (15) $\underline{X}$ may then be used to test H_0 against Ω_5 and, being a likelihood ratio test, this test may be expected to give better results than $\underline{U}$, $\underline{V}$ or $\underline{W}$ for classes of admissible hypotheses resembling Ω_5.[3])

4. Following the proofs of E. S. Pearson (1938) we consider the function

~~(16)~~ ~~[illegible] $\varphi(a, d)$~~

$$(16) \qquad \varphi(a, p) = (a+1)^m p^a \qquad (-1 < a \leq 0;\ 0 \leq p \leq 1),$$

where m is a given positive integer. It is easy to prove, that for given p this function has a maximum

$$(17) \qquad \psi(p) \overset{\text{def}}{=} \operatorname*{Max}_{-1 < a \leq 0} \varphi(a, p) = \begin{cases} 1 & \text{if } p > e^{-m} \\ \frac{1}{p}\left(\frac{m e^{-1}}{-\ln p}\right)^m & \text{if } p \leq e^{-m}, \end{cases}$$

and that

$$(18) \qquad \psi(p_1) \leq \psi(p_2) \quad \text{if} \quad p_1 \geq p_2.$$

Denoting the simultaneous probability density of $\underline{u}_1, \dots, \underline{u}_m$ under hypothesis H^* by $f(u_1, \dots, u_m \mid H^*)$, we have, when H^* is a hypothesis of Ω_4:

$$(19) \qquad f(u_1, \dots, u_m \mid H^*) = (a+1)^m \left\{ \prod_{i=1}^{m} u_i \right\}^a \qquad (-1 < a \leq 0),$$

---

3) The test in this form was proposed to the author by Mr. J. W. E. Vos in a personal communication, which formed the inducement to this study.

where z_i, for every i separately, is either equal to u_i or to v_i.
The likelihood ratio is

$$(20) \qquad \lambda = \frac{f(u_1,\dots,u_m \mid H_0)}{\underset{H^*\in\Omega_4}{\mathrm{Max}}\, f(u_1,\dots,u_m \mid H^*)}.$$

The numerator of this ratio is equal to 1 (cf. (5)) and the denominator ~~may~~ must be computed by choosing a suitable value of a in (19) and by taking z_i equal to u_i or v_i in such a way, that the ~~right hand~~ second member of (19) is maximized. According to (17) and (18) this is accomplished ~~b~~ by minimizing

$$\prod_{i=1}^{m} z_i$$

by means of the choice of z_i. ~~According to definition (4), this minimum is equal to~~ ~~$2^{-m}\prod$~~

$$2^{-m}\prod_{i=1}^{m} w_i = 2^{-m}\, W,$$

~~and (cf (4)) this is equal~~ ~~as~~ with W defined by (4).
We thus have

$$(21) \qquad \lambda = 2^{m}\cdot \cancel{W}\,\{\max_{-1<a\le 0} (a+1)^m W^{a}\}$$

and ∟ according to (18) λ does not increase when W does. ~~As small values of λ are critical, the state we must take large~~ the critical region of a ~~test~~ likelihood ratio test is always of the form

$$(22) \qquad \lambda \le \lambda_0,$$

~~we~~ we must take large values of $\underline{W}$ as critical values. ~~to~~ This completes the proof for the test with Ω_4 as class of admissible hypotheses.

Now consider Ω_5. We then have

(23) $f(u_1,\dots,u_m \mid H^{**})$ is either equal to $(a+1)^m\{\prod_{i=1}^{m} u_i\}^a$ or to $(a+1)^m\{\prod_{i=1}^{m} v_i\}^a$.

The numerator of λ is again equal to 1 and we must now maximize $f(u_1,\dots,u_m \mid H^{**})$ by choosing H^{**} suitably from ~~Ω_1 or Ω_2~~ Ω_5, i.e. from Ω_1 or Ω_2. It follows from (17) and (18), that we should take the smallest of the two products $\prod_{i=1}^{m} u_i$ and $\prod_{i=1}^{m} v_i$. ~~According to (4) this means, that~~

~~X provides a likelihood ratio test, again with large values as critical values, for~~

According to definition (4) this minimum is equal to

$$2^{-m} \prod_{i=1}^{m} w_i = 2^{-m} e^{-\frac{1}{2}W}$$

as may be seen from (11). We thus have

$$(21) \quad \lambda = 2^{m} \left\{ \underset{-1<a\leq 0}{\mathrm{Max}} (a+1)^{m} e^{-\frac{1}{2}aW} \right\}$$

and, according to (18) this function is a non-increasing function of W. As the critical region of a likelihood ratio test is always of the form $\lambda \leq \lambda_0$ we must take large values of W as critical values. This completes the proof for the test with Ω_4 as class of admissable hypotheses.

Now consider Ω_5. The numerator of λ is again equal to 1 and we must now maximize $f(u_1, \dots, u_m | H^{**})$ by choosing H^{**} suitably from Ω_5. This probability density is ~~equal~~ either equal to

$$(a+1)^{m} \left\{ \prod_{i=1}^{m} u_i \right\}^{a}$$

or to

$$(a+1)^{m} \left\{ \prod_{i=1}^{m} v_i \right\}^{a}$$

and ~~it follo~~ again consulting (17) and (18) we see, that we must take the smallest of the two products $\prod u_i$ and $\prod v_i$ to reach the maximum. ~~Now~~ The definitions (9) and (10) may also be written as

$$\prod_{i=1}^{m} u_i = e^{-\frac{1}{2}U} \quad ; \quad \prod_{i=1}^{m} v_i = e^{-\frac{1}{2}V},$$

and this gives

$$\mathrm{Min} \left\{ \prod u_i, \prod v_i \right\} = \mathrm{Min} \left\{ e^{-\frac{1}{2}U}, e^{-\frac{1}{2}V} \right\} = e^{-\frac{1}{2}X}$$

according to (14). Therefore we find

$$(22) \quad \lambda = \underset{-1<a\leq 0}{\mathrm{Max}} \left\{ (a+1)^{m} e^{-\frac{1}{2}aX} \right\},$$

~~and~~ which means, that X, with large values critical, yields a likelihood ratio test for H_0 against Ω_5.

4. We shall now prove that $\underline{X}$ (cf. (12)) provides a likelihood ratio test for H_0, when Ω_5 is the class of admissable hypotheses.

In this case, we must maximize $f(u_1,\dots,u_m|H)$ by choosing a suitable H from Ω_1 or Ω_2. ~~Restricting ourselves to Ω_1~~ This means, that $f(u_1,\dots,u_m|H)$ is either equal to

$$(a+1)^m\Big\{\prod_{i=1}^m u_i\Big\}^a$$

or to

$$(a+1)^m\Big\{\prod_{i=1}^m v_i\Big\}^a.$$

It follows from ~~According to~~ (16) that the maximum will be reached by choosing an alternative H from Ω_1 or Ω_2 ~~respectively~~, according as $\prod u_i < \prod v_i$ or $\prod u_i > \prod v_i$ respectively. This means, however, that $\underline{X}$, ~~with~~ defined by (12), provides a likelihood ratio test with large values as critical values.

~~The proof of (16) It now remains to~~
~~We now have to prove (16), to make the test complete.~~
~~This is very easy~~ The probability distribution of $\underline{X}$, under H_0, satisfies (15); this may be seen as follows:

$$P[\underline{X} \geqq d(\alpha)|H_0] = P[\underline{U} \geqq d(\alpha) \text{ or } \underline{V} \geqq d(\alpha)|H_0] \leqq$$

$$\leqq P[\underline{U} \geqq d(\alpha)|H_0] + P[\underline{V} \geqq d(\alpha)|H_0] = 2\alpha,$$

$\underline{U}$ and $\underline{V}$ having a χ^2 distribution with $2m$ degrees of freedom, if H_0 is true. As a matter of fact the probability, that $\underline{U} \geqq d(\alpha)$ and $\underline{V} \geqq d(\alpha)$ is positive and we may replace $\leqq$ by $<$. The difference between the true level of significance and 2α is very small, when α is not large. This is illustrated by fig. 2, where the critical region has been sketched for the case $m=2$.

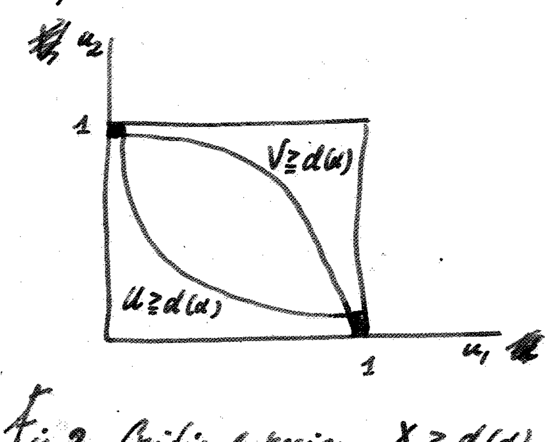


Fig. 2. Critical region $X \geqq d(\alpha)$.
$U = -2\ln u_1u_2$; $V = -2\ln v_1v_2$.

5. It is of interest to compare both tests for the two-sided case, based on $\underline{W}$ and $\underline{X}$ respectively. ~~For some numerical cas~~ ~~This has been done for some~~ ~~numerical cas~~

Ω_5 being a subclass of Ω_4, this comparison is only of interest for hypotheses belonging to Ω_5. For alternative hypotheses, which do not belong to Ω_5, $\underline{X}$ should not be used. Restricting ourselves to Ω_5, we are especially interested in cases, where all values $u_1, \dots, u_m$ found in an experiment are smaller than $\frac{1}{2}$ (or all larger; this comes to the same thing). Then (4) yields

$$w_i = 2u_i$$

and (11) gives

$$W = -2 \ln \prod_{i=1}^{m} w_i = -2 \ln \prod_{i=1}^{m} u_i - 2m \ln 2 = X - 2m \ln 2$$

because of (14). This relation

(23) $$W = X - 2m \ln 2$$

is not generally valid, of course; in fact it has been derived under the supposition, that $u_1, \dots, u_m$ are all smaller than $\frac{1}{2}$. Thus it applies only to ~~those cases~~ values of $u_1, \dots, u_m$ satisfying that condition and the random variables $\underline{W}$ and $\underline{X}$ are not connected by a relation like (23).

With this restriction ~~Restricting ourselves to cases where the values of $u_1, \dots, u_m$, found in the experiment, are all smaller than $\frac{1}{2}$,~~ we may compare the ~~two~~ two sided tail probabilities

(24) $$p_X \overset{\text{def}}{=} 2P[\underline{X} \geq x \mid H_0]$$

and

(25) $$p_W \overset{\text{def}}{=} P[\underline{W} \geq x - 2m \ln 2 \mid H_0],$$

where $x = -2 \ln \prod_{i=1}^{m} u_i$ denotes the value of $\underline{X}$ found in the experiment; deriving p_W and p_X from a table of the χ^2-distribution with $2m$ degrees of freedom.

This gives e.g. the following numerical results.

m	$x - 2m \ln 2$	p_W	p_X
2	6	0.2	0.13
2	9	0.06	0.04
2	12	0.016	0.01
5	16	0.10	0.024
5	20	0.03	0.006
10	30	0.07	0.003

It is clear, that in such cases X yields a far better result than W. It must be remarked, however, that this has not been proved for experiments, where ~~not~~ the values $u_1, \dots, u_m$ do not all lie at the same side of $\frac{1}{2}$.

6. The case of discontinuously distributed statistics t_i (cf. section 1) has been treated by W.A. Wallis (1942), who proved, that [then] the tests given by Pearson [7, when applied in the original form,] result in an overestimation of the size of the critical region. This overestimation may be considerable, ~~whi~~ which means, that the power of the tests also is much ~~smaller~~ diminished ~~for~~ in the discontinuous ~~t_i than for con~~ case. This result of Wallis applies to the test based on X too. Methods to avoid this ~~disad~~ disadvantage have been developed by H.O. Lancaster (1949) and ~~by~~ E.S. Pearson (1950).

References.

Fisher, R.A. (1932). Statistical Methods for Research Workers, 4th ed.

Lancaster, H.O. (1949). The combination of probabilities arising from data in discrete distributions. Biometrika, 36, 370.

Neyman, J. and E.S. Pearson (1933). On the problem of the most efficient tests of statistical hypotheses. Phil. Trans., A, 231, 2

Pearson, E.S. (1938). The probability integral transformation for testing goodness of fit and combining independent tests of significance. Biometrika, 30, 134.

Pearson, E.S. (1950). On questions raised by the combination of tests based on discontinuous distributions. Biometrika, 37, 383.

Pearson, K. (1934). On a new method of determining goodness of fit. Biometrika, 26, 425.

Wallis, W.A. (1942). Compounding probabilities from independent significance tests. Econometrica, 10, 229.