STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

S 252

Some remarks on K; MATHER's test for heterogeneity.

by Dr. Constance van Eeden

and

Ir. A.R. Bloemena.

April 1959

MATHEMATISCH CENTRUM
Statistische Afdeling

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

Report S 252

Some remarks on K. MATHER's test for heterogeneity

by

Dr Constance van Eeden

and

Ir A.R. Bloemena

April 1959

1. Introduction

In his book "The measurement of linkage in heredity" K. MATHER (1951) describes the following χ^2-test for heterogeneity (Chapter II, p. 17-25).

Consider k independent series of independent trials, each trial resulting in a success or a failure. For $i = 1,\dots,k$, the i^{th} series consists of n_i trials with a_i successes and b_i failures ($a_i + b_i = n_i$). Further, for $i = 1,\dots,k$, p_i denotes the probability of a success for each trial of the i^{th} series and $q_i = 1 - p_i$.

These observations may be summarized as follows

series nr	number of successes	number of failures	total
1	a_1	b_1	n_1
.	.	.	.
.	.	.	.
.	.	.	.
k	a_k	b_k	n_k
total	a	b	n

where $a = \sum_{i=1}^{k} a_i$, $b = \sum_{i=1}^{k} b_i$, $n = a+b = \sum_{i=1}^{k} n_i$.

On account of these observations MATHER gives a test for the hypothesis

(1;1) $$H_0:\ p_1 = \dots = p_k.$$

2. Description of the test

For each $i = 1,\dots,k$, the hypothesis

(2;1) $$H_{0,i}:\ p_i = p_0,$$

where p_0 is a given number[1], may be tested by means of a χ^2-test

(2;2) $$\chi_i^2 = \frac{(a_i-n_ip_0)^2}{n_ip_0} + \frac{(b_i-n_iq_0)^2}{n_iq_0} = \frac{(a_i-n_ip_0)^2}{n_ip_0q_0} \quad (q_0 = 1-p_0).$$

----

1) In his example MATHER takes $p_0 = \frac{1}{2}$.

If $H_{0,i}$ is true χ_i^2 possesses, for large values of n_i, approximately a χ^2-distribution with one degree of freedom. Thus if the hypothesis

$$H_0' : p_1 = \ldots = p_k = p_0 \tag{2;3}$$

is true, the statistic

$$\sum_{i=1}^{k} \chi_i^2 = \sum_{i=1}^{k} \frac{(a_i - n_i p_0)^2}{n_i p_0 q_0} \tag{2;4}$$

possesses, for large values of $n_1, \ldots, n_k$, approximately a χ^2-distribution with k degrees of freedom.

According to MATHER ((1951),p.18)"this total χ^2 may be considered as comprising two parts

a) a portion concerned with the grand deviation of all groups taken together from the expected $p_0 : q_0$,
b) a portion concerned with the disagreement among the groups when allowance has been made for the grand deviation."

The former portion may be calculated from the total number of successes and failures in all series taken together:

$$\chi^2 = \frac{(a-n\,p_0)^2}{n\,p_0} + \frac{(b-n\,q_0)^2}{n\,q_0} = \frac{(a-n\,p_0)^2}{n\,p_0 q_0} \ . \tag{2;5}$$

If H_0' is true χ^2 possesses, for large values of n, approximately a χ^2-distribution with one degree of freedom.

Then MATHER remarks (p.18): "The difference of the total χ^2 for k degrees of freedom and this χ^2 for one degree of freedom will be the second or heterogeneity χ^2, testing the agreement between the k groups. This heterogeneity χ^2 must have k-1 degrees of freedom."

Now we have

$$\sum_{i=1}^{k} \chi_i^2 - \chi^2 = \sum_{i=1}^{k} \frac{(a_i - n_i p_0)^2}{n_i p_0 q_0} - \frac{(a-n\,p_0)^2}{n\,p_0 q_0} = \frac{1}{p_0 q_0}\left[\sum_{i=1}^{k} \frac{a_i^2}{n_i} - \frac{a^2}{n}\right] =$$

$$= \frac{1}{p_0 q_0}\left[\sum_{i=1}^{k} \frac{a_i^2}{n_i} - \frac{a}{n} \sum_{i=1}^{k} a_i\right] = \frac{1}{p_0 q_0} \sum_{i=1}^{k} \frac{a_i}{n\,n_i}(a_i n - a\,n_i) \ .$$

Further

$$\sum_{i=1}^{k} (a_i n - a\, n_i) = n \sum_{i=1}^{k} a_i - a \sum_{i=1}^{k} n_i = 0,$$

consequently

$$(2;6) \quad \sum_{i=1}^{k} \chi_i^2 - \chi^2 = \frac{1}{p_0 q_0}\left[\sum_{i=1}^{k} \frac{a_i}{n\, n_i}(a_i n - a\, n_i) - \frac{a}{n^2}\sum_{i=1}^{k}(a_i n - a\, n_i)\right] =$$

$$= \frac{1}{n^2 p_0 q_0} \sum_{i=1}^{k} \frac{(a_i n - a\, n_i)^2}{n_i} \quad .$$

MATHER further remarks (p.19), that this test for H_0 is not absolutely accurate if the χ^2 with one degree of freedom, for testing the grand deviation for all groups taken together, is significant. If this χ^2 for the grand deviation is significant, MATHER uses the ordinary χ^2-test for a 2 x k - table, i.e. he uses the test-statistic

$$\sum_{i=1}^{k} \left\{ \frac{\left(a_i - \frac{a\, n_i}{n}\right)^2}{\frac{a\, n_i}{n}} + \frac{\left(b_i - \frac{b\, n_i}{n}\right)^2}{\frac{b\, n_i}{n}} \right\} = \frac{1}{ab} \sum_{i=1}^{k} \frac{(a_i n - a\, n_i)^2}{n_i} \quad ,$$

possessing, under the hypothesis H_0, for large values of n approximately a χ^2-distribution with k-1 degrees of freedom. Consequently MATHER uses, for testing the hypothesis H_0,

$$\chi'^2 = \frac{1}{n^2 p_0 q_0} \sum_{i=1}^{k} \frac{(a_i n - a\, n_i)^2}{n_i} \quad \text{if } \chi^2 \text{ is not significant}$$

$$\chi''^2 = \frac{1}{ab} \sum_{i=1}^{k} \frac{(a_i n - a\, n_i)^2}{n_i} \quad \text{if } \chi^2 \text{ is significant .}$$

He further remarks that in actual practice a considerably greater value of χ^2 for the grand deviation is necessary for seriously invalidating the heterogeneity χ^2 (p.19) and that in cases where χ^2 for the grand deviation is significant, the heterogeneity χ^2 is somewhat too low (p.38).

In the next section MATHER's test based on χ'^2 and χ''^2 will be compared with the ordinary χ^2-test for a 2xk - table, i.e. with the test, based on χ''^2 alone.

3. Comparison of MATHER's test with the χ^2-test for a 2xk - table.

In the first place we remark that MATHER's test depends on p_0. However p_0 does not occur in the problem: testing the hypothesis H_0 that all p_i are equal. Consequently MATHER's test is not the right test, for testing H_0. This may also be seen as follows. Suppose H_0 is true and let $p_1 = \dots = p_k = p$ (say). Then

$$(3;1) \qquad \mathcal{E}\,\underline{\chi}'^2 = \frac{1}{p_0q_0}\left[\sum_{i=1}^{k} \mathcal{E}\,\frac{\underline{a}_i^2}{n_i} - \mathcal{E}\,\frac{\underline{a}^2}{n}\right] = \frac{pq}{p_0q_0}(k-1) \; .$$

Now for a χ^2-distribution with ν degrees of freedom, we have $\mathcal{E}\,\underline{\chi}^2 = \nu$. Thus MATHER approximates the distribution of $\underline{\chi}'^2$ by means of a χ^2-distribution with mean k-1, whereas the mean of $\underline{\chi}'^2$ is $\frac{pq}{p_0q_0}$ (k-1).

Another question is if the difference between MATHER's test and the test based on χ''^2 alone is of any practical importance. Asymptotically MATHER's test has level of significance α, if for both χ'^2 and χ''^2 a level α is used. This may be seen as follows. If H_0 is true then we have

1. $p = p_0$. Then χ'^2 and χ''^2 both possess asymptotically a χ^2-distribution with k-1 degrees of freedom. Thus in this case the level of significance is asymptotically equal to α.

2. $p \neq p_0$. Then the probability that $\underline{\chi}^2$ is significant can be made arbitrarily close to 1 by choosing n sufficiently large. Thus for sufficiently large n we almost surely use χ''^2, thus a level α . This means that the difference between MATHER's test and the test based on χ''^2 alone is not important for large values of n.
However the following points should be noted

1. MATHER states that, if χ^2 is significant, χ'^2 will be somewhat too low. Thus he states that

$$\frac{1}{n^2p_0q_0} \leq \frac{1}{ab} \quad \text{or} \quad \frac{a}{n}\,\frac{b}{n} \leq p_0q_0 \; .$$

This is only true for all values of a and b if $p_0 = \frac{1}{2}$. If $p_0 \neq \frac{1}{2}$

χ'^2 may be too large, which means that H_0 is rejected too often.

2. MATHER's remark that χ'^2 is not accurate if χ^2 is significant is not clearly stated. He himself uses χ'^2 in cases where χ^2 is significant, but in those cases he adds a few words of caution. Still this may lead to applications of χ'^2 when χ''^2 should be used (cf. e.g. J.H. van der Veen (1957), p. 34), and then the level of significance will be $\neq \alpha$ if $p \neq p_0$. It will be larger than α if (cf.(3;1)) $pq > p_0 q_0$ and smaller than α if $pq < p_0 q_0$.

3. For small values of n the difference between MATHER's test and the test based on χ''^2 alone may be of practical importance. In such cases the approximations of the distributions by means of χ^2-distributions are perhaps not accurate, but then an exact test is possible. Let e.g.

series nr	number of		total
	successes	failures	
1	3	12	15
2	8	7	15
total	11	19	30

and $p_0 = 0,2$.

Then $\chi^2 = 5,2$; the exact two-sided tailprobability may be found by means of a table of the binomial probability distribution for $p_0 = 0,2$; it is found to be 0,05. For $\alpha = 0,05$ this is barely significant, thus we use MATHER's test.

We have

$$\chi'^2 = \frac{1}{900 \times 0,2 \times 0,8} \; \frac{75^2 + 75^2}{15} = 5,2.$$

In this case the exact tailprobability is found by computing (by means of a table of the binomial distribution for $p_0 = 0,2$) the exact distribution of

$$\underline{\chi}' = \frac{\underline{a}_1 n_2 - \underline{a}_2 n_1}{\sqrt{n\, n_1 n_2\, p_0 q_0}}$$

This tailprobability is found to be 0,038; consequently by means of MATHER's test we reject the hypothesis that all p_i are equal. But, as stated above, this is not a correct test, p_0 occurring in the expressions for the test-statistic. The correct test is based on

$$\chi''^2 = \frac{1}{11 \times 19} \times \frac{75^2 + 75^2}{15} = 3,59 .$$

The exact tailprobability is found by computing the exact distribution of

$$\chi'' = \frac{\underline{a}_1 n - a\ n_1}{\sqrt{\frac{ab\ n_1 n_2}{n}}} .$$

Under the condition $\underline{a} = a$, $\underline{a}_1$ possesses a hypergeometric distribution. In this case the tailprobability is 0,12. Consequently the hypothesis H_0 cannot be rejected.
Thus in this case, with a small value of n, the difference between MATHER's test and the test based on χ''^2 is of practical importance.

Remark

The same remarks can be made about a test for the detection of linkage described by MATHER in Chapter IV of his book.

References

Mather, K, The measurement of linkage in heredity, Methuen's monographs on biological subjects, Methuen and Co. Ltd., London, 1951.

Van der Veen, J.H., Studies on the inheritance of leaf shape in nicotiana tabacum. l, thesis Wageningen (Holland) 1957.