NOTES

This section is devoted to brief research and expository articles and other short items.

TESTS OF INDEPENDENCE IN CONTINGENCY TABLES AS UNCONDITIONAL TESTS

By A. M. Mood

Iowa State College[1]

Summary and introduction. Since the ordinary tests for independence in contingency tables use test criteria whose distributions depend on unknown parameters, the justification for the tests is usually made either by an appeal to asymptotic theory or by interpreting the tests as conditional tests. The latter approach employs the conditional distribution of the cell frequencies given the marginal totals, and was first described by Fisher [1]. The purpose of the present note is to show how these tests may be regarded as unconditional tests even though the parameters are unknown by augmenting the test criterion to include estimates of the unknown parameters. We present no new tests, merely a new setting for the old tests which seems to put them in a little better light.

1. Certain conditional tests. A variate or set of variates x has a probability density function $f(x; \theta)$ under a null hypothesis involving a parameter or set of parameters θ. When the parameters have a set of sufficient estimators $\hat{\theta}$, the joint density function of a random sample of size n may be put in the form

$$\prod_{i=1}^{n} f(x_i ; \theta) = g(x_1, x_2, \cdots, x_n \mid \hat{\theta}) h(\hat{\theta}; \theta). \tag{1}$$

It is assumed that n exceeds the number of parameters. We shall be concerned with the class of test criteria which are not functions of the estimators alone. Let $\lambda(x_1, x_2, \cdots, x_n)$ be a test criterion which may not be put in the form $\lambda(\hat{\theta})$. The joint density function for λ and $\hat{\theta}$, obtained by summing (1) for fixed λ and $\hat{\theta}$, will be of the form

$$k(\lambda \mid \hat{\theta}) h(\hat{\theta}; \theta). \tag{2}$$

The marginal distribution of λ will be denoted by $m(\lambda; \theta)$, the result of summing (2) over $\hat{\theta}$ for fixed λ.

In order to test the hypothesis in question one would like to divide the λ space into two regions, an acceptance region S_a and a critical region S_c in such a way that S_c would have a prescribed size α under the null hypothesis. One would of course set up other specifications to be fulfilled by S_c, but we are

[1] The author is now with *The RAND Corporation*, Santa Monica, California.

interested here only in the fact that the size of S_c cannot be determined because of the presence of the unknown parameters θ in $m(\lambda; \theta)$.

One can set up a conditional test by using the conditional distribution $k(\lambda \mid \hat{\theta})$. That is, for fixed $\hat{\theta}$, the measure of any region $R(\hat{\theta})$ (which is measurable relative to $k(\lambda \mid \hat{\theta})$, say, in the Lebesgue-Stieltjes sense) of the λ space is known because the $\hat{\theta}$ are known in any given instance. Thus a conditional test can be made with a critical region $R_c(\hat{\theta})$ of prescribed size.

The conditional test may be interpreted as an unconditional test in the present instance in the following manner: the unconditional test is made by using the double criterion $(\lambda, \hat{\theta})$. The $(\lambda, \hat{\theta})$ space is divided into two regions, T_a for acceptance and T_c for rejection. The critical region T_c consists of all points $(\lambda, \hat{\theta})$ such that λ is contained in $R_c(\hat{\theta})$. If the size of $R_c(\hat{\theta})$ is α for all $\hat{\theta}$, then the size of T_c is also α, for

$$\begin{aligned} \int_{T_c}\int k(\lambda \mid \hat{\theta}) h(\hat{\theta}; \theta)\, d\lambda\, d\hat{\theta} &= \int_{-\infty}^{\infty}\left[\int_{R_c(\hat{\theta})} k(\lambda \mid \hat{\theta})\, d\lambda\right] h(\hat{\theta}; \theta)\, d\hat{\theta} \\ &= \int_{-\infty}^{\infty} \alpha h(\hat{\theta}; \theta)\, d\hat{\theta} \\ &= \alpha. \end{aligned} \tag{3}$$

In this way one can make an unconditional test of the hypothesis with a critical region of prescribed size; of course one does not have complete freedom to specify the shape of T_c, but he can control it to the extent that $R_c(\hat{\theta})$ may be chosen arbitrarily for every $\hat{\theta}$. T_c is of course a similar region in the sense of Neyman and Pearson [2, 3, 4] for the augmented criterion, and the construction of T_c is essentially the same as that used by Neyman and Pearson to test parameters with sufficient estimators.

2. Application to contingency tables. As an illustration we shall follow Wilks' [5] treatment of a two-way table with r rows and c columns; the cell frequencies are n_{ij} and the cell probabilities are p_{ij} with

$$\sum n_{ij} = n; \quad \sum p_{ij} = 1; \quad i = 1, 2, \cdots, r; \quad j = 1, 2, \cdots, c. \tag{4}$$

The sample is thus regarded as having come from a multinomial population. We let

$$p_{i\cdot} = \sum_j p_{ij}; \qquad p_{\cdot j} = \sum_i p_{ij}; \qquad n_{i\cdot} = \sum_j n_{ij}; \qquad n_{\cdot j} = \sum_i n_{ij}. \tag{5}$$

The null hypothesis H_0 (of independence) corresponds to the subspace for which

$$p_{ij} = p_i q_j; \qquad \Sigma p_i = 1 = \Sigma q_j \tag{6}$$

in the parameter space of the p_{ij}. The likelihood ratio criterion for testing H_0 is

$$\lambda = \frac{(\Pi n_{i\cdot}^{\;n_{i\cdot}})(\Pi n_{\cdot j}^{\;n_{\cdot j}})}{n^n \Pi n_{ij}^{\;n_{ij}}} \tag{7}$$

and its distribution depends on the unknown parameters p_i and q_j. However the parameters have sufficient estimators

$$\hat{p}_i = n_{i\cdot}/n, \qquad \hat{q}_j = n_{\cdot j}/n \tag{8}$$

for the marginal distribution of the $n_{i\cdot}$ and $n_{\cdot j}$ is

$$\frac{(n!)^2}{(\Pi n_{i\cdot}!)(\Pi n_{\cdot j}!)} (\Pi p_i^{n_{i\cdot}})(\Pi q_j^{n_{\cdot j}}) \tag{9}$$

and when this is divided into the distribution of the n_{ij} (under the null hypothesis) one finds the conditional distribution of the n_{ij} to be

$$g(n_{11}, n_{12}, \cdots, n_{rc} \mid n_{1\cdot}, n_{2\cdot}, \cdots, n_{\cdot c}) = \frac{(\Pi n_{i\cdot}!)(\Pi n_{\cdot j}!)}{n!\Pi n_{ij}!} \tag{10}$$

which is independent of the parameters. The distribution (10) is just the combinatorial distribution used ordinarily in deriving the distribution of λ for small samples. The test for independence is therefore a conditional test which however may be interpreted as an unconditional test if the criterion λ is augmented by the estimators of the parameters under the null hypothesis. Instead of the likelihood ratio criterion Karl Pearson's Chi-square criterion could just as well have been used since its conditional distribution is also determined by (10).

The usual difficulty due to discreteness arises in this application to contingency tables. It is not possible to make the significance level exactly α. In terms of the notation of the first section, $R_c(\hat{\theta})$ cannot be chosen so that it will have size exactly equal to α for all $\hat{\theta}$. One would ordinarily replace the equalities by inequalities. The $R_c(\hat{\theta})$ would be chosen to have size less than but as close to α as possible. The size of T_c is then unspecified and one can only state that his significance level is less than α. This difficulty is not particularly serious in practice unless the test criterion has only one degree of freedom.

REFERENCES

[1] R. A. Fisher, *Statistical Methods for Research Workers*, Oliver and Boyd, London, 1946, pp. 96, 97.

[2] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses," *Roy. Soc. Phil. Trans.*, Series A, Vol. 231 (1933), p. 289.

[3] J. Neyman and E. S. Pearson, "Sufficient statistics and uniformly most powerful tests of statistical hypotheses," *Stat. Res. Memoirs*, Vol. 1 (1936), p. 113.

[4] J. Neyman, "Outline of a theory of statistical estimation based on the classical theory of probability," *Roy. Soc. Phil. Trans.*, Series A, Vol. 236 (1937), p. 364.

[5] S. S. Wilks, *Mathematical Statistics*, Princeton University Press, 1943, pp. 213–220.