A NOTE ON THE EQUIVALENCE OF SOME TEST CRITERIA

by

V. P. Bhapkar
University of North Carolina
and
University of Poona

Institute of Statistics Mimeo Series No. 421

February 1965

This research was supported by the Mathematics Division of the Air Force Office of Scientific Research Contract No. AF-AFOSR 760-65.

DEPARTMENT OF STATISTICS
UNIVERSITY OF NORTH CAROLINA
Chapel Hill, N. C.

A NOTE ON THE EQUIVALENCE OF SOME TEST CRITERIA

by

V. P. Bhapkar
University of North Carolina
and
University of Poona

---

1. Summary. It is shown in this note that for testing linear hypotheses in categorical data Wald's method [1943] is equivalent to the minimum $-\chi_1^2$ method due to Neyman [1949], and for testing nonlinear hypotheses it is equivalent to using Neyman's linearization technique along with the minimum $-\chi_1^2$ statistic.

2. Introduction. Suppose s independent random samples of experimental units are taken from s populations, n_{oj} is the size of the sample from the j th population and n_{ij} is the observed frequency in the i th category of the j th sample, $i = 1,2,\ldots,r$, $j = 1,2,\ldots, s$. If we assume that p_{ij} is the probability that an experimental unit drawn at random from the j th population belongs to the i th category, and that either the sampling is with replacement or, if without replacement, sampling fractions are negligible, then the probability distribution of the observed frequencies is given by

$$(2.1)\qquad \phi = \prod_{j=1}^{s} \left(\frac{n_{oj}!}{\prod_{i=1}^{r} n_{ij}!} \prod_{i=1}^{r} p_{ij}^{\ n_{ij}} \right) ,$$

where $\Sigma_i\, n_{ij} = n_{oj}$, a given integer, and $\Sigma_i p_{ij} = 1$.

Let

$$(2.2.)\qquad q_{ij} = n_{ij}/n_{oj} , \quad N = \Sigma_j\, n_{oj} \quad \text{and} \quad r_j = n_{oj}/N .$$

Suppose now we have to test the hypothesis

$$(2.3) \qquad H_o: \; F_k(\underset{\sim}{p}) = 0 \quad , \qquad k = 1,2, \ldots, t \, ,$$

where F's are t independent given functions of p's and $\underset{\sim}{p}' = [p_{11}, p_{21}, \ldots, p_{(r-1)1}, \ldots; p_{1s}, p_{2s}, \ldots, p_{(r-1)s}]$. It is assumed that F's possess continuous partial derivatives up to the second order with respect to p's and that the rank of the $t \times (rs-s)$ matrix $[\partial F_k(\underset{\sim}{p})/\partial p_{ij}]$ is $t \leq (rs-s)$. It is then well-known (e.g. refer to [2]) that H_o can be tested in various asymptotically equivalent ways by using either the minimum $-\chi^2$, minimum $-\chi_1^2$ or the likelihood-ratio statistic, together with any set of BAN [2] estimates of p's obtained subject to constraints (2.3). In general, equations giving maximum likelihood or minimum $-\chi^2$ (or χ_1^2) estimates are fairly complicated to solve and iterative methods have to be used; the minimum $-\chi_1^2$ estimates, though, can be obtained fairly easily by solving only linear equations whenever the constraints (2.3) are linear in p's. If the functions F_k are not linear, Neyman [2] has shown that the problem can still be reduced to the linear case by constructing the minimum $-\chi_1^2$ statistic using the minimum $-\chi_1^2$ estimate subject to the 'linearized' constraints

$$(2.4) \qquad F_k^*(\underset{\sim}{p}) \equiv F_k(\underset{\sim}{q}) + \sum_{j=1}^{s} \sum_{i=1}^{r-1} \left(\frac{\partial F_k(\underset{\sim}{p})}{\partial p_{ij}} \right)_{\underset{\sim}{p}=\underset{\sim}{q}} (p_{ij} - q_{ij}) = 0$$
$$k = 1,2,\ldots,t.$$

The purpose of this note is to point out that the minimum $-\chi_1^2$ statistic, in the linear and also in the nonlinear case (using 'linearization'), is equivalent to the statistic proposed by Wald [3] for a much more general problem but as adapted to the categorical set up.

3. Equivalence of Wald and Neyman Statistics:

We shall assume that the constraints (2.3), together with the basic constraints $\Sigma_i p_{ij} = 1$, have at least one set of solutions $\{p_{ij}\}$ for which $p_{ij} > 0$ for all $i = 1, \ldots, r$ and $j = 1, \ldots, s$. If so, the event $\{n_{ij} > 0, \text{ all } i,j\}$ has probability approaching one as $N \to \infty$ with $r_{\cdot j}$'s remaining fixed (or, more generally, $_j \to a_j$ with $a_j > 0$ and $\Sigma_j a_j = 1$). For asymptotic purposes, then, we assume that all n_{ij} are positive. If F's are linear, say,

$$(3.1) \qquad F_k(\boldsymbol{p}) \equiv \sum_{j=1}^{s} \sum_{i=1}^{r} f_{kij} p_{ij} + f_k \equiv \sum_{j=1}^{s} \sum_{i=1}^{r-1} (f_{kij} - f_{krj}) p_{ij} + f_k + \sum_{j=1}^{s} f_{krj} = 0$$

$$k = 1, 2, \ldots, t,$$

it is known (e.g. see [1]) that the minimum $-\chi_1^2$ statistic is

$$(3.2) \qquad \boldsymbol{c}' \boldsymbol{G}^{-1} \boldsymbol{c} \qquad ,$$

where

$$c_k = \sum_{j=1}^{s} \sum_{i=1}^{r} f_{kij} q_{ij} + f_k \quad , \qquad \boldsymbol{c}' = [c_1, c_2, \ldots, c_t]$$

$$(3.3) \qquad g_{kl} = \sum_{j=1}^{s} n_{oj}^{-1} \{ \sum_{i=1}^{r} f_{kij} f_{\ell ij} q_{ij} - (\sum_{i=1}^{r} f_{kij} q_{ij})(\sum_{i=1}^{r} f_{\ell ij} q_{ij}) \}$$

and $\qquad \boldsymbol{G} = [g_{k\ell}] \; , \quad k, \ell = 1, 2, \ldots, t \; .$

If the constraints F_k are not linear, let

$$(3.4) \qquad h_{kij} = \left(\frac{\partial F_k(\boldsymbol{p})}{\partial p_{ij}} \right)_{\boldsymbol{p} = \boldsymbol{q}} \qquad \text{if } i = 1, \ldots, r-1 \text{ and } 0 \text{ for } i = r.$$

Then the 'linearized' constraint (2.4) is linear of the form (3.1) with f's replaced by h's in (3.4) and

$$f_k = F_k(q) - \sum_{j=1}^{s} \sum_{i=1}^{r-1} h_{kij}\, q_{ij} \quad .$$

The minimum $-\chi_1^2$, subject to the linearized constraints (2.4), is then

$$(3.5) \qquad \underset{\sim}{h}'\underset{\sim}{G}^{-1}\underset{\sim}{h}$$

with $\underset{\sim}{h}' = [h_1, h_2, \ldots, h_t]$, $h_k = F_k(\underset{\sim}{q})$ and $\underset{\sim}{G}$ given by (3.3) with f's replaced by h's in (3.4).

Now Wald's general theorem [3] (under some regularity conditions) is as follows:

Let $\phi(x_1, \ldots, x_n;\ \theta_1, \ldots, \theta_u)$ be the joint probability distribution of the random variables x's, involving unknown parameters θ's and

$$\underset{\sim}{B}(\underset{\sim}{\theta}) = \left[-\frac{1}{n}\, \mathcal{E}_{\underset{\sim}{\theta}}\ \frac{\partial^2 \log \phi}{\partial\theta_\alpha \partial\theta_\beta} \right] \quad \alpha, \beta = 1, 2, \ldots, u.$$

which is assumed to be positive definite. Then if $\hat{\underset{\sim}{\theta}}$ is the maximum likelihood estimate of $\underset{\sim}{\theta}$, the statistic

$$(3.6) \qquad n\, \underset{\sim}{h}'(\hat{\underset{\sim}{\theta}})\, [\underset{\sim}{H}(\hat{\underset{\sim}{\theta}})\, \underset{\sim}{B}^{-1}(\hat{\underset{\sim}{\theta}}) \underset{\sim}{H}'(\hat{\underset{\sim}{\theta}})\,]^{-1} \underset{\sim}{h}\, (\hat{\underset{\sim}{\theta}}) \quad ,$$

where

$$\underset{\sim}{h}'(\underset{\sim}{\theta}) = [F_1(\underset{\sim}{\theta})\ , \ \ldots,\ F_t(\underset{\sim}{\theta})]$$

and

$$\underset{\sim}{H}(\underset{\sim}{\theta}) = \left[\frac{\partial F_k(\underset{\sim}{\theta})}{\partial\theta_\alpha} \right]_{t \times u} \qquad (k=1, \ldots, t;\ \alpha=1, \ldots, u) ,$$

has a limiting chi-square distribution with t degrees of freedom as $n \to \infty$ when the t independent conditions

$$F_k(\underset{\sim}{\theta}) = 0 \qquad\qquad k = 1, 2, \ldots, t ,$$

are satisfied.

Now for the categorical data with ϕ given by (2.1), $\underset{\sim}{\theta} = \underset{\sim}{p}$, $\hat{\underset{\sim}{\theta}} = \underset{\sim}{q}$, $n = N$ and $u = rs-s$. It is easy to verify that

$$\underset{\sim}{B}(\underset{\sim}{p}) = \frac{1}{N}\begin{bmatrix} n_{01}(\underset{\sim}{P}_1^{-1} + p_{r1}^{-1}\underset{\sim}{L}) & \underset{\sim}{0} & & \underset{\sim}{0} \\ \underset{\sim}{0} & n_{02}(\underset{\sim}{P}_2^{-1} + p_{r2}^{-1}\underset{\sim}{L}) & \cdots & \underset{\sim}{0} \\ & \cdots & & \\ \underset{\sim}{0} & \underset{\sim}{0} & & n_{0s}(\underset{\sim}{P}_s^{-1} + p_{rs}^{-1}\underset{\sim}{L}) \end{bmatrix}$$

where

$$\underset{\sim}{P}_j = \text{diagonal}\,(p_{1j},\ldots, p_{r-1,j}) \quad \text{and} \quad \underset{\sim}{L} = [1]_{(r-1)x(r-1)} \,.$$

Then $\underset{\sim}{B}^{-1}(\underset{\sim}{p}) = N\,\underset{\sim}{E}(\underset{\sim}{p})$, where

$$(3.7) \qquad \underset{\sim}{E}(\underset{\sim}{p}) = \begin{bmatrix} n_{01}^{-1}(\underset{\sim}{P}_1 - \underset{\sim}{p}_1\underset{\sim}{p}_1') & \underset{\sim}{0} & & \underset{\sim}{0} \\ \underset{\sim}{0} & n_{02}^{-1}(\underset{\sim}{P}_2 - \underset{\sim}{p}_2\underset{\sim}{p}_2') & \cdots & \underset{\sim}{0} \\ & \cdots & & \\ \underset{\sim}{0} & \underset{\sim}{0} & & n_{0s}^{-1}(\underset{\sim}{P}_s - \underset{\sim}{p}_s\underset{\sim}{p}_s') \end{bmatrix}$$

and $\underset{\sim}{p}_j' = [p_{1j} , \ldots, p_{r-1,j}]$.

So to test the hypothesis H_o given by (2.3), Wald's statistic, in view of (3.6), is

$$(3.8) \qquad \underset{\sim}{h}'\,[\underset{\sim}{H}(\underset{\sim}{q})\underset{\sim}{E}(\underset{\sim}{q})\underset{\sim}{H}'(\underset{\sim}{q})]^{-1}\underset{\sim}{h} \;,$$

with $\underset{\sim}{h}$ given as in (3.5), and

$$\underset{\sim}{H}(\underset{\sim}{q}) \equiv \left[\frac{\partial F_k(\underset{\sim}{p})}{\partial p_{ij}}\right]_{\underset{\sim}{p} = \underset{\sim}{q}} = \left[h_{kij}\right]_{tx(rs-s)} \,.$$

Hence

$$\underset{\sim}{H}'(\underset{\sim}{q}) = [\underset{\sim}{h}_1(\underset{\sim}{q}), \underset{\sim}{h}_2(\underset{\sim}{q}), \ldots, \underset{\sim}{h}_t(\underset{\sim}{q})] ,$$

where

$$\underset{\sim}{h}_k'(\underset{\sim}{q}) = [\underset{\sim}{h}_{k1}', \underset{\sim}{h}_{k2}', \ldots, \underset{\sim}{h}_{ks}'] ,$$

and $\underset{\sim}{h}_{kj}' = [h_{k1j}, \ldots, h_{k(r-1)j}]$. Therefore

$$\underset{\sim}{H}(\underset{\sim}{q})\, \underset{\sim}{E}(\underset{\sim}{q})\, \underset{\sim}{H}'(\underset{\sim}{q}) = [\underset{\sim}{h}_k'(\underset{\sim}{q})\, \underset{\sim}{E}(\underset{\sim}{q})\, \underset{\sim}{h}_\ell(\underset{\sim}{q})]_{t \times t}$$

$$(3.9) \qquad = [\Sigma_j n_{oj}^{-1}\, \underset{\sim}{h}_{kj}'\, (\underset{\sim}{Q}_j - \underset{\sim}{q}_j \underset{\sim}{q}_j')\, \underset{\sim}{h}_{\ell j}]_{t \times t} ,$$

where $\underset{\sim}{Q}_j$ = diagonal $(q_{1j}, \ldots, q_{r-1,j})$ and $\underset{\sim}{q}_j' = [q_{1j}, \ldots, q_{r-1,j})$.

It is then easy to see that Wald's statistic (3.8) thus reduces to Neyman's statistic (3.5) in view of (3.3) and (3.4).

If the hypothesis is linear, given by (3.1), $\underset{\sim}{h} = \underset{\sim}{c}$, $h_{kij} = f_{kij} - f_{krj}$, so that

$$\underset{\sim}{h}_{kj}' = \underset{\sim}{f}_{kj}' - \underset{\sim}{f}_{krj}\, \underset{\sim}{\ell}' ,$$

where $\underset{\sim}{f}_{kj}' = [f_{k1j}, \ldots, f_{k(r-1)j}]$ and $\underset{\sim}{\ell}' = [1]_{1x(r-1)}$.

The matrix (3.9) is then equal to the matrix $\underset{\sim}{G}$ in (3.3) and the assertion follows.

REFERENCES

[1] V. P. Bhapkar, "Some tests for categorical data," Ann. Math. Stat., Vol. 32 (1961), pp. 72-83.

[2] J. Neyman, "Contributions to the theory of χ^2 test," Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, 1949, pp. 239-273.

[3] A. Wald, "Tests of statistical hypotheses concerning several parameters when the number of observations is large," Trans. Amer. Math.Soci., Vol. 54 (1943), pp. 426-482.