Scientific
Research

A Multivariate Test for Three-Factor Interaction in 3-Way Contingency Table under the Multiplicative Model

Njoku O. Ama

Department of Statistics, University of Botswana, Gaborone, Botswana
Email: njoku52@gmail.com, amano@mopipi.ub.bw

Abstract

Two test statistics that have been commonly used in analysing interactions in contingency table are the Pearson's Chi-square statistic, χ^2, and likelihood ratio test statistic, G^2. Both test statistics, in tables with sufficiently large sample size, have an asymptotic chi-square distribution with degrees of freedom (df) equal to the number of free parameters in the saturated model. For example under the hypothesis of independence of the row and column conditioned on the layer in an $I \times J \times K$ contingency table, the df is $K(I-1)(J-1)$. These test statistics, in large sized tables, will have less power since they have large degrees of freedom. This paper proposes a product effect model, which combines the advantages of the multiplicative models over the additive, for analysing the interaction between the row and column of the 3-way table conditioned on the layer. The derived statistics is shown to be asymptotically chi-square with a small degree of freedom, $K-1$, for the $I \times J \times K$ contingency table. The performance of the developed statistic is compared with the Pearson's chi-square statistic and the likelihood ratio statistic test using an illustrative example. The results show that the product effect test can detect interaction even when some of the main effects are not significant and can perform better than the other competitors having smaller degree of freedom in large sized tables.

Keywords

Contingency Tables, Product Effect, Models, Interaction

1. Introduction

A 3-way contingency table is a cross-classification of observations by the levels of three categorical variables—*A*,

B, and *C*. The levels can be ordinal or nominal. If *n* units in a sample are independently and identically distributed (*IID*); that is, if they constitute a random sample, then the vector of cell counts $\underline{x}' = \{n_{111}, n_{112}, \cdots, n_{IJK}\}$ has a multinomial distribution with index $n = n_{+++}$ and a parameter, π_{ijk}, where $\pi_{ijk} = P(A = i, B = j, C = k)$ probability that a randomly selected unit falls into the $(i, j, k)^{th}$ cell of the contingency table with variables *A*, *B* and *C*. The probability distribution $\{\pi_{ijk}\}$ is the joint distribution of *A*, *B*, and *C*.

Interaction in the 3-way contingency has been tested using the chi-square statistic and the likelihood ratio test statistic with $(I-1)(J-1)(K-1)$ degree of freedom [1]-[3]. Grizzle *et al.* [4], Darroch [5] and Johnson and Graybill [6] have modelled interaction as a product of the marginal effects or components of the ways of classification of the table. Tukey [7] in order to overcome the difficulty in testing for interaction in the two-factorial experiment with one observation per cell modelled the two-factor interaction as a product of the effects of the two factors and developed a one degree of freedom F-test for analysing the interaction between the factors. Drawing an analogy from the two factorial experiments we can view the 2-way contingency table as a two-factorial experiment with one observation per cell similar to the Poisson modelling of 2-way contingency table where the cell observations are seen as the mean number of occurrences of the event within a defined infinitesimal interval. The $I \times J \times K$ contingency table can be viewed as *K* 2-way contingency table. The present paper argues that the transformation of the contingency table applicable in the likelihood ratio tests [1] [8] is often unnecessary but that the data can be analysed in a manner similar to the Pearson's chi-square which does not transform the data. A multivariate approach is adopted in analysing the interaction in an $I \times J \times K$ table, under the product effect model, where the three-factor interaction is defined as a product of the effects of the ways of classification of the table. The advantage of the proposed model is that it gives rise to chi-square tests with smaller degrees of freedom, irrespective of the size of the table, which is conjectured to have greater power than other tests with larger degrees of freedom. It has been shown [9] [10] that the power of the noncentral chi-square statistics, for a given value of non-centrality parameter and level of significance, increases as the degree of freedom decreases. Our results will be of some practical value to researchers who are involved in analysing mutual independence in higher order tables as their results will be based on small degrees of freedom leaving extra degrees of freedom for further decomposition of other forms of independence in the data. Extension of the proposed method to higher order tables is straightforward.

Model for 3-Factor Interaction

Let us assume that we have an $I \times J \times K$ 3-way contingency, representing respectively the row, column and layer classifications of the table, and that the *K*-dimensional vector $\{n_{ijk}\}_{K_c}$ of frequency layers available in the $(i, j)^{th}$ cell has associated it with a *K*-dimensional vector $\{\pi_{ijk}\}_K$ of unknown cell probabilities such that $\sum_k \pi_{ijk} = 1$. In addition if we assume that $\{n_{ijk}\}_K$ follows a multinomial probability distribution given by

$$P\{(n_{ijk})_K; \pi_{ijk}, n_{ij.}\} = \frac{n_{ij.}!}{n_{ij1}! n_{ij2}! \cdots n_{ijK}!} \pi_{ij1}^{n_{ij1}} \pi_{ij2}^{n_{ij2}} \cdots \pi_{ijK}^{n_{ijK}} \tag{1.1}$$

$n_{ij.} = \sum_k n_{ijk}$ is fixed, $n_{ijk} = 1, 2, \cdots, n_{ij}$; $0 < \pi_{ijk} < 1$, $\hat{\pi}_{ijk} = n_{ijk} / n_{ij}$.

$$Cov(n_{ijk}, n_{ijk'}) = \begin{cases} n_{ij.} \pi_{ijk} (1 - \pi_{ijk}), & k = k' \\ -n_{ij.} \pi_{ijk} \pi_{ijk'}, & k \neq k' \end{cases}$$

$\{n_{ijk}, n_{ijk'}\}$ are independent for all $k \neq k'$.

The 3-way contingency table under the multinomial structure described above is similar to the layout of a three-factorial experiment with one observation per cell. In the spirit of [7] and drawing an analogy from the factorial experimental structure, a linear additive model for the observed cell probability in the (i, j, k)-cell can be written as in (1.2). The interest is in the consideration of models where these probabilities depend on a vector x_i of covariates associated with the $(i, j, k)^{th}$ individual or group.

Under the assumption (1.1), and given the kth layer, we have an identity relation

$$\pi_{ijk} = \pi_{..k} + \pi_{i.k} - \pi_{..k} + \pi_{.jk} - \pi_{..k} + \pi_{ijk} - \pi_{i.k} - \pi_{.jk} + \pi_{..k} \tag{1.2}$$

where, $\pi_{..k}$, $\pi_{i.k}$ and $\pi_{.jk}$ denote respectively the overall probability for the kth layer, i^{th} row, and j^{th} column

for the kth layer. Reasoning by analogy from the regular analysis of variance, we get a linear additive model for (1.2) as

$$E\left(\hat{\pi}_{ijk}\right) = \mu_k + \tau_{ik} + \beta_{jk} + \lambda_{ijk} \tag{1.3}$$

where, $\mu_k = \pi_{..k}$ is the overall probability of an observation belonging to the kth layer of the 3-way contingency table; $\tau_{ik} = \pi_{i.k} - \pi_{..k}$ is the effect of the ith row of the table for the kth layer; $\beta_{jk} = \pi_{.jk} - \pi_{..k}$ is the effect of the jth column of the table for the kth layer; $\lambda_{ijk} = \pi_{ijk} - \pi_{i.k} - \pi_{.jk} + \pi_{..k}$ is the interaction between the ith row and j^{th} column for the kth layer of the table. These parameters are subject to the restrictions:

$$\sum_i n_{ij.}\tau_{ik} = \sum_j n_{ij.}\beta_{jk} = 0; \qquad \sum_i n_{ij.}\lambda_{ijk} = \sum_j n_{ij.}\lambda_{ijk} = \sum_{ij} n_{ij.}\lambda_{ijk} = 0 \tag{1.4}$$

and are independent of the kth layer.

The relation (1.3) can be recast in vector notation as

$$E\left(\underline{n}_{ij}\right) = n_{ij}\left(\underline{\mu} + \underline{\tau}_i + \underline{\beta}_j + \underline{\lambda}_{ij}\right) \tag{1.5}$$

Or

$$E\left(\underline{\hat{\pi}}_{ij}\right) = \left(\underline{\mu} + \underline{\tau}_i + \underline{\beta}_j + \underline{\lambda}_{ij}\right) \tag{1.6}$$

where,

$$\underline{\mu}' = \left(\mu_1, \mu_2, \cdots, \mu_K\right); \quad \underline{\tau}_i' = \left(\tau_{i1}, \tau_{i2}, \cdots, \tau_{iK}\right); \quad \underline{\beta}_j' = \left(\beta_{j1}, \beta_{j2}, \cdots, \beta_{jK}\right); \quad \underline{\lambda}_{ij}' = \left(\lambda_{ij1}, \lambda_{ij2,}\cdots, \lambda_{ijK}\right)$$

$$Cov\left(\underline{n}_{ij}\right) = \begin{cases} n_{ij}\left(D_{ij} - \underline{\pi}_{ij}\underline{\pi}_{ij}'\right) & \text{in the main diagonal} \\ -n_{ij}\underline{\pi}_{ij}\underline{\pi}_{ij}' & \text{in the off diagonal positions} \end{cases} \tag{1.7}$$

$$D_{ij} = \text{diag}\left(\pi_{ij1}, \pi_{ij2}, \cdots, \pi_{ijK}\right); \quad \underline{\pi}_{ij}' = \left(\pi_{ij1}, \pi_{ij2}, \cdots, \pi_{ijK}\right).$$

Estimation of the parameters of this model (1.6) by maximum likelihood proceeds by maximization of the multinomial likelihood (1.1) with the probabilities π_{ijk} viewed as functions of the parameters $\mu_{..k}$, $\tau_{i.k}$, $\beta_{.jk}$, and λ_{ijk} in the Equation (1.3) and yields

$$\hat{\underline{\mu}} = \hat{\underline{\pi}}; \quad \hat{\underline{\tau}}_i = \hat{\underline{\pi}}_i - \hat{\underline{\pi}}; \quad \hat{\underline{\beta}}_j = \hat{\underline{\pi}}_j - \hat{\underline{\pi}}; \quad \hat{\underline{\lambda}}_{ij} = \hat{\underline{\pi}}_{ij} - \hat{\underline{\pi}}_i - \hat{\underline{\pi}}_j + \hat{\underline{\pi}} \tag{1.8}$$

where,

$$\hat{\underline{\pi}}_{ij} = \underline{n}_{ij}/n_{ij.}; \quad \hat{\underline{\pi}}_i = \underline{n}_i/n_{i..}; \quad \hat{\underline{\pi}}_j = \underline{n}_j/n_{.j.}; \quad \hat{\underline{\pi}}' = \left(\hat{\pi}_{..1}, \hat{\pi}_{..2}, \cdots, \hat{\pi}_{..K}\right); \quad \hat{\underline{\pi}}_i' = \left(\hat{\pi}_{i.1}, \hat{\pi}_{i.2}, \cdots, \hat{\pi}_{i.K}\right)$$

$$\hat{\underline{\pi}}_j' = \left(\hat{\pi}_{.j1}, \hat{\pi}_{.j2}, \cdots, \hat{\pi}_{.jK}\right); \quad \hat{\underline{\pi}}_{ij}' = \left(\hat{\pi}_{ij1}, \hat{\pi}_{ij2}, \cdots, \hat{\pi}_{ijK}\right); \quad \hat{\underline{\pi}}' = \underline{n}/n_{...}; \quad \underline{n}' = \left(n_{..1}, n_{..2}, \cdots, n_{..K}\right)$$

$$n_{...} = \sum_{ij} n_{ij} = \sum_i n_{i..} = \sum_j n_{.j.} = \sum_{ijk} n_{ijk}.$$

The matrix $\overline{n}$ is given as

$$\overline{n}_{K \times IJ} = \begin{bmatrix} n_{111}, n_{121}, \cdots, n_{1J1} & n_{211}, n_{221}, \cdots, n_{2J1} & \cdots & n_{I11}, n_{I21}, \cdots, n_{IJ1} \\ n_{112}, n_{122}, \cdots, n_{1J2} & n_{212}, n_{222}, \cdots, n_{2J2} & \cdots & n_{I12}, n_{I22}, \cdots, n_{IJ2} \\ \cdots & \cdots & \cdots & \cdots \\ n_{11K}, n_{12K}, \cdots, n_{1JK} & n_{21K}, n_{22K}, \cdots, n_{2JK} & \cdots & n_{I1K}, n_{I2K}, \cdots, n_{IJK} \end{bmatrix}$$

For the kth layer, the interaction between the ith row and j^{th} column is defined multiplicatively as being proportional to the ith row effect and j^{th} column effect and given as

$$\lambda_{ijk} = c_k \tau_{ik} \beta_{jk} \tag{1.9}$$

where, c_k is an unknown constant for the layer; τ_{ik} and β_{jk} are respectively the effect of the ith row and j^{th} column for the kth layer. The model (1.9) is referred to as the product effect model [2] [3]. The classical method of partitioning the chi-squares for the 3-way $\left(I \times J \times K\right)$ contingency table does not provide a convenient test of the null hypothesis that the 3-way interaction is zero [11]. The model indicates that the three-factor interac-

tion in the contingency table and for the kth layer response is proportional to the product of the effects of ith row classification and the jth column classification of the table. Darroch [12] has demonstrated the advantages of the multiplicative interaction models over the additive.

2. Development of Test Statistics Based on the Model $\lambda_{ijk} = c_k \tau_{ik} \beta_{jk}$

Rewriting (1.9), the model for the two-factor interaction for the kth layer response, in vector notation,

$$\underline{\lambda}_{ij} = \underline{c} D_\tau D_\beta \tag{2.1}$$

where $\underline{c} = (c_1, c_2, \cdots, c_K)$; $D_\tau = \mathrm{diag}(\tau_{i1}, \tau_{i2}, \cdots, \tau_{iK})$; $D_\beta = \mathrm{diag}(\beta_{j1}, \beta_{j2}, \cdots, \beta_{jK})$.

The matrix of interaction Λ can be written as

$$\underset{K \times IJ}{\Lambda_c} = \begin{bmatrix} c_1\tau_{11}\beta_{11}, c_1\tau_{11}\beta_{21}, \cdots, c_1\tau_{11}\beta_{J1} & c_1\tau_{21}\beta_{11}, c_1\tau_{21}\beta_{21}, \cdots, c_1\tau_{21}\beta_{J1} & \cdots & c_1\tau_{I1}\beta_{11}, c_1\tau_{I1}\beta_{21}, \cdots, c_1\tau_{I1}\beta_{J1} \\ c_2\tau_{12}\beta_{12}, c_2\tau_{12}\beta_{22}, \cdots, c_2\tau_{12}\beta_{J2} & c_2\tau_{22}\beta_{12}, c_2\tau_{22}\beta_{22}, \cdots, c_2\tau_{22}\beta_{J2} & \cdots & c_2\tau_{I2}\beta_{12}, c_2\tau_{I2}\beta_{22}, \cdots, c_2\tau_{I2}\beta_{J2} \\ \cdots & \cdots & \cdots & \cdots \\ c_K\tau_{1K}\beta_{1K}, c_K\tau_{1K}\beta_{2K}, \cdots, c_K\tau_{1K}\beta_{JK} & c_K\tau_{2K}\beta_{1K}, c_K\tau_{2K}\beta_{2K}, \cdots, c_K\tau_{2K}\beta_{JK} & \cdots & c_K\tau_{IK}\beta_{1K}, c_K\tau_{IK}\beta_{2K}, \cdots, c_K\tau_{IK}\beta_{JK} \end{bmatrix} \tag{2.2}$$

From (1.5) or (1.6) the residual after substituting (2.1) becomes

$$E(\underline{Z}_{ij}) = n_{ij}\underline{\lambda}_{ij} = n_{ij}(\underline{c}D_\tau D_\beta) \tag{2.3}$$

$$\underline{Z}_{ij} = \underline{\hat{\pi}}_{ij} - (\underline{\hat{\mu}} + \underline{\hat{\tau}}_i + \underline{\hat{\beta}}_j) = \underline{n}_{ij} - n_{ij}(\underline{\hat{\mu}} + \underline{\hat{\tau}}_i + \underline{\hat{\beta}}_j) \tag{2.4}$$

$$\underline{Z}'_{ij} = (Z_{ij1}, Z_{ij2}, \cdots, Z_{ijK}) $$

This gives the least square estimate $\underline{\hat{c}}$ of $\underline{c}$ as

$$\hat{c}_k = \frac{\sum_{ij} n_{ijk}\hat{\tau}_{ik}\hat{\beta}_{jk}}{\sum_{ij} n_{ij}\hat{\tau}_{ik}^2\hat{\beta}_{ij}^2} = \left(\frac{\underline{\hat{\tau}}_i}{\sqrt{n_{ij}\underline{\hat{\tau}}'_i\underline{\hat{\tau}}_i}} \otimes \frac{\underline{\hat{\beta}}_j}{\sqrt{n_{ij}\underline{\hat{\beta}}'_j\underline{\hat{\beta}}_j}} \right)' \underline{n}_{ij} \tag{2.5}$$

The matrix of sum of squares sum of product (SS-SP) for interaction from (2.3) is

$$\Lambda_c\Lambda'_c = H_c = \begin{cases} \hat{c}_k^2\sum_{ij} n_{ij}\hat{\tau}_{ij}^2\hat{\beta}_{jk}^2 & \text{at the } k\text{-th diagonal position} \\ \hat{c}_k\hat{c}_{k'}\sum_{ij} n_{ij}\hat{\tau}_{ik}\hat{\tau}_{ik'}\hat{\beta}_{jk}\hat{\beta}_{jk'} & \text{at the } (k,k')\text{th off diagonal position} \end{cases} \tag{2.6}$$

with expectation

$$E(H_c|\hat{\tau}_{ik}, \hat{\beta}_{jk})$$

$$= \begin{cases} \left[\sum_{ij} n_{ij}\pi_{ijk}(1-\pi_{ijk})\hat{\tau}_{ik}^2\hat{\beta}_{jk}^2 + c_k^2\left(\sum_{ij} \tau_{ik}\hat{\tau}_{ik}\beta_{jk}\hat{\beta}_{jk}\right)^2 \right] \bigg/ \sum_{ij} n_{ij}\hat{\tau}_{ik}^2\hat{\beta}_{jk}^2 & \text{at the } k\text{-th diagonal position} \\ -\frac{\sum_{ij} n_{ij}^2\pi_{ijk}\pi_{ijk'}\hat{\tau}_{ik}^2\hat{\tau}_{ik'}^2\hat{\beta}_{jk}^2\hat{\beta}_{jk'}^2}{\sum_{ij} n_{ij}^2\hat{\tau}_{ik}^2\hat{\tau}_{ik'}^2\hat{\beta}_{jk}^2\hat{\beta}_{jk'}^2} + c_kc_{k'}\sum_{ij} n_{ij}\hat{\tau}_{ik}\hat{\tau}_{ik'}\hat{\beta}_{jk}\hat{\beta}_{jk'} & \text{at the } (k,k') \text{ of diagonal position} \end{cases} \tag{2.7}$$

The total sum of squares and cross product (SS-SP) is given as

$$H_T = n(D_T - \underline{\hat{\pi}}\underline{\hat{\pi}}') \tag{2.8}$$

where,

$$D_T = \mathrm{diag}(\hat{\pi}_{.1}, \hat{\pi}_{.2}, \cdots, \hat{\pi}_{.K}) \tag{2.9}$$

The expectation of H_T is

$$E(H_T) = (n-1)V$$

$$V = \begin{cases} 1 - \pi_{..k} & \text{at the } k\text{-th diagonal position} \\ -\pi_{..k}\pi_{..k'} & \text{at the } (k,k')^{th} \text{ off diagonal position} \end{cases} \tag{2.10}$$

The total SS-SP matrix H_T can be partitioned into unit SS-SP, H_u, SS-SP due to the row effect, H_τ, SS-SP due to the column effect, H_β, and SS-SP due to the residual, H_z, namely

$$H_T = H_u + H_\tau + H_\beta + H_z \tag{2.11}$$

The unit SS-SP matrix H_u is given by

$$H_u = \sum_{ij} n_{ij}\left(D_u - \hat{\underline{\pi}}_{ij}\hat{\underline{\pi}}_{ij}'\right) \tag{2.12}$$

with $D_u = \text{diag}\left(\hat{\pi}_{ij1}, \hat{\pi}_{ij2}, \cdots, \hat{\pi}_{ijK}\right)$.

The expectation of H_u is given by

$$E(H_u) = \sum_{ij}(n_{ij}-1)\begin{cases} \pi_{ijk}\left(1-\pi_{ijk}\right) & \text{at the } k\text{-th diagonal position} \\ -\pi_{ijk}\pi_{ijk'} & \text{at the } (k,k')^{th} \text{ off diagonal position} \end{cases} \tag{2.13}$$

The matrix of SS-SP for the row effect, H_τ, is

$$H_\tau = \sum_i n_{i..}\hat{\underline{\tau}}_i\hat{\underline{\tau}}_i' = \begin{cases} \sum_i n_{i..}\hat{\tau}_{ik}^2 & \text{at the } k\text{-th diagonal position} \\ \sum_i n_{i..}\hat{\tau}_{ik}\hat{\tau}_{ik'} & \text{at the } (k,k') \text{ off diagonal position} \end{cases} \tag{2.14}$$

With expectation,

$$E(H_\tau) = D_\tau + V_\tau \tag{2.15}$$

$D_\tau = \text{diag}\left(b_1, b_2, \cdots, b_K\right)$ and

$$V_\tau = \begin{cases} -\sum_i \pi_{i.k}\left(1-\pi_{i.k}\right) + \pi_{..k}\left(1-\pi_{..k}\right) & \text{at the } k\text{-th diagonal position} \\ \sum_i \pi_{i.k}\pi_{i.k'} + \pi_{..k}\pi_{..k'} + \sum_i n_{i..}\pi_{i.k}\pi_{i.k'} - n\pi_{..k}\pi_{..k'} & \text{at the } (k,k') \text{ off diagonal position} \end{cases} \tag{2.16}$$

$$b_k = \sum_i n_{i..}\tau_{ik}^2$$

The matrix of SS-SP, H_β, due to the column effect is

$$H_\beta = \sum_j n_{.j.}\hat{\underline{\beta}}_j\hat{\underline{\beta}}_j' = \begin{cases} \sum_j n_{.j.}\hat{\beta}_{jk}^2 & \text{at the } k\text{-th diagonal position} \\ \sum_j n_{.j.}\hat{\beta}_{jk}\hat{\beta}_{jk'} & \text{at the } (k,k') \text{ off diagonal position} \end{cases} \tag{2.17}$$

With expectation,

$$E(H_\beta) = D_\beta + V_\beta \tag{2.18}$$

$$D_\beta = \text{diag}\left(t_1, t_2, \cdots, t_K\right); \quad t_k = \sum_j n_{.j.}\hat{\beta}_{jk}^2$$

and

$$V_\beta = \begin{cases} -\sum_i \pi_{.jk}\left(1-\pi_{.jk}\right) + \pi_{..k}\left(1-\pi_{..k}\right) & \text{at the } k\text{-th diagonal position} \\ -\sum_j \pi_{.jk}\pi_{.jk'} + \pi_{..k}\pi_{..k'} + \sum_j n_{.j.}\pi_{.jk}\pi_{.jk'} - n\pi_{..k}\pi_{..k'} & \text{at the } (k,k') \text{ off diagonal position} \end{cases} \tag{2.19}$$

The matrix of SS-SP for the residual (2.3) is

$$H_z = \sum_{ij} n_{ij} \underline{z}_{ij} \underline{z}'_{ij} = \sum_{ij} n_{ij} \left[\underline{\hat{\pi}}_{ij} - \underline{\hat{\pi}} - \underline{\hat{\tau}}_i - \underline{\hat{\beta}}_j \right] \left[\underline{\hat{\pi}}_{ij} - \underline{\hat{\pi}} - \underline{\hat{\tau}}_i - \underline{\hat{\beta}}_j \right]'$$

$$= \sum_{ij} n_{ij} \left(\underline{\hat{\pi}}_{ij} - \underline{\hat{\pi}} \right) \left(\underline{\hat{\pi}}_{ij} - \underline{\hat{\pi}} \right)' - \sum_{ij} n_{ij} \left(\underline{\hat{\pi}}_{ij} - \underline{\hat{\pi}} \right) \underline{\hat{\tau}}'_i - \sum_{ij} n_{ij} \left(\underline{\hat{\pi}}_{ij} - \underline{\hat{\pi}} \right) \underline{\hat{\beta}}'_j - \sum_{ij} n_{ij} \underline{\hat{\tau}}_i \left(\underline{\hat{\pi}}_{ij} - \underline{\hat{\pi}} \right)' \quad (2.20)$$

$$- \sum_{ij} n_{ij} \underline{\hat{\beta}}_j \left(\underline{\hat{\pi}}_{ij} - \underline{\hat{\pi}} \right)' + \sum_{ij} n_{ij} \underline{\hat{\tau}}_i \underline{\hat{\tau}}'_i + \sum_{ij} n_{ij} \underline{\hat{\tau}}_i \underline{\hat{\beta}}'_j + \sum_{ij} n_{ij} \underline{\hat{\beta}}_j \underline{\hat{\tau}}'_i + \sum_{ij} n_{ij} \underline{\hat{\beta}}_j \underline{\hat{\beta}}'_j.$$

With expectation

$$E(H_z) = \sum_{ij} n_{ij} E \left(\underline{\hat{\pi}}_{ij} - \underline{\hat{\pi}} \right) \left(\underline{\hat{\pi}}_{ij} - \underline{\hat{\pi}} \right)' + \sum_{ij} n_{ij} E \left(\underline{\hat{\tau}}_i \underline{\hat{\tau}}'_i \right) + \sum_{ij} n_{ij} E \left(\underline{\hat{\beta}}_j \underline{\hat{\beta}}'_j \right) \quad (2.21)$$

Since the cross-product terms will vanish on taking expectation because of independence and restriction in (1.4)

$$E \left(\underline{\hat{\pi}}_{ij} - \underline{\hat{\pi}} \right) \left(\underline{\hat{\pi}}_{ij} - \underline{\hat{\pi}} \right)' = \begin{cases} E \left(\hat{\pi}_{ijk} - \hat{\pi}_{..k} \right)^2 & \text{at the } k^{th} \text{ diagonal position} \\ E \left(\hat{\pi}_{ijk} - \hat{\pi}_{..k} \right) \left(\hat{\pi}_{ijk'} - \hat{\pi}_{..k'} \right) & \text{at the } (k,k') \text{ off diagonal position} \end{cases} \quad (2.22)$$

Hence,

$$E(H_z) = D_z + V_z \quad (2.23)$$

where,

$$D_z = D_\theta + D_\tau + D_\beta \quad (2.24)$$

and

$$D_\theta = \text{diag}(q_1, q_2, \cdots, q_K), \text{ with } q_k = \sum_{ij} n_{ij} \left(\pi_{ijk} - \pi_{..k} \right)^2$$

$V_z = V_\theta + V_\tau + V_\beta$, with

$$V_\theta = \begin{cases} -\sum_i \pi_{ijk} \left(1 - \pi_{ijk} \right) - \pi_{..k} \left(1 - \pi_{..k} \right) & \text{at the } k^{th} \text{ diagonal position} \\ -\sum_{ij} \left(n_{ij} - 1 \right) \pi_{ijk} \pi_{ijk'} + (n-1) \pi_{..k} \pi_{..k'} & \text{at the } (k,k') \text{ off diagonal position} \end{cases} \quad (2.25)$$

The hypothesis of no interaction, $H_0 : \lambda_{ijk} = 0$, for all k, implies that either $c_k = 0$ or $\tau_{ik} = 0$ or $\beta_{jk} = 0$ for all k.

Hence

$$H_0 : \underline{\lambda}_{ij} = \underline{0} \Rightarrow H_0 : \underline{c} = \underline{0} \text{ or } H_0 : \underline{\tau}_i = \underline{0} \text{ or } H_0 : \underline{\beta}_j = \underline{0}. \quad (2.26)$$

Under the null hypothesis (2.26), in which case $\pi_{ijk} = \pi_{..k}$, and reasoning from (2.13) for the kth layer, $E(H_u) = (n - IJ)V$.

Similarly, $\pi_{i.k} = \pi_{..k}$; $\pi_{.jk} = \pi_{..k}$ (see Section 1.1)

$$E(H_\tau) = (I-1)V; \ E(H_\beta) = (J-1)V, \ E(H_c) = V \text{ and } E(H_z) = (I-1)(J-1)V \quad (2.27)$$

However, whether or not $H_0 : \underline{c} = \underline{0}$ is true,

$$E(H_T) = (n-1)V$$

where V is as defined in (2.10). Each of the quantities $(n-IJ)^{-1} H_u$, $(I-1)^{-1} H_\tau$, $(J-1)^{-1} H_\beta$, $\left[(I-1)(J-1) \right]^{-1} H_z$, $(n-1)^{-1} H_T$ and H_c provides an estimate of V and can be employed in the construction of tests of significance of the row, column effects and interaction provided that they are independent.

Independence of H_T, H_c, H_τ, H_β

By appealing to the following theorem [13], it can be shown that the quadratic forms H_T, H_c, H_τ and H_β, are independent.

Theorem 2.1. Let $\underline{Y}$ be distributed $N(\mu, I)$, the set of positive semi-definite quadratic forms $\underline{Y}'B_1\underline{Y}$, $\underline{Y}'B_2\underline{Y}$, $\cdots$, $\underline{Y}'B_K\underline{Y}$ are jointly independent if and only if $B_iB_j = 0$, the null matrix for all $i \neq j$.

Theorem 2.2. As $n_{ij} \to \infty$, the matrices H_T, H_c, H_τ and H_β are independent.

By theorem 1, the joint independence of H_T, H_c, H_τ and H_β implies pairwise independence.

3. Construction of Test Statistic for the Hypothesis $H_0 : \underline{c} = \underline{0}$

Recall that n_{ijk} follows $M(n_{ij.}, \pi_{ijk})$. As π_{ijk} tends to a constant, say $\dfrac{1}{K}$, then $\underline{n}_{ij}$ will follow asymptotic multivariate normal distribution with mean $\underline{n}_{ij}\underline{\pi}_{ij}$ and variance, $n_{ij}V_{ij}$, where V_{ij} is a singular matrix given by (1.7). Therefore $\underline{n}_{ij}$ has a singular normal distribution.

Under the hypothesis, $H_0 : \underline{c} = \underline{0}$, the matrix H_c has a pseudo Wishart distribution with parameter 1 and $V = \sum_{ij} n_{ij}V_{ij}$. The random matrix H_T follows the Wishart distribution with parameter $(n-1)$ and V and independent of H_c. They can be used in constructing the determinant based test statistic for the hypothesis $H_0 : \underline{c} = \underline{0}$

Since the matrix V is nonsingular, by generating the matrix of contrasts, say B and pre- and post-multiplying each of them by B and B transpose, V can be made non-singular.

Let

$$\Lambda_c = (\underline{m}_1', \underline{m}_2', \cdots, \underline{m}_K') \tag{3.1}$$

$\underline{m}_k'$ is an IJ column vector of independent variables for the kth response. Also define $\underline{\overline{X}}' = (\underline{X}_1', \underline{X}_2', \cdots, \underline{X}_{K-1}')$ such that

$$\underline{X}_1 = \underline{m}_1 - \underline{m}_K$$
$$\underline{X}_2 = \underline{m}_2 - \underline{m}_K$$
$$\vdots$$
$$\underline{X}_{K-1} = \underline{m}_{K-1} - \underline{m}_K$$

Then

$$\underline{\overline{X}} = B\Lambda_c \tag{3.2}$$

where

$$\underset{K-1 \times K}{\boldsymbol{B}} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & -1 \\ 0 & 1 & 0 & \cdots & 0 & -1 \\ 0 & 0 & 1 & \cdots & 0 & -1 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 \end{bmatrix} = (I_d; -1) \tag{3.3}$$

The matrix $\boldsymbol{B}$ is of full rank, $(K-1)$ and I_d is a $(K-1) \times (K-1)$ identity matrix.

Certainly,

$$H_c^* = \underline{\overline{X}}\,\underline{\overline{X}}' = B\Lambda_c\Lambda_c'B' = BH_cB' \tag{3.4}$$

is a non-singular transformation of the matrix H_c and so also is the matrix

$$\underset{(K-1) \times (K-1)}{H_T^*} = BH_TB' \tag{3.5}$$

The hypothesis

$$H_c : \underline{c} = \underline{0}$$

is similarly transformed to

$$H_c^* : B\underline{c} = \underline{0} \quad \text{or} \quad H_c^* : \underline{c}^* = \underline{0}$$

However, [14] and [15] have discussed the equivalence between H_0 and H_C^* and the invariant property of the Wilks Λ criterion under such transformation as above. Also the quadratic forms H_T^* and H_C^* (3.4 and 3.5) are independent Wishart distributed matrices with same degrees of freedom as H_T and H_c respectively and variance-covariance matrix $V^* = BVB'$. Hence the analogue of the Wilks criterion can be used in testing the hypotheses and is given by

$$W_c = \frac{\left|H_T^*\right|}{\left|H_T^* + H_C^*\right|} \approx \Lambda(n, K-1, 1), \text{ asymptotically} \tag{3.6}$$

where $\Lambda(\ ;\ ;\ )$ defines the Wilks distribution with parameters $(\ ;\ ;\ )$. It has been shown (see e.g. Kshirsagar, 1972) that

$$W_c = \prod_{i=1}^{K-1}\left(1-r_i^2\right) \tag{3.7}$$

where r_i^2 is the square of the ith sample canonical correlation and the root of the determinantal equation

$$\left|-r^2\left(H_T^* + H_C^*\right) + H_C^*\right| = 0 \tag{3.8}$$

and r^2 is related to λ the root of the determinantal equation

$$\left|(n-k-1)H_T^{*-1}H_c^* - I_{K-1} - \lambda I_{K-1}\right| = 0 \tag{3.9}$$

by the relation

$$r^2 = (\lambda+1)/(n-K+\lambda) \tag{3.10}$$

Under the null hypothesis, $H_c^* : \underline{c}^* = \underline{0}$ and using (3.9), that is $\lambda = 0$, $r^2 = \dfrac{1}{n-K}$.

It has been shown, [16], that

$$-m\log W_c \approx \chi_{pq}^2$$

where $pq = K-1$, using the notation in this paper.

Thus,

$$-m\log W_c = -m\log\prod_{i=1}^{K-1}\left(1-r_i^2\right) = -m\sum_{i=1}^{K-1}\log\left(1-r_i^2\right) = -m(K-1)\log\left(1-r^2\right) \tag{3.11}$$

under H_0.

Asymptotically as $r^2 \to 0$,

$$\log\left(1-r^2\right) = -r^2 \text{ and } -m\log W_c = m(K-1)r^2 = \frac{m(K-1)}{n-K} \tag{3.12}$$

Hence

$$\frac{m(K-1)}{n-K} \cong \chi_{K-1}^2 \tag{3.13}$$

The best value of m for the expectation on both sides of (3.13) to be equal is $(n-K)$.

Therefore,

$$W_c = \frac{\left|H_T^*\right|}{\left|H_T^* + H_C^*\right|} \cong \chi_{K-1}^2 \tag{3.14}$$

and can provide a test criterion for the rejection or non-rejection of $H_0 : \underline{c} = \underline{0}$.

The test rejects the hull hypothesis if $W_c \geq \chi^2_{K-1}(\alpha)$ at an α-level of significance.

4. Illustrative Example

The application of the developed test makes use of data taken from [17] (see **Table 1**). The data represent the attitude of 333 undergraduate students of University of Nigeria towards taking up teaching as a profession after graduation. The students were sampled from three groups of faculties, F_1, F_2 and F_3. The responses Y (yes), N (no), U (undecided) indicates willing, not willing and undecided respectively.

The estimates of the parameters in (1.5) are:

$$\hat{\mu}' = (0.4444, 0.4655, 0.0901); \quad \hat{\tau}_1' = (0.07538, -0.0822, 0.006826); \quad \hat{\tau}_2' = (0.16144, 0.17604, -0.01462);$$

$$\hat{\beta}_1' = (-0.00966, -0.00170, 0.01136); \quad \hat{\beta}_2' = (0.00703, -0.00133, -0.00570);$$

$$\hat{\beta}_3' = (-0.03704, 0.01602, 0.02102); \quad \hat{c}' = (3.44857, -111.16428, -267.82464);$$

$$\hat{\lambda}_{11}' = (-0.00251, -0.01583, -0.02072); \quad \hat{\lambda}_{12}' = (0.00183, -0.01248, 0.01042);$$

$$\hat{\lambda}_{13}' = (-0.00963, 0.146101, -0.03836); \quad \hat{\lambda}_{21}' = (0.00538, 0.03389, 0.04446);$$

$$\hat{\lambda}_{22}' = (-0.00392, 0.02671, -0.02238); \quad \hat{\lambda}_{23}' = (0.02062, -0.31268, 0.08232)$$

The matrix of SS-SP due to interaction, H_c, is

$$H_c = \begin{bmatrix} 0.00815 & -0.08882 & 0.03952 \\ -0.08882 & 1.38108 & -0.34844 \\ 0.03952 & -0.3484 & 0.20826 \end{bmatrix}; \quad B = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix}$$

Therefore,

$$H_c^* = \begin{bmatrix} 0.13737 & 0.42836 \\ 0.42836 & 2.28622 \end{bmatrix}; \quad H_T^* = \begin{bmatrix} 136.18734 & -14.28710 \\ -14.28710 & 138.08711 \end{bmatrix};$$

$$H_\tau^* = \begin{bmatrix} 4.351971 & 2.172384 \\ 2.172384 & 5.652110 \end{bmatrix}; \quad H_\beta^* = \begin{bmatrix} 0.159903 & 0.039964 \\ 0.039964 & 0.016973 \end{bmatrix};$$

$$\left| H_T^* + H_c^* \right| = 18944.28795; \quad \left| H_T^* \right| = 18601.59497$$

Similarly,

$$\left| H_T^* + H_\tau^* \right| = 20054.2446$$

$$\left| H_T^* + H_\beta^* \right| = 18627.1301$$

These values are summarized in the **Table 2**.

Table 1. Attitude of university students towards the teaching profession.

Sex (i)	Faculty (j)											
	(F₁)				(F₂)				(F₃)			
	Response (k)			n_{i1}	Response (k)			n_{i2}	Response (k)			n_{i3}
	Y	N	U		Y	N	U		Y	N	U	
Male	23	19	5	47	88	58	16	162	7	10	1	18
Female	7	13	2	22	19	52	4	75	4	3	2	9
n.jk	30	32	7		107	110	20		11	13	3	
n.j		69				237				27		

Table 2. Multanova of categorical data for attitude of students towards teaching.

SV	d.f	SS-SP	Ratio of determinants	Hypothesis	$-m\log_e W_c$	Decision
Sex	1	H_τ^*	$\dfrac{\left\|H_T^*\right\|}{\left\|H_T^* + H_\tau^*\right\|} = 0.9115$	$H_0 : \underline{\tau}_i^* = \underline{0}$	30.579	H_0 is rejected
Faculty	2	H_β^*	$\dfrac{\left\|H_T^*\right\|}{\left\|H_T^* + H_\beta^*\right\|} = 0.999$	$H_0 : \underline{\beta}_j^* = \underline{0}$	0.330	H_0 is not rejected
Interaction (faculty × sex)	2	H_c^*	$\dfrac{\left\|H_T^*\right\|}{\left\|H_T^* + H_c^*\right\|} = 0.981$	$H_0 : \underline{c}^* = \underline{0}$	6.33	H_0 is rejected
Response	327	H_u^*	-			
Total	332	H_T^*	-			

$m = 330$.

The Pearson's chi-square for testing the hypothesis of no interaction (independence of the row and column for the kth response), $H_0 : \pi_{ijk} = \pi_{i.k}\pi_{.jk}$ gives the computed value of the test statistic as, $X^2 = 8.214$ based on 6 d.f while the likelihood ratio test statistic, G^2 for testing H_0 is calculated as $G^2 = 7.804$. Both test statistics are based on 6 degrees of freedom and show that interaction is not significant.

5. Conclusion

The results of the analysis show that while the effect of the sex and interaction are significant in the data, the effect of faculty is not significant. Thus, the proposed test for interaction based on the product effect model and based on 2 degrees of freedom $(K-1)$ can produce significant results even when one of the factors in the interaction is not significant. The test performs better than the traditional tests—the Pearson's chi square and the likelihood ratio tests, and could still out perform them in having greater power in larger 3-way contingency tables since it will have smaller degree of freedom. [9] has shown that the power of the non-central chi-square test at a given level of significance and non-centrality parameter increases as the degree of freedom decreases.

References

[1] Cheng, P.E., Liou, M. and Aston, J.A.D. (2009) Likelihood Ratio Tests with Three-Way Tables. Institute of Statistical Science, Academia Sinica 2CRiSM, Department of Statistics, The University of Warwick, UK. http://www3.stat.sinica.edu.tw/library/c_tec_rep/2007-3-20090206.pdf

[2] Agresti, A. (2002) Categorical Data Analysis. 2nd Edition, John Wiley and Sons Inc., Hoboken, 132. http://dx.doi.org/10.1002/0471249688

[3] Christensen, R. (1997) Loglinear Models and Logistic Regression. 2nd Edition, Springer-Verlag, New York Inc., New York.

[4] Grizzle, J.E., Starmer, F. and Koch, G.G. (1969) Analysis of Categorical Data by Linear Models. *Biometrics*, **25**, 489-504. http://dx.doi.org/10.2307/2528901

[5] Darroch, J.N. (1962) Interactions in Multi-Factor Contingency Tables. *Journal of the Royal Statistical Society*, **B24**, 251-263.

[6] Johnson, D.E. and Graybill, F.A. (1972) An Analysis of a Two-Way Model with Interaction and No Replication. *Journal of the American Statistical Association*, **67**, 862-868. http://dx.doi.org/10.1080/01621459.1972.10481307

[7] Tukey, T.W. (1949) One Degree of Freedom for Non-Additivity. *Biometrics*, **5**, 232-242. http://dx.doi.org/10.2307/3001938

[8] Cheng, P.E., Liou, M. and Aston, J.A.D. (2010) Likelihood Ratio Tests with Three-Way Tables. *Journal of the American Statistical Association*, **105**, 740-749. http://dx.doi.org/10.1198/jasa.2010.tm09061

[9] Ama, N.O. (1991) On Multiplicative Models for Interaction in Contingency Tables. Unpublished Ph.D. Thesis, University of Nigeria, Nsukka.

[10] Hélie, S. (2007) Understanding Statistical Power Using Noncentral Probability Distributions: Chi-Squared, G-Squared, and ANOVA. *Tutorials in Quantitative Methods for Psychology*, **3**, 63-69.

[11] Lancaster, H.O. (1969) The Chi-Squared Distribution. Wiley & Sons, Inc., New York.

[12] Darroch, J.N. (1974) Multiplicative and Additive Interaction in Contingency Tables. *Biometrika*, **61**, 207-214.
http://dx.doi.org/10.1093/biomet/61.1.207

[13] Graybill, F.A. (1961) An Introduction to Linear Statistical Models. McGraw-Hill, New York.

[14] Mardia, K.V., Kent, J.T. and Bibby, J.M. (1979) Multivariate Analysis. Academic Press, London.

[15] Kshirsagar, A.M. (1972) Multivariate Analysis. Marcel Dekker, Inc., New York.

[16] Bartlett, M.S. (1938) Further Aspects of the Theory of Multiple Regression. *Mathematical Proceedings of the Cambridge Philosophical Society*, **34**, 33-40.
http://dx.doi.org/10.1017/S0305004100019897

[17] Onukogu, I.B. (1985) An Analysis of Variance of Nominal Data. *Statistica*, anno, XLIV, No. 1, 87-96.

Scientific Research

Scientific Research Publishing (SCIRP) is one of the largest Open Access journal publishers. It is currently publishing more than 200 open access, online, peer-reviewed journals covering a wide range of academic disciplines. SCIRP serves the worldwide academic communities and contributes to the progress and application of science with its publication.

Other selected journals from SCIRP are listed as below. Submit your manuscript to us via either submit@scirp.org or Online Submission Portal.