ASYMPTOTIC POWER AND INDEPENDENCE OF CERTAIN CLASSES OF TESTS ON CATEGORICAL DATA

by

Earl Louis Diamond

University of North Carolina

This research was supported by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command, under Contract No. AF 18(600)-83. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Institute of Statistics
Mimeograph Series No. 196
April, 1958

ACKNOWLEDGMENT

I wish to express my deep gratitude to Professor S. N. Roy for his guidance and encouragement throughout the course of this investigation. I also wish to thank Dr. J. Ogawa for making several useful suggestions and for reviewing the manuscript.

Finally, I wish to thank Mrs. Marion Smith for her careful typing of the manuscript and Miss Marianne Byrd for preparing this material to be mimeographed.

TABLE OF CONTENTS

ASYMPTOTIC POWER AND INDEPENDENCE OF CERTAIN CLASSES OF TESTS ON CATEGORICAL DATA[1]

by

Earl Louis Diamond

INTRODUCTION

Although many papers on the analysis of categorical data have appeared since Karl Pearson [14][2] introduced the chi-square test criterion in 1900, only a few which have been published recently consider the approach that is followed here. Barnard [1] and E. S. Pearson [13] introduce this approach for the simple 2 x 2 table. Mitra [7], Roy and Mitra [18] and Ogawa [11] extend it to more general cases.

This approach has two primary features. The first is that one is careful to distinguish between the case in which the sample is assumed to be drawn from a single multinomial population and the case in which the sample is assumed to be drawn from several independent multinomial populations. In the first case all marginal totals of the resulting table of data are stochastic variates. In the second case some marginals are fixed by the manner in which the sample is drawn

---

1 This research was supported by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command, under Contract No. AF 18(600)-83. Reproduction in whole or in part is permitted for any purpose of the United States Government.

2 The numbers in square brackets refer to the bibliography listed at the end of the paper.

while the rest are stochastic variates. A third case in which all the marginal totals are fixed will not be considered here.

The second important feature is the careful definition of the hypothesis to be tested and its alternative. This is essential to the mathematical derivation of the test statistic and its null and non-null limiting distributions and is certainly necessary for any meaningful discussion of the results of a specific analysis of experimental data.

Two properties of the various types of tests which are appropriate to the two sampling situations described above will be considered in this report. These are (1) the asymptotic power of the test and (2) the necessary and sufficient conditions for the asymptotic independence in probability of two tests of a given type.

The mathematical methods used here are essentially extensions of those introduced by Cramer [5]. These methods have also been used by Mitra [7] and Ogawa [11] in extending Cramer's results on the existence of unique consistent estimates and limiting distributions in the null case. In addition, Mitra has derived the limiting non-null distribution for two of the cases considered here.

In relation to the theorems presented in this paper, these previous results may be summarized briefly as follows:

> Cramer [5] proves the existence of a unique consistent estimate and gives the limiting distribution in the null case of theorem 1 of this paper.
>
> Mitra [7] proves the existence of a unique consistent estimate for theorems 2 and 3 and gives the

limiting distribution in the null case for theorems 2, 3 and 6. He also derives the limiting power functions of theorems 1 and 3 and the independence condition of section 4.4.

Ogawa [11] repeats Cramer's theorem (theorem 1) in greater detail, proves the existence of a unique consistent estimate for theorems 2, 3 and 4 and finds the limiting null distribution for theorems 2, 3, 4, 6 and 8 and the special cases in sections 4.5 and 5.5.

Wald [20] derives the large sample distribution of the likelihood ratio under a hypothesis and alternative which are identical in form to those considered in this paper. Considering Cramer's statement ([5], p. 426) that the modified chi-square minimum method is identical with the maximum likelihood method, it seems that we might apply Wald's results to the chi-square tests considered here and thereby arrive at the same results.

Cochran [2] states that this can be done. However, Wald's results are derived under very strong assumptions. These assumptions are difficult to verify under the theorems as stated in this paper and some points of Wald's very difficult proof are not clear.

Cochran [2] describes the asymptotic power function of the simple goodness of fit test, which is not considered here. Mitra [7] derives the asymptotic power functions of the tests given in theorems 1 and 3 of this paper. Both follow Pitman's [15] definition of limiting power for a consistent test and this definition will be followed here.

The notion of asymptotic independence in probability of two chi-square tests is introduced by Mitra [7] and he derived a necessary and sufficient condition for the asymptotic independence in probability of two tests of the type considered in Chapter IV of this paper.

For each of the tests considered here it has been shown that under the hypothesis the test statistic asymptotically follows the chi-square distribution. Here it is shown that under the alternative each test statistic asymptotically follows the non-central chi-square distribution with a certain non-centrality parameter. For completeness each theorem is stated so as to include the previous results and it will be shown that these results on the null distributions are special cases of the results derived here for the non-null distributions. In addition, a necessary and sufficient condition for the asymptotic independence in probability of two tests under the respective null hypotheses is derived for each type of test under consideration.

In Chapter I a single-multinomial sample is defined as one having a distribution of the form $n! \prod_{j=1}^{s} p_j^{n_j} / \prod_{j=1}^{s} n_j!$, where $\sum_{j=1}^{s} p_j = 1$ and $\sum_{j=1}^{s} n_j = n$, and a product-multinomial sample is defined as one having a distribution of the form $\prod_{i=1}^{r} [n_{io}! \prod_{j=1}^{s} p_{ij}^{n_{ij}} / \prod_{j=1}^{s} n_{ij}!]$, where $\sum_{j=1}^{s} p_{ij} = 1$ and $\sum_{j=1}^{s} n_{ij} = n_{io}$ for $i = 1, \ldots, r$. The types of tests to be considered are described and definitions of asymptotic power and asymptotic independence are given. A method of finding confidence limits for the non-centrality parameter of a non-central chi-square distribution is derived.

In Chapter II, given a single-multinomial sample, we consider the test of a hypothesis of the form H_o: $p_j = p_j(\theta_1,\ldots,\theta_t)$ for $j = 1,\ldots,s$ $(t \leq s-1)$ subject to $f_m(\theta_1,\ldots,\theta_t) = 0$ for $m = 1,\ldots,u < t$ where $p_j(\theta_1,\ldots,\theta_t)$ and $f_m(\theta_1,\ldots,\theta_t)$ are known functions of the unknown parameters $\theta_1,\ldots,\theta_t$ defined over the whole space of the θ_k and have certain properties (described in Chapter I) which are supposed to hold over a non-degenerate open interval $\underline{(\overline{\ })}$ in the t-dimensional space of the θ_k. We suppose that the alternative is of the form H_n^*: $p_{jn} = p_j(\theta_1,\ldots,\theta_t) + d_j/\sqrt{n}$ subject to $f_m(\theta_1,\ldots,\theta_t) = 0$ where $\sum_{j=1}^{s} d_j = 0$ but not all $d_j = 0$, and where $p_{jn} \neq p_j(\theta_1,\ldots,\theta_t)$ subject to $f_m(\theta_1,\ldots,\theta_t) = 0$ for any $(\theta_1,\ldots,\theta_t)$ in $\underline{(\overline{\ })}$. In order to obtain the asymptotic power it is necessary to consider a simple alternative of the form H_n: $p_{jn} = p_j(\theta_1^o,\ldots,\theta_t^o) + d_j/\sqrt{n}$ subject to $f_m(\theta_1^o,\ldots,\theta_t^o) = 0$ where $(\theta_1^o,\ldots,\theta_t^o)$ is supposed to be the unknown true parameter point in $\underline{(\overline{\ })}$ under H_o.

It is shown that, under H_n, the test statistic follows, in the limit, the non-central chi-square distribution with s-1-t+u degrees of freedom and a certain non-centrality parameter. A necessary and sufficient condition for the asymptotic independence in probability of two such tests under the respective null hypotheses is derived. Several specific examples of such tests are given.

In Chapter III, given a product-multinomial sample, we consider the test of a hypothesis of the form H_o: $p_{ij} = p_{ij}(\theta_1,\ldots,\theta_t)$ for $i = 1,\ldots,r$ and $j = 1,\ldots,s$ $(t \leq rs-r)$ subject to $f_m(\theta_1,\ldots,\theta_t) = 0$ for $m = 1,\ldots,u < t$ where $p_{ij}(\theta_1,\ldots,\theta_t)$ and $f_m(\theta_1,\ldots,\theta_t)$ are known functions of the unknown parameters $\theta_1,\ldots,\theta_t$ defined over the whole space of the θ_k and have certain properties (described in Chapter I) which are supposed to hold over a non-degenerate open interval $\underline{(\bar{\ })}$ in the t-dimensional space of the θ_k. We suppose that the alternative is of the form H_n^*: $p_{ijn} = p_{ij}(\theta_1,\ldots,\theta_t) + d_{ij}/\sqrt{n}$ subject to $f_m(\theta_1,\ldots,\theta_t) = 0$ where $\sum_{j=1}^{s} d_{ij} = 0$ for $i = 1,\ldots,r$ but not all $d_{ij} = 0$, and where $p_{ijn} \neq p_{ij}(\theta_1,\ldots,\theta_t)$ subject to $f_m(\theta_1,\ldots,\theta_t) = 0$ for any $(\theta_1,\ldots,\theta_t)$ in $\underline{(\bar{\ })}$. In order to obtain the asymptotic power it is necessary to consider a simple alternative of the form H_n: $p_{ijn} = p_{ij}(\theta_1^o,\ldots,\theta_t^o) + d_{ij}/\sqrt{n}$ subject to $f_m(\theta_1^o,\ldots,\theta_t^o) = 0$ where $(\theta_1^o,\ldots,\theta_t^o)$ is supposed to be the unknown true parameter point in $\underline{(\bar{\ })}$ under H_o.

It is shown that, under H_n, the test statistic follows, in the limit, the non-central chi-square distribution with rs-r-t+u degrees of freedom and a certain non-centrality parameter. A necessary and sufficient condition for the asymptotic independence in probability of two such tests under the respective null hypotheses is derived. Several specific examples are given.

In Chapter IV we are given a single-multinomial sample and that $p_j = p_j(\theta_1,\ldots,\theta_t)$ for $j = 1,\ldots,s$ $(t \leq s-1)$ where $p_j(\theta_1,\ldots,\theta_t)$ are known functions of the unknown parameters $\theta_1,\ldots,\theta_t$ defined over the whole space of the θ_k and have certain properties (described in Chapter I) which are supposed to hold over a non-degenerate open interval $\underline{(\bar{\ })}$ in the t-dimensional space of the θ_k. We consider the test of a hypothesis of the form H_o: $f_m(\theta_1,\ldots,\theta_t) = 0$ for $m = 1,\ldots,u < t$ where

$f_m(\theta_1,\ldots,\theta_t)$ are known functions of the unknown parameters $\theta_1,\ldots,\theta_t$ defined over the whole space of the θ_k and have certain properties (described in Chapter I) which are supposed to hold over Ω. We suppose that the alternative is of the form H_n^*: $f_m(\theta_1,\ldots,\theta_t) = c_m / \sqrt{n}$ where not all $c_m = 0$. In order to obtain the asymptotic power it is necessary to consider a simple alternative of the form H_n: $f_m(\theta_1^o,\ldots,\theta_t^o) = c_m / \sqrt{n}$ where $(\theta_1^o,\ldots,\theta_t^o)$ is supposed to be the unknown true parameter point in Ω.

It is shown that, under H_n, the test statistic follows, in the limit, the non-central chi-square distribution with u degrees of freedom and a certain non-centrality parameter. A necessary and sufficient condition for the asymptotic independence in probability of two such tests under the respective null hypotheses is derived. It is shown that the test of a hypothesis of the form H_o: $f_m(p_1,\ldots,p_s) = 0$, $m = 1,\ldots,u \leq s-1$, is a special case and a specific example is given.

In Chapter V we are given a product-multinomial sample and that $p_{ij} = p_{ij}(\theta_1,\ldots,\theta_t)$ for $i = 1,\ldots,r$ and $j = 1,\ldots,s$ $(t \leq rs-r)$ where $p_{ij}(\theta_1,\ldots,\theta_t)$ are known functions of the unknown parameters $\theta_1,\ldots,\theta_t$ defined over the whole space of the θ_k and have certain properties (described in Chapter I) which are supposed to hold over a non-degenerate open interval Ω in the t-dimensional space of the θ_k. We consider the test of a hypothesis of the form H_o: $f_m(\theta_1,\ldots,\theta_t) = 0$ for $m = 1,\ldots,u<t$ where $f_m(\theta_1,\ldots,\theta_t)$ are known functions of the unknown parameters $\theta_1,\ldots,\theta_t$ defined over the whole space of the θ_k and have certain properties (described in Chapter I) which are supposed to hold over Ω. We

suppose that the alternative is of the form H_n^*: $f_m(\theta_1,\ldots,\theta_t) = c_m/\sqrt{n}$ where not all $c_m = 0$. In order to obtain the asymptotic power it is necessary to consider a simple alternative of the form H_n: $f_m(\theta_1^0,\ldots,\theta_t^0) = c_m/\sqrt{n}$ where $(\theta_1^0,\ldots,\theta_t^0)$ is supposed to be the unknown true parameter point in Ω.

It is shown that under H_n the test statistic follows, in the limit, the non-central chi-square distribution with u degrees of freedom and a certain non-centrality parameter. A necessary and sufficient condition for the asymptotic independence in probability of two such tests under the respective null hypotheses is derived. It is shown that the test of a hypothesis of the form H_o: $f_m(p_{11},\ldots,p_{1s},\ldots,p_{r1},\ldots,p_{rs}) = 0$, $m = 1,\ldots,u \leq rs-r$ is a special case and two specific examples are given.

The bibliography at the end of the paper does not include all papers on the analysis of categorical data but only those which are referred to in the text or which were used by the author in the preparation of this report. In this sense the bibliography is incomplete and should be supplemented by those given at the end of Cochran's papers ([2], [3]).

NOTATION

As far as possible the following notation and convention will be used, all departures being clearly indicated at the proper places.

Capital letters will denote matrices, small letters underscored will denote column vectors and row vectors if they are primed. The transpose of a matrix A is denoted by a prime, A'. To indicate its structure a matrix M with p rows and q columns will sometimes be written as M (pxq) or $\underset{pxq}{M}$. Matrices formed by adjoining 2 or more matrices will be denoted (M : P) and dimensions will sometimes be indicated as r $(\underset{s}{M} : \underset{t}{P})$.

d.f. denotes degrees of freedom.

$N(\mu, \sigma)$ denotes a random variable having the normal distribution with mean μ and standard deviation σ.

$N(\underline{\mu}, \Sigma)$ denotes a random variable having the multivariate normal distribution with mean vector $\underline{\mu}$ and variance-covariance matrix Σ.

Where there are multiple subscripts as in p_{ijk}, a zero in place of a subscript indicates the result of summation over that subscript. Thus $p_{ijo} = \sum_{k=1}^{t} p_{ijk}$ or $n_{io} = \sum_{j=1}^{s} n_{ij}$. However, $n_{oo} = n$.

χ^2 is used to denote the test statistic which is calculated from the data when a chi-square test is to be performed.

$\xrightarrow[p]{}$ denotes convergence in probability.

Chapter I

BASIC PROBLEMS, PRINCIPLES, AND DEFINITIONS

1.1 Sampling Schemes

In the statistical analysis of categorical data the probability model upon which the analysis is based is entirely dependent upon the manner in which the sample is drawn. Roy and Mitra [18] state:

> "It is clear that a p-variate body of data arranged in a q-way classification will formally look like a (p+q)-dimensional conteingency table, but a distinction can be made between a 'variate' and a 'way of classification,' in that along the direction of a 'variate' the marginal frequencies are supposed to be stochastic variates while along a 'way of classification' the marginal frequencies are supposed to be fixed."

Consider the case in which the sample is drawn with replacement or the sample is large, but small in relation to the size of the population.

A single-multinomial sample may be formally described as follows:

(1.1.1) Let the possible results of a certain random experiment E be divided into s mutually exclusive groups and suppose that the probability of obtaining a result belonging to the j-th group is $p_j \geq c^2 > 0$. Let n_j denote the number of results belonging to the j-th group which occur in a sequence of n repetitions of E, so that $\sum_{j=1}^{s} n_j = n$. That is, a sample with a distribution of the form

$$\frac{n!}{\prod_{j=1}^{s} n_j!} \prod_{j=1}^{s} p_j^{n_j} ,$$

where $\sum_{j=1}^{s} p_j = 1$ and $\sum_{j=1}^{s} n_j = n$.

In this case we shall say that 'j' is a variate and if 'j' itself is a multiple subscript, say $j_1\, j_2 \ldots j_k$, we shall say that 'j' is multivariate.

A <u>product-multinomial</u> sample may be formally described as follows:

(1.1.2) Suppose that there are r mutually independent random experiments $E_1, \ldots, E_r$ and the possible outcomes in each experiment are divided into s mutually exclusive groups. Let the probability of obtaining a result belonging to the j-th group of E_i be $p_{ij} \geq c^2 > 0$. Let n_{ij} denote the number of results belonging to the j-th group in n_{io} repetitions of E_i so that $\sum_{j=1}^{s} n_{ij} = n_{io}$, and let $\sum_{i=1}^{r} n_{io} = n$. That is, a sample with a distribution of the form

$$\prod_{i=1}^{r} \left[\frac{n_{io}!}{\prod_{j=1}^{s} n_{ij}!} \prod_{j=1}^{s} p_{ij}^{\;n_{ij}} \right] ,$$

where $\sum_{j=1}^{s} p_{ij} = 1$ and $\sum_{j=1}^{s} n_{ij} = n_{io}$ for $i = 1,\ldots,r$. In this case we shall say that 'j' is a variate (single or multi-, according as 'j' is a single or a multiple subscript) and 'i' is a way of classification (single or multi-, according as 'i' is a single or multiple subscript).

Although the single-multinomial sample is a special case of the product-multinomial sample (i.e., when $r = 1$), for the purposes of this paper it is convenient to define the two separately. As given here the single-multinomial seems to refer only to one-way tables and the product-multinomial to two-way tables. However, there is nothing to prevent the subscript i or the subscript j, or both, from representing multiple subscripts. Indeed, the product-multinomial sample can be expressed in complete generality as follows:

Let $i = i_1 i_2 \ldots i_u$ and $j = j_1 j_2 \ldots j_v$, where $i_1 \in (r_1)_{i_2 i_3 \ldots i_u}$ (a subset of 1, 2, ..., r_1 depending on the subscript set $i_2 i_3 \ldots i_u$), $i_2 \in (r_2)_{i_3 \ldots i_u}$, ..., $i_{u-1} \in (r_{u-1})_{i_u}$ and $i_u = 1,2,\ldots,r_u$, and where $j_1 = 1,2,\ldots,s_1,\ldots,j_v = 1,2,\ldots,s_v$.

With this sampling scheme each independent multinomial sample, denoted by the subscripts $i_1 i_2 \ldots i_u$, could be an element of a complicated analysis of variance design (e.g., factorial or partially balanced incomplete block) and there would be $n_{i_1 \ldots i_u o \ldots o}$ observations for the $(i_1 \ldots i_u)$-th element of the design. Each observation would be classified into one of the $\prod_{k=1}^{v} s_k$ cells of a v-way table, and the data from the entire experiment could be presented in a (u+v)-way table with 'holes' in it, the number and location of these 'holes' depending upon the experimental design.

In order to avoid undue complication of the notation only the simpler descriptions, (1.1.1) and (1.1.2), will be considered. Subsequent theorems can easily be extended to the more complicated case by use of multiple subscripts as above.

If the sampling is without replacement from a finite population we have a hypergeometric or product of hypergeometric distributions according as all dimensions represent 'variates' or some represent 'variates' and some represent 'ways of classification'. However, in order to develop tests of hypotheses, it is necessary that we restrict ourselves to the multinomial distribution, and this may be regarded as an approximation to the hypergeometric distribution.

1.2 Asymptotic Power

Each of the tests presented in the following chapters is a chi-square test, and since this is a consistent test it would not be fruitful to find the power in the usual sense in that it would not provide a comparison between the powers of two rival tests. However, following the definition given by Pitman [15], it is possible to find the asymptotic power. This definition requires that the alternative depend upon the sample size in the sense that as the sample size increases the alternative gets closer and closer to the hypothesis. It will be shown that this type of alternative provides us with a comparison between the large sample powers of different rival consistent tests. It will be shown that the asymptotic power in each case follows the non-central chi-square distribution with a certain non-centrality parameter which depends upon the sampling scheme, the hypothesis and the alternative.

1.3 Hypotheses and Alternatives for Single-Multinomial Samples

In order to define appropriate hypotheses and alternatives for single-multinomial samples the following definitions are needed:

(1.3.1) Suppose we are given s functions $p_1(\theta_1,\ldots,\theta_t)$, $p_2(\theta_1,\ldots,\theta_t)$, $\ldots$, $p_s(\theta_1,\ldots,\theta_t)$, of $t \leq s-1$ unknown parameters $\theta_1,\ldots,\theta_t$ such that, for all points of a non-degenerate open interval $\underline{(\bar{\ })}$ in the t-dimensional space of the θ_k, the $p_j(\underline{\theta})$ where $\underline{\theta}' = (\theta_1,\ldots,\theta_t)$ satisfy the following conditions:

(a) $\sum_{j=1}^{s} p_j(\underline{\theta}) = 1$.

(b) $p_j(\underline{\theta}) \geq c^2 > 0$ for all j.

(c) Every $p_j(\underline{\theta})$ has continuous first and second order derivates with respect to the θ_k.

(d) The matrix $\left(\dfrac{\partial p_j}{\partial \theta_k}\right)$ where $j = 1,\ldots,s$ and $k = 1,\ldots,t$ is of rank t.

(1.3.2) Suppose we are given $u < t$ real-valued functions, $f_1(\underline{\theta}),\ldots,f_u(\theta)$, of $\theta_1,\ldots,\theta_t$ such that, for every point of $\underline{(\bar{\ })}$,

(e) every $f_m(\underline{\theta})$ has continuous first and second order derivatives with respect to the θ_k, and

(f) the matrix $\left(\dfrac{\partial f_m}{\partial \theta_k}\right)$ where $m = 1,\ldots,u$ and $k = 1,\ldots,t$ is of rank u.

(1.3.3) Let $\underline{\theta}_o$, where $\underline{\theta}_o' = (\theta_1^o,\ldots,\theta_t^o)$, be a point of $\underline{(\bar{\ })}$.

For single-multinomial samples the following types of hypotheses and alternatives will be considered:

(1.3.4) H_o: $p_j = p_j(\underline{\Theta})$ for $j = 1,\ldots,s$.

H_n^*: $p_{jn} = p_j(\underline{\Theta}) + d_j / \sqrt{n}$ where $\sum_{j=1}^{s} d_j = 0$ but not all $d_j = 0$ and where $p_{jn} \neq p_j(\underline{\Theta})$ for any $\underline{\Theta}$ in $\underline{\bigcap}$.

In order to obtain the asymptotic power it is necessary to consider the simple alternative

H_n: $p_{jn} = p_j(\underline{\Theta}_o) + d_j / \sqrt{n}$ where $\underline{\Theta}_o$ is supposed to be the unknown true parameter point in $\underline{\bigcap}$ under H_o.

(1.3.5) H_o: $p_j = p_j(\underline{\Theta})$ for $j = 1,\ldots,s$ subject to $f_m(\underline{\Theta}) = 0$ for $m = 1,\ldots,u$.

H_n^*: $p_{jn} = p_j(\underline{\Theta}) + d_j / \sqrt{n}$ subject to $f_m(\Theta) = 0$ where $\sum_{j=1}^{s} d_j = 0$ but not all $d_j = 0$ and where $p_{jn} \neq p_j(\underline{\Theta})$ for any $\underline{\Theta}$ in $\underline{\bigcap}$.

In order to obtain the asymptotic power it is necessary to consider the simple alternative

H_n: $p_{jn} = p_j(\underline{\Theta}_o) + d_j / \sqrt{n}$ subject to $f_m(\underline{\Theta}_o) = 0$ where $\underline{\Theta}_o$ is supposed to be the unknown true parameter point in $\underline{\bigcap}$ under H_o.

(1.3.6) It is given that $p_j = p_j(\underline{\Theta})$ for $j = 1,\ldots,s$.

H_o: $f_m(\underline{\Theta}) = 0$ for $m = 1,\ldots,u$.

H_n^*: $f_m(\underline{\Theta}) = c_m / \sqrt{n}$ where not all $c_m = 0$.

In order to obtain the asymptotic power it is necessary to consider the simple alternative

H_n: $f_m(\underline{\Theta}_o) = c_m / \sqrt{n}$ where $\underline{\Theta}_o$ is supposed to be the unknown true parameter point in $\underline{\bigcap}$.

1.4 Hypotheses and Alternatives for Product-Multinomial Samples

In order to define appropriate hypotheses and alternatives for product-multinomial samples the following definitions are needed:

(1.4.1) Suppose we are given rs functions $p_{ij}(\theta_1,\ldots,\theta_t)$ where $i = 1,\ldots,r$ and $j = 1,\ldots,s$, of $t \leq rs-r$ unknown parameters $\theta_1,\ldots,\theta_t$ such that, for all points of a non-degenerate open interval $\underline{(\ )}$ in the t-dimensional space of the θ_k, the $p_{ij}(\underline{\theta})$ where $\underline{\theta}' = (\theta_1,\ldots,\theta_t)$ satisfy the following conditions:

(a) $\sum_{j=1}^{s} p_{ij}(\underline{\theta}) = 1$ for $i = 1,\ldots,r$.

(b) $p_{ij}(\underline{\theta}) \geq c^2 > 0$ for all i and j.

(c) Every $p_{ij}(\underline{\theta})$ has continuous first and second order derivatives with respect to the θ_k.

(d) The matrix $\left(\frac{\partial p_{ij}}{\partial \theta_k}\right)$, where $i = 1,\ldots,r$, $j = 1,\ldots,s$ and $k = 1,\ldots,t$, is of rank t.

Note that the p_{ij} are to be arranged lexicographically (i.e., $p_{11}, p_{12}, \ldots, p_{1s}; p_{21}, p_{22}, \ldots, p_{2s}; \ldots ; p_{r1}, p_{r2}, \ldots, p_{rs}$) and the matrix $\left(\frac{\partial p_{ij}}{\partial \theta_k}\right)$ has rs rows and t columns.

(1.4.2) Suppose we are given $u < t$ functions $f_1(\underline{\theta}),\ldots,f_u(\underline{\theta})$ as in (1.3.2).

(1.4.3) Let $\underline{\theta}_o$ be as in (1.3.3).

For product-multinomial samples the following types of hypotheses and alternatives will be considered:

(1.4.4) H_o: $p_{ij} = p_{ij}(\underline{\theta})$ for $i = 1,\dots,r$ and $j = 1,\dots,s$.

H_n^*: $p_{ijn} = p_{ij}(\underline{\theta}) + d_{ij}/\sqrt{n}$ where $\sum_{j=1}^{s} d_{ij} = 0$ for $i = 1,\dots,r$ but not all $d_{ij} = 0$ and where $p_{ijn} \neq p_{ij}(\underline{\theta})$ for any $\underline{\theta}$ in $\underline{\Omega}$.

In order to obtain the asymptotic power it is necessary to consider the simple alternative

H_n: $p_{ijn} = p_{ij}(\underline{\theta}_o) + d_{ij}/\sqrt{n}$ where $\underline{\theta}_o$ is supposed to be the unknown true parameter point in $\underline{\Omega}$ under H_o.

(1.4.5) H_o: $p_{ij} = p_{ij}(\underline{\theta})$ for $i = 1,\dots,r$ and $j = 1,\dots,s$ subject to $f_m(\underline{\theta}) = 0$ for $m = 1,\dots,u$.

H_n^*: $p_{ijn} = p_{ij}(\underline{\theta}) + d_{ij}/\sqrt{n}$ subject to $f_m(\underline{\theta}) = 0$ where $\sum_{j=1}^{s} d_{ij} = 0$ for $i = 1,\dots,r$ but not all $d_{ij} = 0$ and where $p_{ijn} \neq p_{ij}(\underline{\theta})$ for any $\underline{\theta}$ in $\underline{\Omega}$.

In order to obtain the asymptotic power it is necessary to consider the simple alternative

H_n: $p_{ijn} = p_{ij}(\underline{\theta}_o) + d_j/\sqrt{n}$ subject to $f_m(\underline{\theta}_o) = 0$ where $\underline{\theta}_o$ is supposed to be the unknown true parameter point in $\underline{\Omega}$ under H_o.

(1.4.6) It is given that $p_{ij} = p_{ij}(\underline{\theta})$ for $i = 1,\dots,r$ and $j = 1,\dots,s$.

H_o: $f_m(\underline{\theta}) = 0$ for $m = 1,\dots,u$.

H_n^*: $f_m(\underline{\theta}) = c_m/\sqrt{n}$ where not all $c_m = 0$.

In order to obtain the asymptotic power it is necessary to consider the simple alternative

H_n: $f_m(\underset{\sim}{\theta}_o) = c_m / \sqrt{n}$ where $\underset{\sim}{\theta}_o$ is supposed to be the unknown true parameter point in $\underset{\sim}{\Omega}$.

1.5 Asymptotic Independence

In some situations it may be of interest to determine whether or not two tests are independent. A particular hypothesis H_o with an associated X^2 may be the intersection of several hypotheses H_{01}, H_{02}, ..., with associated X_1^2, X_2^2, ..., and it may happen that $X^2 = X_1^2 + X_2^2 + \ldots$, and that X_1^2, X_2^2, ..., are independently distributed. However, the additivity is not in the usual algebraic sense but is only in probability and asymptotically as $n \longrightarrow \infty$, and the independence is also in the asymptotic sense (Mitra [7]). A necessary and sufficient condition for such independence and additivity will be developed subsequently and for that purpose the following theorem due to Ogawa [9] will be referred to:

(1.5.1) If Q_1 and Q_2 are the symmetric matrices of two homogeneous quadratic forms in m variates which are jointly distributed as $N(\underset{\sim}{\mu}, \Sigma)$, a necessary and sufficient condition for the independence in probability of the two forms is that

$$Q_1 \Sigma Q_2 = 0.$$

1.6 Limiting Processes

In the chapters that follow a limiting process will be used which, for single-multinomial samples, may be stated as "in the limit as $n \longrightarrow \infty$" and for product-multinomial samples may be stated as "in the limit as $n \longrightarrow \infty$ subject to $\frac{n_{io}}{n}$ being held fixed for $i = 1,\ldots,r$." For convenience this process will be denoted by the phrase "in the limit" in both cases.

1.7 Degrees of Freedom

In the theorems that follow, the number of degrees of freedom associated with each type of test becomes readily apparent. However, in a given practical situation, there is a simpler method of determining the number of degrees of freedom.

If we define the model as the sampling scheme and any additional restrictions on the p_j's or p_{ij}'s before the test, it is seen that the number of degrees of freedom equal the number of independent parameters under the model less the number of independent parameters under the hypothesis. Thus, the number of degrees of freedom might also be considered to be the same as the number of independent parameters under the alternative.

For the types of hypotheses and alternatives given in sections 1.3 and 1.4 we may construct the following table:

Hypothesis and Alternative	Number of Independent Parameters		
	Under Model	Under H_o	Under H_n (d.f.)
1.3.4	s-1	t	s-1-t
1.3.5	s-1	t-u	s-1-t+u
1.3.6	t	t-u	u
1.4.4	rs-r	t	rs-r-t
1.4.5	rs-r	t-u	rs-r-t+u
1.4.6	t	t-u	u

1.8 Confidence Limits for Non-Centrality Parameters

In this thesis we are concerned with test statistics, χ_o^2 say, which have the property $\chi_o^2 - \underline{w}'\underline{w} \xrightarrow[p]{} 0$ where $\underline{w}' = (w_1, \ldots, w_m)$ with m being equal to the number of degrees of freedom associated with the test statistic and where the w_i are independent N(0, 1) under the hypothesis H_o and are independent $N(d_i, 1)$ under the alternative H_n. Then χ_o^2 is asymptotically distributed under H_o as a chi-square variate with m d.f. and under H_n as a non-central chi-square variate with m d.f. and a non-centrality parameter $\Delta = \underline{d}'\underline{d}$ where $\underline{d}' = (d_1, \ldots, d_m)$.

Let χ_α^2 be the tabulated $100(1-\alpha)\%$ point of the chi-square distribution with m d.f. Then

$$\Pr[\underline{w}'\underline{w} \leq \chi_\alpha^2 \mid H_o] = 1 - \alpha \quad \text{and}$$

$$\Pr[(\underline{w}' - \underline{d}')(\underline{w} - \underline{d}) \leq \chi_\alpha^2 \mid H_n] = 1 - \alpha.$$

By a theorem due to Roy [16],

$$\Pr[-\chi_\alpha \leq \underline{a}'(\underline{w} - \underline{d}) \leq \chi_\alpha \mid H_n] = 1 - \alpha$$

for all $\underline{a}$ subject to $\underline{a}'\underline{a} = 1$ where $\underline{a}' = (a_1, \ldots, a_m)$. Therefore,

$$\underline{a}'\underline{w} - \chi_\alpha \leq \underline{a}'\underline{d} \leq \underline{a}'\underline{w} + \chi_\alpha \text{ with Prob. } 1 - \alpha.$$

Considering the inequality on the right, if $\underline{a}'\underline{d}$ is maximized with respect to $\underline{a}$, it is seen that $[\underline{d}'\underline{d}]^{1/2} \leq \underline{a}_1'\underline{w} + \chi_\alpha$ where $\underline{a}_1$ is the value of $\underline{a}$ that maximizes $\underline{a}'\underline{d}$. By Schwarz's inequality $\underline{a}_1'\underline{w} \leq [\underline{a}_1'\underline{a}_1]^{1/2} [\underline{w}'\underline{w}]^{1/2}$ and since $\underline{a}_1'\underline{a}_1 = 1$,

$$[\underline{d}'\underline{d}]^{1/2} \leq [\underline{w}'\underline{w}]^{1/2} + \chi_\alpha.$$

In the inequality on the left, maximizing $\underline{a}'\underline{w}$ with respect to $\underline{a}$, we get $[\underline{w}'\underline{w}]^{1/2} - \chi_\alpha \leq \underline{a}_2'\underline{d}$ where $\underline{a}_2$ is the value of $\underline{a}$ that maximizes $\underline{a}'\underline{w}$. By Schwarz's inequality

$$\underline{a}_2'\underline{d} \leq [\underline{a}_2'\underline{a}_2]^{1/2} [\underline{d}'\underline{d}]^{1/2} \text{ and since } \underline{a}_2'\underline{a}_2 = 1,$$

$$[\underline{w}'\underline{w}]^{1/2} - \chi_\alpha \leq [\underline{d}'\underline{d}]^{1/2}.$$

Thus, with probability $\geq 1 - \alpha$

$$[\underline{w}'\underline{w}]^{1/2} - \chi_\alpha \leq [\underline{d}'\underline{d}]^{1/2} \leq [\underline{w}'\underline{w}]^{1/2} + \chi_\alpha.$$

However, $\underline{d}'\underline{d} = \Delta$ and $\chi_o^2 - \underline{w}'\underline{w} \xrightarrow[p]{} 0$. Therefore, in the limit, with probability $\geq 1 - \alpha$,

$$\chi_o - \chi_\alpha \leq \Delta^{1/2} \leq \chi_o + \chi_\alpha.$$

Chapter II

CONTINGENCY-LIKE TESTS FOR SINGLE-MULTINOMIAL SAMPLES

2.1 Introduction

For single-multinomial samples, tests of hypotheses of types (1.3.4) and (1.3.5) are referred to as 'contingency-like' tests since they include all the usual so-called contingency tests on samples in which all the marginal totals are stochastic variables. However, many more tests of these types may be constructed and the theorems presented in this chapter for such tests give a test statistic and its null and non-null distributions.

Two theorems are presented. The first is preliminary in nature, in that the type of test it covers would seldom if ever be used in practice, but is necessary for the development of future theorems. The second theorem covers the contingency-like tests for single-multinomial samples that are most likely to arise in practice. In addition a necessary and sufficient condition is developed for the asymptotic independence in probability of two tests of the type covered by theorem 2. Finally, several examples of specific tests are given.

2.2 Theorem 1

Given a single-multinomial sample (1.1.1) and s functions $p_j(\underline{\theta})$ of t unknown parameters $\theta_1,\ldots,\theta_t$ as in (1.3.1), consider the hypothesis

$$H_o\colon p_j = p_j(\underline{\theta}) \text{ for } j = 1,\ldots,s$$

and the alternative

H_n^*: $p_{jn} = p_j(\underline{\theta}) + d_j / \sqrt{n}$ where $\sum_{j=1}^{s} d_j = 0$ but not all $d_j = 0$ and where $p_{jn} \neq p_j(\underline{\theta})$ for any $\underline{\theta}$ in $\underline{\Gamma}$.

In order to obtain the asymptotic power we consider the simple alternative

H_n: $p_{jn} = p_j(\underline{\theta}_o) + d_j / \sqrt{n}$ where $\underline{\theta}_o$ is supposed to be the unknown true parameter point in $\underline{\Gamma}$ under H_o. (Notice however that H_o itself does not specify $\underline{\theta}_o$.)

Then (1) the equations

$$\sum_{j=1}^{s} \frac{n_j - np_j(\underline{\theta})}{p_j(\underline{\theta})} \cdot \frac{\partial\, p_j(\underline{\theta})}{\partial\, \theta_k} = 0, \; k = 1, \ldots, t, \tag{2.2.1}$$

for minimizing chi-square in the modified sense (Cramer [5]) have exactly one system of solutions $\hat{\underline{\theta}}' = (\hat{\theta}_1, \ldots, \hat{\theta}_t)$ such that $\hat{\underline{\theta}} \longrightarrow \underline{\theta}_o$ in probability as $n \longrightarrow \infty$, where $\underline{\theta}_o$ is the unknown true parameter point in $\underline{\Gamma}$ under H_o and is the point in terms of which the convergence in probability is set up, and (2) the statistic

$$\chi^2 = \sum_{j=1}^{s} \frac{(n_j - np_j(\hat{\underline{\theta}}))^2}{np_j(\hat{\underline{\theta}})} \tag{2.2.2}$$

is, in the limit, distributed under H_o as a central chi-square variate with s-1-t d.f. and under H_n as a non-central chi-square variate with s-1-t d.f. and a non-centrality parameter

$$\Delta = \underline{d}' [I - A(A'A)^{-1} A'] \underline{d} , \tag{2.2.3}$$

where
$$\underset{s\times 1}{\underline{d}'} = \left(\frac{d_1}{\sqrt{p_1^o}}, \ldots, \frac{d_s}{\sqrt{p_s^o}} \right),$$

$$p_j^o = p_j(\underline{\theta}_o), \quad \text{and}$$

$$\underset{s\times t}{A} = \left[\frac{1}{\sqrt{p_j^o}} \left(\frac{\partial p_j}{\partial \theta_k} \right)_o \right]$$

where $\left(\frac{\partial p_j}{\partial \theta_k} \right)_o$ is $\left(\frac{\partial p_j}{\partial \theta_k} \right)$ evaluated at the point $\underline{\theta}_o$.

<u>Proof</u>

Part (1) has been proved by Cramer [5]. Part (2) has been proved by Mitra [7] but the proof is repeated here for completeness.

Let
$$y_j = \frac{n_j - np_j(\hat{\underline{\theta}})}{\sqrt{np_j(\hat{\underline{\theta}})}}, \quad j = 1, \ldots, s,$$

$$\underline{y}' = (y_1, \ldots, y_s),$$

$$x_j = \frac{n_j - np_j^o}{\sqrt{np_j^o}}, \quad j = 1, \ldots, s, \text{ and}$$

$$\underline{x}' = (x_1, \ldots, x_s).$$

It is seen that $X^2 = \underline{y}'\underline{y}$.

Cramer [5] shows that $\underline{y} - [I - A(A'A)^{-1} A'] \underline{x} \xrightarrow[p]{} 0$ under H_o and following the identical line of proof it can be shown (see appendix) that the same result holds under H_n. Since $I - A(A'A)^{-1} A'$ is idempotent,

$$\chi^2 = \underline{x}'[I - A(A'A)^{-1} A']\underline{x} \xrightarrow[p]{} 0. \tag{2.2.4}$$

Cramer also shows that under H_o $\underline{x}$ is asymptotically $N(\underline{0}, I - \underline{pp}')$ where $\underline{p}' = (\sqrt{p_1^o}, \ldots, \sqrt{p_s^o})$. It is easily seen that under H_n $\underline{x}$ is asymptotically $N(\underline{d}, I - \underline{pp}')$.

Let $\underline{w}' = \underline{x}' L$ where $\underset{s\times s}{L} = (\underset{1}{\underline{p}} : \underset{t}{L_1} : \underset{s-1-t}{L_2})$s such that L is orthogonal and $L_1L_1' = A(A'A)^{-1}A'$. Then $w_1 = 0$ with probability one and $w_2, \ldots, w_s$ are asymptotically independent $N(0,1)$ under H_o and under H_n are asymptotically independent normal with common variance unity and means given by $E(\underline{w}') = \underline{d}'L$.

Thus, in the limit

$$\begin{aligned}
\chi^2 &= \underline{x}'[I - L_1L_1']\underline{x}, \\
&= \underline{x}'[\underline{pp}' + L_2L_2']\underline{x}, \\
&= w_1^2 + w_{t+2}^2 + \ldots + w_s^2 \\
&= w_{t+2}^2 + \ldots + w_s^2.
\end{aligned}$$

Therefore, in the limit, χ^2 is distributed under H_o as a central chi-square variate with s-t-1 d.f. and under H_n as a non-central chi-square variate with s-t-1 d.f. and a non-centrality parameter

$$\Delta = \underline{d}' L_2 L_2' \underline{d} .$$

But $\underline{d}'\underline{p} = \sum_{j=1}^{s} d_j = 0$ and we may put

$$\begin{aligned}
\Delta &= \underline{d}'[\underline{pp}' + L_2L_2']\underline{d}, \\
&= \underline{d}'[I - A(A'A)^{-1}A']\underline{d}.
\end{aligned}$$

2.3 Some Remarks on Theorem 1

The convergence in probability ($\hat{\underline{\theta}} \longrightarrow \underline{\theta}_o$) in part (1) of theorem 1 may be formally expressed as

$$\lim_{n \to \infty} \sum_{|\hat{\underline{\theta}}-\underline{\theta}_o| \leq \epsilon} \frac{n!}{\prod_j n_j!} \prod_j p_j^{n_j} (\underline{\theta}_o) = 1$$

no matter where $\underline{\theta}_o$ is in $\underline{\Gamma}$. It is seen that the equations (2.2.1) and their solution do not depend upon $\underline{\theta}_o$ although this is the point in terms of which the convergence in probability is set up. This point is called the unknown true parameter point and is the reference point upon which the asymptotic power is based.

It should be noted that the possibility of zero degrees of freedom has not been excluded in theorem 1. In this case it is seen that χ^2 asymptotically equals zero and that the chi-square distribution with zero d.f. is singular with all the probability concentrated at the origin. Although this degenerate case is of no practical use in itself, it must be included because of the use of χ^2 in theorems 5 and 6. In these theorems the possibility that the d.f. associated with χ^2 is zero cannot be excluded and, indeed, this is always the case in the special case of theorem 6 considered in section 4.5.

Although we are interested in testing a composite hypothesis H_o against a composite alternative H_n^* in theorem 1, it is necessary to consider a simple alternative in order to obtain the asymptotic power. It should be realized that we are dealing with three spaces: the (s-1)-dimensional sample space, the (s-1)-dimensional space of the p_j's, and the interval Ω in the t-dimensional space of the θ_k's. The hypothesis forms a t-dimensional surface in the space of the p_j's.

Corresponding to a specific point $\underline{\theta}_o$ in Ω we have a point $\underline{p}^o$ on the t-dimensional hypothesis surface in the space of the p_j's. We reject, at a level of significance α, the hypothesis that $\underline{p} = \underline{p}^o$ if

$$\chi^2 = \sum_{j=1}^{s} \frac{(n_j - np_j(\hat{\underline{\theta}}))^2}{np_j(\hat{\underline{\theta}})} \geq \chi^2_\alpha(s-1-t) \tag{2.3.1}$$

where $\chi^2_\alpha(s-1-t)$ is the tabulated $100(1-\alpha)\%$ point of the chi-square distribution with s-1-t d.f., and accept it otherwise. $\hat{\underline{\theta}}$ are the modified chi-square minimum estimates subject to H_o of $\underline{\theta}$. Notice that the domain (2.3.1) is determined independently of the unknown true parameter point $\underline{\theta}_o$. Thus we have a partition of the sample space into two portions defined by (2.3.1) and the opposite inequality.

The composite alternative H^*_n is expressed in terms of deviation from the hypothesis surface in the space of the p_j's and the simple alternative is expressed in terms of deviation from a given point on the hypothesis surface. This hypothesis surface may be considered to be a mapping from Ω to the space of the p_j's and since there is some true value of the probabilities if H_o is true there is implied a true parameter point in Ω. This point is assumed to be $\underline{\theta}_o$ and the true probabilities under H_o are $p_j(\underline{\theta}_o)$. Under the simple alternative H_n the true values of the probabilities are supposed to be $p_j(\underline{\theta}_o) + d_j/\sqrt{n}$ with the d_j's being a specific set of deviation parameters where $\sum_{j=1}^{s} d_j = 0$ but not all $d_j = 0$ and where $p_j(\underline{\theta}_o) + d_j/\sqrt{n} \neq p_j(\underline{\theta})$ for any $\underline{\theta}$ in Ω.

Under H_o we have the domain of rejection given by (2.3.1) which implies that

$$\lim_{n \to \infty} \sum_{\chi^2 \geq \chi^2_\alpha} \frac{n!}{\prod_{j=1}^{s} n_j!} \prod_{j=1}^{s} p_j^{n_j}(\underline{\theta}) = \alpha$$

no matter where $\underline{\theta}$ may be in $\underline{\Gamma}$.

Under H_n, to obtain the asymptotic power we have to compute

$$\lim_{n \to \infty} \Pr[\chi^2 \geq \chi^2_\alpha \mid H_n] \text{ or}$$

$$\lim_{n \to \infty} \sum_{\chi^2 \geq \chi^2_\alpha} \frac{n!}{\prod_{j=1}^{s} n_j!} \prod_{j=1}^{s} [p_j(\underline{\theta}_o) + d_j / \sqrt{n}]^{n_j}$$

This, however, turns out to be dependent not only on the d_j's but also on $\underline{\theta}_o$.

This situation is not unique to this type of test. With Student's t-test we may test a composite hypothesis against a composite alternative but we can find the power for some assumed simple alternative only since the non-central t-distribution depends upon the unknown true values of μ and σ.

This discussion will be seen to apply in essence to the more complex theorems that follow.

2.4 Theorem 2

Given a single-multinomial sample (1.1.1), s functions $p_j(\underline{\theta})$ of t unknown parameters $\theta_1, \ldots, \theta_t$ as in (1.3.1) and u functions $f_m(\underline{\theta})$

as in (1.3.2), consider the hypothesis

H_o: $p_j = p_j(\underline{\theta})$ for $j = 1,\ldots,s$ subject to $f_m(\underline{\theta}) = 0$ for $m = 1,\ldots,u$,

and the alternative

H_n^*: $p_{jn} = p_j(\underline{\theta}) + d_j/\sqrt{n}$ subject to $f_m(\underline{\theta}) = 0$ where $\sum_{j=1}^{s} d_j = 0$ but not all $d_j = 0$ and where $p_{jn} \neq p_j(\underline{\theta})$ for any $\underline{\theta}$ in Θ.

In order to obtain the asymptotic power we consider the simple alternative

H_n: $p_{jn} = p_j(\underline{\theta}_o) + d_j/\sqrt{n}$ subject to $f_m(\underline{\theta}_o) = 0$ where $\underline{\theta}_o$ is supposed to be the unknown true parameter point in Θ under H_o.

Then (1) the equations

$$\sum_{j=1}^{s} \frac{n_j - np_j(\underline{\theta})}{p_j(\underline{\theta})} \cdot \frac{\partial p_j(\underline{\theta})}{\partial \theta_k} + \sum_{m=1}^{u} \lambda_m \frac{\partial f_m(\underline{\theta})}{\partial \theta_k} = 0, \tag{2.4.1}$$

$$k = 1,\ldots,t,\ f_m(\underline{\theta}) = 0,\ m = 1,\ldots,u,$$

for minimizing chi-square in the modified sense subject to H_o have exactly one system of solutions $\hat{\hat{\underline{\theta}}}' = (\hat{\hat{\theta}}_1,\ldots,\hat{\hat{\theta}}_t)$, $\hat{\hat{\underline{\lambda}}}' = (\hat{\hat{\lambda}}_1,\ldots,\hat{\hat{\lambda}}_u)$ such that $\hat{\hat{\underline{\theta}}} \longrightarrow \underline{\theta}_o$ in probability as $n \longrightarrow \infty$, and (2) the statistic

$$\chi_*^2 = \sum_{j=1}^{s} \frac{(n_j - np_j(\hat{\hat{\underline{\theta}}}))^2}{np_j(\hat{\hat{\underline{\theta}}})} \tag{2.4.2}$$

is, in the limit, distributed under H_o as a central chi-square variate with $s-1-t+u$ d.f. and under H_n as a non-central chi-square variate with s-1-t+u d.f. and a non-centrality parameter

$$\Delta^* = \underline{d}'[I - A_*(A_*'A_*)^{-1}A_*']\underline{d}, \tag{2.4.3}$$

where $\underset{s\times t-u}{A_*} = \left[\frac{1}{\sqrt{p_j^o}}\left(\frac{\partial p_j^*}{\partial \theta_{k''}}\right)_o\right]$, with p_j^* being $p_j(\theta)$ expressed in terms of the t-u independent θ_k and $\theta_{k''}$ are the t-u independent θ_k.

Proof

Part (1) has been proved by Mitra [7] and by Ogawa [11]. Part (2) will be proved here.

Let $y_j^* = \dfrac{n_j - np_j(\hat{\underline{\hat{\theta}}})}{\sqrt{np_j(\hat{\underline{\hat{\theta}}})}}$, $j = 1,\ldots,s$, and

$\underline{y}^{*\prime} = (y_1^*,\ldots,y_s^*)$.

Mitra and Ogawa show that $\underline{y}^* - [I - A_*(A_*'A_*)^{-1}A_*']\underline{x} \xrightarrow[p]{} 0$ under H_o and in a manner similar to that of theorem 1 it can be shown (see appendix) that the same is true under H_n. It follows that

$$\chi^2 - \underline{x}'[I - A_*(A_*'A_*)^{-1}A_*']\underline{x} \xrightarrow[p]{} 0. \tag{2.4.4}$$

Under H_o $\underline{x}$ is asymptotically $N(\underline{0}, I - \underline{pp}')$ and under H_n $\underline{x}$ is asymptotically $N(\underline{d}, I - \underline{pp}')$.

Let $\underline{w}^{*\prime} = \underline{x}'L^*$ where $\underset{s\times s}{L^*} = (\underset{1}{\underline{p}} : \underset{t-u}{L_1^*} : \underset{s-1-t+u}{L_2^*})$s such that L^* is orthogonal and $L_1^*L_1^{*\prime} = A_*(A_*'A_*)^{-1}A_*'$. Then $w_1^* = 0$ with probability one and $w_2^*,\ldots,w_s^*$ are asymptotically independent $N(0,1)$ under H_o and under H_n are asymptotically independent normal with common variance unity and means given by $E(\underline{w}^{*\prime}) = \underline{d}'L^*$.

Thus, in the limit,

$$\begin{aligned}\chi^2_* &= \underline{x}'[I - L_1^*L_1^{*\prime}]\,\underline{x},\\ &= \underline{x}'[\underline{p}\underline{p}' + L_2^*L_2^{*\prime}]\,\underline{x},\\ &= w_1^{*2} + w_{t-u+2}^{*2} + \ldots + w_s^{*2},\\ &= w_{t-u+2}^{*2} + \ldots + w_s^{*2}.\end{aligned}$$

Therefore, in the limit, χ^2_* is distributed under H_o as a central chi-square variate with s-1-t+u d.f. and under H_n as a non-central chi-square variate with s-1-t+u d.f. and a non-centrality parameter

$$\Delta^* = \underline{d}'\,L_2^*L_2^{*\prime}\,\underline{d}.$$

But $\underline{d}'\underline{p} = \sum_{j=1}^{s} d_j = 0$ and we may put

$$\begin{aligned}\Delta^* &= \underline{d}'[\underline{p}\underline{p}' + L_2^*L_2^{*\prime}]\,\underline{d},\\ &= \underline{d}'[I - A_*(A_*'A_*)^{-1}A_*']\,\underline{d}.\end{aligned}$$

2.5 Asymptotic Independence of Tests

Suppose we have two tests of the type in theorem 2 to be made on the same data. Denote the two hypotheses by H_{01} and H_{02} where

H_{01}: $p_j = p_{1j}\,(\theta_{11},\ldots,\theta_{1t_1})$ for $j = 1,\ldots,s$,

subject to $f_{1m}\,(\theta_{11},\ldots,\theta_{1t_1}) = 0$ for $m_1 = 1,\ldots,u_1$,

H_{02}: $p_j = p_{2j}\,(\theta_{21},\ldots,\theta_{2t_2})$ for $j = 1,\ldots,s$,

subject to $f_{2m_2}\,(\theta_{21},\ldots,\theta_{2t_2}) = 0$ for $m_2 = 1,\ldots,u_2$

and where $u_1 < t_1$, $u_2 < t_2$ and $(t_1 + t_2) - (u_1 + u_2) < s$.

Define A_{1*} $(sx\overline{t_1 - u_1})$ and A_{2*} $(sx\overline{t_2 - u_2})$ for H_{01} and H_{02} respectively as in theorem 2 and let the test statistic for each test be X^2_{1*} and X^2_{2*}, defined as in theorem 2.

Then, when both H_{01} and H_{02} hold,

$$X_1^2 - \underline{x}'[I - A_{1*}(A_{1*}'A_{1*})^{-1}A_{1*}']\,\underline{x} \xrightarrow[p]{} 0, \text{ and}$$

$$X_2^2 - \underline{x}'[I - A_{2*}(A_{2*}'A_{2*})^{-1}A_{2*}']\,\underline{x} \xrightarrow[p]{} 0.$$

By the theorem in section 1.5, when H_{01} and H_{02} are true, X^2_{1*} and X^2_{2*} are, in the limit, independent in probability if and only if

$$[I - A_{1*}(A_{1*}'A_{1*})^{-1}A_{1*}']\,[I - \underline{pp}']\,[I - A_{2*}(A_{2*}'A_{2*})^{-1}A_{2*}'] = 0.$$

But $\underline{p}'A_{1*} = 0$ and $\underline{p}'A_{2*} = 0$.

Therefore, a necessary and sufficient condition for the asymptotic independence in probability of X^2_{1*} and X^2_{2*} when H_{01} and H_{02} are true is that

$$[I - A_{1*}(A_{1*}'A_{1*})^{-1}A_{1*}']\,[I - A_{2*}(A_{2*}'A_{2*})^{-1}A_{2*}'] = \underline{pp}' .$$

2.6 Examples of Specific Tests

Here are some specific tests given for two-way and three-way tables in which all dimensions are 'variates' (Roy and Mitra [18]). For two-way tables the dimensions are denoted by 'i' and 'j' and for three-way tables by 'i', 'j' and 'k'. In each example the hypothesis, alternative, test statistic, degrees of freedom and non-centrality parameter are listed.

(2.6.1) Two-way: hypothesis of independence between 'i' and 'j'.

H_o: $p_{ij} = p_{io}p_{oj}$ for $i = 1,\ldots,r$ and $j = 1,\ldots,s$.

$$H_n:\ p_{ij} = p^o_{io}p^o_{oj} + \frac{d_{ij}}{\sqrt{n}} \text{ where } \sum_{ij} d_{ij} = 0.$$

H_o and H_n subject to $\sum_i p_{io} = 1$ and $\sum_j p_{oj} = 1$.

$$\chi^2 = \sum_{ij} \frac{\left(n_{ij} - \frac{n_{io}n_{oj}}{n}\right)^2}{\frac{n_{io}n_{oj}}{n}} .$$

d.f. = (r-1)(s-1).

$$\Delta = \sum_{ij} \frac{1}{p^o_{io}p^o_{oj}} (d_{ij} - p^o_{oj}d_{io} - p^o_{io}d_{oj})^2.$$

If $d_{io} = d_{oj} = 0$, $\quad \Delta = \sum_{ij} \frac{d^2_{ij}}{p^o_{io}p^o_{oj}}$.

(2.6.2) Three-way: Hypothesis of conditional independence between 'i' and 'j' for fixed 'k'.

$$H_o:\ p_{ijk} = \frac{p_{iok}p_{ojk}}{p_{ook}} \text{ for } i = 1,\ldots,r,\ j = 1,\ldots,s \text{ and } k = 1,\ldots,t.$$

$$H_n:\ p_{ijk} = \frac{p^o_{iok}p^o_{ojk}}{p^o_{ook}} + \frac{d_{ijk}}{\sqrt{n}} \text{ where } \sum_{ijk} d_{ijk} = 0.$$

H_o and H_n subject to $\sum_i p_{iok} = \sum_j p_{ojk} = p_{ook}$ and $\sum_k p_{ook} = 1$.

$$\chi^2 = \sum_{ijk} \frac{\left(n_{ijk} - \frac{n_{iok}n_{ojk}}{n_{ook}}\right)^2}{\frac{n_{iok}n_{ojk}}{n_{ook}}} .$$

d.f. = (r-1)(s-1)t.

$$\chi^2 = \sum_{ijk} \frac{\left(n_{ijk} - \dfrac{n_{ioo}n_{ojo}n_{ook}}{n^2}\right)^2}{\dfrac{n_{ioo}n_{ojo}n_{ook}}{n^2}} .$$

d.f. = rst - r - s - t + 2.

$$\Delta = \sum_{ijk} \frac{1}{p^o_{ioo}p^o_{ojo}p^o_{ook}} (d_{ijk}-p^o_{ojo}p^o_{ook}d_{ioo}-p^o_{ioo}p^o_{ook}d_{ojo}-p^o_{ioo}p^o_{ojo}d_{ook})^2.$$

If $d_{ioo} = d_{ojo} = d_{ook} = 0$, $\quad \Delta = \sum_{ijk} \frac{d^2_{ijk}}{p^o_{ioo}p^o_{ojo}p^o_{ook}}$.

Chapter III

CONTINGENCY-LIKE TESTS FOR PRODUCT-MULTINOMIAL SAMPLES

3.1 Introduction

For product-multinomial samples, tests of hypotheses of types (1.4.4) and (1.4.5) are referred to as "contingency-like tests" since they include all the usual so-called contingency tests for samples in which some of the marginals are stochastic variates and others are fixed by the sampling scheme. As in the single-multinomial case, many more tests of these types may be constructed and the theorems in this chapter give a test statistic and its null and non-null asymptotic distributions.

Two theorems are presented. As in Chapter II, the first is of a preliminary nature while the second theorem covers those tests that are most likely to be useful in practice. A necessary and sufficient condition for asymptotic independence in probability of two tests is derived. Several examples of specific tests are given.

3.2 Theorem 3

Given a product-multinomial sample (1.1.2) and rs functions $p_{ij}(\underline{\theta})$ of t unknown parameters $\theta_1,\ldots,\theta_t$ as in (1.4.1), consider the hypothesis

$$H_o\text{: } p_{ij} = p_{ij}(\underline{\theta}) \text{ for } i = 1,\ldots,r \text{ and } j = 1,\ldots,s$$

and the alternative

$$H_n^*\text{: } p_{ijn} = p_{ij}(\underline{\theta}) + d_{ij}/\sqrt{n} \text{ where } \sum_{j=1}^{s} d_{ij} = 0$$

but not all $d_{ij} = 0$ for $i = 1,\ldots,r$ and where

$p_{ijn} \neq p_{ij}(\underline{\theta})$ for any $\underline{\theta}$ in $\underline{\Gamma}$.

In order to obtain the asymptotic power we consider the simple alternative

H_n: $p_{ijn} = p_{ij}(\underline{\theta}_o) + d_{ij}/\sqrt{n}$ where $\underline{\theta}_o$ is supposed to be the unknown true parameter point in $\underline{(\ )}$ under H_o.

Then (1) the equations

$$(3.2.1) \qquad \sum_{i=1}^{r} \sum_{j=1}^{s} \frac{n_{ij}-n_{io}p_{ij}(\underline{\theta})}{p_{ij}(\underline{\theta})} \cdot \frac{\partial\, p_{ij}(\underline{\theta})}{\partial\, \theta_k} = 0,\ k = 1,\ldots,t,$$

for minimizing chi-square in the modified sense subject to H_o have exactly one system of solutions $\hat{\underline{\theta}}' = (\hat{\theta}_1,\ldots,\hat{\theta}_t)$ such that $\hat{\underline{\theta}} \longrightarrow \underline{\theta}_o$ in probability as $n \longrightarrow \infty$ subject to the ratios n_{io}/n being held fixed, and (2) the statistic

$$(3.2.2) \qquad \chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{s} \frac{(n_{ij}-n_{io}p_{ij}(\hat{\underline{\theta}}))^2}{n_{io}p_{ij}(\hat{\underline{\theta}})}$$

is, in the limit, distributed under H_o as a central chi-square variate with rs-r-t d.f. and under H_n as a non-central chi-square variate with rs-r-t d.f. and a non-centrality parameter

$$(3.2.3) \qquad \Delta = \underline{d}'\left[I - B(B'B)^{-1}B'\right]\underline{d},$$

where $\underset{rsx1}{\underline{d}} = \left(\frac{d_{ij}}{\sqrt{p^o_{ij}}} \sqrt{\frac{n_{io}}{n}} \right)$, $i = 1,\ldots,r$ and $j = 1,\ldots,s$

(lexicographically),

$p^o_{ij} = p_{ij}(\underline{\theta}_o)$, and

$$\underset{rsxt}{B} = \left[\frac{1}{\sqrt{p^o_{ij}}} \sqrt{\frac{n_{io}}{n}} \left(\frac{\partial\, p_{ij}}{\partial\, \theta_k} \right)_o \right]$$

where $\left(\frac{\partial p_{ij}}{\partial \theta_k}\right)_o$ is $\left(\frac{\partial p_{ij}}{\partial \theta_k}\right)$ evaluated at the point $\underline{\theta}_o$.

Proof

Part (1) has been proved by Mitra [7] and Ogawa [11]. Part (2) has been proved by Mitra but the proof is repeated here for completeness.

Let $y_{ij} = \dfrac{n_{ij}-n_{io}p_{ij}(\hat{\underline{\theta}})}{\sqrt{n_{io}p_{ij}(\hat{\underline{\theta}})}}$, $i = 1,\ldots,r$ and $j = 1,\ldots,s$,

$\underline{y}' = (y_{11},\ldots,y_{1s},\ldots,y_{r1},\ldots,y_{rs})$,

$x_{ij} = \dfrac{n_{ij}-n_{io}p_{ij}(\underline{\theta}_o)}{\sqrt{n_{io}p_{ij}(\underline{\theta}_o)}}$, $i = 1,\ldots,r$ and $j = 1,\ldots,s$, and

$\underline{x}' = (x_{11},\ldots,x_{1s},\ldots,x_{r1},\ldots,x_{rs})$.

Ogawa [11] shows that $\underline{y} - [I - B(B'B)^{-1}B']\,\underline{x} \xrightarrow[p]{} 0$ under H_o and in a manner similar to that of theorem 1 it can be shown (see appendix) that the same is true under H_n. Thus

(3.2.4) $$\chi^2 = \underline{x}'[I - B(B'B)^{-1}B']\,\underline{x} \xrightarrow[p]{} 0.$$

Ogawa also shows that under H_o $\underline{x}$ is asymptotically $N(\underline{0}, I - PP')$ and under H_n it is easily seen that $\underline{x}$ is asymptotically $N(\underline{d}, I - PP')$, where

$$\underset{rs \times r}{P} = \begin{bmatrix} \sqrt{p^o_{11}} & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ \sqrt{p^o_{1s}} & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sqrt{p^o_{r1}} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sqrt{p^o_{rs}} \end{bmatrix} .$$

Let $\underline{w}' = \underline{x}'L$ where $\underset{rs \times rs}{L} = (\underset{r}{P} : \underset{t}{L_1} : \underset{rs-r-t}{L_2})$ rs such that L is orthogonal and $L_1L_1' = B(B'B)^{-1}B'$. Then $w_1, \ldots, w_r$ equal zero with probability one and $w_{r+1}, \ldots, w_{rs}$ are asymptotically distributed under H_o as independent N(0, 1) variates and under H_n as independent normal variates with common variance unity and means given by $E(\underline{w}') = \underline{d}'L$.

Thus, in the limit,

$$\begin{aligned} \chi^2 &= \underline{x}' [I - L_1L_1'] \underline{x}, \\ &= \underline{x}' [PP' + L_2L_2'] \underline{x}, \\ &= w_1^2 + \ldots + w_r^2 + w_{r+t+1}^2 + \ldots + w_{rs}^2 , \\ &= w_{r+t+1}^2 + \ldots + w_{rs}^2 . \end{aligned}$$

Therefore, in the limit, χ^2 is distributed under H_o as a central chi-square variate with rs-r-t d.f. and under H_n as a non-central chi-square variate with rs-r-t d.f. and a non-centrality parameter

$$\Delta = \underline{d}' L_2 L_2' \underline{d}.$$

But $\underline{d}'P = \left(\sqrt{\frac{n_{io}}{n}} \sum_j d_{1j}, \ldots, \sqrt{\frac{n_{ro}}{n}} \sum_j d_{rj}\right) = (0,\ldots,0)$

and we may put

$$\Delta = \underline{d}' [PP' + L_2 L_2'] \underline{d},$$

$$= \underline{d}' [I - B(B'B)^{-1}B'] \underline{d}.$$

3.3 Theorem 4

Given a product-multinomial sample (1.1.2), rs functions $p_{ij}(\underline{\theta})$ of t unknown parameters $\theta_1,\ldots,\theta_t$ as in (1.4.1) and u functions $f_m(\underline{\theta})$ as in (1.4.2), consider the hypothesis

H_o: $p_{ij} = p_{ij}(\underline{\theta})$ for $i = 1,\ldots,r$ and $j = 1,\ldots,s$

subject to $f_m(\underline{\theta}) = 0$ for $m = 1,\ldots,u$,

and the alternative

H_n^*: $p_{ijn} = p_{ij}(\underline{\theta}) + d_{ij} / \sqrt{n}$ subject to $f_m(\underline{\theta}) = 0$

where $\sum_{j=1}^{s} d_{ij} = 0$ but not all $d_{ij} = 0$ for $i = 1,\ldots,r$

and where $p_{ijn} \neq p_{ij}(\underline{\theta})$ for any $\underline{\theta}$ in Ω.

In order to obtain the asymptotic power we consider the simple alternative

H_n: $p_{ijn} = p_{ij}(\underline{\theta}_o) + d_{ij} / \sqrt{n}$ subject to $f_m(\underline{\theta}_o) = 0$

where $\underline{\theta}_o$ is supposed to be the unknown true parameter point in Ω under H_o.

Then (1) the equations

$$(3.3.1) \qquad \sum_{i=1}^{r} \sum_{j=1}^{s} \frac{n_{ij}-n_{io}p_{ij}(\underline{\theta})}{p_{ij}(\underline{\theta})} \cdot \frac{\partial \; p_{ij}(\underline{\theta})}{\partial \; \theta_k} + \sum_{m=1}^{u} \lambda_m \frac{\partial \; f_m(\underline{\theta})}{\partial \; \theta_k} = 0,$$

$$k = 1,\ldots,t, \; f_m(\underline{\theta}) = 0, \; m = 1,\ldots,u,$$

for minimizing chi-square in the modified sense subject to H_o have exactly one system of solutions $\hat{\hat{\underline{\theta}}}' = (\hat{\hat{\theta}}_1,\ldots,\hat{\hat{\theta}}_t)$, $\hat{\hat{\underline{\lambda}}}' = (\hat{\hat{\lambda}}_1,\ldots,\hat{\hat{\lambda}}_u)$ such that $\hat{\hat{\underline{\theta}}} \longrightarrow \underline{\theta}_o$ in probability as $n \longrightarrow \infty$ subject to the ratios n_{io}/n being held fixed and (2) the statistic

$$(3.3.2) \qquad \chi_*^2 = \sum_{i=1}^{r} \sum_{j=1}^{s} \frac{(n_{ij}-n_{io}p_{ij}(\hat{\hat{\underline{\theta}}}))^2}{n_{io}p_{ij}(\hat{\hat{\underline{\theta}}})}$$

is, in the limit, distributed under H_o as a central chi-square variate with rs-r-t+u d.f. and under H_n as a non-central chi-square variate with rs-r-t+u d.f. and a non-centrality parameter

$$(3.3.3) \qquad \Delta^* = \underline{d}' [I - B_*(B_*'B_*)^{-1}B_*'] \, \underline{d},$$

where $\underset{rs \times \overline{t-u}}{B_*} = \left[\frac{1}{\sqrt{p_{ij}^o}} \sqrt{\frac{n_{io}}{n}} \left(\frac{\partial \; p_{ij}^*}{\partial \; \theta_{k''}} \right)_o \right]$ with p_{ij}^* being $p_{ij}(\underline{\theta})$ expressed in terms of the t-u independent θ_k and $\theta_{k''}$ are the t-u independent θ_k.

Proof

Part (1) has been proved by Ogawa [11]. Part (2) will be proved here.

$$\text{Let } y_{ij}^* = \frac{n_{ij}-n_{io}p_{ij}(\hat{\hat{\underline{\theta}}})}{\sqrt{n_{io}p_{ij}(\hat{\hat{\underline{\theta}}})}}, \quad i = 1,\ldots,r, \; j = 1,\ldots,s,$$

and $\underline{y}^{*\prime} = (y_{11}^*,\ldots,y_{1s}^*,\ldots,y_{r1}^*,\ldots,y_{rs}^*)$.

Ogawa shows that $\underline{y} - [I - B_*(B_*'B_*)^{-1}B_*'] \underline{x} \xrightarrow[p]{} 0$ under H_o and in a manner similar to that of theorem 1 it can be shown (see appendix) that the same is true under H_n. It follows that

(3.3.4) $$\chi_*^2 - \underline{x}'[I - B_*(B_*'B_*)^{-1}B_*'] \underline{x} \xrightarrow[p]{} 0.$$

Under H_o $\underline{x}$ is asymptotically $N(\underline{0}, I - PP')$ and under H_n $\underline{x}$ is asymptotically $N(\underline{d}, I - PP')$.

Let $\underline{w}^{*\prime} = \underline{x}'L^*$ where $\underset{rs\times rs}{L^*} = (\underset{r}{P} : \underset{t-u}{L_1^*} : \underset{rs-r-t+u}{L_2^*})$ such that L^* is orthogonal and $L_1^*L_1^{*\prime} = B_*(B_*'B_*)^{-1}B_*'$. Then $w_1^*, \ldots, w_r^*$ equal zero with probability one and $w_{r+1}^*, \ldots, w_{rs}^*$ are asymptotically distributed under H_o as independent $N(0,1)$ variates and under H_n as independent normal variates with common variance unity and means given by $E(\underline{w}^{*\prime}) = \underline{d}'L^*$.

Thus, in the limit,

$$\begin{aligned}\chi_*^2 &= \underline{x}'[I - L_1^*L_1^{*\prime}] \underline{x},\\ &= \underline{x}'[PP' + L_2^*L_2^{*\prime}] \underline{x},\\ &= w_1^{*2} + \ldots + w_r^{*2} + w_{r+t-u+1}^{*2} + \ldots + w_{rs}^{*2},\\ &= w_{r+t-u+1}^{*2} + \ldots + w_{rs}^{*2}.\end{aligned}$$

Therefore, in the limit, χ_*^2 is distributed under H_o as a central chi-square variate with rs-r-t+u d.f. and under H_n as a non-central chi-square variate with rs-r-t+u d.f. and a non-centrality parameter

$$\Delta^* = \underline{d}'\, L_2^*L_2^{*\prime}\, \underline{d}.$$

But $\underline{d}'P = (0,\ldots,0)$ and we may put

$$\Delta^* = \underline{d}' \left[PP' + L_2^* L_2^{*\prime} \right] \underline{d},$$

$$= \underline{d}' \left[I - B_*(B_*' B_*)^{-1} B_*' \right] \underline{d}.$$

3.4 Asymptotic Independence of Tests

Suppose we have two hypotheses

H_{01}: $p_{ij} = p_{1ij}(\theta_{11}, \ldots, \theta_{1t_1})$ for $i = 1, \ldots, r$ and $j = 1, \ldots, s$ subject to $f_{1m_1}(\theta_{11}, \ldots, \theta_{1t_1}) = 0$ for $m_1 = 1, \ldots, u_1$, and H_{02}: $p_{ij} = p_{2ij}(\theta_{21}, \ldots, \theta_{2t_2})$ for $i = 1, \ldots, r$ and $j = 1, \ldots, s$ subject to $f_{2m_2}(\theta_{11}, \ldots, \theta_{1t_1}) = 0$ for $m_2 = 1, \ldots, u_2$, where $u_1 < t_1$, $u_2 < t_2$ and $(t_1 + t_2) - (u_1 + u_2) \leq rs-r$.

Let X_{1*}^2 and X_{2*}^2 be the test statistics for H_{01} and H_{02} respectively and define B_{1*} and B_{2*} for H_{01} and H_{02} respectively as in theorem 4. Then in a manner exactly analogous to that of section 2.5 it can be shown that a necessary and sufficient condition for the asymptotic independence in probability of X_{1*}^2 and X_{2*}^2 when H_{01} and H_{02} are true is that

$$\left[I - B_{1*}(B_{1*}' B_{1*})^{-1} B_{1*}' \right] \left[I - B_{2*}(B_{2*}' B_{2*})^{-1} B_{2*}' \right] = PP'.$$

3.5 Examples of Specific Tests

Some specific tests are given here for two-way and three-way tables in which some dimensions represent 'variates' and others represent 'ways of classification' (Roy and Mitra [18]). As in section 2.6, 'i' and 'j' or 'i', 'j' and 'k' denote the dimensions. In each example the hypothesis, alternative, test statistic, degrees of freedom and non-centrality parameter are listed.

(3.5.1) Two-way; 'i' a 'way of classification' and 'j' a 'variate'; hypothesis that for any j, p_{ij} is independent of 'i'.

H_o: $p_{ij} = q_{*j}$ (say) for $i = 1,\ldots,r$ and $j = 1,\ldots,s$.

H_n: $p_{ij} = q^o_{*j} + \frac{d_{ij}}{\sqrt{n}}$ where $\sum_j d_{ij} = 0$ for $i = 1,\ldots,r$.

H_o and H_n subject to $\sum_j q_{*j} = 1$.

$$\chi^2 = \sum_{ij} \frac{\left(n_{ij} - \frac{n_{io}n_{oj}}{n}\right)^2}{\frac{n_{io}n_{oj}}{n}} .$$

d.f. = (r-1)(s-1)

$$\Delta = \sum_{ij} \frac{n_{io}}{n} \cdot \frac{1}{q^o_{*j}} \cdot (d_{ij} - d_{*j})^2 \text{ where } d_{*j} = \sum_i \frac{n_{io}}{n} d_{ij}.$$

(3.5.2) Three-way; 'i' a 'way of classification' and 'j' and 'k' 'variates'; hypothesis of independence between 'j' and 'k' for each 'i'.

H_o: $p_{ijk} = p_{ijo}p_{iok}$ for $i = 1,\ldots,r$, $j = 1,\ldots,s$ and $k = 1,\ldots,t$.

H_n: $p_{ijk} = p^o_{ijo}p^o_{iok} + \frac{d_{ijk}}{\sqrt{n}}$ where $\sum_{jk} d_{ijk} = 0$ for $i = 1,\ldots,r$.

H_o and H_n subject to $\sum_j p_{ijo} = 1$ and $\sum_k p_{iok} = 1$ for $i = 1,\ldots,r$.

$$\chi^2 = \sum_{ijk} \frac{\left(n_{ijk} - \frac{n_{ijo}n_{iok}}{n_{ioo}}\right)^2}{\frac{n_{ijo}n_{iok}}{n_{ioo}}} .$$

d.f. = r(s-1)(t-1).

$$\Delta = \sum_{ijk} \frac{n_{ioo}}{n} \cdot \frac{1}{p^o_{ijo}p^o_{iok}} \cdot (d_{ijk}-p^o_{ijo}d_{iok}-p^o_{iok}d_{ijo})^2 .$$

If $d_{ijo} = d_{iok} = 0$, $\quad \Delta = \sum_{ijk} \frac{n_{ioo}}{n} \cdot \frac{d^2_{ijk}}{p^o_{ijo}p^o_{iok}}$.

(3.5.3) Three-way; 'i' a 'way of classification' and 'j' and 'k' 'variates'; hypothesis that p_{ijk} is independent of 'i' for all j and k.

H_o: $p_{ijk} = q_{*jk}$ (say) for $i = 1,\ldots,r$, $j = 1,\ldots,s$ and $k = 1,\ldots,t$.

H_n: $p_{ijk} = q^o_{*jk} + \frac{d_{ijk}}{\sqrt{n}}$ where $\sum_{jk} d_{ijk} = 0$ for $i = 1,\ldots,r$.

H_o and H_n subject to $\sum_{jk} q_{*jk} = 1$.

$$\chi^2 = \sum_{ijk} \frac{\left(n_{ijk} - \frac{n_{ioo}n_{ojk}}{n}\right)^2}{\frac{n_{ioo}n_{ojk}}{n}} .$$

d.f. = (r-1)(st-1).

$$\Delta = \sum_{ijk} \frac{n_{ioo}}{n} \cdot \frac{1}{q^o_{*jk}} (d_{ijk}-d_{*jk})^2 \text{ where } d_{*jk} = \sum_i \frac{n_{ioo}}{n} \; d_{ijk} .$$

(3.5.4) Three-way; 'i' and 'j' 'ways of classification' and 'k' a 'variate'; hypothesis that for any 'j', p_{ijk} is independent of 'i'.

H_o: $p_{ijk} = q_{*jk}$ (say) for $i = 1,\ldots,r$, $j = 1,\ldots,s$ and $k = 1,\ldots,t$.

H_n: $p_{ijk} = q^o_{*jk} + \dfrac{d_{ijk}}{\sqrt{n}}$ where $\sum\limits_k d_{ijk} = 0$ for $i = 1,\ldots,r$ and $j = 1,\ldots,s$.

H_o and H_n subject to $\sum\limits_k q_{*jk} = 1$ for $j = 1,\ldots,s$.

$$\chi^2 = \sum_{ijk} \frac{\left(n_{ijk} - \dfrac{n_{ijo}n_{ojk}}{n_{ojo}}\right)^2}{\dfrac{n_{ijo}n_{ojk}}{n_{ojo}}},$$

d.f. = $(r-1)\,s\,(t-1)$.

$$\Delta = \sum_{ijk} \frac{n_{ijo}}{n} \frac{1}{q^o_{*jk}} (d_{ijk}-d_{*jk})^2 \text{ where } d_{*jk} = \sum_i \frac{n_{ijo}}{n_{ojo}} d_{ijk}.$$

(3.5.5) Three-way; 'i' and 'j' 'ways of classification' and 'k' a 'variate'; hypothesis that for any 'i', p_{ijk} is independent of 'j'.

H_o: $p_{ijk} = q_{i*k}$ (say) for $i = 1,\ldots,r$, $j = 1,\ldots,s$ and $k = 1,\ldots,t$.

H_n: $p_{ijk} = q^o_{i*k} + \dfrac{d_{ijk}}{\sqrt{n}}$ where $\sum\limits_k d_{ijk} = 0$ for $i = 1,\ldots,r$ and $j = 1,\ldots,s$.

H_o and H_n subject to $\sum\limits_k q_{i*k} = 1$ for $i = 1,\ldots,r$.

$$\chi^2 = \sum_{ijk} \frac{\left(n_{ijk} - \frac{n_{ijo}n_{iok}}{n_{ioo}}\right)^2}{\frac{n_{ijo}n_{iok}}{n_{ioo}}} \quad .$$

d.f. = r(s-1)(t-1).

$$\Delta = \sum_{ijk} \frac{n_{ijo}}{n} \cdot \frac{1}{q^o_{i*k}} (d_{ijk}-d_{i*k})^2 ,$$

where $d_{i*k} = \sum_j \frac{n_{ijo}}{n_{ioo}} d_{ijk}$.

(3.5.6) Three-way; 'i' and 'j' 'ways of classification' and 'k' a 'variate'; hypothesis that for any k, p_{ijk} is independent of 'ij'.

H_o: $p_{ijk} = q_{**k}$ (say) for i = 1,...,r, j = 1,...,s and k = 1,...,t.

H_n: $p_{ijk} = q^o_{**k} + \frac{d_{ijk}}{\sqrt{n}}$ where $\sum_k d_{ijk} = 0$ for i = 1,...,r and j = 1,...,s.

H_o and H_n subject to $\sum_k q_{**k} = 1$.

$$\chi^2 = \sum_{ijk} \frac{\left(n_{ijk} - \frac{n_{ijo}n_{ook}}{n}\right)^2}{\frac{n_{ijo}n_{ook}}{n}} \quad .$$

d.f. = (rs-1)(t-1).

$$\Delta = \sum_{ijk} \frac{n_{ijo}}{n} \cdot \frac{1}{q^o_{**k}} (d_{ijk}-d_{**k})^2 \text{ where } d_{**k} = \sum_{ij} \frac{n_{ijo}}{n} d_{ijk} \quad .$$

Chapter IV

ANALOGUES TO NORMAL UNIVARIATE AND MULTIVARIATE ANALYSIS OV VARIANCE TESTS FOR SINGLE-MULTINOMIAL SAMPLES

4.1 Introduction

For single-multinomial samples tests of hypotheses of the type given in (1.3.6) are called analogous to normal univariate and multivariate analysis of variance tests since we are given that the basic parameters (the p_j's) are known functions of a smaller number of unknown parameters (the θ_k's) and we wish to test the hypothesis that some given function of this smaller number of unknown parameters is equal to zero. Few examples of this type of test have been used thus far in the analysis of categorical data, but one application is given here and others can be found.

Two theorems on this type of test are given, the first preliminary and the second very general in nature. The special case which is given is most applicable in practice and the example of a specific test which is included fits this special case. Here, as in the previous chapters, a necessary and sufficient condition for the asymptotic independence in probability of two tests is developed.

4.2 Theorem 5

Given a single-multinomial sample (1.1.1) and s functions $p_j(\underline{\theta})$ of t unknown parameters $\theta_1,\ldots,\theta_t$ as in (1.3.1), suppose it is also given that $p_j = p_j(\underline{\theta})$ for $j = 1,\ldots,s$. Consider the hypothesis

$$H_o\colon \theta_{k'} = h_{k'},\ k' = 1,\ldots,u < t,$$

and the alternative

H_n: $\theta_{k'} = h_{k'} + c_{k'} / \sqrt{n} = \theta^o_{k'}$ (say) where not all $c_{k'} = 0$ and where $\underline{\theta}'_o = (\theta^o_1, \ldots, \theta^o_u, \theta^o_{u+1}, \ldots, \theta^o_t)$ is the unknown true parameter point in $\underline{(\,)}$.

Let $\hat{\underline{\theta}}$ be the unique consistent solution of the equations given by (2.2.1) and let χ^2 be the corresponding test statistic given by (2.2.2)

Let $\hat{\hat{\underline{\theta}}}$ be the unique consistent solution of the formal equations, subject to H_o,

$$\sum_{j=1}^{s} \frac{n_j - np_j(\underline{\theta})}{p_j(\underline{\theta})} \cdot \frac{\partial\, p_j(\underline{\theta})}{\partial\, \theta_k} + \sum_{k'=1}^{u} \lambda_{k'} \frac{\partial\, \theta_{k'}}{\partial\, \theta_k} = 0, \tag{4.2.1}$$

$$k = 1, \ldots, t, \quad \theta_{k'} = h_{k'}, \quad k' = 1, \ldots, u,$$

and let χ^2_* be the corresponding test statistic as given by (2.3.2). Note that $\frac{\partial\, \theta_{k'}}{\partial\, \theta_k} = 1$ or 0 according as $k' = k$ or $k' \neq k$ and hence (4.2.1) may be reduced to t-u equations in $\theta_{u+1}, \ldots, \theta_t$.

Then $\chi^2_o = \chi^2_* - \chi^2$ is in the limit, independent of χ^2 in probability and distributed under H_o as a central chi-square variate with u d.f. and under H_n as a non-central chi-square variate with u d.f. and a non-centrality parameter.

$$\Delta = \underline{c}' A_1' \left[I - A_2 (A_2' A_2)^{-1} A_2' \right] A_1' \underline{c}, \tag{4.2.2}$$

where $\underline{c}' = (c_1, \ldots, c_u)$,

$$\underset{s \times u}{A_1} = \left[\frac{1}{\sqrt{p^o_j}} \left(\frac{\partial p_j}{\partial \theta_{k'}} \right)_o \right], \quad j = 1, \ldots, s \text{ and } k' = 1, \ldots, u, \text{ and}$$

$$\underset{s\times t-u}{A_2} = \left[\frac{1}{\sqrt{p_j^o}} \left(\frac{\partial p_j}{\partial \theta_{k''}} \right) \right]_o , \; j = 1,\ldots,s \text{ and } k'' = u+1,\ldots,t.$$

Proof

Define $y_j (j = 1,\ldots,s)$, $\underline{y}$, y_j^*, $\underline{y}^*$, x_j and $\underline{x}$ as in theorems 1 and 2. Cramer [5] has shown that, in the limit, the equations (2.2.1) may be written

$$\sum_{k=1}^{t} (\theta_k - \theta_k^o) \sum_{j=1}^{s} \frac{1}{p_j^o} \left(\frac{\partial p_j}{\partial \theta_k} \right)_o \left(\frac{\partial p_j}{\partial \theta_m} \right)_o = \sum_{j=1}^{s} \frac{n_j - np_j^o}{np_j^o} \left(\frac{\partial p_j}{\partial \theta_m} \right)_o ,$$

$m = 1,\ldots,t.$

In matrix form, this becomes

$$A'A(\underline{\theta} - \underline{\theta}_o) = \frac{1}{\sqrt{n}} A'\underline{x} \text{ where } \underset{s\times t}{A} = [\underset{u}{A_1} : \underset{t-u}{A_2}] \, s.$$

Similarly, in the limit, the equations (4.2.1) may be written

$$\sum_{k'=1}^{u} (h_{k'} - \theta_{k'}^o) \sum_{j=1}^{s} \frac{1}{p_j^o} \left(\frac{\partial p_j}{\partial \theta_{k'}} \right)_o \left(\frac{\partial p_j}{\partial \theta_{m''}} \right)_o$$

$$+ \sum_{k''=u+1}^{t} (\theta_{k''} - \theta_{k''}^o) \sum_{j=1}^{s} \frac{1}{p_j^o} \left(\frac{\partial p_j}{\partial \theta_{k''}} \right)_o \left(\frac{\partial p_j}{\partial \theta_{m''}} \right)_o$$

$$= \sum_{j=1}^{s} \frac{n_j - np_j^o}{np_j^o} \left(\frac{\partial p_j}{\partial \theta_{m''}} \right)_o , \quad m'' = u+1,\ldots,t,$$

and in matrix notation become

$$A_2'A_1(\underline{h} - \underline{\theta}_1^o) + A_2'A_2(\underline{\theta}_2 - \underline{\theta}_2^o) = \frac{1}{\sqrt{n}} A'_2\underline{x}$$

where $\underline{h}' = (h_1, \ldots, h_u)$, $\underline{\theta}_1' = (\theta_1, \ldots, \theta_u)$ and $\underline{\theta}_2' = (\theta_{u+1}, \ldots, \theta_t)$.

As a solution of equations (2.2.1) we have, in the limit,

$$\hat{\underline{\theta}} - \underline{\theta}_o = \frac{1}{\sqrt{n}} (A'A)^{-1} A\underline{x},$$

and as a solution of equations (4.2.1) we have, in the limit,

$$(\hat{\hat{\underline{\theta}}}_2 - \underline{\theta}_2^o) = \frac{1}{\sqrt{n}} (A_2'A_2)^{-1}A_2'\underline{x} - (A_2'A_2)^{-1}A_2'A_1(\underline{h} - \underline{\theta}_1^o).$$

Cramer [5] has shown that

$$\underline{y} - [\underline{x} - \sqrt{n}\; A(\hat{\underline{\theta}} - \underline{\theta}_o)] \xrightarrow[p]{} 0, \text{ or}$$

$$\underline{y} - [I - A(A'A)^{-1}A'] \underline{x} \xrightarrow[p]{} 0.$$

Similarly,

$$\underline{y}^* - [\underline{x} - \sqrt{n}\, A_1(\underline{h} - \underline{\theta}_1^o) - \sqrt{n}\, A_2(\hat{\hat{\underline{\theta}}}_2 - \underline{\theta}_2^o)] \xrightarrow[p]{} 0, \quad \text{or}$$

$$\underline{y}^* - [I - A_2(A_2'A_2)^{-1} A_2'] (\underline{x} + A_1\underline{c}) \xrightarrow[p]{} 0.$$

We can choose a matrix K such that

$$C = AK, \text{ where } C = [\underset{u}{C_1} : \underset{t-u}{C_2}]\text{s},$$

$$K = \begin{array}{cc} \begin{bmatrix} R & O \\ Q & S \end{bmatrix} & \begin{array}{c} u \\ t-u \end{array} \\ \begin{array}{cc} u & t-u \end{array} & \end{array}$$, where R and S are non-singular, and such that $C_1'C_2 = 0$ (Ogawa [10]). Then $C_1 = A_1R + A_3Q$, $C_2 = A_2S$, and $A_2'A_1R + A_2'A_2Q = 0$ $(\overline{t-u}xu)$. Since K is non-singular we may find its inverse

$$K^{-1} = \begin{array}{cc} \left[\begin{array}{cc} R^{-1} & 0 \\ -S^{-1}QR^{-1} & S^{-1} \end{array}\right] & \begin{array}{c} u \\ t-u \end{array} \\ \begin{array}{cc} u & t-u \end{array} & \end{array} .$$

Then $A = CK^{-1}$, $A_1 = C_1R^{-1} - C_2S^{-1}QR^{-1}$ and $A_2 = C_2S^{-1}$. Thus, in the limit,

$$\underline{y} = [I - CK^{-1}(K'^{-1}C'CK^{-1})^{-1}K'^{-1}C'] \underline{x},$$

$$= [I - C(C'C)^{-1}C'] \underline{x}$$

$$= [I - C_1(C_1'C_1)^{-1}C_1' - C_2(C_2'C_2)^{-1}C_2'] \underline{x}, \text{ and}$$

$$\underline{y}^* = [I - C_2S^{-1}(S'^{-1}C_2'C_2S^{-1})^{-1}S'^{-1}C_2'] \cdot$$

$$[\underline{x} + (C_1R^{-1} - C_2S^{-1}QR^{-1})\underline{c}] ,$$

$$= [I - C_2(C_2'C_2)^{-1}C_2'][\underline{x} + (C_1R^{-1} - C_2S^{-1}QR^{-1})\underline{c}] ,$$

$$= [I - C_2(C_2'C_2)^{-1}C_2'] \underline{x} + C_1R^{-1}\underline{c} .$$

Let $\underline{w} = M'\underline{x}$ where $M = [\underset{1}{\underline{p}} : \underset{u}{M_1} : \underset{t-u}{M_2} : \underset{s-t-1}{M_3}]\, s$

such that M is orthogonal and where

$$\underline{p}' = (\sqrt{p_1^o} , \ldots, \sqrt{p_s^o}) ,$$

$$M_1M_1' = C_1(C_1'C_1)^{-1}C_1', \text{ and}$$

$$M_2M_2' = C_2(C_2'C_2)^{-1}C_2' .$$

Since $\underline{x}$ is asymptotically $N(\underline{0}, I - \underline{pp}')$, $w_1 = 0$ with probability one and $w_2, \ldots, w_s$ are asymptotically independent $N(0, 1)$.

$$\underline{x} = M_1\underline{w}_1 + M_2\underline{w}_2 + M_3\underline{w}_3, \text{ where}$$

$$\underline{w}_1' = (w_2, \ldots, w_{u+1}),\ \underline{w}_2' = (w_{u+2}, \ldots, w_{t+1}) \text{ and}$$

$$\underline{w}_3' = (w_{t+2}, \ldots, w_s).$$

Then, in the limit,

$$\begin{aligned}
\underline{y} &= (I - M_1M_1' - M_2M_2')(M_1\underline{w}_1 + M_2\underline{w}_2 + M_3\underline{w}_3), \\
&= M_1\underline{w}_1 + M_2\underline{w}_2 + M_3\underline{w}_3 - M_1\underline{w}_1 - M_2\underline{w}_3, \\
&= M_3\underline{w}_3, \text{ and}
\end{aligned}$$

$$\begin{aligned}
\underline{y}^* &= (I - M_2M_2')(M_1\underline{w}_1 + M_2\underline{w}_2 + M_3\underline{w}_3) + C_1R^{-1}\underline{c}, \\
&= M_1\underline{w}_1 + M_2\underline{w}_2 + M_3\underline{w}_3 - M_2\underline{w}_2 + C_1R^{-1}\underline{c}, \\
&= M_1\underline{w}_1 + M_3\underline{w}_3 + C_1R^{-1}\underline{c}.
\end{aligned}$$

Therefore,

$$X^2 - \underline{w}_3'M_3'M_3\underline{w}_3 \xrightarrow[p]{} 0 \text{ or } X^2 - \underline{w}_3'\underline{w}_3 \xrightarrow[p]{} 0, \text{ and}$$

$$\begin{aligned}
X_*^2 - [&\underline{w}_1'M_1'M_1\underline{w}_1 + \underline{w}_1'M_1'C_1R^{-1}\underline{c} + \underline{c}'R'^{-1}C_1'M_1\underline{w}_1 \\
&+ \underline{w}_3'M_3'M_3\underline{w}_3 + \underline{w}_3'M_3'C_1R^{-1}\underline{c} + \underline{c}'R'^{-1}C_1'M_3\underline{w}_3 \\
&+ \underline{c}'R'^{-1}C_1'C_1R^{-1}\underline{c}] \xrightarrow[p]{} 0.
\end{aligned}$$

But, $C_1'M_3 = (C_1'C_1)^{1/2}(C_1'C_1)^{-1/2}C_1'M_3 = (C_1'C_1)^{1/2}M_1'M_3 = 0$, and $C_1'M_1 = C_1'C_1(C_1'C_1)^{-1/2} = (C_1'C_1)^{1/2}$, where the square root of a symmetric positive definite matrix is defined as in Ogawa [10].

Thus

$$\begin{aligned}
X_*^2 - [&\underline{w}_1'\underline{w}_1 + 2\underline{c}'R'^{-1}(C_1'C_1)^{1/2}\underline{w}_1 \\
&+ \underline{c}'R'^{-1}C_1'C_1R^{-1}\underline{c} + \underline{w}_3'\underline{w}_3] \xrightarrow[p]{} 0.
\end{aligned}$$

Finally, in the limit, we have that

$$\chi_o^2 = \chi_*^2 - \chi^2 = \underline{w}_1'\underline{w}_1 + 2\underline{c}'R'^{-1}(C_1'C_1)^{1/2}\underline{w}_1$$
$$+ \underline{c}'R'^{-1}C_1'C_1R^{-1}\underline{c} ,$$
$$= [\underline{w}_1 + (C_1'C_1)^{1/2}R^{-1}\underline{c}]' [\underline{w}_1 + (C_1'C_1)^{1/2}R^{-1}\underline{c}] ,$$
$$= \underline{v}'\underline{v} \text{ (say) where } \underline{v}' = (v_1, \ldots, v_u).$$

Since $\chi^2 - \underline{w}_3'\underline{w}_3 \xrightarrow[p]{} 0$ and χ_o^2 does not involve $\underline{w}_3$, it is seen that χ_o^2 is asymptotically independent of χ^2.

Since $\underline{w}_1$ is asymptotically $N(\underline{0}, I)$, $\underline{v}$ is asymptotically normal with mean vector $(C_1'C_1)^{1/2}R^{-1}\underline{c}$ and variance-covariance matrix $I(u\mathrm{x}u)$.

Under H_o, $\underline{c} = \underline{0}$, and χ_o^2 is, in the limit, distributed as a central chi-square variate with u d.f.

Under H_n, χ_o^2 is, in the limit, distributed as a non-central chi-square variate with u d.f. and a non-centrality parameter

$$\Delta = \underline{c}'R'^{-1}(C_1'C_1)R^{-1}\underline{c} .$$

But $R'^{-1}C_1'C_1R^{-1} = R'^{-1}(R'A_1' + Q'A_2')(A_1R + A_2Q)R^{-1}$,

$$= (A_1' + R'^{-1}Q'A_2')(A_1 + A_2QR^{-1}),$$
$$= A_1'A_1 + A_1'A_2QR^{-1} + R'^{-1}Q'A_2'A_1 + R'^{-1}Q'A_2'A_2QR^{-1},$$
$$= A_1'A_1 + A_1'A_2QR^{-1} + R'^{-1}Q'A_2'A_1 - R'^{-1}Q'A_2'A_1,$$
$$= A_1'A_1 + A_1'A_2QR^{-1},$$
$$= A_1'A_1 - A_1'A_2(A_2'A_2)^{-1}A_2'A_1.$$

Therefore, $\Delta = \underline{c}'A_1'[I - A_2(A_2'A_2)^{-1}A_2'] A_1\underline{c}.$

4.3 Theorem 6

Given a single-multinomial sample (1.1.1), s functions $p_j(\underline{\theta})$ of t unknown parameters $\theta_1,\ldots,\theta_t$ as in (1.3.1) and u functions $f_{m'}(\underline{\theta})$ as in (1.3.2), suppose it is also given that $p_j = p_j(\underline{\theta})$ for $j = 1,\ldots,s$. Consider the hypothesis

$$H_o\colon f_{m'}(\underline{\theta}) = 0 \text{ for } m' = 1,\ldots,u,$$

and the alternative

$$H_n^*\colon f_{m'}(\underline{\theta}) = c_{m'}/\sqrt{n} \quad \text{where not all } c_{m'} = 0.$$

In order to obtain the asymptotic power we consider the simple alternative

$$H_n\colon f_{m'}(\underline{\theta}_o) = c_{m'}/\sqrt{n} \quad \text{where } \underline{\theta}_o \text{ is supposed to be the unknown true parameter point in } \Omega.$$

Let $\hat{\underline{\theta}}$ be the unique consistent solution of equations (2.2.1) and let χ^2 be the corresponding test statistic defined as in (2.2.2).

Let $\hat{\hat{\underline{\theta}}}$ be the unique consistent solution of the equations (2.4.1) and let χ^2_* be the corresponding test statistic defined as in (2.4.2).

Then $\chi^2_o = \chi^2_* - \chi^2$ is, in the limit, independent of χ^2 in probability and distributed under H_o as a central chi-square variate with u d.f. and under H_n as a non-central chi-square variate with u d.f. and a non-centrality parameter

$$\Delta = \underline{c}' F_1'^{-1} A_1' [I - A_*(A_*'A_*)^{-1}A_*'] A_1 F_1^{-1} \underline{c},$$

where $\underline{c}' = (c_1,\ldots,c_u)$,

$$\underset{u\times u}{F_1} = \left(\frac{\partial f_{m'}}{\partial \theta_{k'}}\right)_o ,$$

$$\underset{s\times u}{A_1} = \left[\frac{1}{\sqrt{p_j^o}} \left(\frac{\partial\, p_j}{\partial\, \theta_{k'}} \right)_o \right] ,$$

$$\underset{s\times \overline{t-u}}{A_*} = \left[\frac{1}{\sqrt{p_j^o}} \left(\frac{\partial\, p_j^*}{\partial\, \theta_{k''}} \right)_o \right] ,$$

and where $\theta_{k'}$ denotes the u θ_k made dependent under H_o, $\theta_{k''}$ denotes the t-u θ_k independent under H_o and p_j^* is $p_j(\underline{\theta})$ expressed in terms of the t-u independent θ_k under H_o. For convenience of notation we may assume, without loss of generality, that $k' = 1,\ldots,u$ and $k'' = u+1,\ldots,t$.

<u>Proof</u>

Denote by $F = [\underset{u}{F_1} : \underset{t-u}{F_2}]$ u the matrix $\left(\frac{\partial\, f_{m'}}{\partial\, \theta_k} \right)_o$ where $F_1 = \left(\frac{\partial\, f_{m'}}{\partial\, \theta_{k'}} \right)_o$ and $F_2 = \left(\frac{\partial\, f_{m'}}{\partial\, \theta_{k''}} \right)_o$. Since rank $(F) = u$, we may assume without loss of generality that $|F_1| \neq 0$. Define $f_{m''}(\underline{\theta}) = \theta_{m''}$ for $m'' = u+1,\ldots,t$ so that $f_m(\underline{\theta})$ where $m = 1,\ldots,t$ are continuously differentiable and the jacobian of the transformation from $\underline{\theta}' = (\theta_1,\ldots,\theta_t)$ to $(f_1(\underline{\theta}),\ldots,f_t(\underline{\theta}))$ at $\underline{\theta}_o$ is given by $|F_1|^{-1}$.

By a well known theorem on the inversion of a transformation (Courant [4], p. 152) it is seen that, in a neighborhood Θ of $\underline{\theta}_o$, the system of equations $a_m = f_m(\underline{\theta})$ for $m = 1,\ldots,t$ has a unique inverse

$$\theta_k = g_k(a_1,\ldots,a_t) \text{ for } k = 1,\ldots,t,$$

and if $a_m^o = f_m(\theta_1^o,\ldots,\theta_t^o)$, then in a neighborhood Φ of $\underline{a}^{o\prime} = (a_1^o,\ldots,a_t^o)$

the inverse functions g_k possess continuous first and second order derivatives and the first order derivatives are given by

$$\left(\frac{\partial \theta_k}{\partial a_m}\right) = \left(\frac{\partial a_m}{\partial \theta_k}\right)^{-1} \quad \text{for k and m} = 1,\ldots,t.$$

Put $p_j(\theta_1,\ldots,\theta_t) = p_j(g_1(\underline{a}),\ldots,g_t(\underline{a})) = q_j(a_1,\ldots,a_t)$ for $j = 1,\ldots,s$.

Mitra [7] proves that the equations

$$\sum_{j=1}^{s} \frac{n_j - nq_j}{q_j} \cdot \frac{\partial q_j}{\partial a_m} = 0 \text{ for } m = 1,\ldots,t$$

have a unique solution $\hat{\underline{a}}$ such that $\hat{\underline{a}} \longrightarrow \underline{a}^o$ in probability as $n \longrightarrow \infty$ and that $\hat{\underline{\theta}} = \underline{g}(\hat{\underline{a}})$ is asymptotically a solution of the equations (2.2.1). Mitra also proves that the equations

$$\sum_{j=1}^{s} \frac{n_j - nq_j}{q_j} \cdot \frac{\partial q_j}{\partial a_m} + \sum_{m'=1}^{u} \mu_{m'} \frac{\partial a_{m'}}{\partial a_m} = 0,$$

$$m = 1,\ldots,t,\ a_{m'} = 0,\ m' = 1,\ldots,u,$$

have a unique solution $\hat{\hat{\underline{a}}}' = [\underset{u}{\underline{0}'} : \underset{t-u}{\hat{\hat{\underline{a}}}_2'}]1$, where $\hat{\hat{\underline{a}}}_2' = (\hat{\hat{a}}_{u+1},\ldots,\hat{\hat{a}}_t)$, such that $\hat{\hat{\underline{a}}}' \longrightarrow [\underline{0}' : \underline{a}_2^{o'}]$ in probability as $n \longrightarrow \infty$ and that $\hat{\hat{\underline{\theta}}} = \underline{g}(\hat{\hat{\underline{a}}})$ is asymptotically a solution of the equations (2.4.1).

Thus, it is seen that the problem of testing the hypothesis H_o: $f_{m'}(\underline{\theta}) = 0$ formally reduces to that of testing

$$H_o\text{: } a_{m'} = 0 \text{ for } m' = 1,\ldots,u$$

where $\underline{a}^{o'} = (a_1^o,\ldots,a_t^o)$ is now the true parameter point with respect to the new frame of reference $a_1,\ldots,a_t$.

Define $D = \underset{u}{[D_1} : \underset{t-u}{D_2}]\,s$ where $D_1 = \left[\frac{1}{\sqrt{q_j^o}} \left(\frac{\partial q_j}{\partial a_{m'}} \right)_o \right]$,

$m' = 1, \ldots, u$, and $D_2 = \left[\frac{1}{\sqrt{q_j^o}} \left(\frac{\partial q_j}{\partial a_{m''}} \right)_o \right]$, $m'' = u+1, \ldots, t$. Then

by theorem 5, we have that $\chi_o^2 = \chi_*^2 - \chi^2$ is, in the limit, independent of χ^2 in probability and distributed under H_o as a central chi-square variate with u d.f. and under H_n as a non-central chi-square variate with u d.f. and a non-centrality parameter

$$\Delta = \underline{c}'D_1'[I - D_2(D_2'D_2)^{-1}D_2']\,D_1\underline{c}.$$

Let $A = \underset{u}{[A_1} : \underset{t-u}{A_2}]\,s$ where A_1 is defined as before and

$A_2 = \left[\frac{1}{\sqrt{p_j^o}} \left(\frac{\partial p_j}{\partial \theta_{k''}} \right)_o \right]$. It is seen that $A = D \cdot \left(\frac{\partial a_m}{\partial \theta_k} \right)_o$

where $m; k = 1, \ldots, t$, or $\underset{u}{[A_1} : \underset{t-u}{A_2}]\,s =$

$$s\,\underset{u}{[D_1} : \underset{t-u}{D_2}] \cdot \begin{array}{cc} \left[\begin{array}{cc} \left(\frac{\partial a_{m'}}{\partial \theta_{k'}} \right)_o & \left(\frac{\partial a_{m'}}{\partial \theta_{k''}} \right)_o \\ \left(\frac{\partial a_{m''}}{\partial \theta_{k'}} \right)_o & \left(\frac{\partial a_{m''}}{\partial \theta_{k''}} \right)_o \end{array} \right] & \begin{array}{c} u \\ t-u \end{array} \\ \begin{array}{cc} u & \quad t-u \end{array} & \end{array},$$

$$= s\,\underset{u}{[D_1} : \underset{t-u}{D_2}] \cdot \begin{array}{cc} \left[\begin{array}{cc} F_1 & F_2 \\ 0 & I \end{array} \right] & \begin{array}{c} u \\ t-u \end{array} \\ \begin{array}{cc} u & t-u \end{array} & \end{array}.$$

Inverting this, we get

$$[D_1 : D_2] = [A_1 : A_2] \cdot \begin{bmatrix} F_1 & F_2 \\ 0 & I \end{bmatrix}^{-1},$$

$$= [A_1 : A_2] \begin{bmatrix} F_1^{-1} & -F_1^{-1} \cdot F_2 \\ 0 & I \end{bmatrix}$$

or $D_1 = A_1F_1^{-1}$ and $D_2 = A_2 - A_1F_1^{-1}F_2$. However, from a theorem on jacobians due to Roy [17],

$$A_* = \left[\frac{1}{\sqrt{p_j^o}} \left(\frac{\partial\, p_j^*}{\partial\, \theta_{k''}} \right)_o \right],$$

$$= \left[\frac{1}{\sqrt{p_j^o}} \left(\frac{\partial\, p_j}{\partial\, \theta_{k''}} \right)_o \right] - \left[\frac{1}{\sqrt{p_j^o}} \left(\frac{\partial\, p_j}{\partial\, \theta_{k'}} \right)_o \right] \left(\frac{\partial\, f_{m'}}{\partial\, \theta_{k'}} \right)_o^{-1} \left(\frac{\partial\, f_{m'}}{\partial\, \theta_{k''}} \right)_o,$$

$$= A_2 - A_1F_1^{-1}F_2,$$

$$= D_2.$$

Therefore the non-centrality parameter may be expressed as

$$\Delta = \underline{c}'F_1'^{-1}A_1'[I - A_*(A_*'A_*)^{-1}A_*']A_1F_1^{-1}\underline{c}.$$

4.4 Asymptotic Independence of Tests

Suppose we have two hypotheses of the type considered in theorem 6:

H_{01}: $f_{1m_1'}(\theta_1,\ldots,\theta_t) = 0$ where $m_1' = 1,\ldots,u_1$, and

H_{02}: $f_{2m_2'}(\theta_1,\ldots,\theta_t) = 0$ where $m_2' = 1,\ldots,u_2$

with $u_1 + u_2 \leq t$.

Let χ^2_{01} and χ^2_{02} be the test statistics for H_{01} and H_{02} respectively and define $\underset{sxt}{A} = \left[\frac{1}{\sqrt{p_j^o}} \left(\frac{\partial p_j}{\partial \theta_k} \right)_o \right]$. Define A_* for each hypothesis and let this be denoted by A_{1*} and A_{2*} for H_{01} and H_{02} respectively.

When H_{01} is true

$$\chi^2_{01} - \underline{x}' [A(A'A)^{-1}A' - A_{1*}(A_{1*}'A_{1*})^{-1}A_{1*}'] \underline{x} \xrightarrow[p]{} 0,$$

and when H_{02} is true

$$\chi^2_{02} - \underline{x}' [A(A'A)^{-1}A' - A_{2*}(A_{2*}'A_{2*})^{-1}A_{2*}'] \underline{x} \longrightarrow 0.$$

Then, by the theorem in section 1.5, in the limit, when H_{01} and H_{02} are true, χ^2_{01} and χ^2_{02} are independent in probability if and only if

$$[A(A'A)^{-1}A' - A_{1*}(A_{1*}'A_{1*})^{-1}A_{1*}'] \cdot [I - \underline{pp}'] \cdot$$

$$[A(A'A)^{-1}A' - A_{2*}(A_{2*}'A_{2*})A_{2*}'] = 0.$$

But $\underline{p}'A = 0$, $\underline{p}'A_{1*} = 0$ and $\underline{p}'A_{2*} = 0$, and the condition becomes

$$[A(A'A)^{-1}A' - A_{1*}(A_{1*}'A_{1*})^{-1}A_{1*}'][A(A'A)^{-1}A' - A_{2*}(A_{2*}'A_{2*})^{-1}A_{2*}'] = 0.$$

By simple algebraic manipulation, this condition may be shown to be identical to a necessary and sufficient condition derived by Mitra [7]. That is

$$[I - A_{1*}(A_{1*}'A_{1*})^{-1}A_{1*}'] [I - A_{2*}(A_{2*}'A_{2*})^{-1}A_{2*}'] = I - A(A'A)^{-1}A'.$$

4.5 A Special Case

Of particular interest, and perhaps widest applicability, is the case in which, for a single-multinomial sample, we wish to test a hypothesis of the form $f_m(p_1,\dots,p_s) = 0$ for $m = 1,\dots,u \leq s-1$.

In this case, we are given that $p_1 = p_1,\dots,p_{s-1} = p_{s-1}$ and $p_s = 1 - p_1 - \dots - p_{s-1}$. The hypothesis is

$$H_o\colon f_m(p_1,\dots,p_s) = 0 \text{ for } m = 1,\dots,u \leq s-1,$$

and the alternative is

$$H_n\colon f_m(p_1,\dots,p_s) = c_m / \sqrt{n} \text{ where not all } c_m = 0.$$

Here the solution of the equations (2.2.1) is

$$\hat{p}_j = \frac{n_j}{n} \text{ for } j = 1,\dots,s \text{ and } \chi^2 = \sum_{j=1}^{s} \frac{(n_j - n\hat{p}_j)^2}{n\hat{p}_j} = 0.$$

The equations (2.4.1) may be written

$$\sum_{j=1}^{s} \frac{n_j - np_j}{p_j} \cdot \frac{\partial p_j}{\partial p_k} + \sum_{m=1}^{u} \lambda_m \frac{\partial f_m^*}{\partial p_k} = 0,$$

$$k = 1,\dots,s-1,\ f_m^*(p_1,\dots,p_{s-1}) = 0,\ m = 1,\dots,u,$$

where f_m^* is f_m expressed in terms of the s-1 independent parameters $p_1, p_2, \dots, p_{s-1}$. Then the equations (2.4.1) reduce to

$$\frac{n_j - np_j}{p_j} - \frac{n_s - np_s}{p_s} + \sum_{m=1}^{u} \lambda_m \left(\frac{\partial f_m}{\partial p_j} - \frac{\partial f_m}{\partial p_s} \right) = 0,$$

$$j = 1,\dots,s-1,\ f_m(p_1,\dots,p_s) = 0,\ m = 1,\dots,u,$$

$$\sum_{j=1}^{s} p_j = 1.$$

These equations have a unique consistent solution $\hat{\hat{p}}_j$ and we may form

$$\chi_*^2 = \sum_{j=1}^{s} \frac{(n_j - n\hat{\hat{p}}_j)^2}{n\hat{\hat{p}}_j} \quad .$$

It is seen in this case that

$$\chi_o^2 = \chi_*^2 \quad .$$

We can assume, without loss of generality, that $p_1, \ldots, p_u$ are made dependent under H_o and $p_{u+1}, \ldots, p_{s-1}$ remain independent under H_o. Then

$$\underset{s\times u}{A_1} = \begin{array}{c@{\;}l} \left[\begin{array}{cccc} \frac{1}{\sqrt{p_1^o}} & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sqrt{p_u^o}} \\ \hline & \mathbf{0} & & \\ \hline -\frac{1}{\sqrt{p_s^o}} & -\frac{1}{\sqrt{p_s^o}} & \cdots & -\frac{1}{\sqrt{p_s^o}} \end{array}\right] & \begin{array}{l} u \\ \\ s-1-u \\ 1 \end{array} \end{array}$$

$$\text{and } \underset{s\times s-1-u}{A_2} = \begin{array}{c@{\;}l} \left[\begin{array}{cccc} & \mathbf{0} & & \\ \hline \frac{1}{\sqrt{p_{u+1}^o}} & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sqrt{p_{s-1}^o}} \\ \hline -\frac{1}{\sqrt{p_s^o}} & -\frac{1}{\sqrt{p_s^o}} & \cdots & -\frac{1}{\sqrt{p_s^o}} \end{array}\right] & \begin{array}{l} u \\ \\ s-1-u \\ 1 \end{array} \end{array}$$

Note that A_1 and A_2 are fixed for all hypotheses of this type.

$$F_1 = \left(\frac{\partial f^*_m}{\partial p_{k'}}\right)_o = \left(\frac{\partial f_m}{\partial p_{k'}} - \frac{\partial f_m}{\partial p_s}\right)_o \quad \text{for } m = 1,\dots,u \text{ and } k' = 1,\dots,u,$$

and

$$F_2 = \left(\frac{\partial f^*_m}{\partial p_{k''}}\right)_o = \left(\frac{\partial f_m}{\partial p_{k''}} - \frac{\partial f_m}{\partial p_s}\right)_o \quad \text{for } m = 1,\dots,u \text{ and}$$

$$k'' = u+1,\dots,s-1.$$

Then $A_* = A_2 - A_1 F_1^{-1} F_2$ and we may form the non-centrality parameter Δ.

4.6 A Specific Example

As a specific example of the preceding special case consider the hypothesis that $p_1 = p_2 = \dots = p_{u+1}$ where $u < s-1$.

Then the hypothesis is

H_o: $p_m - p_{u+1} = 0$ for $m = 1,\dots,u < s-1$,

and the alternative is

H_n: $p_m - p_{u+1} = c_m / \sqrt{n}$ where not all $c_m = 0$.

The modified chi-square minimum equations subject to H_o are

$$\frac{n_1 - np_1}{p_1} - \frac{n_s - np_s}{p_s} + \lambda_1 = 0,$$

$$\vdots$$

$$\frac{n_u - np_u}{p_u} - \frac{n_s - np_s}{p_s} + \lambda_u = 0,$$

$$\frac{n_{u+1} - np_{u+1}}{p_{u+1}} - \frac{n_s - np_s}{p_s} - \lambda_1 - \dots - \lambda_u = 0,$$

$$\frac{n_{u+2} - np_{u+2}}{p_{u+2}} - \frac{n_s - np_s}{p_s} = 0,$$

$$\vdots$$

$$\frac{n_{s-1} - np_{s-1}}{p_{s-1}} - \frac{n_s - np_s}{p_s} = 0,$$

$$p_1 - p_{u+1} = 0,$$

$$\vdots$$

$$p_u - p_{u+1} = 0,$$

$$p_1 + \dots + p_s = 1.$$

The solution of these equations is

$$\hat{\hat{p}}_1 = \dots = \hat{\hat{p}}_{u+1} = \frac{n_1 + \dots + n_{u+1}}{n(u+1)} ,$$

$$\hat{\hat{p}}_{u+2} = \frac{n_{u+2}}{n}, \dots, \hat{\hat{p}}_s = \frac{n_s}{n} , \text{ and}$$

$$\chi_o^2 = \chi_*^2 = \sum_{j=1}^{u+1} \frac{\left(n_j - \frac{n_1 + \dots + n_{u+1}}{u+1}\right)^2}{\frac{n_1 + \dots + n_{u+1}}{u+1}} .$$

A_1 and A_2 are as in section 4.5.

$$F_1 = I_{(u \times u)} \text{ and } \underset{u \times \overline{s-1-u}}{F_2} = \begin{bmatrix} -1 & 0 & \dots & 0 \\ -1 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ -1 & 0 & \dots & 0 \end{bmatrix} .$$

$$\underset{s\times \overline{s-u-1}}{A_*} = \begin{bmatrix} \frac{1}{\sqrt{p_1^o}} & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ \frac{1}{\sqrt{p_u^o}} & 0 & \cdots & 0 \\ \frac{1}{\sqrt{p_{u+1}^o}} & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sqrt{p_{s-1}^o}} \\ -\frac{u+1}{\sqrt{p_s^o}} & -\frac{1}{\sqrt{p_s^o}} & \cdots & -\frac{1}{\sqrt{p_s^o}} \end{bmatrix} .$$

$$\Delta = \frac{1}{p_{u+1}^o} \left[\sum_{m=1}^{u} c_m^2 - \frac{1}{u+1} \left(\sum_{m=1}^{u} c_m \right)^2 \right] .$$

If $\sum_{m=1}^{u} c_m = 0$, $\quad \Delta = \frac{1}{p_{u+1}^o} \sum_{m=1}^{u} c_m^2$.

Thus, χ_o^2 is, in the limit, distributed under H_o as a central chi-square variate with u d.f. and under H_n as non-central chi-square variate with u d.f. and a non-centrality parameter given by Δ above.

Chapter V

ANALOGUES TO NORMAL UNIVARIATE AND MULTIVARIATE ANALYSIS OF VARIANCE TESTS FOR PRODUCT-MULTINOMIAL SAMPLES

5.1 Introduction

For product-multinomial samples, tests of hypotheses of the type given in (1.4.6) are called analogous to normal univariate and multivariate analysis of variance tests for the same reasons as given in section 4.1. As in the previous chapter few examples of this type of test are found in practice, but two applications are given here. Because we are here dealing with the case where some marginals are 'variates' and some are 'ways of classification,' this situation most closely approximates the kind of sampling scheme found in normal analysis of variance. Here we have the advantage that we need not be restricted to linear models and linear hypotheses.

Two theorems are given, the first being preliminary and the second very general. A special case is given which seems to be most applicable in practice and two examples of specific tests in this special case are included. In addition, a necessary and sufficient condition for the asymptotic independence in probability of two tests is developed.

5.2 Theorem 7

Given a product-multinomial sample (1.1.2) and rs functions $p_{ij}(\underline{\theta})$ of t unknown parameters $\theta_1,\ldots,\theta_t$ as in (1.4.1), suppose it is also given that $p_{ij} = p_{ij}(\underline{\theta})$ for $i = 1,\ldots,r$ and $j = 1,\ldots,s$. Consider the hypothesis

$$H_o\colon \theta_{k'} = h_{k'},\ k' = 1,\ldots,u < t,$$

and the alternative

$H_n\colon \theta_{k'} = h_{k'} + c_{k'} / \sqrt{n} = \theta^o_{k'}$ (say) where not all $c_{k'} = 0$ and where $\underline{\theta}'_o = (\theta^o_1,\ldots,\theta^o_u,\ \theta^o_{u+1},\ldots,\theta^o_t)$ is supposed to be the unknown true parameter point in $\underline{(\ )}$.

Let $\hat{\underline{\theta}}$ be the unique consistent solution of the equations given by (3.2.1) and let χ^2 be the corresponding test statistic.

Let $\hat{\hat{\underline{\theta}}}$ be the unique consistent solution of the equations, subject to H_o,

$$(5.2.1)\qquad \sum_{i=1}^{r}\sum_{j=1}^{s} \frac{n_{ij}-n_{io}p_{ij}(\underline{\theta})}{p_{ij}(\underline{\theta})} \cdot \frac{\partial\, p_{ij}(\underline{\theta})}{\partial\, \theta_k} + \sum_{k'=1}^{u} \lambda_{k'} \frac{\partial\, \theta_{k'}}{\partial\, \theta_k} = 0,$$

$$k = 1,\ldots,t,\ \theta_{k'} = h_{k'},\ k' = 1,\ldots,u,$$

and let χ^2_* be the corresponding test statistic.

Then $\chi^2_o = \chi^2_* - \chi^2$ is, in the limit, independent of χ^2 in probability and distributed under H_o as a central chi-square variate with u d.f. and under H_n as a non-central chi-square variate with u d.f. and a non-centrality parameter

$$\Delta = \underline{c}'B_1'\left[I - B_2(B_2'B_2)^{-1}B_2'\right] B_1\underline{c}$$

where $\underline{c}' = (c_1,\ldots,c_u)$,

$$\underset{r s \times u}{B_1} = \left[\frac{1}{\sqrt{p^o_{ij}}}\sqrt{\frac{n_{io}}{n}}\left(\frac{\partial\, p_{ij}}{\partial\, \theta_{k'}}\right)\right]_o \qquad \text{where } k' = 1,\ldots,u,$$

$$\underset{rsx\overline{t-u}}{B_2} = \left[\frac{1}{\sqrt{p^o_{ij}}} \sqrt{\frac{n_{io}}{n}} \left(\frac{\partial\, p_{ij}}{\partial\, \theta_{k''}} \right) \right]_o \quad \text{and where } k'' = u+1,\ldots,t.$$

<u>Proof</u>

Define y_{ij}, $\underline{y}$, y^*_{ij}, $\underline{y}^*$, x_{ij} and $\underline{x}$ as in theorems 3 and 4. In a manner analogous to that of theorem 5, Ogawa [11] has shown that the equations (3.2.1) become, in the limit,

$$B'B(\underline{\theta} - \underline{\theta}_o) = \frac{1}{\sqrt{n}} B'\underline{x}$$

where $B = [B_1 : B_2]$. Similarly the equations (5.2.1) become, in the limit,

$$B_2'B_1(\underline{h} - \underline{\theta}_1^o) + B_2'B_2(\underline{\theta}_2 - \underline{\theta}_2^o) = \frac{1}{\sqrt{n}} B_2'\underline{x}$$

where $\underline{h}' = (h_1,\ldots,h_u)$, $\underline{\theta}_1' = (\theta_1,\ldots,\theta_u)$ and $\underline{\theta}_2' = (\theta_{u+1},\ldots,\theta_t)$.

As a solution of equations (3.2.1) we have, in the limit,

$$(\hat{\underline{\theta}} - \underline{\theta}_o) = \frac{1}{\sqrt{n}} (B'B)^{-1}B'\underline{x}$$

and as a solution of equations (5.2.1), in the limit,

$$(\hat{\hat{\underline{\theta}}}_2 - \underline{\theta}_2^o) = \frac{1}{\sqrt{n}} (B_2'B_2)^{-1}B_2'\underline{x} - (B_2'B_2)^{-1}B_2'B_1(\underline{h} - \underline{\theta}_1^o).$$

Ogawa [11] has shown that $\underline{y} - [\underline{x} - \sqrt{n}\, B(\hat{\underline{\theta}} - \underline{\theta}_o)] \xrightarrow[p]{} 0$, or $\underline{y} - [I - B(B'B)^{-1}B'] \underline{x} \xrightarrow[p]{} 0$. Similarly,

$$\underline{y}^* - [\underline{x} - \sqrt{n}\, B_1(\underline{h} - \underline{\theta}_1^o) - \sqrt{n}\, B_2(\hat{\hat{\underline{\theta}}}_2 - \underline{\theta}_2^o)] \xrightarrow[p]{} 0,$$

or

$$\underline{y}^* - [I - B_2(B_2'B_2)^{-1}B_2'] (\underline{x} + B_1\underline{c}) \xrightarrow[p]{} 0.$$

We can find a matrix K such that $C = BK$ where $C = \begin{array}{c} [\, C_1 \;:\; C_2 \,]\, rs \\ \;\; u \qquad t-u \end{array}$

and $K = \begin{array}{cc} \begin{bmatrix} R & 0 \\ Q & S \end{bmatrix} & \begin{array}{c} u \\ t-u \end{array} \\ \begin{array}{cc} u & t-u \end{array} & \end{array}$, where R and S are non-singular and such that

$C_1'C_2 = 0$.

Then, as in theorem 5,

$$\underline{y} - [\, I - C_1(C_1'C_1)^{-1}C_1' - C_2(C_2'C_2)^{-1}C_2' \,]\, \underline{x} \xrightarrow[p]{} 0, \text{ and}$$

$$\underline{y}^* - [\, I - C_2(C_2'C_2)^{-1}C_2' \,]\, \underline{x} - C_1R^{-1}\underline{c} \xrightarrow[p]{} 0.$$

Let $\underline{w} = M'\underline{x}$ where $M = [\, \underset{r}{P} : \underset{u}{M_1} : \underset{t-u}{M_2} : \underset{rs-r-t}{M_3} \,]\, rs$, where P is defined as in theorem 3, such that M is orthogonal,

$$M_1M_1' = C_1(C_1'C_1)^{-1}C_1', \text{ and}$$

$$M_2M_2' = C_2(C_2'C_2)^{-1}C_2'.$$

Since $\underline{x}$ is asymptotically $N(\underline{0}, I - PP)$, $w_1, \ldots, w_r$ equal zero with probability one and $w_{r+1}, \ldots, w_{rs}$ are asymptotically independent $N(0, 1)$. Thus $\underline{x} = M_1\underline{w}_1 + M_2\underline{w}_2 + M_3\underline{w}_3$, where $\underline{w}_1' = (w_{r+1}, \ldots, w_{r+u})$, $\underline{w}_2' = (w_{r+u+1}, \ldots, w_{r+t})$ and $\underline{w}_3' = (w_{r+t+1}, \ldots, w_{rs})$.

Then, as in theorem 5,

$$\underline{y} - M_3\underline{w}_3 \xrightarrow[p]{} 0 \text{ and}$$

$$\underline{y}^* - M_1\underline{w}_1 - M_3\underline{w}_3 - C_1R^{-1}\underline{c} \xrightarrow[p]{} 0, \text{ and therefore}$$

$$\chi^2 - \underline{w}_3'\underline{w}_3 \xrightarrow[p]{} 0 \text{ and}$$

$$\chi_*^2 - \underline{w}_3'\underline{w}_3 - \underline{w}_1'\underline{w}_1 - 2\underline{c}'R'^{-1}(C_1'C_1)^{1/2}\underline{w}_1$$

$$- \underline{c}'R'^{-1}C_1'C_1R^{-1}\underline{c} \xrightarrow[p]{} 0.$$

Thus, in the limit,

$$\chi_o^2 = \chi_*^2 - \chi^2 = [\underline{w}_1 + (C_1'C_1)^{1/2}R^{-1}\underline{c}]' \cdot$$

$$[\underline{w}_1 + (C_1'C_1)^{1/2}R^{-1}\underline{c}],$$

$$= \underline{v}'\underline{v} \text{ (say) where } \underline{v}' = (v_1, \ldots, v_u) .$$

Since $\chi^2 - \underline{w}_3'\underline{w}_3 \xrightarrow[p]{} 0$ while χ_o^2 does not involve $\underline{w}_3$, it is seen that χ_o^2 is asymptotically independent of χ^2 in probability whether H_o or H_n is true.

It is seen that $\underline{v}$ is asymptotically $N[(C_1'C_1)^{1/2}R^{-1}\underline{c};\ I]$. But under H_o, $\underline{c} = \underline{0}$.

Thus, under H_o, χ_o^2 is, in the limit, distributed as a central chi-square variate with u d.f., and under H_n as a non-central chi-square variate with u d.f. and a non-centrality parameter $\Delta = \underline{c}'R'^{-1}(C_1'C_1)R^{-1}\underline{c}$.

But, as in theorem 5,

$$R'^{-1}(C_1'C_1)R^{-1} = B_1'B_1' - B_1'B_2(B_2'B_2)^{-1}B_2'B_1 .$$

Therefore, $$\Delta = \underline{c}'B_1'[I - B_2(B_2'B_2)^{-1}B_2']B_1\underline{c} .$$

5.3 <u>Theorem 8</u>

Given a product-multinomial sample (1.1.2), rs functions $p_{ij}(\underline{\theta})$ of t unknown parameters $\theta_1, \ldots, \theta_t$ as in (1.4.1) and u functions $f_{m'}(\underline{\theta})$ as in (1.4.2), suppose it is also given that $p_{ij} = p_{ij}(\underline{\theta})$ for $i = 1, \ldots, r$ and $j = 1, \ldots, s$. Consider the hypothesis

$$H_o: f_{m'}(\underline{\theta}) = 0 \text{ for } m' = 1,\ldots,u,$$

and the alternative

$$H_n^*: f_{m'}(\underline{\theta}) = c_{m'} / \sqrt{n} \text{ where not all } c_{m'} = 0.$$

In order to obtain the asymptotic power we consider the simple alternative

$$H_n: f_{m'}(\underline{\theta}_o) = c_{m'} / \sqrt{n} \text{ where } \underline{\theta}_o \text{ is supposed to be}$$
the unknown true parameter point in $\underline{(\overline{\;})}$.

Let $\hat{\underline{\theta}}$ be the unique consistent solution of equations (3.2.1) and let X^2 be the corresponding test statistic defined as in (3.2.2).

Let $\hat{\hat{\underline{\theta}}}$ be the unique consistent solution of the equations (3.3.1) and let X_*^2 be the corresponding test statistic defined as in (3.3.2).

Then $X_o^2 = X_*^2 - X^2$ is, in the limit, independent of X^2 in probability and distributed under H_o as a chi-square variate with u d.f. and under H_n as a non-central chi-square variate with u d.f. and a non-centrality parameter

$$\Delta = \underline{c}' F_1'^{-1} B_1' [I - B_*(B_*'B_*)^{-1}B_*'] B_1 F_1^{-1} \underline{c}$$

where $\underline{c}' = (c_1,\ldots,c_u)$,

$$\underset{u \times u}{F_1} = \left(\frac{\partial f_{m'}}{\partial \theta_{k'}}\right)_o ,$$

$$\underset{rs \times u}{B_1} = \left[\frac{1}{\sqrt{p_{ij}^o}} \sqrt{\frac{n_{io}}{n}} \left(\frac{\partial p_{ij}}{\partial \theta_{k'}}\right)_o\right] , \text{ and}$$

$$\underset{rs \times \overline{t-u}}{B_*} = \left[\frac{1}{\sqrt{p_{ij}^o}} \sqrt{\frac{n_{io}}{n}} \left(\frac{\partial p_{ij}^*}{\partial \theta_{k''}}\right)_o\right] ,$$

and where $\theta_{k'}$ denotes the u θ_k made dependent under H_o, $\theta_{k''}$ denotes the t-u θ_k remaining independent under H_o and p^*_{ij} is $p_{ij}(\underline{\theta})$ expressed in terms of the t-u independent θ_k under H_o. For convenience of notation we may assume, without loss of generality, that $k' = 1,\ldots,u$ and $k'' = u+1,\ldots,t$.

Proof

Define $F = [F_1 : F_2]$ as in theorem 6 and let $f_{m''}(\theta_1,\ldots,\theta_t) = \theta_{m''}$ for $m'' = u+1,\ldots,t$ so that the $f_m(\underline{\theta})$, where $m = 1,\ldots,t$, are continuously differentiable and the jacobian of the transformation from $\underline{\theta}' = (\theta_1,\ldots,\theta_t)$ to $(f_1(\underline{\theta}),\ldots,f_t(\underline{\theta}))$ at $\underline{\theta}_o$ is given by $|F_1|^{-1}$.

As in theorem 6, if we put $a_m = f_m(\underline{\theta})$ for $m = 1,\ldots,t$ and define $q_{ij}(a_1,\ldots,a_t) = p_{ij}(\theta_1,\ldots,\theta_t)$, the problem is reduced to that of testing the hypothesis H_o: $a_{m'} = 0$, for $m' = 1,\ldots,u$, where $\underline{a}_o = (a_1^o,\ldots,a_t^o)$ is now the true parameter point with respect to the new frame of reference $a_1,\ldots,a_t$.

Define $G = [\underset{u}{G_1} : \underset{t-u}{G_2}]$ rs, where

$$\underset{rs\times u}{G_1} = \left[\frac{1}{\sqrt{q^o_{ij}}}\sqrt{\frac{n_{io}}{n}}\left(\frac{\partial q_{ij}}{\partial a_{m'}}\right)_o\right] \text{ where } m' = 1,\ldots,u \text{ and}$$

$$\underset{rs\times t-u}{G_2} = \left[\frac{1}{\sqrt{q^o_{ij}}}\sqrt{\frac{n_{io}}{n}}\left(\frac{\partial q_{ij}}{\partial a_{m''}}\right)_o\right] \text{ where } m'' = u+1,\ldots,t.$$

Then, by theorem 7, $X_o^2 = X_*^2 - X^2$ is, in the limit, independent of X^2 in probability and distributed under H_o as a central chi-square variate with u d.f. and under H_n as a non-central chi-square variate with u d.f. and a non-centrality parameter

$$\Delta = \underline{c}'G_1'[I - G_2(G_2'G_2)^{-1}G_2'] G_1\underline{c} .$$

Let $B = [\underset{u}{B_1} : \underset{t-u}{B_2}]$rs where B_1 is defined as before and

$$B_2 = \left[\frac{1}{\sqrt{p^o_{ij}}} \sqrt{\frac{n_{io}}{n}} \left(\frac{\partial p_{ij}}{\partial \theta_{k''}}\right)_o\right]$$ where $k'' = u+1,\ldots,t$. It is seen that

$$B = G \left(\frac{\partial a_m}{\partial \theta_k}\right)_o$$ for m and $k = 1,\ldots,t$ or

$$[B_1 : B_2] = [G_1 : G_2] \cdot \begin{bmatrix} \left(\frac{\partial a_{m'}}{\partial \theta_{k'}}\right)_o & \left(\frac{\partial a_{m'}}{\partial \theta_{k''}}\right)_o \\ \left(\frac{\partial a_{m''}}{\partial \theta_{k'}}\right)_o & \left(\frac{\partial a_{m''}}{\partial \theta_{k''}}\right)_o \end{bmatrix}$$

$$= [G_1 : G_2] \cdot \begin{bmatrix} F_1 & F_2 \\ 0 & I \end{bmatrix} .$$

Inverting, we get

$$[G_1 : G_2] = [B_1 : B_2] \cdot \begin{bmatrix} F_1 & F_2 \\ 0 & I \end{bmatrix}^{-1}$$

$$= [B_1 : B_2] \cdot \begin{bmatrix} F_1^{-1} & -F_1^{-1} \cdot F_2 \\ 0 & I \end{bmatrix}$$

or $G_1 = B_1F_1^{-1}$ and $G_2 = B_2 - B_1F_1^{-1}F_2$.

However, as in theorem 5,

$$B_* = \left[\frac{1}{\sqrt{p^o_{ij}}} \sqrt{\frac{n_{io}}{n}} \left(\frac{\partial\, p^*_{ij}}{\partial\, \theta_{k''}} \right) \right]_o = \left[\frac{1}{\sqrt{p^o_{ij}}} \sqrt{\frac{n_{io}}{n}} \left(\frac{\partial\, p_{ij}}{\partial\, \theta_{k''}} \right) \right]_o -$$

$$\left[\frac{1}{\sqrt{p^o_{ij}}} \sqrt{\frac{n_{io}}{n}} \left(\frac{\partial\, p_{ij}}{\partial\, \theta_{k'}} \right) \right]_o \left(\frac{\partial\, f_{m'}}{\partial\, \theta_{k'}} \right)^{-1}_o \left(\frac{\partial\, f_{m'}}{\partial\, \theta_{k''}} \right)_o$$

$$= B_2 - B_1 F_1^{-1} F_2 = G_2 .$$

Therefore the non-centrality parameter may be expressed as

$$\Delta = \underline{c}' F_1'^{-1} B_1' \left[I - B_*(B_*' B_*)^{-1} B_*' \right] B_1 F_1^{-1} \underline{c} .$$

5.4 Asymptotic Independence of Tests

Suppose we have two hypotheses of the type considered in theorem 8:

H_{01}: $f_{1m_1'}(\theta_1, \ldots, \theta_t) = 0$ where $m_1' = 1, \ldots, u_1$, and

H_{02}: $f_{2m_2'}(\theta_1, \ldots, \theta_t) = 0$ where $m_2' = 1, \ldots, u_2$,

with $u_1 + u_2 \leq t$.

Let χ^2_{01} and χ^2_{02} be the test statistics for H_{01} and H_{02} respectively and define $B_{rsxt} = \left[\frac{1}{\sqrt{p^o_{ij}}} \sqrt{\frac{n_{io}}{n}} \left(\frac{\partial\, p_{ij}}{\partial\, \theta_k} \right) \right]_o$. Define B_* for H_{01} and H_{02} and let this be denoted by B_{1*} and B_{2*} respectively.

In a manner exactly analogous to that of section 4.4 it can be shown that when H_{01} and H_{02} are true, a necessary and sufficient condition for the asymptotic independence in probability of χ^2_{01} and χ^2_{02} is that

$$[B(B'B)^{-1}B' - B_{1*}(B_{1*}'B_{1*})^{-1}B_{1*}'][B(B'B)^{-1}B' - B_{2*}(B_{2*}'B_{2*})^{-1}B_{2*}'] = 0 \ (rsxrs),$$

$$\text{or } [I - B_{1*}(B_{1*}'B_{1*})^{-1}B_{1*}'][I - B_{2*}(B_{2*}'B_{2*})^{-1}B_{2*}'] = I - B(B'B)^{-1}B'.$$

5.5 A Special Case

Of particular interest because of its wide applicability is the case in which, for a product-multinomial sample, we wish to test a hypothesis of the form $f_m(p_{11},\dots,p_{1s},\dots,p_{r1},\dots,p_{rs}) = 0$ for $m = 1,\dots,u \leq rs-r$.

In this case we are given that

$$p_{i1} = p_{i1},\dots,p_{is-1} = p_{is-1} \text{ and } p_{is-1} = 1 - p_{i1} - \dots - p_{is-1},$$
$$\text{for } i = 1,\dots,r.$$

The hypothesis is

$$H_o\colon f_m(p_{11},\dots,p_{1s},\dots,p_{r1},\dots,p_{rs}) = 0 \text{ for } m = 1,\dots,u, \text{ and}$$

the alternative is

$$H_n\colon f_m(p_{11},\dots,p_{1s},\dots,p_{r1},\dots,p_{rs}) = c_m/\sqrt{n}\ .$$

Here the solution of the equations (3.2.1) is

$$\hat{p}_{ij} = \frac{n_{ij}}{n_{io}} \quad \text{for } i = 1,\dots,r \text{ and } j = 1,\dots,s \text{ and}$$

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{s} \frac{(n_{ij} - n_{io}\hat{p}_{ij})^2}{n_{io}\hat{p}_{ij}} = 0.$$

The equations (3.3.1) may be written

$$\sum_{i=1}^{r}\sum_{j=1}^{s} \frac{n_{ij} - n_{io}p_{ij}}{p_{ij}} \cdot \frac{\partial p_{ij}}{\partial p_{hk}} + \sum_{m=1}^{u} \lambda_m \frac{\partial f_m^*}{\partial p_{hk}} = 0,$$

$$h = 1,\dots,r,\ k = 1,\dots,s-1,$$

$$f^*_m(p_{11},\dots,p_{1s-1},\dots,p_{r1},\dots,p_{rs-1}) = 0,\ m = 1,\dots,u,$$

where f^*_m is f_m expressed in terms of the rs-r independent parameters $p_{11},\dots,p_{1s-1},\dots,p_{r1},\dots,p_{rs-1}$. Then the equations (3.3.1) reduce to

$$\frac{n_{ij} - n_{io}p_{ij}}{p_{ij}} - \frac{n_{is} - n_{io}p_{is}}{p_{is}} + \sum_{m=1}^{u} \lambda_m \left(\frac{\partial f_m}{\partial p_{ij}} - \frac{\partial f_m}{\partial p_{is}}\right) = 0,$$

$$i = 1,\dots,r,\ j = 1,\dots,s-1,$$

$$f_m(p_{11},\dots,p_{1s},\dots,p_{r1},\dots,p_{rs}) = 0,\ m = 1,\dots,u,$$

$$\sum_{j=1}^{s} p_{ij} = 1,\ i = 1,\dots,r.$$

Define $B\ (rs \times \overline{rs-r}) = \left[\frac{1}{\sqrt{p^o_{ij}}} \sqrt{\frac{n_{io}}{n}} \left(\frac{\partial p_{ij}}{\partial p_{hk}}\right)\right]_o$, $h = 1,\dots,r,$ $k = 1,\dots,s-1$, and let B_1 (rsxu) be the matrix formed by taking the u columns of B which correspond to the p_{hk} made dependent by H_o and let $B_2\ (rs \times \overline{rs-r-u})$ be the matrix formed by taking the rs-r-u columns of B which correspond to the p_{hk} left independent by H_o. Define $F(u \times \overline{rs-r}) = \left(\frac{\partial f^*_m}{\partial p_{hk}}\right)_o$, $h = 1,\dots,r,\ k = 1,\dots,s-1$, and let F_1(uxu) be the matrix formed by taking the u columns of F corresponding to the p_{hk} made dependent by H_o and let $F_2\ (u \times \overline{rs-r-u})$ be the matrix formed by taking the rs-r-u columns of F corresponding to the p_{hk} left independent by H_o.

Then $B_* = B_2 - B_1F_1^{-1}F_2$ and we may form the non-centrality parameter Δ.

5.6 Two Specific Examples

As a specific example of the preceding special case, consider the situation in which we wish to test whether all r multinomial distributions are equal.

The hypothesis is

$$H_o: p_{ij} - p_{rj} = 0 \text{ for } i = 1,\ldots,r-1 \text{ and } j = 1,\ldots,s-1,$$

and the alternative is

$$H_n: p_{ij} - p_{rj} = c_{ij} / \sqrt{n} \text{ for } i = 1,\ldots,r-1 \text{ and } j = 1,\ldots,s-1.$$

The dependent (under H_o) parameters are

$$p_{11},\ldots,p_{1s-1},\ldots,p_{r-11},\ldots,p_{r-1s-1},$$

and the independent (under H_o) parameters are

$$p_{r1},\ldots,p_{rs-1}.$$

The modified chi-square minimum equations subject to H_o are

$$\frac{n_{ij} - n_{io}p_{ij}}{p_{ij}} - \frac{n_{is} - n_{io}p_{is}}{p_{is}} + \lambda_{ij} = 0, \; i = 1,\ldots,r-1, \; j = 1,\ldots,s-1,$$

$$\frac{n_{rj} - n_{ro}p_{rj}}{p_{rj}} - \frac{n_{rs} - n_{ro}p_{rs}}{p_{rs}} - \sum_{i=1}^{r-1} \lambda_{ij} = 0, \; j = 1,\ldots,s-1,$$

$$p_{ij} - p_{rj} = 0, \; i = 1,\ldots,r-1, \; j = 1,\ldots,s-1,$$

$$\sum_{j=1}^{s} p_{ij} = 1, \; i = 1,\ldots,r.$$

The solution of these equations is $\hat{\hat{p}}_{ij} = \frac{n_{oj}}{n}$ for $i = 1,\ldots,r$ and $j = 1,\ldots,s$ and $\chi_o^2 = \sum_{i=1}^{r} \sum_{j=1}^{s} \frac{\left(n_{ij} - \frac{n_{io}n_{oj}}{n}\right)^2}{\frac{n_{io}n_{oj}}{n}}$ with $(r-1)(s-1)$ d.f.

When B_1, B_2, F_1, F_2 and B_* are formed as defined in section 5.5, we find that the non-centrality parameter is

$$\Delta = \sum_{j=1}^{s} \frac{1}{p^o_{rj}} \left[\sum_{i=1}^{r-1} \frac{n_{io}}{n} c^2_{ij} - \left(\sum_{i=1}^{r-1} \frac{n_{io}}{n} c_{ij}\right)^2 \right]$$

where $c_{is} = -\sum_{j=1}^{s-1} c_{ij}$ for $i = 1,\ldots,r-1$ (i.e., $\sum_{j=1}^{s} c_{ij} = 0$ for $i = 1,\ldots,r-1$).

Thus χ^2_o is, in the limit, distributed under H_o as a central chi-square variate with $(r-1)(s-1)$ d.f. and under H_n as a non-central chi-square variate with $(r-1)(s-1)$ d.f. and a non-centrality parameter given by Δ above.

Note that this hypothesis is really the same as that given in (3.5.1) and that the test statistic and the degrees of freedom are identical. If the deviation parameters in the two alternatives are identified with each other ($c_{ij} = d_{ij} - d_{rj}$ for $i = 1,\ldots,r-1$ and $j = 1,\ldots,s$) it can be shown that the two non-centrality parameters are identical.

We may also test whether the first (say) $u+1 \leq r$ out of the r multinomial distributions are equal. In this case the hypothesis is

H_o: $p_{ij} - p_{u+1j} = 0$ for $i = 1,\ldots,u < r$ and $j = 1,\ldots,s-1$,

and the alternative is

H_n: $p_{ij} - p_{u+1j} = c_{ij}/\sqrt{n}$ for $i = 1,\ldots,u$ and $j = 1,\ldots,s-1$.

Here $\hat{\hat{p}}_{ij} = \dfrac{n_{1j} + \ldots + n_{u+1j}}{n_{1o} + \ldots + n_{u+1o}}$ for $i = 1,\ldots,u+1$ and $j = 1,\ldots,s$ and

$\hat{\hat{p}}_{ij} = \frac{n_{ij}}{n_{io}}$ for $i = u+2,\ldots,r$ and $j = 1,\ldots,s$.

The test statistic is $\chi_o^2 = \sum_{i=1}^{u+1} \sum_{j=1}^{s} \frac{\left(n_{ij} - n_{io}\frac{n_{1j} + \ldots + n_{u+1j}}{n_{1o} + \ldots + n_{u+1o}}\right)^2}{n_{io}\frac{n_{1j} + \ldots + n_{u+1j}}{n_{1o} + \ldots + n_{u+1o}}}$

with $u(s-1)$ d.f.

If we define $c_{is} = -\sum_{j=1}^{s-1} c_{ij}$ for $i = 1,\ldots,u$

(i.e., $\sum_{j=1}^{s} c_{ij} = 0$ for $i = 1,\ldots,u$), the non-centrality parameter is found to be

$$\Delta = \sum_{j=1}^{s} \frac{1}{p^o_{u+1j}} \left[\sum_{i=1}^{u} \frac{n_{io}}{n} c_{ij}^2 - \frac{n}{n_{1o} + \ldots + n_{u+1o}} \left(\sum_{i=1}^{u} \frac{n_{io}}{n} c_{ij} \right)^2 \right] .$$

Note that the first example is a special case of this (when $u = r-1$).

APPENDIX

Cramer's Chi-square Theorem

In Cramer's theorem ([5], page 426) a probability measure $\underline{p}(\theta)$ where $\underline{p}'(\underline{\theta}) = [p_1(\underline{\theta}),\dots,p_s(\underline{\theta})]$ is hypothesized and this measure is assumed to satisfy conditions (a)-(d) of (1.3.1) of this paper. The main point of Cramer's proof is that if, under the hypothesized probability measure, we have with a probability which approaches one as $n \longrightarrow \infty$

$$|x_j| < \frac{w}{c} \quad \text{for all } j = 1,\dots,s,$$

where w denotes a function of n such that $w \longrightarrow \infty$ as $n \longrightarrow \infty$ while $w^2/\sqrt{n} \longrightarrow 0$, c is defined in (1.3.1), (b) and x_j is as in theorem 1, then

(1) with a probability which approaches one as $n \longrightarrow \infty$ the equations (2.2.1) of the modified chi-square minimum method have exactly one set of solutions $\hat{\underline{\theta}}$ such that $|\hat{\underline{\theta}} - \underline{\theta}_o| < \epsilon$ where $\underline{\theta}_o$ is the true parameter point, and

(2) with a probability which approaches one as $n \longrightarrow \infty$ we have

$$|\ \underline{y} - [I - A(A'A)^{-1}A']\ \underline{x}\ | < \epsilon$$

where $\underline{y}$, $\underline{x}$ and A are defined as in theorem 1.

A study of Cramer's proof reveals that when conditions (a)-(d) of (1.3.1) hold for <u>any</u> probability measure $\underline{q}(\underline{\theta})$ (say) and if, under the measure $\underline{q}(\underline{\theta})$, with a probability which approaches one as $n \longrightarrow \infty$ we have

$$| x_j | < \frac{w}{c} \text{ for all } j = 1,\ldots,s,$$

then results (1) and (2) above will follow with the difference that the convergence in probability is now with respect to the new probability measure $\underline{q}(\underline{\theta})$.

Now consider the probability measure $\underline{p}_n$ under H_n in theorem 1, where $\underline{p}_n' = (p_{1n},\ldots,p_{sn})$ with $p_{jn} = p_j(\underline{\theta}) + d_j / \sqrt{n}$ where $\sum_{j=1}^{n} d_j = 0$ but not all $d_j = 0$. Conditions (a), (c) and (d) of (1.3.1) are obviously satisfied and if we take $c_1^2 = c^2/2$ then for sufficiently large n, $p_{jn} \geq c_1^2 > 0$ for all $j = 1,\ldots,s$ and condition (b) is satisfied.

We observe that for any $w > d = \max (|d_1|,\ldots,|d_s|)$, for any given j, the statement $|n_j - np_j(\underline{\theta}_o)| \geq w \sqrt{n}$ implies

$$|n_j - np_{jn}| \geq \min(w-d_j, w+d_j) \sqrt{n} \geq (w-d) \sqrt{n}.$$

Then, by Techebyscheff's inequality,

$$\text{Prob} \left[|n_j-np_j(\underline{\theta}_o)| \geq w \sqrt{n} \middle| \underline{p}_n \right] \leq \frac{p_{jn}(1-p_{jn})}{(w-d)^2} < \frac{p_{jn}}{(w-d)^2} .$$

It follows that $|n_j - np_j(\underline{\theta}_o)| \geq w \sqrt{n}$ for at least one value of j with probability smaller than $(w-d)^{-2} \sum_j p_{jn} = (w-d)^{-2}$, and conversely, with a probability greater than $1 - (w-d)^{-2}$ we have $|n_j - np_j(\underline{\theta}_o)| < w \sqrt{n}$ for all $j = 1,\ldots,s$.

Since $\min[p_1(\underline{\theta}_o),\ldots,p_s(\underline{\theta}_o)] \geq c^2$, we have $|x_j| < \frac{w}{c}$ for all j with probability greater than $1 - (w-d)^{-2}$ and this probability approaches one as $n \longrightarrow \infty$.

Therefore, (1) and (2) hold under H_n but the convergence in probability is with respect to $\underline{p}_n$, the probability measure under H_n.

This proof is an amplification of that presented by Mitra [7]. With suitable adjustment in notation the same can be shown for theorems 2, 3 and 4.

BIBLIOGRAPHY

[1] Barnard, G. A., "Significance tests for 2x2 tables," Biometrika, Vol. XXXIV (1947), pp. 123-138.

[2] Cochran, W. G., "The X^2 test of goodness of fit," Annals of Mathematical Statistics, Vol. XXIII (1952), pp. 314-325.

[3] Cochran, W. G., "Some methods of strengthening the common X^2 tests," Biometrics, Vol. X (1954), pp. 417-451.

[4] Courant, R., (translated by E. J. McShane), Differential and Integral Calculus, Vol. II, New York, Interscience Publishers, Inc., 1951.

[5] Cramer, H., Mathematical Methods of Statistics, Princeton, Princeton University Press, 1946.

[6] Gnanadesikan, R., "Contributions to multivariate analysis including univariate and multivariate variance components analysis and factor analysis," North Carolina Institute of Statistics Mimeograph Series, No. 158, June 1956.

[7] Mitra, S. K., "Contributions to the statistical analysis of categorical data," North Carolina Institute of Statistics Mimeograph Series, No. 142, December 1955.

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

[9] Ogawa, J., "On the independence of quadratic forms in a non-central normal system," Osaka Mathematical Journal, Vol. II, (1950), pp. 151-159.

[10] Ogawa, J., "On the sampling distributions of classical statistics in multivariate analysis," Osaka Mathematical Journal, Vol. V (1953), pp. 13-52.

[11] Ogawa, J., "On the mathematical principles underlying the theory of the X^2 test," North Carolina Institute of Statistics Mimeograph Series, No. 162, January 1957.

[12] Patnaik, P. B., "The non-central X^2 and F distributions and their applications," Biometrika, Vol. XXXVI (1949), pp. 202-232.

[13] Pearson, E. S., "The choice of statistical tests illustrated on the interpretation of data classed in a 2x2 table," Biometrika, Vol. XXXIV (1947), pp. 139-167.

[14] Pearson, K., "On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling," Philosophical Magazine, Series 5, Vol. L (1900), pp. 157-172.

[15] Pitman, E. J. G., Unpublished lecture notes on non-parametric statistical inference, 1948.

[16] Roy, S. N., "A report on some aspects of multivariate analysis," North Carolina Institute of Statistics Mimeograph Series, No. 121, December 1954.

[17] Roy, S. N. and Gnanadesikan, R., "Further contribution to multivariate confidence bounds," North Carolina Institute of Statistics Mimeograph Series, No. 155, August 1956.

[18] Roy, S. N. and Mitra, S. K., "An introduction to some non-parametric generalizations of analysis of variance and multivariate analysis," Biometrika, Vol. XLIV (1956), pp. 361-376.

[19] Roy, S. N. and Sarhan, A. E., "On inverting a class of patterned matrices," Biometrika, Vol. XLIV (1956), pp. 227-231.

[20] Wald, A., "Tests of statistical hypotheses concerning several parameters when the number of observations is large," Transactions of the American Mathematical Society, Vol. LIV (1943), pp. 426-482.