ASYMPTOTIC SIMULTANEOUS CONFIDENCE INTERVALS FOR THE PROBABILITIES OF A MULTINOMIAL DISTRIBUTION

by

Charles Price Quesenberry III

Thesis submitted to the Graduate Faculty of the
Virginia Polytechnic Institute
in candidacy for the degree of

MASTER OF SCIENCE
in
STATISTICS

APPROVED:

Director of Graduate Studies

APPROVED:

Head of Department

Dean of the School of Applied Science and Business Administration

Major Professor

November, 1958

Blacksburg, Virginia

TABLE OF CONTENTS

ASYMPTOTIC SIMULTANEOUS CONFIDENCE INTERVALS FOR THE PROBABILITIES OF A MULTINOMIAL DISTRIBUTION

I. Introduction

Given a random sample from a multinomial distribution, what can be said about the parameters of the distribution? This is a question of interest, for samples from multinomial distributions arise in many situations. The following example illustrates a situation in which we are concerned with a sample from such a distribution.

An agency purchases copies of a certain "component" to be installed on equipments. Any "component" which fails during the inspection and assembling operations of the agency is termed a "failure," and the data recorded are the number of failures for each possible type of defect. Data from such a situation can be tabulated as follows: (The defect types have been ranked in order of increasing number of failures.)

Defect Type	1	2	3	4	5	6	7	8	9	10
Failures	5	10	20	30	60	65	90	120	170	300

What conclusions can be drawn from these frequencies? What can be said about the proportion of failures which is attributable to any particular type of defect?

We would, of course, be interested in isolating types of defects which are significant problem areas. Can we make probability statements as to the relative importance of different types of defects, i.e., does one type of defect account for more (or less) of the total number of failures than some (or any) other type of defect? These are the kinds of questions which we will provide a way of answering. A manager of a defect reporting system or an unsatisfactory condition reporting system will note that "Type of Defect" could be replaced by "Month", "Area Found", "Severity", or any of many similar categorical variables. The same flexibility holds as regards the item, part, or equipment considered.

II. Review of Literature

In this section a review of the literature that is concerned with the statistical concepts and procedures that are germane to the theme of this paper shall be given. In particular, we shall consider the Chi-square goodness-of-fit statistic, its applications, and simultaneous confidence intervals.

1. The χ^2 Goodness of Fit Statistic.

Let a random sample of N observations be drawn from a population, and suppose that these observations can be classified into k mutually exclusive classes, or cells

by some criteria. Let π_i ($i = 1, 2, \ldots, k$) denote the population probability that any individual observation will fall into the i^{th} cell. The sample of observed cell frequencies, $(n_1, n_2, \ldots, n_k)$, will have a multinomial distribution, i.e., the joint distribution function of the observed cell frequencies is given by

$$(1) \qquad \frac{N!\pi_1^{n_1}\pi_2^{n_2} \cdots \pi_k^{n_k}}{n_1!n_2! \cdots n_k!},$$

where

$$\sum_{i=1}^{k} n_i = N,$$

and

$$\sum_{i=1}^{k} \pi_i = 1.$$

Our purpose in drawing a sample from a multinomial distribution is to obtain information concerning the values of the parameters, $(\pi_1, \ldots, \pi_k)$, of the distribution. Many previous writers have shown that:

$$E(p_i) = E(n_i/N) = \pi_i, \qquad i = 1, \ldots, k$$

$$\operatorname{var}(p_i) = \pi_i(1 - \pi_i)/N,$$

$$\operatorname{cov}(p_i, p_j) = -\pi_i\pi_j/N, \qquad i \neq j$$

and that p_i is the maximum likelihood estimator of π_i. It should be noted that p_i is the classical relative frequency estimator of the proportion of the population in

cell i.

Naturally, we would like to obtain much more information about the parameters. One thing that we might be interested in knowing would be whether or not the parameters have some specified set of values. Pearson (1900) has given a method for testing to determine whether or not a specified set of values for the parameters is consistent with the set of observed cell frequencies obtained by a random sample. He has shown that for sufficiently large values of the expected cell frequencies, $N\pi_i$ $(i = 1, \ldots, k)$, the statistic

$$(2) \qquad X^2 \equiv \sum_{i=1}^{k} \frac{(n_i - N\pi_i)^2}{N\pi_i}$$

is distributed as a Chi-square variate with k - 1 degrees of freedom. That is, the distribution function of X^2 is approximated by

$$(3) \qquad \frac{1}{\left(\frac{k-3}{2}\right)!\, 2^{\frac{k-1}{2}}} (\chi^2)^{\frac{k-3}{2}} e^{-\frac{\chi^2}{2}}$$

as $N \longrightarrow \infty$.

Therefore, this Chi-square approximation to the distribution of the quantity X^2 can be used to make an asymptotic test of the null hypothesis that the parameters have a specified set of values, i.e.,

$$H_o : (\pi_1, \pi_2, \ldots, \pi_k) = (\pi_{10}, \pi_{20}, \ldots, \pi_{k0}).$$

(The goodness of this approximation will be discussed in the appendix.) It should be noted that if this hypothesis is rejected the alternative is composite, and we can only conclude that one or more of the hypothesized values of the parameters are incorrect. Also, if the hypothesis is not rejected the experimenter should be aware that in general there exists an infinity of sets of values for the parameters, which could be specified, which would not lead to rejection either.

As useful as this test is, there still remain questions about the parameters which the experimenter would like to be able to answer. Methods for obtaining more information about the parameters from the sample are quite scarce in the literature. Most efforts towards obtaining more information have been made by partitioning the quantity X^2 into expressions corresponding to the individual degrees of freedom of X^2, and attempting to use these expressions to make tests in some manner so that inferences can be made about the parameters of the multinomial, or about the phenomena occurring in the physical situation under study. (This approach seems to have been suggested by the partitioning of the treatment sum of squares into individual degrees of freedom in the analysis of variance.)

This was, in fact, the first approach considered by this author.

Cochran (1952) shows that X^2 can be partitioned in general in the following manner.

Let

$$z_i = \sum_{j=1}^{k} b_{ij} n_j ,$$

where

$$\sum_{j=1}^{k} b_{ij} \pi_j = 0 ,$$

and

$$\sum_{j=1}^{k} b_{ij} b_{hj} \pi_j = 0, \quad \text{if } h \neq i$$

$$= 1, \quad \text{if } h = i.$$

Then

$$X^2 = \sum_{i=1}^{k} \frac{(n_i - N\pi_i)^2}{N\pi_i} = \sum_{j=1}^{k-1} z_j^2$$

and for N sufficiently large the individual terms on the right are asymptotically distributed as a Chi-square variate with 1 degree of freedom each.

Lancaster (1949) has performed an approximate partitioning of X^2 in the following manner. The multinomial distribution function can be written in an algebraically equivalent form as the product of k - 1 binomial distribution functions, i.e.,

$$\frac{N!\pi_1^{n_1}\pi_2^{n_2}\cdots\pi_k^{n_k}}{n_1!n_2!\cdots n_k!} = \frac{(n_1 + n_2)!}{n_1!n_2!}\left(\frac{\pi_2}{\pi_1 + \pi_2}\right)^{n_2}\left(\frac{\pi_1}{\pi_1 + \pi_2}\right)^{n_1}$$

$$\times\frac{(n_1 + n_2 + n_3)!}{(n_1 + n_2)!n_3!}\left(\frac{\pi_3}{\pi_1 + \pi_2 + \pi_3}\right)^{n_3}\left(\frac{\pi_1 + \pi_2}{\pi_1 + \pi_2 + \pi_3}\right)^{n_1+n_2}$$

$$\vdots$$

(4)

$$\times\frac{(n_1+n_2+n_3+\cdots+n_i)!}{(n_1+n_2+\cdots+n_{i-1})!n_i!}\left(\frac{\pi_i}{\pi_1+\cdots+\pi_i}\right)^{n_i}\left(\frac{\pi_1+\cdots+\pi_{i-1}}{\pi_1+\cdots+\pi_i}\right)^{n_1+\cdots+n_{i-1}}$$

$$\vdots$$

$$\times\frac{N!}{(N - n_k)!n_k!}\,\pi_k^{n_k}\,(1 - \pi_k)^{N-n_k}$$

The binomial variates on the right hand side of this equation are uncorrelated. (For a proof of this, see Lancaster (1949).) The (i-1)st factor on the right gives the probability that n_i observations will fall into the i^{th} cell, conditional upon the sum of the cell frequencies of the first i cells, i.e.,

$$\Pr(n_i \mid n_1 + \cdots + n_i) = \frac{(n_1 + \cdots + n_i)!}{(n_1 + \cdots + n_{i-1})!n_i!}\left(\frac{\pi_i}{\pi_1+\cdots+\pi_i}\right)^{n_i}$$

$$\times\left(\frac{\pi_1 + \cdots + \pi_{i-1}}{\pi_1 + \cdots + \pi_i}\right)^{n_1+\cdots+n_{i-1}}.$$

Since the binomial distribution approaches the normal distribution as sample size increases, if as N becomes

large $n_1 + n_2 + \cdots + n_i$ $(i = 1, 2, \ldots, k)$ becomes large also, then the general binomial variate above can be approximated by a normal variate, and since the binomial variates are uncorrelated the normal variates will be asymptotically independent. The mean and variance of the general normal variate is

$$\text{mean} = (n_1 + \cdots + n_i)\frac{\pi_i}{\pi_1 + \cdots + \pi_i}$$

and

$$\text{variance} = (n_1 + \cdots + n_i)\frac{\pi_i(\pi_1 + \cdots + \pi_{i-1})}{(\pi_1 + \cdots + \pi_i)^2} .$$

Let

$$z_i = \left(n_i - \frac{(n_1+\cdots+n_i)\pi_i}{\pi_1+\cdots+\pi_i}\right) \Bigg/ \sqrt{\frac{(n_1+\cdots+n_i)\pi_i(\pi_1+\cdots+\pi_{i-1})}{(\pi_1+\cdots+\pi_i)^2}}$$

and then $z_i \sim NI(0,1)$, approximately, so that $z_i^2 \sim \chi^2_{(1)}$, approximately.

Since there are $k - 1$ of the z_i's with 1 degree of freedom each, then the z_i^2's correspond to the individual degrees of freedom of χ^2. Also,

$$\chi^2 = \sum_{i=2}^{k} z_i^2 ,$$

approximately.

The purpose in obtaining these z_i's was to use them to obtain information about the parameters of the multinomial distribution. Let us consider using them to test

hypotheses involving the parameters of the multinomial. In order to test a hypothesis it must be such that it will give a constant value to the expression

$$\frac{\pi_i}{\pi_1 + \cdots + \pi_i},$$

for this is necessary and sufficient to remove all of the parameters from the expression for z_i, i.e.,

$$(5) \quad z_i = \frac{n_i - (n_1 + \cdots + n_i)\left(\frac{\pi_i}{\pi_1 + \cdots + \pi_i}\right)}{\sqrt{(n_1+\cdots+n_i)\left(\frac{\pi_i}{\pi_1+\cdots+\pi_i}\right)\left(1 - \frac{\pi_i}{\pi_1+\cdots+\pi_i}\right)}}, \quad i=2,\cdots,k.$$

One such hypothesis of this type would be a hypothesis which specifies the value of each of the first i parameters, i.e., of the form

$$H_o : (\pi_1, \pi_2, \cdots, \pi_i) = (\pi_{10}, \pi_{20}, \cdots, \pi_{i0}).$$

However, this hypothesis can be tested by the usual χ^2 test without using the z_i's. The above hypothesis is equivalent to

$$H_o : (\pi_1, \pi_2, \cdots, \pi_i, 1 - \pi_1 - \pi_2 - \cdots - \pi_i) =$$

$$(\pi_{10}, \pi_{20}, \cdots, \pi_{i0}, 1 - \pi_{10} - \pi_{20} - \cdots - \pi_{i0}),$$

because of the restriction $\sum_{i=1}^{k} \pi_i = 1$, and this hypothesis can be tested by the usual χ^2 test with i degrees of freedom.

Any other hypothesis to be tested would be of the form

$$H_o : \frac{\pi_i}{\pi_1 + \cdots + \pi_i} = \text{constant} = \frac{1}{C} \text{ , say.}$$

If C is chosen equal to i, then

$$\pi_i = \frac{\pi_1 + \cdots + \pi_i}{i} \Longrightarrow \pi_i = \frac{\pi_1 + \cdots + \pi_{i-1}}{i - 1}$$

so we consider the null hypothesis

$$(6) \qquad H_o : \pi_i = \frac{\pi_1 + \cdots + \pi_{i-1}}{i - 1} \text{ .}$$

(Note that the hypothesis

$$H_o : \pi_1 = \pi_2 = \cdots = \pi_i$$

is a special case because

$$\pi_1 = \pi_2 = \cdots = \pi_i \Longrightarrow \pi_i = \frac{\pi_1 + \cdots + \pi_{i-1}}{i - 1} \text{ ,}$$

but not conversely.) Under (6) we have

$$(7) \qquad z_i^2 = \frac{[(i-1)n_i - (n_1 + \cdots + n_{i-1})]^2}{(n_1 + \cdots + n_i)(i-1)}, \quad i=2,\ldots,k.$$

which is the test statistic to be computed from the sample.

This test purportedly gives a comparison of π_i with the i - 1 preceeding parameters. The value of the test is limited, however, in that it compares π_i with the mean of other parameters which are themselves unknown. So regardless of whether the hypothesis is rejected or not, no definite information is obtained about the magnitude of π_i or any other parameter.

Another criticism of this test is that the results will in general be different if the cells are taken in different orderings. As an illustration of this, consider the following example.

Example 1: For $N = 100$, $k = 4$, and the n_i as follows:

$$n_1 = 10,\ n_2 = 20,\ n_3 = 30,\ n_4 = 40.$$

The z_i^2 for the cells taken in the order 1, 2, 3, 4, are

$$z_2^2 = (20 - 10)^2/30 = 3.333,$$

$$z_3^2 = (60 - 30)^2/2(60) = 7.5,$$

$$z_4^2 = (120 - 60)^2/3(100) = 12,$$

$$\sum_{i=2}^{4} z_i^2 = 22.833.$$

From these z_i^2's we can reject the hypotheses

$$H_o : \pi_3' = (\pi_1 + \pi_2)/2$$

and

$$H_o : \pi_4' = (\pi_1 + \pi_2 + \pi_3)/3$$

at the .05 level of significance.

Now consider taking the cells in the order 2, 3, 4, 1. Then

$$z_2^2 = (30 - 20)^2/50 = 2,$$

$$z_3^2 = (80 - 50)^2/180 = 5$$

$$z_4^2 = (30 - 90)^2/300 = 12$$

$$\sum_{i=2}^{4} z_i^2 = 19.$$

From these z_i^2's we can reject the hypotheses

$$H_o : \pi_4' = (\pi_2 + \pi_3)/2$$

and

$$H_o : \pi_1' = (\pi_2 + \pi_3 + \pi_4)/3$$

at the .05 level of significance.

By comparison of these two results it is obvious that different orderings of the cells lead to quite different results. It should also be noted that $\sum_{i=2}^{4} z_i^2$ is different for different orderings.

The above discussion is not intended to imply that this method is of no value, but merely to point out its limitations. For a situation where the order of the cells is determined by other considerations it may be quite useful. This is the situation in the following example taken from Lancaster (1949), which illustrates a situation in which this method can yield useful results.

Example 2: "Measured constant amounts of a liquid suspension of a bacterial culture are mixed with an equal quantity of disinfectant solution of known concentration, and a plate is poured and the number of colonies developing are noted. For each plate the concentration of disinfectant used is given by a series such as 1, Y, Y^2, Y^3, ... where Y is some factor such as 2 or 1.5. In such a case the following results might be obtained."

"Number of colonies (n_i) developing in successive plates 427, 440, 494, 422, 409, 310, 302. We are interested in finding the point at which the disinfectant began to inhibit growth."

The z_i^2 components taken for this ordering are

z_2^2	z_3^2	z_4^2	z_5^2	z_6^2	z_7^2
0.195	5.379	1.687	2.465	32.947	28.299

Of the components, only z_3^2, z_6^2, and z_7^2 are significant at the .05 level, so that we can reject the hypotheses

$$H_o : \pi_3' = (\pi_1 + \pi_2)/2$$
$$H_o : \pi_6' = (\pi_1 + \ldots + \pi_5)/5$$
$$H_o : \pi_7' = (\pi_1 + \ldots + \pi_6)/6$$

From these results we cannot say that π_6 or π_7 are greater than any of the other parameters (except π_1), but the evidence does seem to indicate that growth is beginning to be inhibited at plate 6.

Irwin (1949) has presented a method for partitioning X^2 exactly into z_i^2's. However, the same kind of difficulty is encountered in attempting to use his z_i's to make tests of hypotheses in such a way as to make inferences about the individual parameters. The difficulty is again that the individual z_i's involve functions of the first i parameters, and tests on these functions do not lead to inferences about the individual parameters.

For either Lancaster's or Irwin's partitioning, procedures can be developed for obtaining asymptotic confidence intervals for the functions of the parameters $\pi_i/(\pi_1 + \pi_2 + \ldots + \pi_i)$. However, confidence intervals for these functions do not give definite information about the individual parameters.

Two methods have been mentioned for partitioning X^2 into expressions corresponding to individual degrees of freedom of $X^2_{(k-1)}$. It seems that there may be other methods which would lead to more fruitful results. However, in partitioning X^2 it seems desirable to maintain the property of asymptotic independence among the z_i^2 components obtained. This writer has attempted such partitionings, but all efforts have either been entirely unsuccessful or have led to the conditional binomial variates given by Lancaster.

The above discussion indicates that the information to be obtained from existing methods of analysis is limited. It is the purpose of this paper to present a method of analysis that will yield the experimenter more information about his cell parameters, (π_i), than can be obtained by present methods. The approach considered here is to find a set of asymptotic simultaneous confidence intervals for the cell parameters.

2. Simultaneous Confidence Intervals.

The theory of confidence intervals is due largely to Neyman (1935, 1937, and others) and Pearson (1934 and others). Other contributors have been Wilks (1938a, 1938b, 1939), Wald (1939, 1942), Welsh (1939), and Bartlett (1936, 1939). No individual discussion of the various papers that are concerned with confidence intervals will be attempted here. This exposition will be limited to a definition of what is meant by a confidence interval for one parameter, and a set of simultaneous confidence intervals for a set of k parameters.

Let the distribution function of an observed sample depend upon k (= 1, 2, 3, ...) parameters, $(\pi_1, \pi_2, \ldots, \pi_k)$. Select one of these parameters, say π_i $(i = 1, 2, \ldots, k)$. Suppose that there is a procedure such that when a sample is drawn this procedure will specify an interval upon the axis of real numbers. If, when a large number of samples are drawn and intervals obtained for each sample by this procedure, the proportion of intervals which contain the parameter π_i tends to a limit, say $1 - \gamma_i$, as the number of samples tends to infinity, then the interval obtained for any individual sample is called a confidence interval for the parameter π_i $(i = 1, \ldots, k)$ with confidence coefficient $1 - \gamma_i$.

Again let the distribution function of an observed sample depend upon k parameters. Denote the parameters by $(\pi_1, \pi_2, \ldots, \pi_k)$ and let $S = (S_1, S_2, \ldots, S_k)$ denote a set of k intervals where S_i corresponds to π_i $(i = 1, \ldots, k)$. Now let the set S be obtained by some procedure from a sample such that if such sets are obtained by this same procedure for a number of samples the proportion of these sets for which <u>every</u> S_i $(i = 1, \ldots, k)$, contains its corresponding parameter, π_i, tends to a limit, say $1 - \alpha$, (as the number of samples increases without bound), the set S obtained for any particular sample is called a set of <u>simultaneous confidence intervals</u> with confidence coefficient $1 - \alpha$. This may be represented as

$$\Pr\left\{\bigcap_{i=1}^{k} (\pi_i \epsilon S_i)\right\} = 1 - \alpha.$$

It should be emphasized that the confidence coefficient for a set of simultaneous confidence intervals refers to the procedure by which the intervals are obtained, and not to whether or not the intervals contain their parameters (for the probability of this event can only be zero or one).

Each of the intervals of a set of simultaneous confidence intervals is a confidence interval for its parameter. However, the relationship between the confidence coeffi-

cient for an individual interval, $1 - \gamma_i$, and the confidence coefficient for the set of simultaneous confidence intervals, $1 - \alpha$, is not in general known unless any two intervals are independent in a probability sense, i.e.,

$$\Pr\left\{(\pi_i \epsilon S_i)\ (\pi_j \epsilon S_j)\right\} = (1 - \gamma_i)(1 - \gamma_j),\ i \neq j.$$

In this case the simultaneous confidence coefficient is equal to the product of the confidence coefficients for the individual intervals, i.e.,

$$(9) \qquad 1 - \alpha = (1 - \gamma_1)(1 - \gamma_2)\ldots(1 - \gamma_k).$$

And, if $\gamma_1 = \gamma_2 = \ldots = \gamma_k = \gamma$, then

$$(10) \qquad 1 - \alpha = (1 - \gamma)^k.$$

For the general case it is true that:

$$1 - \gamma_i \geq (1 - \gamma_1)(1 - \gamma_2)\ldots(1 - \gamma_i)\ldots(1 - \gamma_j) \geq 1 - \alpha$$

for $j = 1, \ldots, k$; and $i \leq j$.

III. Derivation of Intervals.

The following symbols have been used previously but for completeness they are defined here.

N is the number of observations in the sample.

n_i is the number of observations in cell i.

π_i is the probability that any observation will fall in cell i.

k is the number of cells.

$p_i = n_i/N$, and is the proportion of observations that fall

in cell i.

The joint distribution function of a set of observed cell frequencies is as before given by the multinomial distribution function:

$$(1) \qquad \Pr\left\{n_1, n_2, \ldots, n_k \mid N\right\} = \frac{N!\pi_1^{n_1}\pi_2^{n_2} \ldots \pi_k^{n_k}}{n_1!n_2! \ldots n_k!}$$

where

$$\sum_{i=1}^{k} n_i = N,$$

and

$$\sum_{i=1}^{k} \pi_i = 1.$$

It is well known that

$$(2) \qquad E(n_i) = N\pi_i \qquad i = 1, \ldots, k$$

$$(3) \qquad \operatorname{var}(n_i) = N\pi_i(1 - \pi_i)$$

$$(4) \qquad \operatorname{cov}(n_i, n_j) = -N\pi_i\pi_j \qquad j \neq i.$$

Pearson (1900) has shown that

$$X^2 = \sum_{i=1}^{k} \frac{(n_i - N\pi_i)^2}{N\pi_i} = \sum_{i=1}^{k} \frac{Np_i^2}{\pi_i} - N$$

is asymptotically distributed as $\chi^2_{(k-1)}$. In the present application of the statistic X^2 the parameters, $(\pi_i, \ldots, \pi_k)$, will not be assumed known. They are only a set of unknown numbers which satisfy the relation

$$\sum_{i=1}^{k} \pi_i = 1.$$

Using this asymptotic approximation we can make the following probability statement

$$(5) \qquad \Pr\left\{\sum_{i=1}^{k} \frac{Np_i^2}{\pi_i} - N \leq \chi^2_{\alpha,k-1}\right\} \approx 1 - \alpha,$$

where $\chi^2_{\alpha,k-1}$ is the α percentage point of the χ^2 distribution with $k - 1$ degrees of freedom.

For fixed α, k, and N the inequality

$$(6) \qquad \sum_{i=1}^{k} \frac{p_i^2}{\pi_i} \leq \frac{\chi^2_{\alpha,k-1}}{N} + 1 = C$$

will define a region in the k dimensional Euclidean space of the parameters. The probability that this region will "contain" the point $(\pi_1, \ldots, \pi_k)$ is $1 - \alpha$. The associated equation

$$(7) \qquad \sum_{i=1}^{k} \frac{p_i^2}{\pi_i} = C$$

will define the bounding hypersurface of this region. The intersection of this hypersurface with the hyperplane

$$(8) \qquad \sum_{i=1}^{k} \pi_i = 1$$

will be a k dimensional hypercurve, or a $k - 1$ dimensional hypersurface. If maximum and minimum values will constitute a set of asymptotic simultaneous confidence intervals for the set of parameters with confidence coefficient equal to or

greater than $1 - \alpha$.

Equation (7) expresses any one of the parameters in terms of the $k - 1$ others, i.e.,

$$(9) \qquad \pi_i = \frac{p_i^2}{C - \sum_{\substack{j=1 \\ j \neq i}}^{k} \frac{p_i^2}{\pi_j}} \qquad i = 1, 2, \ldots, k.$$

Therefore, the function to be maximized and minimized is

$$(10) \qquad \frac{(1-\lambda)p_i^2}{C - \sum_{\substack{j=1 \\ j \neq i}}^{k} \frac{p_i^2}{\pi_j}} - \lambda\left(\sum_{\substack{j=1 \\ j \neq i}}^{k} \pi_j - 1\right),$$

where λ is the Lagrange multiplier.

Differentiating this function with respect to π_j and π_m and setting these derivatives equal to zero gives the equations

$$(11) \qquad \frac{\lambda p_i^2}{(\lambda-1)\pi_i^2} = \frac{p_j^2}{\pi_j^2} \qquad i, j, m = 1, \ldots, k;\ j \neq i,\ m \neq i,\ m \neq j$$

and

$$\frac{\lambda p_i^2}{(\lambda-1)\pi_i^2} = \frac{p_m^2}{\pi_m^2}.$$

which can be solved to obtain

$$(12) \qquad \pi_m = \frac{p_m \pi_j}{p_j}.$$

This represents $k - 2$ equations, and we can solve these

equations with equations (7) and (8) to obtain the required values of the k parameters.

Using (12) to substitute into (8) gives

$$\sum_{i=1}^{k} \pi_i = \pi_i + \pi_j + \sum_{\substack{m=1 \\ m\neq i \\ m\neq j}}^{k} \frac{p_m \pi_j}{p_j} = 1,$$

and using $\sum_{i=1}^{k} p_i = 1$ and simplifying gives

$$\pi_j = \frac{(1 - \pi_i)p_j}{1 - p_i} \quad (13)$$

Substituting into (7) gives

$$\frac{p_i^2}{\pi_i} + \sum_{\substack{j=1 \\ j\neq i}}^{k} \frac{p_j^2(1 - p_i)}{p_j(1 - \pi_i)} = C$$

or

$$\frac{p_i^2}{\pi_i} + \frac{(1 - p_i)^2}{1 - \pi_i} = C,$$

and solving this quadratic for π_i gives

$$\pi_i = \frac{C + 2p_i - 1 \pm \sqrt{(C + 2p_i - 1)^2 - 4Cp_i^2}}{2C} \quad (14)$$

Or, rewriting for computational convenience

$$\pi_i = \frac{\chi^2 + 2n_i \pm \sqrt{\chi^2(\chi^2 + \frac{4n_i(N - n_i)}{N}}}{2(N + \chi^2)} \quad (14)'$$

where $\chi^2 = \chi^2_{\alpha,k-1}$.

Let S_i represent the set of all numbers y for which

$$\frac{C+2p_i-1-\sqrt{(C+2p_i-1)^2-4Cp_i^2}}{2C} \leq y \leq \frac{C+2p_i-1+\sqrt{(C+2p_i-1)^2-4Cp_i^2}}{2C}$$

Then the set $S = (S_1, S_2, \ldots, S_k)$ is a set of asymptotic simultaneous confidence intervals for the set of parameters $(\pi_1, \pi_2, \ldots, \pi_k)$, with confidence coefficient equal to or greater than $1 - \alpha$. Symbolically this may be written

$$\Pr\left\{(\pi_1 \epsilon S_1)(\pi_2 \epsilon S_2)\ldots(\pi_k \epsilon S_k)\right\} \geq 1 - \alpha.$$

IV. Properties of the Intervals

These intervals have certain desirable properties which will be stated as theorems and proved in this section.

Let

$$U_i = \frac{C + 2p_i - 1 + \sqrt{(C + 2p_i - 1)^2 - 4Cp_i^2}}{2C}$$

and

$$L_i = \frac{C + 2p_i - 1 - \sqrt{(C + 2p_i - 1)^2 - 4Cp_i^2}}{2C}$$

and

$$\chi^2 = \chi^2_{\alpha,k-1}.$$

Theorem 1. U_i and L_i are always real numbers, i.e., the intervals will always exist.

Proof

1) $C + 2p_i - 1 = \chi^2/N + 1 + 2p_i - 1 > 0$

2) $2C = 2(\chi^2/N + 1) > 0$

3) $(C + 2p_i - 1)^2 - 4Cp_i^2 = \frac{\chi^4}{N^2} + 4p_i \frac{\chi^2}{N} - 4 \frac{\chi^2}{N} p_i^2$

$= \frac{\chi^2}{N} \left[\frac{\chi^2}{N} + 4p_i(1 - p_i) \right] > 0$

∴ theorem is proved.

Theorem 2. $\lim_{N \to \infty} U_i = \lim_{N \to \infty} L_i = p_i$. i.e., the intervals will converge upon p_i, the maximum likelihood estimator of π_i, from both the left and the right as the sample size increases.

Proof:

1) $\lim_{N \to \infty} \frac{C + 2p_i - 1}{2C} = \lim_{N \to \infty} \frac{\chi^2/N + 2p_i}{2(\chi^2/N + 1)} = \frac{2p_i}{2} = p_i$

2) $\lim_{N \to \infty} \sqrt{(C + 2p_i - 1)^2 - 4Cp_i^2} \,/2C$

$= \lim_{N \to \infty} \sqrt{\frac{\chi^2}{N} \left[\frac{\chi^2}{N} + 4p_i(1 - p_i) \right]} \; 2(N + \chi^2)$

$= 0.$

∴ since the limit of a sum is the sum of the limits the theorem is proved.

Theorem 3. If $p_j = 1 - p_i$, then $U_j - L_j = U_i - L_i$, i.e., intervals whose corresponding p's are equal distances from $\frac{1}{2}$ will be of equal length.

Proof:

1) $U_j - L_j = \sqrt{\frac{\chi^2}{N} \frac{\chi^2}{N} + 4p_j(1 - p_j)} \Big/ C$

2) $U_i - L_i = \sqrt{\frac{\chi^2}{N} \frac{\chi^2}{N} + 4p_i(1 - p_i)} \Big/ C$

3) But $p_j = 1 - p_i \Longrightarrow p_j(1 - p_j) = p_i(1 - p_i)$

$\therefore U_j - L_j = U_i - L_i$.

Theorem 4. $0 \leq L_i \leq p_i$ and $p_i \leq U_i \leq 1$, i.e., the intervals will always lie entirely in the closed interval [0,1] and contain the point p_i.

This theorem will be proved in four parts.

Proof:

Part 1) To show: $0 \leq L_i$

(i) $\quad 0 \geq -4Cp_i^2$, since $C > 0$ and $p_i \geq 0$

(ii) $\quad (C + 2p_i - 1)^2 \geq (C + 2p_i - 1)^2 - 4Cp_i^2$

(iii) $\quad C + 2p_i - 1 \geq \sqrt{(C + 2p_i - 1)^2 - 4Cp_i^2}$

by definition of a radical sign and parts (1) and (3) of theorem 1.

(iv) $\quad C + 2p_i - 1 - \sqrt{(C + 2p_i - 1)^2 - 4Cp_i^2} \geq 0$

(v) $\therefore \quad \dfrac{C + 2p_i - 1 - \sqrt{(C + 2p_i - 1)^2 - 4Cp_i^2}}{2C} = L_i \geq 0$

since $C > 0$ by (2) of theorem 1.

Part 2) To show: $L_i \leq p_i$

(i) $\quad C - 1 = \dfrac{\chi^2}{N} > 0$

(ii) $\quad 4Cp_i^2(C - 1) + 4Cp_i\sqrt{(C + 2p_i - 1)^2 - 4Cp_i^2} \geq 0$

since $C > 0$, $p_i \geq 0$ and $\sqrt{(C + 2p_i - 1)^2 - 4Cp_i^2} > 0$

(iii) $\quad 4C^2p_i^2 + 4Cp_i\sqrt{(C + 2p_i - 1)^2 - 4Cp_i^2}$

$\qquad + (C + 2p_i - 1)^2 - 4Cp_i^2 \geq (C + 2p_i - 1)^2$

(iv) $2Cp_i + \sqrt{(C + 2p_i - 1)^2 - 4Cp_i^2} \geq C + 2p_i - 1$

since both sides of (iii) are positive.

$\therefore$ $C + 2p_i - 1 - \sqrt{(C + 2p_i - 1)^2 - 4Cp_i^2} \leq 2Cp_i$

or

$$[(C + 2p_i - 1) - \sqrt{(C + 2p_i - 1)^2 - 4Cp_i^2}] /2C \leq p_i$$

or

$$L_i \leq p_i.$$

Part 3) To show: $U_i \leq 1$

(i) $(p_i - 1)^2 = p_i^2 - 2p_i + 1 \geq 0$

(ii) $1 - 2p_i \geq - p_i^2$

(iii) $4C^2 - 4C^2 - 8Cp_i + 4C \geq - 4Cp_i^2$

(iv) $4C^2 - 4C(C + 2p_i - 1) + (C + 2p_i - 1)^2 \geq$
$(C + 2p_i - 1)^2 - 4Cp_i^2$

but

$$(C + 2p_i - 1)^2 - 4Cp_i^2 - 4Cp_i^2 \geq 0$$

by (3) of theorem 1, and $\therefore$ both sides of (iv) are positive.

(v) $2C - (C + 2p_i - 1) \geq \sqrt{(C + 2p_i - 1)^2 - 4Cp_i^2}$

(vi) $C + 2p_i - 1 + \sqrt{(C + 2p_i - 1)^2 - 4Cp_i^2} \leq 2C$

(vii) $[(C + 2p_i - 1) + \sqrt{(C + 2p_i - 1)^2 - 4Cp_i^2}] /2C =$
$U_i \leq 1$

Part 4) To show: $U_i \geq p_i$. We shall consider the three cases for $p_i = \frac{1}{2}$, $p_i < \frac{1}{2}$, $p_i > \frac{1}{2}$.

<u>Case (1)</u> If $p_i = \frac{1}{2}$

(i) $$\frac{C + 2p_i - 1 + \sqrt{(C + 2p_i - 1)^2 - 4Cp_i^2}}{2C} = \frac{C + \sqrt{C^2 - C}}{2C}$$

$$= \frac{1}{2} + \frac{\sqrt{C^2 - C}}{2C} \geq \frac{1}{2}$$

Case (2) If $p_i < \frac{1}{2}$. Then

(i) $$p_i < \tfrac{1}{2} \Longrightarrow (2p_i - 1) < 0$$

$$(2p_i - 1)(C - 1) \leq 0$$

since

$$C - 1 = \frac{\chi^2}{N} \geq 0$$

(ii) $$\sqrt{(C + 2p_i - 1)^2 - 4Cp_i^2} \geq (2p_i - 1)(C - 1) =$$

$$2Cp_i - C - 2p_i + 1$$

since a positive is greater than a negative.

(iii) $$C + 2p_i - 1 + \sqrt{(C + 2p_i^2 - 1)^2 - 4Cp_i^2} \geq 2Cp_i$$

or

$$U_i = [(C + 2p_i - 1) + \sqrt{(C + 2p_i - 1)^2 - 4Cp_i^2}]/2C \geq p_i$$

Case (3) If $p_i > \frac{1}{2}$, then

(i) $$p_i > \tfrac{1}{2} \Longrightarrow (p_i - 1)(1 - C) \geq 0$$

since product of two negatives is a positive.

(ii) $$p_i + C - Cp_i - 1 \geq 0$$

(iii) $$-p_i \geq Cp_i - (C + 2p_i - 1)$$

(iv) $$-4Cp_i^2 \geq 4C^2p_i^2 - 4Cp_i(C + 2p_i - 1)$$

(v) $$(C+2p_i-1)^2-4Cp_i^2 \geq 4C^2p_i^2-4Cp_i(C+2p_i-1)+(C+2p_i-1)^2 =$$

$$= [2Cp_i-(C+2p_i-1)]^2 \geq 0$$

(vi) $$\sqrt{(C + 2p_i - 1)^2 - 4Cp_i^2} \geq 2Cp_i - (C + 2p_i - 1)$$

(vii) $C + 2p_i - 1 + \sqrt{(C + 2p_i - 1)^2 - 4Cp_i^2} \geq 2Cp_i$

(viii) $U_i = [(C + 2p_i - 1) + \sqrt{(C+2p_i-1)^2 - 4Cp_i^2}]/2C \geq p_i$.

Cases (1), (2), and (3) imply $U_i \geq p_i$. Parts (1), (2), (3), and (4) imply that theorem 4 is true.

These theorems establish several desirable properties of the intervals. The first theorem shows that the intervals always exist, which is certainly a necessary property. We must have intervals for all of the cells in order for the concept of simultaneous confidence intervals to be meaningful.

The property established by theorem 2, viz., that the intervals will converge upon the maximum likelihood estimator as the sample size increases, is certainly a satisfying result. Theorem 3 gives a result that is not too surprising. It shows that the lengths of intervals which correspond to p's that are at equal distances from the point $\frac{1}{2}$ are equal. It might be noted from the proof of this theorem that the length of the interval is determined essentially by the estimated variance of p_i. That the length of the interval and the variance of the estimator of π_i are so related is not a totally unexpected result.

The last theorem shows that all intervals will lie entirely in the closed interval [0, 1], which is desirable, for if an interval contained points outside [0, 1] these

points would be meaningless, as π_i could not assume such values.

V. Binomial Case

It is of interest to note that the above confidence intervals can be derived quite simply for the binomial case (k = 2). Consider the graphs of the equations (see figure 1)

(1) $$p_1^2/\pi_1 + p_2^2/\pi_2 = C$$

and

(2) $$\pi_1 + \pi_2 = 1.$$

The graph of the first equation will be a hyperbola since its discriminant is positive, and it will have a vertical asymptote at $\pi_1 = p_1^2/C$ and a horizontal asymptote at $\pi_2 = p_2^2/C$. In this case the maximum and minimum values of π_1 and π_2 are given by the coordinates of the points of intersection of the two curves. Solving the equations simultaneously gives the values

(3) $$\pi_i = [(C + 2p_i - 1) \pm \sqrt{(C + 2p_i - 1)^2 - 4Cp_i^2}]/2C$$
$$i = 1, 2.$$

Since this equation always gives real values for π_1 and π_2 by theorem 1 of the previous section, we are assured that the curves will always intersect. It should be observed that for this case the confidence intervals for π_1 and π_2 are completely dependent, for if the

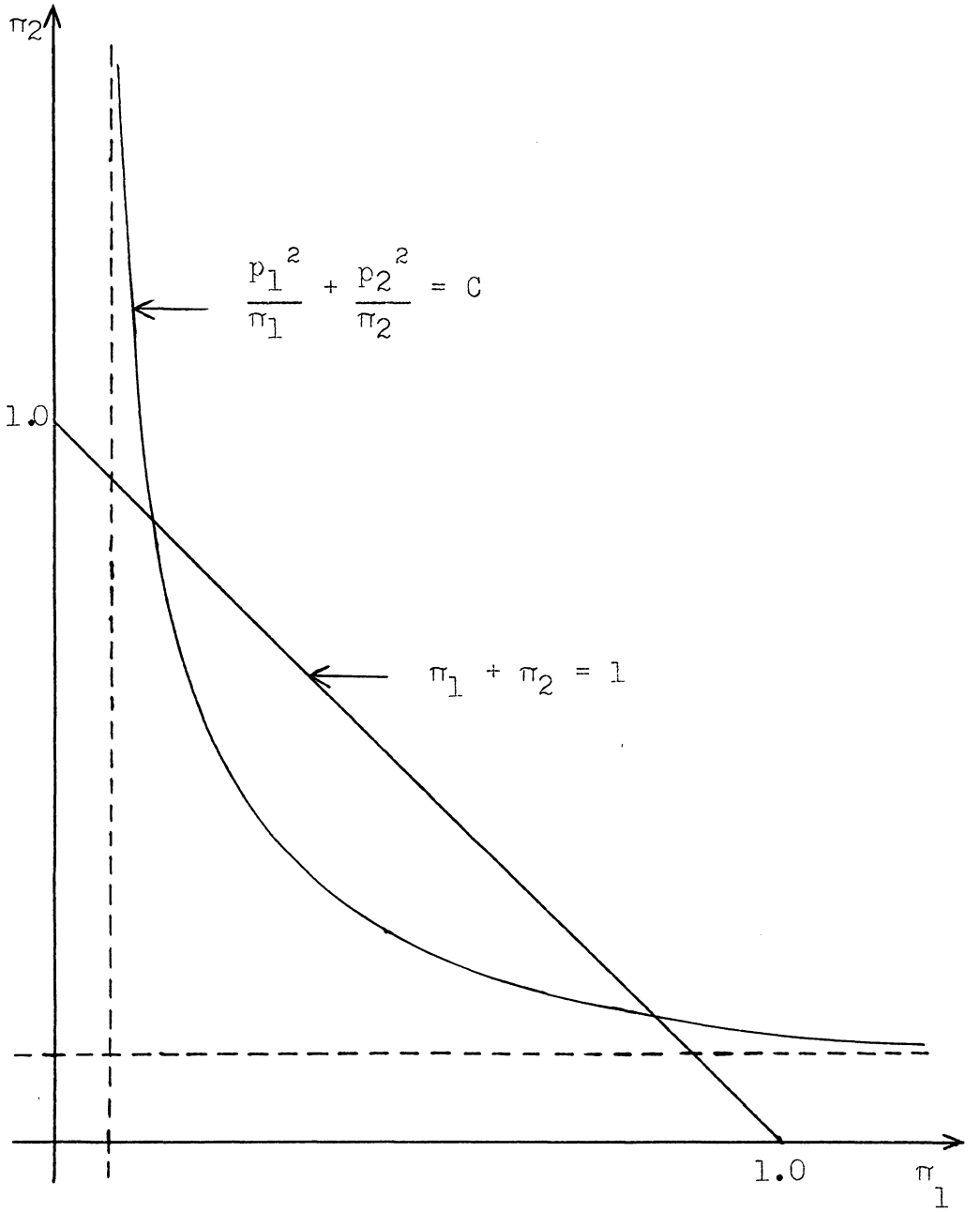

Figure 1. Graphical Solution for Simultaneous Limits for π_1 and π_2.

interval for one parameter is specified then the interval for the other one is completely determined.

Formula (3) can be written as

$$(4) \qquad \pi_i = \frac{\chi^2/N + 2p_i \pm 2\chi\sqrt{p_i(1 - p_i)/N + \chi^2/4N^2}}{2(\chi^2/N + 1)}$$

$$= \frac{N}{N + \chi^2}\left[p_i + \chi^2/2N \pm \chi\sqrt{p_i(1 - p_i)/N + \chi^2/4N^2}\right]$$

and, remembering that a χ^2 variate with 1 degree of freedom is the square of a standardized normal variate, we see that this formula is the same as the formula given by Cramer (1946, p. 515) based on the normal approximation to the binomial distribution, i.e.,

$$(5) \qquad \pi = \frac{N}{N + z^2_{1-\alpha/2}}\left[p + \frac{z^2_{1-\alpha/2}}{2N} \pm z_{1-\alpha/2}\sqrt{\frac{p(1 - p)}{N} + \frac{z^2_{1-\alpha/2}}{4N^2}}\right]$$

where $z_{1-\alpha/2}$ is the $1 - \alpha/2$ percentage point of the normal distribution.

Thus the intervals obtained for the case $k = 2$ (binomial distribution) by this method are exactly the same intervals as those obtained by the normal approximation to the binomial.

VI. Examples

We now apply this method of analysis to the data

given for the example in Section I and to that given for the two examples in Section II. We shall use the equation

$$\pi_i = \frac{\chi^2 + 2n_i \pm \sqrt{\chi^2[\chi^2 + 4n_i(N - n_i)/N]}}{2(\chi^2 + N)}$$

$$\chi^2 = \chi^2_{\alpha,k-1}$$

for computing.

Example 1. Choosing $\alpha = .3$ $(1 - \alpha = .7)$, then $\chi^2_{.3,9} = 10.656$. The computations can be carried out as illustrated in table 1. The intervals obtained are:

$.001 < \pi_1 < .022$	$.050 < \pi_6 < .109$
$.004 < \pi_2 < .031$	$.074 < \pi_7 < .142$
$.011 < \pi_3 < .046$	$.104 < \pi_8 < .180$
$.019 < \pi_4 < .061$	$.155 < \pi_9 < .248$
$.046 < \pi_5 < .103$	$.294 < \pi_{10} < .399$

The confidence coefficient is $1 - \alpha' \geq .70$. This means that if more samples were drawn and intervals computed for each, as the number of samples increases the proportion of the number of sets of intervals computed for which every interval contains its corresponding parameter will approach $1 - \alpha'$. In the same manner we can say with a confidence coefficient greater than .70 that any one of the above intervals will contain its parameter. If we are interested only in comparing the different types of defects we can rank them with a confidence coefficient greater

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
n_i	$N-n_i$	$4n_i(N-n_i)$	(3)/N	$\chi^2[\chi^2+(4)]$	$\sqrt{(5)}$	χ^2+2n_i	(7)-(6)	(7)+(6)	$\frac{(8)}{2(N+\chi^2)}$	$\frac{(9)}{2(N+\chi^2)}$
5	865	17,300	19.885	325.44	18.04	20.66	2.62	38.70	.001	.022
10	860	34,400	39.540	534.89	23.13	30.66	7.53	53.79	.004	.031
20	850	68,000	78.161	946.43	30.76	50.66	19.90	81.42	.011	.046
30	840	100,800	115.862	1348.77	36.74	70.66	33.92	107.40	.019	.061
60	810	194,400	223.448	2494.61	49.95	130.66	80.71	180.61	.046	.103
65	805	209,300	240.574	2677.11	51.74	140.66	88.92	192.40	.050	.109
90	780	280,800	322.759	3552.87	59.61	190.66	131.05	250.17	.074	.142
120	750	360,000	413.793	4522.93	67.25	250.66	183.41	317.91	.104	.180
170	700	476,000	547.126	5943.72	77.09	350.66	273.57	437.75	.155	.248
300	570	684,000	786.207	8491.37	92.15	610.66	518.51	702.81	.294	.399

Table 1. Computational Procedure for the Limits on Example 1, section VI.

than .70 in the following manner,

$$\underline{\pi_1 \quad \pi_2 \quad \pi_3 \quad \pi_4} \quad \pi_5 \quad \pi_6 \quad \pi_7 \quad \pi_8 \quad \pi_9 \quad \pi_{10},$$

where any two parameters are declared significantly different if they are not underscored by the same line.

Example 2. Confidence intervals for the data given in example 1 of section II are as follows:

$$.052 < \pi_1 < .183 \qquad .212 < \pi_3 < .406$$
$$.128 < \pi_2 < .299 \qquad .301 < \pi_4 < .508 ,$$

where $\alpha = .2$ or $1 - \alpha = .8$.

For this method of analysis the criticism advanced against the method of partitioning X^2, viz., that different orderings of the cells gave different results, obviously does not apply.

Example 3. Confidence intervals for the parameters in the second example in section II can be found for $\alpha = .30$ to be

$$.09296 < \pi_7 < .12447$$
$$.09562 < \pi_6 < .12749$$
$$.12885 < \pi_5 < .16470$$
$$.13324 < \pi_4 < .16955$$
$$.13493 < \pi_1 < .17142$$
$$.13933 < \pi_2 < .17627$$
$$.15767 < \pi_3 < .19634.$$

Ranking these parameters for easy comparison gives

$$\underline{\pi_7 \quad \pi_6} \quad \underline{\pi_5 \quad \pi_4 \quad \pi_1 \quad \pi_2 \quad \pi_3}.$$

This shows that the disinfectant began to inhibit growth at the 6th plate. We can also make more definite statements from these confidence intervals, such as plate 6 and plate 7 each have greater inhibiting power than any of the five other plates.

VII. Summary and Discussion

A method has been presented for analyzing a set of cell frequencies obtained by random sampling from a multinomial distribution. This method consists of finding a set of asymptotic simultaneous confidence intervals for the parameters of the distribution. Particular attention should be given to the term "simultaneous confidence intervals" and the associated term "confidence coefficient." The confidence coefficient does not refer to the probability that any set of intervals contain their corresponding parameters, for this probability is either 0 or 1. Rather, the confidence coefficient refers to the method used to obtain the intervals.

By using such a set of simultaneous confidence intervals the investigator obtains the whole class of sets of values of the parameters which are consistent with the sample. He can make statements about the ranges of values

which the individual parameters can assume, and statements about the relative magnitudes of the parameters with a confidence coefficient that is greater than some stated value.

The alternative methods of analysis available are to test hypotheses that the parameters have specified sets of values, or to make tests upon variates derived by partitioning the X^2 quantity into components corresponding to the individual degrees of freedom of χ^2. Any information that can be obtained by these methods is available (and more also) from a set of simultaneous confidence intervals.

The selection of the confidence coefficient has been left completely arbitrary. Further research would be helpful in making a more judicious choice. Such research should be directed towards establishing the relationship between the confidence coefficient for the set of simultaneous confidence intervals, $1 - \alpha$, and the confidence coefficients for the individual intervals, say $1 - \gamma_i$ $(i = 1, \ldots, k)$.

Bibliography

Bartlett, M. S. (1936). "The information available in small samples," Proc. Camb. Phil. Soc., 32, 560.

Bartlett, M. S. (1939). "Complete simultaneous confidence distributions," A.M.S., 10, 129.

Cochran, W. G. (1952). "The χ^2 test of goodness of fit," A.M.S., 23, 315.

Cochran, W. G. (1954). "Some methods for strengthening the common χ^2 tests," Biometrics, 10, 418.

Cramer, H. "Mathematical methods of statistics," Princeton University Press.

Dixon, W. J., and Massey, F. J., Jr. "Introduction to statistical analysis," McGraw-Hill Book Company, Inc.

Fisher, R. A. (1924). "The conditions under which χ^2 measures the discrepancy between observation and hypothesis," Journal of the Royal Statistical Society, 87, 442.

Irwin, J. O. (1949). "A note on the subdivision of χ^2 into components," Biometrics, 35, 130.

Kendall, M. G. "The advanced theory of statistics," Vols. I and II, Charles Griffin and Co., Ltd., London.

Lancaster, H. O. (1949). "The derivation and partition of χ^2 in certain discrete distributions," Biometrika 36, 117.

Neyman, J. (1935). "On the problem of confidence intervals," A.M.S., 6., 111.

Neyman, J. (1937). "Outline of a theory of statistical estimation based on the classical theory of probability," Phil. Trans. of the Roy. Soc., 236. 333.

Pearson, E. S. and Clopper, C. J. (1934). "Confidence or fiducial limits of the binomial," Biometrika, 26, 404.

Pearson, E. S. and Hartley, H. O., "Biometrika tables for statisticians," Vol. I.

Pearson, K. (1900). "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 reasonably be supposed to have arisen from random sampling," Philos. Mag. Series 5, 50, p. 157.

Rao, C. R. (1952). "Advanced statistical methods in biometric research," John Wiley and Sons, Inc.

Wald, A. (1939). "Contributions to theory of statistical estimation and testing hypotheses," A.M.S., 10, 299.

Wald, A. (1942). "Asymptotically shortest confidence intervals" A.M.S., 12, 127.

Welch, B. L. (1939). "On confidence limits and sufficiency with particular reference to parameters of location," A.M.S., 10, 58.

Wilks, S. S. (1939a). "Shortest average confidence intervals from small samples," A.M.S., 9, 60.

Wilks, S. S. (1939b). "Fiducial distributions in fiducial inference," A.M.S., 9, 272.

Wilks, S. S. and Daly, J. F. (1939). "An optimum property of confidence regions associated with the likelihood function," A.M.S., 10, 225.

APPENDIX

In order that the approximation of X^2 by χ^2 be adequate various writers have recommended that the minimum value of the smallest expected cell frequency should not be less than 5, 10, and even as much as 20. Cochran (1954) expresses the opinion that a minimum value of 5 is too conservative, and that the approximation may be adequate even if as many as 20 of the expected cell frequencies are less than 5, and that expected cell frequencies of 1 are allowable. He further warns that eliminating cells with small expected values by grouping such cells may be hazardous for it tends to reduce the power of the test.

ACKNOWLEDGEMENTS

The author wishes to express his sincere appreciation to the faculty of the department of statistics for their many contributions towards making his graduate program successful. Special appreciation is expressed to Professor David C. Hurst for his encouragement and guidance during the performance of this research.

The vita has been removed from the scanned document

ABSTRACT

ASYMPTOTIC SIMULTANEOUS CONFIDENCE INTERVALS FOR THE PROBABILITIES OF A MULTINOMIAL DISTRIBUTION

by

Charles P. Quesenberry

Approximate formulae are derived for obtaining confidence intervals for the probabilities of a multinomial distribution. The approach used is to consider the Chi-square goodness of fit statistic as a function of the population parameters and to invert this function to obtain a set of simultaneous confidence intervals for the parameters.

The confidence coefficient for the set of simultaneous confidence intervals obtained by this procedure is conservative, i.e., the true probability that every interval covers its corresponding parameter will in general be greater than the coefficient obtained by this method. As the sample size increases the intervals will converge on the population parameters and will estimate them exactly in the limit.