SIMULTANEOUS CONFIDENCE INTERVALS FOR CONTRASTS AMONG MULTINOMIAL POPULATIONS[1]

By Leo A. Goodman

University of Chicago

1. Introduction and summary. In the present article, we shall present two different methods for obtaining simultaneous confidence intervals for judging contrasts among multinomial populations, and we shall compare the intervals obtained by these methods with the simultaneous confidence intervals obtained earlier by Gold [5] for all linear functions of multinomial probabilities. One of the methods presented herein is particularly suited to the situation where all contrasts among, say, I multinomial populations may be of interest, where each population consists of, say, J classes. The other method presented herein is suited to certain situations where a specific set of contrasts among these populations is of interest; e.g., where, for each of the $\frac{1}{2}I(I-1)$ pairs of populations, the J contrasts between the corresponding probabilities associated with the two populations in the pair are of interest. For judging all contrasts among the I multinomial populations, the confidence intervals obtained with the first method presented herein have the desirable property that they are shorter than the corresponding intervals obtained with the method presented by Gold [5]. For judging the $\frac{1}{2}I(I-1)J$ pair-wise contrasts between the multinomial populations, the confidence intervals obtained with the second method presented herein have the desirable property that they are shorter than the corresponding intervals obtained with the first method, for the usual probability levels.

In the present paper we shall also solve a problem first mentioned by Gold [5] but left unsolved in the earlier article. Gold took note of the fact that, in the usual analysis of variance context, the simultaneous confidence intervals obtained by Scheffé [14] and by Tukey [15] for judging contrasts among the parameters have the desirable property that rejection of the homogeneity hypothesis by the usual F or Studentized range test implies the existence of at least one relevant contrast for which the corresponding confidence interval does not cover zero (see, for example, [14], pp. 66–77). She also noted that a result analogous to the Scheffé-Tukey result had not yet been obtained for her simultaneous confidence intervals, and she stated that the difficulty seemed to be that the homogeneity test is based on a χ^2 statistic with $(I-1)(J-1)$ degrees of freedom in the case where I multinomial populations, each consisting of J classes, are tested for homogeneity, whereas her confidence intervals were based

Received 9 April 1963; revised 2 January 1964.

[1] This research was supported in part by Research Grant No. NSF-G21058 from the Division of Mathematical, Physical and Engineering Sciences of the National Science Foundation.

upon the χ^2 distribution with $I(J-1)$ degrees of freedom. In the present article, one of the methods we shall present for obtaining simultaneous confidence intervals for the contrasts among the I multinomial populations will be based upon the χ^2 distribution with $(I-1)(J-1)$ degrees of freedom, and these intervals will have desirable properties somewhat analogous to those enjoyed in the analysis of variance by the Scheffé confidence intervals and by the Tukey confidence intervals. A modification of the usual test of the null hypothesis that the I multinomial populations are homogeneous will be presented herein, and we shall show that this modified test will lead to rejection of the null hypothesis if and only if there is at least one contrast, of the kind presented herein, for which the relevant confidence interval does not cover zero.

2. Contrasts among multinomial populations. Let Π_{ij} be the probability that an observation in the ith multinomial population $(i = 1, 2, \cdots, I)$ will fall in the jth class $(j = 1, 2, \cdots, J)$. Let $\sum_{j=1}^{J} \Pi_{ij} = 1$ for all i. We define a contrast among the I multinomial populations to be a linear function of the Π_{ij}, $\sum_{i,j} c_{ij}\Pi_{ij}$, where the c_{ij} are known constants subject to the condition that $\sum_{i=1}^{I} c_{ij} = 0$ for all j. For example, the $(I-1)(J-1)$ functions $\Delta_{ij} = \Pi_{ij} - \Pi_{1j}$ $(i = 2, 3, \cdots, I; j = 2, 3, \cdots, J)$ are contrasts among the I multinomial populations. Note that the contrast $\Pi_{i1} - \Pi_{11}$ can be written as $-\sum_{j=2}^{J} \Delta_{ij}$; i.e., this contrast is a linear function of the Δ_{ij}. More generally, all contrasts are linear functions of the $(I-1)(J-1)$ contrasts Δ_{ij} $(i = 2, 3, \cdots, I; j = 2, 3, \cdots, J)$, and all linear functions of these Δ_{ij} are also contrasts.

In the special case where $c_{ij} = c_{i1}$ for all i, j, then $\sum_{i,j} c_{ij}\Pi_{ij} = \sum_{i=2}^{I} \sum_{j=2}^{J} (c_{ij} - c_{i1})\Delta_{ij} = 0$. For simplicity, we shall exclude this case from our consideration and shall assume throughout that the c_{ij} are subject to the condition that $c_{ij} \neq c_{i1}$ for some i, j. In other words, we shall assume herein that the $(I-1)(J-1) \times 1$ vector of constants $c_{ij} - c_{i1}$ $(i = 2, 3, \cdots, I; j = 2, 3, \cdots, J)$, the coefficients of the Δ_{ij}, is a non-zero vector.

Let n_i be the number of observations from the ith multinomial population, and let n_{ij} denote the number of observations in the jth class in this population. Then $\sum_{j=1}^{J} n_{ij} = n_i$. Denoting n_{ij}/n_i by p_{ij}, the maximum likelihood estimators of Π_{ij} and Δ_{ij} are p_{ij} and d_{ij}, respectively, where $d_{ij} = p_{ij} - p_{1j}$. Denoting the contrast $\sum_{i,j} c_{ij}\Pi_{ij}$ by θ, the maximum likelihood estimator of θ is $\hat{\theta} = \sum_{i,j} c_{ij}p_{ij}$. The variance of p_{ij} is $\Pi_{ij}(1 - \Pi_{ij})/n_i = \nu_{ij}$, the variance of d_{ij} is $\nu_{ij} + \nu_{1j}$, and the variance of $\hat{\theta}$ is

$$\sigma^2(\hat{\theta}) = \sum_{i=1}^{I} n_i^{-1}\{ \sum_{j=1}^{J} c_{ij}^2\Pi_{ij} - [\sum_{j=1}^{J} c_{ij}\Pi_{ij}]^2\}.$$

These three variances can be estimated consistently by $p_{ij}(1 - p_{ij})/n_i = g_{ij}$, $g_{ij} + g_{1j}$, and

$$S^2(\hat{\theta}) = \sum_{i=1}^{I} n_i^{-1}[\sum_{j=1}^{J} c_{ij}^2 p_{ij} - (\sum_{j=1}^{J} c_{ij}p_{ij})^2],$$

respectively. In presenting asymptotic results, we assume that $n_i/n \to \rho_i > 0$ as $\sum_{i=1}^{I} n_i = n \to \infty$. When the c_{ij} have been specified *a priori*, it is well known that as $n \to \infty$ the probability will approach $1 - \alpha$ that

$$\hat{\theta} - S(\hat{\theta})Z_\alpha \leqq \theta \leqq \hat{\theta} + S(\hat{\theta})Z_\alpha,$$

here Z_α is the $100(1 - \frac{1}{2}\alpha)$th percentile of the unit normal distribution. When the c_{ij} have not been specified *a priori*, the following theorem can be applied:

THEOREM 1. *As $n \to \infty$, the probability will approach $1 - \alpha$ that simultaneously for all functions θ*

$$\hat{\theta} - S(\hat{\theta})L \leqq \theta \leqq \hat{\theta} + S(\hat{\theta})L.$$

Here the functions θ are subject to the condition that $\sum_{i=1}^{I} c_{ij} = 0$ for all j, and L is the positive square root of the $100(1 - \alpha)$th percentile of the χ^2 distribution with $(I - 1)(J - 1)$ degrees of freedom.

PROOF. The variance of d_{ij} was given earlier herein, and the covariance between d_{ij} and d_{hk} can be calculated from

$$\begin{aligned} C(d_{ij}, d_{hk}) &= -\Pi_{ij}\Pi_{ik}n_i^{-1} - \Pi_{1j}\Pi_{1k}n_1^{-1}, && \text{for } h = i \text{ and } j \neq k, \\ &= -\Pi_{1j}\Pi_{1k}n_1^{-1}, && \text{for } h \neq i \text{ and } j \neq k, \\ &= \Pi_{1j}(1 - \Pi_{1j})n_1^{-1}, && \text{for } h \neq i \text{ and } j = k. \end{aligned}$$

These variances and covariances can be estimated consistently by replacing the Π_{ij} in the formulae by the correspondng p_{ij}. Writing the $(I - 1)(J - 1)$ contrasts Δ_{ij} $(i = 2, 3, \cdots, I; j = 2, 3, \cdots, J)$ as the $(I - 1)(J - 1) \times 1$ column vector $\mathbf{\Delta}$ and the corresponding $(I - 1)(J - 1)$ estimated contrasts d_{ij} as the $(I - 1)(J - 1) \times 1$ column vector $\mathbf{d}$, then the probability will approach $1 - \alpha$ that $(\mathbf{\Delta} - \mathbf{d})'B^{-1}(\mathbf{\Delta} - \mathbf{d}) \leqq L^2$, where B is the estimated variance-covariance matrix of $\mathbf{d}$. In other words, the probability will approach $1 - \alpha$ that the true parameter point $\mathbf{\Delta}$ lies inside the ellipsoid $(\mathbf{x} - \mathbf{d})'B^{-1}(\mathbf{x} - \mathbf{d}) \leqq L^2$, where $\mathbf{x}$ denotes any point in the $(I - 1)(J - 1)$-dimensional space of possible values of $\mathbf{\Delta}$. But $\mathbf{\Delta}$ lies inside this ellipsoid if and only if for every non-zero vector $\mathbf{h}$ in $(I - 1)(J - 1)$-dimensional space $|\mathbf{h}'(\mathbf{\Delta} - \mathbf{d})| \leqq (\mathbf{h}'M^{-1}\mathbf{h})^{\frac{1}{2}}$, where $M = L^{-2}B^{-1}$ (see [14], p. 69). Hence, the probability will approach $1 - \alpha$ that for all $\mathbf{h}$, $|\mathbf{h}'(\mathbf{\Delta} - \mathbf{d})| \leqq L(\mathbf{h}'B\mathbf{h})^{\frac{1}{2}}$. Note that $\mathbf{h}'B\mathbf{h}$ is the estimated variance of $\mathbf{h}'\mathbf{d}$. Since $\mathbf{h}'\mathbf{\Delta}$ is a contrast among the multinomial populations, and all contrasts among the multinomial populations are of this form, the probability will approach $1 - \alpha$ that simultaneously for all contrasts θ, $|\theta - \hat{\theta}| \leqq LS(\hat{\theta})$. Q.E.D.

The confidence intervals obtained using this theorem are shorter than the corresponding confidence intervals obtained using the method presented by Gold [5], since in the former method L is based upon the percentiles of the χ^2 distribution with $(I - 1)(J - 1)$ degrees of freedom, whereas in the latter method the corresponding quantity is based upon the percentiles of the χ^2 distribution with $I(J - 1)$ degrees of freedom. For judging all contrasts among the multinomial populations, the additional $(J - 1)$ degrees of freedom in the latter method increase the length of the corresponding confidence intervals and are unnecessary. If in addition to judging all contrasts we also wish to obtain simultaneous confidence intervals for all linear functions of the multinomial probabilities (where the c_{ij} are not necessarily subject to the condition that $\sum_{i=1}^{I} c_{ij} = 0$ for all j), then the additional $(J - 1)$ degrees of freedom are needed.

The confidence intervals obtained with Theorem 1 are suited to the case where all contrasts among the I multinomial populations may be of interest. However, if only a specified subset of these contrasts is of interest, it will sometimes be possible to determine simultaneous confidence intervals pertaining to this subset which are shorter than the corresponding intervals obtained by the methods presented above.

For example, in judging the contrasts between I multinomial populations, each consisting of J classes, interest may be restricted to the $\frac{1}{2}I(I - 1)J$ contrasts

$$\Delta_{ijh} = \Pi_{ij} - \Pi_{hj} \qquad (i \neq h, j = 1, 2, \cdots, J),$$

or more generally interest may be restricted to a specified set of, say, G contrasts, $\theta_1, \theta_2, \cdots, \theta_G$, and in this case simultaneous confidence intervals can be based upon the fact that simultaneously for the G contrasts

$$\lim_{n\to\infty} \Pr\{\hat{\theta}_i - S(\hat{\theta}_i)Z_i \leqq \theta_i \leqq \hat{\theta}_i + S(\hat{\theta}_i)Z_i, \text{ for } i = 1, 2, \cdots, G\} \geqq 1 - \alpha$$

where Z_i is the $100(1 - \beta_i)$th percentile of the unit normal distribution, and $\sum_{i=1}^{G} \beta_i = \frac{1}{2}\alpha$. (See [16], p. 291 for related results.) If we take $Z_1 = Z_2 = \cdots = Z_G = Z$, where Z is the $100(1 - \beta)$th percentile of the unit normal distribution and $\beta = \alpha/(2G)$, then the confidence intervals obtained thereby will be shorter than the corresponding confidence intervals obtained using Theorem 1 herein if and only if

$$(1) \qquad Z \leqq L,$$

where L is the $100(1 - \alpha)$th percentile of the χ distribution with $(I - 1)(J - 1)$ degrees of freedom. For a given value of α, whether or not this Condition (1) is satisfied will depend upon the relative magnitudes of G and $(I - 1)(J - 1)$.

Let us for the moment consider the special case where $J = 2$ and where the $\frac{1}{2}I(I - 1)$ pair-wise contrasts $\Delta_{i2h} = \Pi_{i2} - \Pi_{h2}$ are of interest. In this special case, $\Delta_{i1h} = \Pi_{i1} - \Pi_{h1} = -\Delta_{i2h}$, and so confidence intervals for the Δ_{i2h} would also provide confidence intervals for the Δ_{i1h}, though we would actually calculate confidence intervals for only $\frac{1}{2}I(I - 1)$ contrasts. It is easy to see in this case that Condition (1) is satisfied for the usual values of α (viz., $\alpha = .05$ and $.01$), and so the L of Theorem 1 can be replaced here by the value of Z defined above when calculating confidence intervals. More generally, in the case where the $\frac{1}{2}I(I - 1)J$ pair-wise contrasts $\Delta_{ijh} = \Pi_{ij} - \Pi_{hj}$ $(i \neq h; j = 1, 2, \cdots, J)$ are of interest, setting G equal to $\frac{1}{2}I(I - 1)J$ (when $J > 2$), we find in this case that Condition (1) is satisfied for the usual values of α (.05 or .01), and so we would again use Z rather than L when calculating the relevant confidence intervals. (For somewhat related results, see Dunn [2].)

The procedure described above based upon Z takes advantage of the fact that interest may be focused upon a specified set of G contrasts, rather than the set of all contrasts. The Studentized range procedure developed by Tukey, which provides simultaneous confidence intervals in the analysis of variance context (see, for example, [15], [16], p. 294, [14], p. 73), also takes advantage of the fact that interest may be focused upon a specified set of contrasts, but it may be worth noting that this procedure can not be applied in the present context since one of the restrictions of the Tukey method is that the estimators have equal variance, which is not the case for the estimators of the parameters in the multinomial populations.

3. Chi-square tests of homogeneity. The null hypothesis H_0 that the I multinomial populations are homogeneous states that $\Pi_{1j} = \Pi_{2j} = \cdots = \Pi_{Ij} = \Pi_{.j}$ (for $j = 1, 2, \cdots, J$). When H_0 is true, the statistic

$$Y^2 = \sum_{i,j} (n_{ij} - n_i p_j)^2 / n_{ij}$$

will have a χ^2 distribution with $(I - 1)(J - 1)$ degrees of freedom, where p_j is set equal to $\bar{p}_j / \sum_{k=1}^J \bar{p}_k = p_j^*$, and $\bar{p}_j$ is the weighted harmonic mean $n / \sum_{i=1}^I n_i p_{ij}^{-1}$. These values of p_j^* are obtained by minimizing the Y^2 statistic subject to the condition that $\sum_{j=1}^J p_j = 1$. The Y^2 statistic can be written more simply as

$$Y^2 = \sum_{j=1}^J \{ [p_j^*]^2 [\sum_{i=1}^I n_i p_{ij}^{-1}] \} - n = n \{ [\sum_{j=1}^J \bar{p}_j]^{-1} - 1 \}.$$

The statistic Y^2 differs from the usual statistic presented for testing the null hypothesis H_0; viz.,

$$X^2 = \sum_{i,j} (n_{ij} - n_i p_j)^2 / n_i p_j ,$$

where p_j is set equal to the weighted arithmetic mean $\hat{p}_j = \sum_{i=1}^I n_i p_{ij} / n = n_{.j}/n$, and where $n_{.j} = \sum_{i=1}^I n_{ij}$. These values of $\hat{p}_j$ are obtained by minimizing X^2 subject to the condition that $\sum_{j=1}^J p_j = 1$. The X^2 statistic can be written as

$$X^2 = \sum_{j=1}^J \{ [\sum_{i=1}^I n_i p_{ij}^2] / \hat{p}_j \} - n = n \{ [\sum_{j=1}^J \sum_{i=1}^I [n_{ij}^2 / n_i n_{.j}]] - 1 \}.$$

When H_0 is true, the statistics X^2 and Y^2 will be asymptotically equivalent to each other (see [11]).

Writing the $(J - 1)$ probabilities $\{\Pi_{i2}, \Pi_{i3}, \cdots, \Pi_{iJ}\}$ as the $(J - 1) \times 1$ column vector $\mathbf{\Pi}_i$, the null hypothesis H_0 is equivalent to the hypothesis that $\mathbf{\Pi}_1 = \mathbf{\Pi}_2 = \cdots = \mathbf{\Pi}_I = \mathbf{\Pi}$. Writing the $(J - 1)$ estimates $\{p_{i2}, p_{i3}, \cdots, p_{iJ}\}$ as the $(J - 1) \times 1$ column vector $\mathbf{p}_i$, we see that the statistic Y^2 can be written as

$$Y^2 = \sum_{i=1}^I (\mathbf{p}_i - \mathbf{\Pi})' D_i^{-1} (\mathbf{p}_i - \mathbf{\Pi}),$$

where D_i is the estimated variance-covariance matrix of the $\{p_{i2}, p_{i3}, \cdots, p_{iJ}\}$, and $\mathbf{\Pi}$ denotes the point in the $(J - 1)$-dimensional space of possible values of $\{\Pi_{\cdot 2}, \Pi_{\cdot 3}, \cdots, \Pi_{\cdot J}\}$ which minimizes Y^2. To test the null hypothesis of homogeneity using Y^2, we note that the probability will approach $1 - \alpha$ that $Y^2 \leqq L^2$, where L^2 is the $100(1 - \alpha)$th percentile of the χ^2 distribution with $(I - 1) \cdot (J - 1)$ degrees of freedom. The following theorem describes the relationship between this test of homogeneity and the simultaneous confidence intervals presented in Theorem 1 for all contrasts θ:

THEOREM 2. *The test of homogeneity based upon Y^2 will reject H_0 if and only if at least one estimated contrast is significantly different from zero.*

PROOF. The null hypothesis H_0 states that $\mathbf{\Delta}_i = \mathbf{\Pi}_i - \mathbf{\Pi}_1 = \mathbf{0}$ for $i = 2, 3, \cdots, I$, where $\mathbf{0}$ is the $(J - 1) \times 1$ column vector of zeros. We can estimate $\mathbf{\Delta}_i$ by $\mathbf{p}_i - \mathbf{p}_1 = \mathbf{d}_i$. Writing the $\{\mathbf{d}_2, \mathbf{d}_3, \cdots, \mathbf{d}_I\}$ as the $(I - 1) \times 1$ column vector $\mathbf{\Delta}$, we see that $\mathbf{d}$ is actually a $(I - 1)(J - 1) \times 1$ column vector with entries $p_{ij} - p_{1j}$ $(i = 2, 3, \cdots, I; j = 2, 3, \cdots, J)$. We can test H_0 using the statistic

$$W^2 = \mathbf{d}'B^{-1}\mathbf{d},$$

where B is the estimated variance-covariance matrix of $\mathbf{d}$. The statistic W^2 is actually equal to Y^2 (see, for example, [7], [10]). From the proof of Theorem 1 herein, we see that $W^2 \geqq L^2$ if and only if there is at least one estimated contrast, say, $\hat{\theta}$, such that $|\hat{\theta}| \geqq LS(\hat{\theta})$. In this case, $\hat{\theta}$ is significantly different from zero. Q.E.D.

From Theorem 2 we see that the simultaneous confidence intervals presented in Theorem 1 for all contrasts can be used to supplement the test of the null hypothesis of homogeneity based upon the statistic Y^2. If the test based upon Y^2 leads to acceptance of the null hypothesis then all simultaneous confidence intervals would include zero, but if the test leads to rejection of this null hypothesis we could then calculate the simultaneous confidence intervals to determine which particular contrasts are significantly different from zero and thus determine the particular ways in which the multinomial populations are not homogeneous.

With any simultaneous confidence interval procedure we can, of course, associate a testing procedure for any null hypothesis pertaining to the relevant parameters. We would accept the null hypothesis when and only when the simultaneous confidence intervals included the parameter values specified under the null hypothesis. Thus, in testing the null hypothesis of homogeneity using the test procedure associated with the simultaneous confidence intervals of Theorem 1 herein, we would accept the null hypothesis when and only when zero is included in the simultaneous confidence intervals for all contrasts among the multinomial populations. We have now shown that, in testing this null hypothesis, the calculation of the simultaneous confidence intervals for all contrasts can be replaced by the calculation of a single statistic; viz., Y^2. The user will find the calculation of Y^2 easier than the calculation of the simultaneous confidence intervals for all the contrasts, particularly in cases where the null hypothesis is accepted. (Even in cases where the null hypothesis may be rejected, if the user can not say beforehand which contrasts he would expect to be significantly different from zero he may prefer to first calculate Y^2 to determine whether there are any contrasts at all that are significant.) This simplification applies to the first method presented in Section 2 herein for obtaining simultaneous confidence intervals (Theorem 1), but it does not apply to the second method presented in that section, where attention was focused upon a specified set of G contrasts. Although the second method for obtaining simultaneous confidence intervals also provides a test of the null hypothesis of homogeneity (the null hypothesis would be accepted when and only when zero is included in the simultaneous confidence intervals for all G contrasts), this test involves the calculation of G statistics (viz. $\hat{\theta}_i/S(\hat{\theta}_i)$ for $i = 1, 2, \cdots, G$) and it can not be replaced by a simpler but equivalent procedure.

The Y^2 statistic presented herein is a rather simple chi-square-like statistic, though it differs somewhat from the usual chi-square statistic X^2. We have found for Y^2 a confidence interval procedure (see Theorem 1) for which the associated test of homogeneity is equivalent to the test based upon Y^2. To see why results of this kind are readily obtained for Y^2 but not for X^2 it is sufficient to consider the special case where $I = J = 2$. In this case, both X^2 and Y^2 can be written as $(p_{11} - p_{21})^2$ divided by estimates of the variance of $p_{11} - p_{21}$, where for X^2 the estimated variance is consistent only if the null hypothesis is correct, and for Y^2 the estimate is consistent whether or not the null hypothesis is correct. The estimated variance applied in the calculation of Y^2 can also be used in the calculation of the corresponding confidence intervals, whereas the estimated variance applied in the calculation of X^2 can not be so used since it is in general not a consistent estimator. Thus, we do not find a direct relation between X^2 and the confidence interval procedure, whereas we do find such a relation for Y^2.

4. Statistical methods for Markov chains. In testing certain hypotheses about finite Markov chains, Anderson and Goodman [1] and Gold [5] have derived asymptotic χ^2 tests which are analogous to the usual χ^2 tests of homogeneity among multinomial populations or to the corresponding χ^2 tests of independence in contingency tables. For example, considering a Markov chain having K states, and denoting by $P(t)$ the $K \times K$ transition probability matrix associated with transitions from time t to time $t + 1$, the null hypothesis that $P(1) = P(2) = \cdots = P(T)$ can be tested when N observed sequences are obtained from the chain at times $t = 1, 2, \cdots, T + 1$, by forming K "contingency tables," each having T rows and K columns, and using a test statistic having an asymptotic χ^2 distribution with $K(T - 1)(K - 1)$ degrees of freedom (when $N \rightarrow \infty$) appropriate for testing the hypothesis that there is independence between rows and columns in each of the K contingency tables (see [1], [5]). These test statistics are analogous to the statistics presented earlier herein for testing the hypothesis of homogeneity among T multinomial populations, each having K

classes, when these tests are performed separately on K different sets of T multinomial populations and the K corresponding statistics are summed. Proceeding in a manner similar to [5], we shall present a theorem which yields simultaneous confidence intervals for judging all contrasts among the transition probabilities which are relevant to the hypothesis that $P(1) = P(2) = \cdots = P(T)$. For judging these contrasts, the confidence intervals obtained thereby will be smaller than the corresponding confidence intervals given in [5].

Let $\Pi_{ij}(t)$ be the transition probability that an observation will be in state j at time $t + 1$ given that it was in state i at time t. Assume that $\Pi_{ij}(t) > 0$ for $i = 1, 2, \cdots, K; j = 1, 2, \cdots, K; t = 1, 2, \cdots, T$. Let $n_i(t)$ be the number of observations in state i at time t, and let $n_{ij}(t)$ be the observed frequency of transitions from state i at time t to state j at time $t + 1$. In a manner analogous to the definition of the contrasts among multinomial populations presented in Section 2 herein, the contrasts ψ among the transition probabilities, which are relevant to the hypothesis that $P(1) = P(2) = \cdots = P(T)$, are defined as $\psi = \sum_{t=1}^{T} \sum_{i=1}^{K} \sum_{j=1}^{K} b_{ij}(t)\Pi_{ij}(t)$, where the $b_{ij}(t)$ are any constants subject to the condition that $\sum_{t=1}^{T} b_{ij}(t) = 0$ for all i and j, and $b_{ij}(t) \neq b_{i1}(t)$ for some i, j, t. The estimated contrasts are

$$\hat{\psi} = \sum_{t=1}^{T} \sum_{i=1}^{K} \sum_{j=1}^{K} b_{ij}(t)p_{ij}(t), \qquad \text{where } p_{ij}(t) = n_{ij}(t)/n_i(t).$$

Let

$$S^2(\hat{\psi}) = \sum_{i,t} [n_i(t)]^{-1}[\sum_{j=1}^{K} b_{ij}^2 p_{ij}(t) - (\sum_{j=1}^{K} b_{ij}(t)p_{ij}(t))^2].$$

We then have the following theorem analogous to Theorem 1 herein:

THEOREM 3. *As* $N = \sum_{i=1}^{K} n_i(t) \to \infty$, *the probability will approach* $1 - \alpha$ *that simultaneously for all functions* ψ

$$\hat{\psi} - S(\hat{\psi})M \leqq \psi \leqq \hat{\psi} + S(\hat{\psi})M,$$

where M *is the positive square root of the* $100(1 - \alpha)$th *percentile of the* χ^2 *distribution with* $K(T - 1)(K - 1)$ *degrees of freedom.*

We shall not give the proof here since it follows in a quite straightforward fashion from the earlier results. Note that the simultaneous confidence intervals obtained using Theorem 3 are based upon the percentiles of the χ^2 distribution with $K(T - 1)(K - 1)$ degrees of freedom, whereas the corresponding confidence intervals given in [5] are based upon the percentiles of the χ^2 distribution with $KT(K - 1)$ degrees of freedom. For judging the relevant contrasts, the additional $K(K - 1)$ degrees of freedom appearing in the latter method increase the length of the corresponding confidence intervals and are unnecessary.

In addition to testing the hypothesis that $P(1) = P(2) = \cdots = P(T)$, it is possible to test other hypotheses concerning the Markov chain (e.g., hypotheses concerning the order of the chain) using asymptotic χ^2 tests which are analogous to the usual χ^2 tests of homogeneity between multinomial populations (see [1], [5]). Applying the methods presented herein, it is now also possible to obtain simultaneous confidence intervals for judging all contrasts which are relevant to these hypotheses.

In the preceding sections, we presented two different methods for obtaining simultaneous confidence intervals for the contrasts among multinomial populations, and we discussed the equivalence between the test of homogeneity associated with one of these methods and a modification of the usual chi-square test. Theorem 3 provides a method for obtaining simultaneous confidence intervals for contrasts among transition probabilities in the Markov chain which is analogous to one of the methods presented earlier (Theorem 1), and a different method for Markov chains analogous to the second method presented earlier herein can be obtained similarly. In addition, results concerning the equivalence between the test of homogeneity associated with the simultaneous confidence intervals given in Theorem 3 and a modification of the usual chi-square test can also be obtained, but we shall not go into these details here since they are quite straightforward.

5. Further remarks. We can, of course, view the I multinomial populations each consisting of J classes, which were studied earlier herein, as an $I \times J$ contingency table with fixed row marginal frequencies. The contrasts among the multinomial populations, which were defined herein, can be viewed as measures of the association in the $I \times J$ table. To illustrate how these contrasts differ from other measures of association suggested in the literature, let us consider the special case where $I = J = 2$. In this case, the contrasts considered herein reduce to $\Pi_{11} - \Pi_{21}$, whereas the measures of association studied earlier by, for example, Edwards [3], Fisher [4], Goodman [7], [8] or Plackett [12] were based upon the cross-product ratio $\Pi_{11}\Pi_{22}/(\Pi_{12}\Pi_{21}) = \rho$, the logarithm of ρ, or some other function of ρ. Which methods will be suitable in a particular situation will depend upon whether the difference $\Pi_{11} - \Pi_{21}$ or the cross-product ratio is of special interest in that situation. For further discussion of the case where differences are of interest, see [6].

For the $I \times J$ table, whether the methods presented herein or methods based upon the cross-product ratio or the generalized cross-product ratio (see [8]) should be applied in a given situation will depend again on which measures of association are of special interest in that case. In addition to the measures mentioned here, there are still other measures of association in the $I \times J$ table which may also be of interest. We shall not discuss them here but shall instead refer the reader to [9] where methods are given for obtaining confidence intervals for some of these measures.

REFERENCES

[1] Anderson, T. W. and Goodman, Leo A. (1957). Statistical inference about Markov chains. *Ann. Math. Statist.* **28** 89–110.

[2] Dunn, O. J. (1961). Multiple comparisons among means. *J. Amer. Statist. Assoc.* **56** 52–64.

[3] Edwards, A. W. F. (1963). The measure of association in a 2 × 2 table. *J. Roy. Statist. Soc. Ser. A* **126** 109–114.

[4] Fisher, R. A. (1962). Confidence limits for a cross-product ratio. *Austral. J. Statist.* **4** 41.

[5] Gold, Ruth Z. (1963). Tests auxiliary to χ^2 tests in a Markov chain. *Ann. Math. Statist.* **34** 56–74.

[6] Goodman, L. A. (1963). On methods for comparing contingency tables. *J. Roy. Statist. Soc. Ser. A* **126** 94–108.

[7] Goodman, L. A. (1963). On Plackett's test for contingency table interactions. *J. Roy. Statist. Soc. Ser. B* **25** 179–188.

[8] Goodman, L. A. (1964). Simultaneous confidence intervals for cross-product ratios in contingency tables. *J. Roy. Statist. Soc. Ser. B.* **26**.

[9] Goodman, L. A. and Kruskal, W. H. (1963). Measures of association for cross classifications. III: Approximate sampling theory. *J. Amer. Statist. Assoc.* **58** 310–364.

[10] James, G. S. (1954). Tests of linear hypotheses in univariate and multivariate analysis when the ratios of the population variances are known. *Biometrika* **41** 19–43.

[11] Neyman, Jerzy (1949). Contribution to the theory of the χ^2 test. *Proc. Berkeley Symp. Math. Statist. Prob.* 239–274. Univ. of California Press.

[12] Plackett, R. L. (1962). A note on interactions in contingency tables. *J. Roy. Statist. Soc. Ser. B* **24** 162–166.

[13] Scheffé, Henry (1953). A method for judging all contrasts in the analysis of variance. *Biometrika* **40** 87–104.

[14] Scheffé, Henry (1959). *The Analysis of Variance.* Wiley, New York.

[15] Tukey, John W. (1953). The problem of multiple comparisons. Dittoed MS, Princeton Univ.

[16] Wilks, S. S. (1962). *Mathematical Statistics.* Wiley, New York.