ASYMPTOTIC DISTRIBUTIONS OF "PSI-SQUARED" GOODNESS OF FIT CRITERIA FOR m-TH ORDER MARKOV CHAINS[1]

By Leo A. Goodman

University of Chicago

1. Introduction and Summary. Let $\{X_1, X_2, \cdots, X_N\}$ be an observed sequence from a stochastic process, where X_i can take any one of s values 1, 2, $\cdots$, s. Let f_{u} be the frequency of the m-tuple $\mathrm{u} = (u_1, u_2, \cdots, u_m)$ in the sequence. Let H_n' be the composite hypothesis that the process is a Markov chain of order n. Let H_n be any simple hypothesis belonging to H_n'. Let H_n^* be the maximum likelihood H_n. Let the expected value of f_{u} in a new sequence of length N given H_n be $f_{\mathrm{u},n}$, and given H_n^* be $f_{\mathrm{u},n}^*$. Let

$$\psi_{m,n}^2 = \sum_{\mathrm{u}} (f_{\mathrm{u}} - f_{\mathrm{u},n})^2/f_{\mathrm{u},n},$$

$$\psi_{m,n}^{*2} = \sum_{\mathrm{u}} (f_{\mathrm{u}} - f_{\mathrm{u},n}^*)^2/f_{\mathrm{u},n}^*,$$

$$\psi_{n+1,n}^{*2} = 0.$$

Good had proposed in [7] the following two conjectures: (a) that the asymptotic distribution ($N \to \infty$) of $\psi_{m,n}^{*2}$, when H_n' is true, is

$$\mathop{*}_{\lambda=1}^{m-n-1} K_{g(\lambda)}\,(x/\lambda),$$

where $*$ denotes convolution, $g(\lambda) = (s-1)^2 s^{m-1-\lambda}$, and $K_i(x)$ is the χ^2-distribution with i degrees of freedom; (b) that the asymptotic distribution of $\psi_{m,n}^2$, when H_n is true, is

$$\mathop{*}_{\lambda=1}^{m-1} K_{g(\lambda)}(x/\lambda) * K_{s-1}(x/m),$$

mathematically independent of n. Conjectures (a) and (b) were proved by Billingsley [2] for the special case $n = 0$. For the special case $n = -1$ (by convention, H_{-1}' is the hypothesis of equiprobable or perfect randomness (see [7])), Conjecture (b) was proved by Good [5] when s is prime. In the present paper, Conjecture (a) will be proved for the general case $n \geqq -1$; conjecture (b) will be shown to be incorrect for $n > 0$, although a modified version of (b) will be proved for $n \geqq -1$. A third conjecture by Good [6] will also be proved here. It was

Received December 13, 1957.

[1] Research carried out at the Statistical Research Center, University of Chicago, under sponsorship of the Statistics Branch, Office of Naval Research. Reproduction in whole or in part is permitted for any purpose of the United States Government.

I am indebted to I. J. Good for the opportunity to read [7] before its publication, and also for mentioning that he suspected that his conjectures in [7] could be proved with the aid of the results in my earlier paper [10].

assumed in these earlier papers, and it will be assumed here, that all transition probabilities in the Markov chain are positive; the results can be modified accordingly when some of these probabilities are zero (see [1] and [10]).

Let $M_{m,n} = -2 \log \lambda_{n,m-1}$, where $\lambda_{n,m-1}$ is the ratio of the maximum likelihood given H'_n to that given H'_{m-1} (see [6]). For $m = n + 2$, the statistic $\psi^{*2}_{m,n}$ is asymptotically equivalent, when H'_n is true, to the likelihood ratio statistic $M_{m,n}$. For $m > n + 2$, $\psi^{*2}_{m,n}$ is asymptotically equivalent, when H'_n is true, to $\sum_{\lambda=1}^{m-n-1} \lambda M_{m+1-\lambda,m-1-\lambda}$, while $M_{m,n}$ is asymptotically equivalent to

$$\sum_{\lambda=1}^{m-n-1} M_{m+1-\lambda,m-1-\lambda}$$

(see [6], [10]). Thus, $\psi^{*2}_{m,n}$ corresponds asymptotically to a weighted sum of the likelihood ratio statistics $M_{n+2,n}, M_{n+3,n+1}, \cdots, M_{m,m-2}$, with the weights $m - n - 1, m - n - 2, \cdots, 1$, respectively, while $M_{m,n}$ weights these statistics equally (see [13] and reference to [13] in Section 4 herein).

Let $L_{m,n} = -2 \log \mu_{n,m-1}$, where $\mu_{n,m-1}$ is the ratio of the likelihood given H_n to the maximum likelihood given H'_{m-1}. For $m - 1 = n = 0$, the statistic $\psi^2_{m,n}$ is asymptotically equivalent, when H_n is true, to $L_{m,n}$. For $m - 1 > n = 0$, $\psi^2_{m,n}$ is asymptotically equivalent, when H_n is true, to

$$\sum_{\lambda=1}^{m-1} \lambda M_{m+1-\lambda,m-1-\lambda} + mL_{n+1,n},$$

while $L_{m,n}$ is asymptotically equivalent to $\sum_{\lambda=1}^{m-1} M_{m+1-\lambda,m-1-\lambda} + L_{n+1,n}$. For $n > 0$, the relation between $\psi^2_{m,n}$ and the likelihood ratio statistics $L_{m,n}$ and $M_{m,n}$ is not so straightforward. However, a modification $\psi'^2_{m,n}$ of $\psi^2_{m,n}$ (see Section 6 herein) is asymptotically equivalent, when H_n is true, to $L_{m,n}$ for $m = n + 1$, and to $\sum_{\lambda=1}^{m-n-1} \lambda M_{m+1-\lambda,m-1-\lambda} + (m - n)L_{n+1,n}$ for $m > n + 1$; while the likelihood ratio statistic $L_{m,n}$ is asymptotically equivalent to

$$\sum_{\lambda=1}^{m-n-1} M_{m+1-\lambda,m-1-\lambda} + L_{n+1,n}.$$

In [10], the m-tuple u was "split" into an $(m - n - 1)$-tuple, an n-tuple, and a 1-tuple; thus obtaining s^n "contingency tables" $(n \geqq 0)$ each $s^{m-n-1} \times s$ (see [10]). The statistic $M_{m,n}$ can be seen to be asymptotically equivalent to the sum of the "likelihood ratio statistics" (for testing "independence" in each table) for the s^n tables, and the asymptotic distribution, when H'_n is true, of $M_{m,n}$ will be χ^2 with $s^n(s^{m-n-1} - 1)(s - 1) = s^m - s^{m-1} - s^{n+1} + s^n$ degrees of freedom. It is also possible to "split" the m-tuple u into an $(m - n - 1 - r)$-tuple, an n-tuple, and a $(1 + r)$-tuple $(0 \leqq r \leqq m - n - 2)$; thus obtaining s^n "contingency tables," each $s^{m-n-1-r} \times s^{1+r}$ (see [10]). The sum ${}_rM_{m,n}$ of the likelihood ratio (or any equivalent goodness of fit) statistics for the s^n tables will have an asymptotic mean value, when H'_n is true, of

$$s^n(s^{m-n-1-r} - 1)(s^{1+r} - 1) = s^m - s^{m-r-1} - s^{n+1+r} + s^n.$$

but the asymptotic distribution will not be χ^2 unless $r = 0$ or $m - n - 2$. It can be seen, using the methods developed in the present paper, that the statistic ${}_rM_{m,n}$ will be asymptotically equivalent, when H'_n is true, to

$$\sum_{\lambda=1}^{m-n-1} h(\lambda) M_{m+1-\lambda, m-1-\lambda} ,$$

where

$$h(\lambda) = \begin{cases} \lambda & \text{for} \quad 0 < \lambda \leqq v \\ v & \text{for} \quad v \leqq \lambda \leqq m - n - v \\ (m - n - \lambda) & \text{for} \quad m - n - v \leqq \lambda \leqq m - n - 1, \end{cases}$$

and $v = \min\,[r + 1,\ m - n - r - 1]$. Thus, the asymptotic distribution $(N \to \infty)$ of ${}_rM_{m,n}$ (or the corresponding asymptotically equivalent goodness of fit statistics), when H'_n is true, is

$$\mathop{*}_{\lambda=1}^{m-n-1} K_{g(\lambda)}[x/(h(\lambda))] .$$

This result generalizes the earlier published results concerning the asymptotic distribution of the likelihood ratio statistic $M_{m,n}$ (or the corresponding asymptotically equivalent goodness of fit statistics) for testing the null hypothesis H'_n within H'_{m-1}, since ${}_rM_{m,n}$ for $r = 0$ or $m - n - 2$ is asymptotically equivalent to $M_{m,n}$ (see [6], [10]). A proof of this result will not be given since the method of proof is quite similar to that presented here for the asymptotic distribution of $\psi^{*2}_{m,n}$.

2. The Case $n = -1$. Let us first consider the case of equiprobable or perfect randomness ($n = -1$). We have that $H'_{-1} = H_{-1} = H^*_{-1}$, and $\psi^2_{m,-1} = \psi^{*2}_{m,-1}$. Thus, Conjectures (a) and (b) must be in agreement when $n = -1$. For $n = -1$, Conjecture (a) states that the asymptotic distribution of $\psi^{*2}_{m,-1}$ is

$$\mathop{*}_{\lambda=1}^{m} K_{g(\lambda)}(x/\lambda) ,$$

while (b) states that the asymptotic distribution of $\psi^2_{m,-1}$ is

$$\mathop{*}_{\lambda=1}^{m-1} K_{g(\lambda)}(x/\lambda) * K_{s-1}(x/m) .$$

Thus, we must define $K_{g(m)}(x/m)$ as $K_{s-1}(x/m)$; i.e., $K_{(s-1)^2 s^{-1}}(x/m)$ as $K_{s-1}(x/m)$. It should also be mentioned that $\psi^2_{m,n}$ and $\psi^{*2}_{m,n}$ are defined only for $m \geqq n + 1$ (with $m \geqq 1$, for $n = -1$), and the symbol $\mathop{*}_{\lambda=1}^{0} K$ is to be understood as the atomic distribution that has the total probability 1 at the value $x = 0$. Since H'_{-1} is a special case of H'_0, results for $n = -1$ will follow directly from results for $n = 0$.

3. The Case $n = 0$. In the present paper, it will be convenient deal to with circular sequences, so that (for $m = 2$) $\sum_j f_{ij} = \sum_j f_{ji} = f_i$. A more general statement (for $m \geqq 2$) can be seen to hold for circular sequences (see [6]). A method of modifying results obtained for circular sequences so that they can be applied to linear sequences has been given in [9], and this method can be used to indicate that results analogous to those presented in the present paper will hold for linear sequences. The reader is cautioned that formulas for circular sequences can not be applied directly to linear sequences (see [9] and Corrigenda to [6]). It will also be convenient herein to replace $\psi^2_{m,n}$ and $\psi^{*2}_{m,n}$ by their asymptotically equivalent forms (when H_n is true)

$$\psi^2_{m,n} \approx 2\sum_{\text{u}} f_{\text{u}} \log (f_{\text{u}} / f_{\text{u},n}),$$

and $\psi^{*2}_{m,n} \approx 2\sum_{\text{u}} f_{\text{u}} \log (f_{\text{u}} / f^*_{\text{u},n})$, respectively.

Let us first consider Conjecture (a) when $n = 0$. For $m = 1$, this conjecture is obviously correct. For $m = 2$, this conjecture was first stated in [8] and was proved by Dawson and Good [4] and by Goodman [10]. The analogous result for the asymptotically equivalent form of $\psi^{*2}_{2,0}$ was proved by Hoel [11].

For $m = 3$, Conjecture (a) states that

$$\psi^{*2}_{3,0} = \sum_{ijk} (f_{ijk} - f_i f_j f_k/N^2)^2/(f_i f_j f_k/N^2)$$
$$\approx \chi^2_{s(s-1)^2} + 2\chi^2_{(s-1)^2},$$

where the symbols χ^2_i denote independent random variables each having a chi-square distribution with i degrees of freedom. (The $f_if_jf_k/N^2$ used above is not the exact expected value, but is an asymptotic approximation; such asymptotic approximations for expected values will be used throughout.) We have that

$$\psi^{*2}_{3,0} \approx 2 \sum_{ijk} f_{ijk} \log [f_{ijk}/(f_i f_j f_k/N^2)]$$
$$= 2 \sum_{ijk} f_{ijk} \log [f_{ijk}/(f_i f_{jk}/N)] + 2 \sum_{jk} f_{jk} \log [f_{jk}/(f_j f_k/N)].$$

The second term in the sum is asymptotically $\chi^2_{(s-1)^2}$, by the result for $m = 2$. The first term in the sum can be split into two parts, thus obtaining

$$2 \sum_{ijk} f_{ijk} \log [f_{ijk}/(f_{ij} f_{jk}/f_j)] + 2 \sum_{ij} f_{ij} \log [f_{ij} /(f_i f_j/N)].$$

By the results in [10], for the test of H_1' within H_2', the asymptotic distribution of the first part is $\chi^2_{s(s-1)^2}$; the asymptotic distribution of the second part is $\chi^2_{(s-1)^2}$ (by the results for $m = 2$). The first part is asymptotically independent of the second part. This can be seen from the fact that their sum has the same asymptotic behavior, under H_0', as the standard likelihood ratio statistic used in testing independence in an $s^2 \times s$ contingency table (see the test of H_0' within H_2' in [10]), and the two parts in the sum are obtained in the same manner as the partitioning of the likelihood ratio for the contingency table into two independent parts (see p. 439 in [3] and the articles referred to therein; rigorous proofs of some of the published results concerning partitioning of contingency tables are given in [12][2]). The first part is obtained by separating the s^2 rows into s sets of s rows, thus obtaining s contingency tables, each $s \times s$, and using the combined likelihood ratio for the s tables to obtain asymptotically a $\chi^2_{s(s-1)^2}$ distribution (which leads to a test of H_1' within H_2' in [10]); the second part is obtained by combining the s rows in each set to obtain an $s \times s$ contingency table, and using the likelihood ratio for this table to obtain asymptotically a $\chi^2_{(s-1)^2}$ distribution (which leads to a test of H_0' within H_1' in [10]). Since the second part of the first term in $\psi^{*2}_{3,0}$ is equal to the second term in $\psi^{*2}_{3,0}$, their sum is asymptotically $2\chi^2_{(s-1)^2}$. Thus we have proved that $\psi^{*2}_{3,0} \approx \chi^2_{s(s-1)^2} + 2\chi^2_{(s-1)^2}$.

For $m = 4$, Conjecture (a) states that

$$\psi^{*2}_{4,0} = \sum_{ijkl} (f_{ijkl} - f_i f_j f_k f_l/N^3)^2/(f_i f_j f_k f_l/N^3)$$
$$\approx \chi^2_{s^2(s-1)^2} + 2\chi^2_{s(s-1)^2} + 3\chi^2_{(s-1)^2}.$$

We have that

$$\psi^{*2}_{4,0} \approx 2\sum_{ijkl} f_{ijkl} \log [f_{ijkl}/(f_i f_j f_k f_l/N^3)]$$
$$= 2\sum_{ijkl} f_{ijkl} \log [f_{ijkl}/(f_i f_{jkl}/N)] + 2\sum_{jkl} f_{jkl} \log [f_{jkl}/(f_j f_k f_l/N^2)].$$

The second term in the sum is $\psi^{*2}_{3,0}$ and is asymptotically $\chi^2_{s(s-1)^2} + 2\chi^2_{(s-1)^2}$, by Conjecture (a) for $m = 3$. The first term can be split into two parts, thus obtaining

$$2\sum_{ijkl} f_{ijkl} \log [f_{ijkl}/(f_{ijk} f_{jkl}/f_{jk})] + 2\sum_{ijk} f_{ijk} \log [f_{ijk}/(f_i f_{jk}/N)].$$

By the results in [10] for the test of H_2' within H_3', the first part is asymptotically $\chi^2_{s^2(s-1)^2}$; the second part is asymptotically $\chi^2_{(s^2-1)(s-1)}$ (by the results for $m = 3$). The two parts are asymptotically independent. This follows from the fact that their sum has the same asymptotic behavior, under H_0', as the standard likelihood ratio statistic used in testing independence in an $s^3 \times s$ contingency table (see the test of H_0' within H_3' in [10]), and the two parts in the sum are obtained in the same manner as the partitioning of the likelihood ratio for the contingency table into two independent parts. The first part is obtained by separating the s^3 rows into s^2 sets of s rows, thus obtaining s^2 contingency tables, each $s \times s$, and using the combined likelihood ratio for the s^2 tables to obtain $\chi^2_{s^2(s-1)^2}$ (which leads to a test of H_2' within H_3' in [10]); the second part is obtained by combining the s rows in each set to obtain an $s^2 \times s$ contingency table, and using the likelihood ratio for this table we get $\chi^2_{(s^2-1)(s-1)}$ (which leads to a test of H_0' within H_2' in [10]). Since the second part of the first term in $\psi^{*2}_{4,0}$ can be written as $\chi^2_{(s^2-1)(s-1)} = \chi^2_{s(s-1)^2} + \chi^2_{(s-1)^2}$ (see the results for $m = 3$), and since the second term in $\psi^{*2}_{4,0}$ is $\psi^{*2}_{3,0} \approx \chi^2_{s(s-1)^2} + 2\chi^2_{(s-1)^2}$ (where the $\chi^2_{s(s-1)^2}$ and the $\chi^2_{(s-1)^2}$,

[2] I am indebted to T. W. Anderson for bringing [12] to my attention.

expressions are identical with those appearing in the second part of the first term), their sum is asymptotically $2\chi^2_{s(s-1)^2} + 3\chi^2_{(s-1)^2}$. We have thus proved that

$$\psi^{*2}_{4,0} \approx \chi^2_{s^2(s-1)^2} + 2\chi^2_{s(s-1)^2} + 3\chi^2_{(s-1)^2}.$$

For $m = 5, 6, \cdots$, the same method of proof applies for Conjecture (a) when $n = 0$; it is easy to see that $\psi^{*2}_{m,0}$ is asymptotically equivalent, under H'_0, to a weighted sum of asymptotically independent likelihood ratio statistics.

Let us now consider Conjecture (b) when $n = 0$. We have

$$\begin{aligned}\psi^2_{m,n} &\approx 2 \sum_{u} f_u \log (f_u/f_{u,n}) \\ &= 2 \{\sum_u f_u \log (f_u/f^*_{u,n}) + \sum_u f_u \log (f^*_{u,n}/f_{u,n})\} \\ &\approx \psi^{*2}_{m,n} + 2 \sum_u f_u \log (f^*_{u,n}/f_{u,n}).\end{aligned}$$

For $m = 1$, $f^*_{u,0} = f_u$, and the second term is $2\sum_i f_i \log (f_i/Np_i)$, which is asymptotically χ^2_{s-1} by the standard statistical theory for goodness of fit tests. For $m = 2$, the second term is

$$\begin{aligned}2 \sum_u f_u \log (f^*_{u,0}/f_{u,0}) &= 2 \sum_{ij} f_{ij} \log [(f_i f_j/N)/Np_i p_j] \\ &= 4 \sum_i f_i \log (f_i/Np_i),\end{aligned}$$

which is asymptotically $2\chi^2_{(s-1)}$. The first term $\psi^{*2}_{2,0}$ is asymptotically independent of the second. This follows from the fact that the sum of $\psi^{*2}_{2,0}$ and $2\sum f_i \log (f_i/Np_i)$ is the likelihood ratio obtained in testing the null hypothesis H_0 that the transition probabilities for the Markov chain are $p_{ij} = p_j = p^0_j$ (specified) within the hypothesis H'_1 (i.e., $2\sum_{ij} f_{ij} \log (f_{ij} / f_i p_j) \approx \chi^2_{(s-1)^2} + \chi^2_{(s-1)} = \chi^2_{s(s-1)}$ (see [1])), and the two terms in the sum are obtained by partitioning the likelihood ratio into two independent parts (the independence of the two parts follows directly from an examination of the asymptotic behavior of the f_{ij} (see, e.g., [9])). The first part is asymptotically $\chi^2_{(s-1)^2}$ and tests the null hypothesis H'_0 that $p_{ij} = p_j$ (unspecified) within H'_1; the second part is asymptotically $\chi^2_{(s-1)}$ and tests the null hypothesis H_0 that $p_j = p^0_j$ (specified) within H'_0. Thus, $\psi^2_{2,0} \approx \chi^2_{(s-1)^2} + 2\chi^2_{(s-1)}$.

For $m = 3$ the second term is

$$\begin{aligned}2 \sum_u f_u \log(f^*_{u,0}/f_{u,0}) &= 2 \sum_{ijk} f_{ijk} \log [(f_i f_j f_k/N^2)/Np_i p_j p_k] \\ &= 6 \sum_i f_i \log (f_i/Np_i),\end{aligned}$$

which is asymptotically $3\chi^2_{(s-1)}$. The first term is independent of the second, by a similar argument to that presented for $m = 2$. Thus,

$$\psi^2_{3,0} \approx \chi^2_{s(s-1)^2} + 2\chi^{2'}_{(s-1)^2} + 3\chi^2_{(s-1)}.$$

For $m = 4, 5, 6, \cdots$, the same method of proof applies for Conjecture (b) when $n = 0$.

We have thus given an altogether different method for proving the results obtained in [2] for $n = 0$; the results in [2] were based on the theory of finite-dimensional vector spaces. Since H'_{-1} is a special case of H'_0, the results given in the present section also prove that Conjectures (a) and (b), when properly interpreted, are true for $n = -1$, which generalizes the result proved in [5] for $n = -1$ and s prime. The different method presented in the present paper may further the understanding of the results in [2] and [5].

4. The Case $n = 1$. Let us now consider Conjecture (a) when $n = 1$. For $m = 2$, the conjecture is obviously true. For $m = 3$, we have that

$$\psi_{3,1}^{*2} \approx 2 \sum_{ijk} f_{ijk} \log [f_{ijk}/(f_{ij} f_{jk}/f_j)] \approx \chi^2_{s(s-1)^2},$$

by the results in [10] for the test of H'_1 within H'_2. For $m = 4$, we have that

$$\begin{aligned} \psi_{4,1}^{*2} &\approx 2 \sum_{ijkl} f_{ijkl} \log \{f_{ijkl}/[f_{ij} f_{jk} f_{kl}/(f_j f_k)]\} \\ &= 2 \sum_{ijkl} f_{ijkl} \log [f_{ijkl}/(f_{ijk} f_{jkl}/f_{jk})] + 2 \sum_{ijk} f_{ijk} \log [f_{ijk}/(f_{ij} f_{jk}/f_j)] \\ &\qquad + 2 \sum_{jkl} f_{jkl} \log [f_{jkl}/(f_{jk} f_{kl}/f_k)]. \end{aligned}$$

By the results in [10] for the test of H'_2 within H'_3, the first term in the sum is asymptotically $\chi^2_{s^2(s-1)^2}$, and the second term is asymptotically $\chi^2_{s(s-1)^2}$ (see $m = 3$). The first term is asymptotically independent of the second. This follows from the fact that their sum can be regarded as the combined likelihood ratio used in testing independence in s contingency tables, each $s^2 \times s$ (see the test of H'_1 within H'_3 in [10]), and the two terms in the sum are obtained by partitioning the likelihood ratio for each of the s tables into two independent parts. For each of the s tables, the first part is obtained by separating the s^2 rows into s sets of s rows, thus obtaining s new tables, each $s \times s$, and using the combined likelihood ratio for the total of s^2 tables to obtain $\chi^2_{s^2(s-1)^2}$ (which is a test of H'_2 within H'_3 in [10]); the second part, for each of the original s tables, is obtained by combining the s rows in each new table to obtain an $s \times s$ table, and using the likelihood ratio for this table (there are s such tables) we get $\chi^2_{s(s-1)^2}$ (which is a test of H'_1 within H'_2 in [10]). The third term in the sum is asymptotically $\chi^2_{s(s-1)^2}$ (see $m = 3$), and it is equal to the second term in the sum. Thus we have $\psi_{4,1}^{*2} \approx \chi^2_{s^2(s-1)^2} + 2\chi^2_{s(s-1)^2}$.

For $m = 5, 6, \cdots$, the same method of proof applies for Conjecture (a) when $n = 1$; $\psi_{m,1}^{*2}$ is asymptotically equivalent to a weighted sum of asymptotically independent likelihood ratio statistics, under H'_1.

Let us now consider Conjecture (b) when $n = 1$. We have that

$$\psi_{m,1}^{2} \approx \psi_{m,1}^{*2} + 2 \sum_{u} f_u \log (f^*_{u,1}/f_{u,1}).$$

For $m = 2$, $f^*_{u,1} = f_u$; thus, the first term $\psi^{*2}_{m,1} = 0$, and the second term is $2\sum_{ij} f_{ij} \log (f_{ij} / Np_i p_{ij})$, where the p_i are the stationary probabilities for the first order Markov chain with constant transition matrix $P = [p_{ij}]$. Conjecture (b) states that $\psi^2_{2,1} \approx \chi^2_{(s-1)^2} + 2\chi^2_{(s-1)}$. We could write

$$\psi^2_{2,1} \approx 2 \sum_{ij} f_{ij} \log [f_{ij}/(f_i f_j/N)] + 2 \sum_{ij} f_{ij} \log [(f_i f_j/N)/(Np_i p_{ij})].$$

The first is not asymptotically $\chi^2_{(s-1)^2}$, except when $n = 0$; and the second term is not asymptotically $2\chi^2_{(s-1)}$, except when $n = 0$. It is easy to see that Conjecture (b) will not hold true for $n = 1$, nor for $n > 1$.

Conjecture (b) will now be modified and this modified version will be proved true. This modification, for the special case $n = 1$, was first mentioned to the author by P. Billingsley in a private communication. In this communication, he mentioned that he had also obtained independently a proof of Conjecture (a), for the case $n = 1$, by very different methods from those used in the present paper, and that his results for Conjecture (a) and the modified Conjecture (b), when $n = 1$, could be extended to the case when $n > 1$, although the detailed asymptotic distributions were not given in the more general case [13].

Let $\psi'^2_{m,1} = \sum_u (f_u - f'_{u,1})^2/f'_{u,1}$, where $f'_{u,1}$ is the expected value of f_u in a new sequence of length N given H_1 and f_{u_1}; i.e., $f'_{u,1} = f_{u_1}\Pi_{i=1}^{m-1} p_{u_i u_{i+1}}$. Then

$$\psi'^2_{m,1} \approx \psi^{*2}_{m,1} + 2 \sum_u f_u \log (f^*_{u,1}/f'_{u,1}).$$

When $m = 2$, the first term $\psi^{*2}_{2,1}$ in the sum is zero and the second term is

$$2 \sum_{ij} f_{ij} \log (f_{ij}/f_i p_{ij}),$$

which is asymptotically $\chi^2_{s(s-1)}$ (see [1]). Thus, the asymptotic distribution of $\psi'^2_{2,1}$ is $\chi^2_{s(s-1)}$.

When $m = 3$, the first term $\psi^*_{3,1}$ is asymptotically $\chi^2_{s(s-1)^2}$, and the second term is

$$2 \sum_{ijk} f_{ijk} \log (f_{ij} f_{jk}/f_j f_i p_{ij} p_{jk}) = 4 \sum_{ij} f_{ij} \log (f_{ij}/f_i p_{ij}),$$

which is asymptotically $2\chi^2_{s(s-1)}$. The first term leads to a test of H'_1 within H'_2, and the second term leads to a test of H_1 within H'_1; it can be seen that the two terms are asymptotically independent under H_1. Thus, for $m = 3$, the asymptotic distribution of $\psi'^2_{m,1}$, when H_1 is true, is

$$\mathop{*}_{\lambda=1}^{m-2} K_{g(\lambda)}(x/\lambda) * K_{s(s-1)}[x/(m - 1)].$$

This result can be proved for $m \geqq 3$ by the same method as given here for $m = 3$. Thus, a modified version of Conjecture (b) holds true for $n = 1$.

5. The Case $n = 2$. Let us now consider Conjecture (a) when $n = 2$. For $m = 3$, the conjecture is obviously true. For $m = 4$, we have

$$\psi^{*2}_{4,2} \approx \sum_{ijkl} f_{ijkl} \log [f_{ijkl}/(f_{ijk} f_{jkl}/f_{jk})] \approx \chi^2_{s^2(s-1)^2},$$

by the results in [10]. For $m = 5$, we have

$$\begin{aligned}\psi_{5,2}^{*2} &\approx 2 \sum_{ijklm} f_{ijklm} \log [f_{ijklm}/(f_{ijk} f_{jkl} f_{klm}/f_{jk} f_{kl})] \\ &= 2 \sum_{ijklm} f_{ijklm} \log [f_{ijklm}/(f_{ijkl} f_{jklm}/f_{jkl})] + 2 \sum_{ijkl} f_{ijkl} \log [f_{ijkl}/(f_{ijk} f_{jkl}/f_{jk})] \\ &\qquad + 2 \sum_{jklm} f_{jklm} \log [f_{jklm}/(f_{jkl} f_{klm}/f_{kl})].\end{aligned}$$

By the results in [10], the first term in the sum is asymptotically $\chi^2_{s^3(s-1)^2}$, and the second term is asymptotically $\chi^2_{s^2(s-1)^2}$ (see $m = 4$). The first term is asymptotically independent of the second; this follows by an argument similar to those appearing earlier here. The third term in the sum is asymptotically

$$\chi^2_{s^2(s-1)^2}$$

(see $m = 4$), and it is equal to the second term in the sum. Thus, we have $\psi_{5,2}^{*2} \approx \chi^2_{s^3(s-1)^2} + 2\chi^2_{s^2(s-1)^2}$.

For $m = 6, 7, \cdots$, the same method of proof applies for Conjecture (a) when $n = 2$. Conjecture (b) will not be true for $n = 2$, as it was not for $n = 1$. A modification of Conjecture (b) for $n = 2$ will now be given, which is similar to, although different from, Billingsley's modification of this conjecture for the special case $n = 1$.

Let $\psi_{m,2}^{\prime 2} = \sum_u (f_u - f'_{u,2})^2/f'_{u,2}$, where $f'_{u,2}$ is the expected value of f_u in a new sequence of length N given H_2 and $f_{u_1u_2}$; i.e., $f'_{u,2} = f_{u_1u_2} \prod_{i=1}^{m-2} p_{u_iu_{i+1}u_{i+2}}$ where $p_{u_1u_2u_3}$ is the second order transition probability that $X_t = u_3$, given that $X_{t-1} = u_2$ and $X_{t-2} = u_1$. Then $\psi_{m,2}^{\prime 2} \approx \psi_{m,2}^{*2} + 2 \sum_u f_u \log (f^*_{u,2}/f'_{u,2})$. When $m = 3$, the first term $\psi_{3,2}^{*2}$ in the sum is zero, and the second term is $2 \sum_{ijk} f_{ijk} \log (f_{ijk}/f_{ij}p_{ijk})$, which is asymptotically $\chi^2_{s^2(s-1)}$ (see [1]). Thus, the asymptotic distribution of $\psi_{3,2}^{\prime 2}$ is $\chi^2_{s^2(s-1)}$.

When $m = 4$, the first term $\psi_{4,2}^{*2}$ is asymptotically $\chi^2_{s^2(s-1)^2}$, and the second term is

$$2 \sum_{ijkl} f_{ijkl} \log (f_{ijk} f_{jkl}/f_{jk} f_{ij} p_{ijk} p_{jkl}) = 4 \sum_{ijk} f_{ijk} \log (f_{ijk}/f_{ij} p_{ijk}),$$

which is asymptotically $2\chi^2_{s^2(s-1)}$. The first term leads to a test of H'_2 within H'_3, and the second term leads to a test of H_2 within H'_2; it can be seen that the two terms are asymptotically independent under H_2. Thus, for $m = 4$, the asymptotic distribution of $\psi_{m,2}^{\prime 2}$, when H_2 is true, is

$$\mathop{*}_{\lambda=1}^{m-3} K_{g(\lambda)}(x/\lambda) * K_{s^2(s-1)}[x/(m-2)].$$

This result can be proved for $m \geqq 4$ by the same method as given here for $m = 4$. Thus, a modified version of Conjecture (b) holds true for $n = 2$.

6. The General Case. The method of proof used in the preceding sections for $n = -1, 0, 1, 2$ can also be applied when $n = 3, 4, \cdots$. In this way, Conjecture (a) can be proved in the general case $n \geqq -1$ and the following modification of

Conjecture (b) also holds in the general case. Let $\psi'^2_{m,n} = \sum_u (f_u - f'_{u,n})^2/f'_{u,n}$, where $f'_{u,n}$ is the expected value of f_u in a new sequence of length N given H_n and $f_{u_1u_2\cdots u_n}$ $(n \geqq 1)$. Then, the asymptotic distribution of $\psi'^2_{m,n}$, when H_n is true, is

$$\mathop{*}_{\lambda=1}^{m-n-1} K_{g(\lambda)}(x/\lambda)*K_{s^n(s-1)}[x/(m-n)].$$

If we define $\psi'^2_{m,0}$ as $\psi^2_{m,0}$, then Conjecture (b) for $n = 0$, is identical with the modified version, and it also holds true. For $n = -1$, H_{-1} is a special case of H'_0, and the modified version of Conjecture (b) can be applied with n taken as zero. The reader will note that the asymptotic distribution of $\psi'^2_{m,n}$ is not mathematically independent of n; neither was the asymptotic distribution of $\psi^{*2}_{m,n}$. The result presented here for $\psi'^2_{m,n}$ generalizes Billingsley's result for $n = 1$.

A direct proof of these results could be given for the general case; this was not done here, since the proof proceeds along the same lines as the earlier discussion herein, and the results may be simpler to understand by considering first $n = 0$, $m = 1, 2, 3, 4, \cdots$; $n = 1$, $m = 2, 3, 4, \cdots$; $n = 2$, $m = 3, 4, \cdots$; etc.

In closing, we mention another conjecture by I. J. Good. In [6], the author conjectures that, when H'_{m-1} is true, the variables $-2 \log \lambda_{m-1,m}$ $(m = 0, 1, 2, \cdots)$ are asymptotically independent, where $\lambda_{m-1,m}$ is the ratio of the maximum likelihood given H'_{m-1} to that given H'_m. If this conjecture were true, than an elegant proof of some results for testing H'_m within H'_n would be available (see [6]). We have that $-2 \log \lambda_{m-1,m} \approx \psi^{*2}_{m+1,m-1}$, when H'_{m-1} is true. The asymptotic independence of the likelihood ratios follows by the same kind of argument presented earlier in the present paper for the independence of some of the statistics considered (see, e.g., the reason why $\psi^{*2}_{4,2}$ and $\psi^{*2}_{3,1}$ are asymptotically independent, given $n = 1$, in the discussion here of the case $m = 4$ and $n = 1$).

The reader is referred to [13] for results that are closely related to some of those presented here, although the general approach and methods are very different. Also, some of the work in [14], [15], and [16] has some (but not much) relation to the present paper.

REFERENCES

[1] M. S. Bartlett, "The frequency goodness of fit test for probability chains," *Biometrika*, Vol. 47 (1951), pp. 86–95.

[2] P. Billingsley, "Asymptotic distributions of two goodness of fit criteria," *Ann. Math. Stat.*, Vol. 27 (1956), pp. 1123–9.

[3] W. G. Cochran, "Some methods for strengthening the common χ^2 tests," *Biometrics*, Vol. 10 (1954), pp. 417–51.

[4] R. Dawson and I. J. Good, "Exact Markov probabilities from oriented linear graphs," *Ann. Math. Stat.*, Vol. 28 (1957), pp. 838–946.

[5] I. J. Good, "The serial test for sampling numbers and other tests for randomness," *Proc. Cambridge Philos. Soc.*, Vol. 49 (1953), pp. 276–84.

[6] I. J. Good, "The likelihood ratio test for Markoff chains," *Biometrika*, Vol. 42 (1955), pp. 531–3; Corrigenda, Vol. 44 (1957), p. 301.

[7] I. J. Good, "Review of P. Billingsley's 'Asymptotic distributions of two goodness of fit criteria'," *Mathematical Reviews*, Vol. 18 (1957), p. 607.

[8] LEO A. GOODMAN, "On the statistical analysis of Markov chains" (abstract), *Ann. Math. Stat.*, Vol. 26 (1955), p. 771.

[9] LEO A. GOODMAN, "Simplified runs tests and likelihood ratio tests for Markoff chains," *Biometrika*, Vol. 45 (1958), pp. 181–97.

[10] LEO A. GOODMAN, "Exact probabilities and asymptotic relationships for some statistics from m-th order Markov chains," *Ann. Math. Stat.*, Vol. 29 (1958), pp. 476–90.

[11] P. G. HOEL, "A test for Markoff chains," *Biometrika*, Vol. 41 (1954), pp. 430–3.

[12] MARILYN SPENCER, "The two-way contingency table and approximate chi-square distributions," Master's Essay, Columbia University Library, 1957.

[13] P. BILLINGSLEY, "On testing Markov Chains," unpublished report.

[14] G. P. BASHARIN, "On using the test of goodness of fit (χ^2 agreement) as a test of independence of trials," *Teoriya Veroyatœnostei i ee Primereniya*, Vol. 2 (1957), pp. 141–2.

[15] N. V. SMIRNOV, "The statistical estimation of transition probabilities in Markov chains," *Vestnik Leningradskoyo Universiteta*, Vol. 1 (1955), pp. 47–8.

[16] V. E. STEPANOW, "Some statistical tests for Markov chains," *Teoriya Veroyatnostei i ee Primeneniya*, Vol. 2 (1957), pp. 143–4.