Small Sample Comparison of Likelihood-ratio and Pearson Chi-square Statistics for the Null Distribution

by Kinley Larntz*

Technical Report No. 218

August 1973

*Kinley Larntz is Assistant Professor, Department of Applied Statistics, University of Minnesota, St. Paul, Minnesota 55101. Research for this paper was supported in part by a Grant-in-Aid from the Graduate School, University of Minnesota. This paper is based on part of the author's doctoral dissertation written under Professor Stephen E. Fienberg at the University of Chicago. Support for this research at the University of Chicago was provided by a National Defense Education Act Fellowship and a Warner-Lambert Research Institute Fellowship, with additional support from the Shell Companies Foundation, Inc. The author is indebted to S.E. Fienberg for guidance and to R.R. Bahadur, S.J. Haberman, and D.L. Wallace for valuable comments.

1. INTRODUCTION

Several statistics are commonly used to judge the goodness-of-fit for counted data models. In this paper, two of these statistics will be compared with respect to their small sample properties under the null hypothesis. The usual chi-square statistic (Pearson statistic) is defined by

$$X^2 = \sum_{\text{all cells}} \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}} \quad . \tag{1.1}$$

A suggested alternative statistic that has some asymptotically optimal properties (see [1] and [13]) is the likelihood-ratio statistic

$$G^2 = 2 \sum_{\text{all cells}} \text{Observed} \log_e(\text{Observed}/\text{Expected}) \quad . \tag{1.2}$$

Many statisticians prefer the use of one or the other of these statistics, although among everyday users the Pearson statistic is far more popular. Also, some statisticians follow the practice of reporting both statistics (see for example [11]), but little guidance is available concerning the occurrance of large discrepancies between the two statistics.

In Section 2 we introduce the model used for comparison. While one particular model was chosen for comparison, it is still believed that the results hold for most other parametric models. Section 3 provides the small sample comparison of the statistics, and in Section 4 the properties of the likelihood-ratio statistic are examined in more detail. Conclusions are given in the final section.

2. THE MODEL

Comparisons between the statistics are made for a particular parametric model that arises naturally in a group helping situation. Individuals or groups are given the opportunity to help another individual in distress. The degree of help is graded I, II, or III: I for not helping, III for actively helping, and II for an intermediate action. Further details can be found in [9] or [17]. Similar models are also used in component testing problems (see [7]).

Data were gathered for individuals and groups of size two. Let p_1, p_2, and p_3 be the probabilities of observing an individual with help graded I, II, and III, respectively. Then if the individuals in a group act independently and if only the higher grade of help is scored, p_1^2, $p_2^2 + 2p_1p_2$, and $p_3^2 + 2p_1p_3 + 2p_2p_3$ are the respective probabilities of observing I, II, and III for groups of size two.

Suppose N_1 individuals and N_2 groups are tested. The results can be summarized in a 3 x 2 contingency table with column totals fixed as in Table 1.

<u>Table 1 goes about here</u>

Under the above assumptions, (n_{11}, n_{21}, n_{31}) follows a multinomial distribution with probability vector (p_1, p_2, p_3), and (n_{12}, n_{22}, n_{32}) follows a multinomial distribution with probability vector (g_1, g_2, g_3) where

$$\begin{aligned} g_1 &= p_1^2 \\ g_2 &= p_2^2 + 2p_1p_2 \\ g_3 &= p_3^2 + 2p_1p_3 + 2p_2p_3 \; . \end{aligned} \tag{2.1}$$

For this case the unique maximum likelihood estimates for (p_1, p_2, p_3) can be written down directly as

$$\hat{p}_1 = (-n_{31} + \sqrt{n_{31}^2 + 4ac})/2a$$
$$\hat{p}_2 = r\hat{p}_1 \qquad (2.2)$$
$$\hat{p}_3 = 1 - (1+r)\hat{p}_1$$

where

$$r = \frac{n_{21} - 2n_{11} - 4n_{12} + \sqrt{(2n_{11}+4n_{12}-n_{21})^2 + 8(n_{21}+n_{22})(n_{11}+2n_{12})}}{2(n_{11}+2n_{12})} , \qquad (2.3)$$

$$a = (1+r)[(n_{11}+2n_{12})(1+r) + (n_{31}+2n_{32}) + 2n_{22}(1+r)/(2+r)] , \qquad (2.4)$$

and

$$c = n_{11} + 2n_{12} + 2n_{22}/(2+r) . \qquad (2.5)$$

(If the i-th row total (R_i in Table 1) is zero, the maximum likelihood estimates are derived conditional on the zero total. The estimates in such a case are just the extension by continuity of the estimates given by (2.2) through (2.5).)

The selection of this model for making comparisons between the likelihood-ratio and Pearson chi-squares provides several advantages:

(a) The model depends on two parameters, p_1 and p_2, and thus the goodness-of-fit test for the null hypothesis involves the estimation of these parameters. Comparisons are therefore made for a composite null hypothesis.

(b) Since the maximum likelihood estimates can be written down in closed form, iteration is not necessary for finding the estimates. This is important when considering the feasibility of doing large amounts of computation.

(c) Examining (2.1), note that the probability of Help Grade I for groups is p_1^2. When p_1 is small, p_1^2 is quite small. Thus the selection of this model allows for comparisons of very skew multinomials, which means comparisons can be made for small as well as moderate minimum cell expectations. Previous studies ([5],[18],[19]) have indicated that, for small expected values, the Pearson statistic does not follow the chi-square distribution well, while some suggestion has been indicated (cf. [3], p.38) that the likelihood-ratio statistic would be better in such situations.

In the next two sections, the properties of the likelihood-ratio and Pearson chi-square statistics for this model will be presented. Although the results apply directly only to this model, it is believed that similar results hold for other parametric models.

3. SMALL SAMPLE PROPERTIES UNDER THE NULL HYPOTHESIS

Under the null hypothesis the goodness-of-fit statistics, X^2 and G^2, have asymptotic chi-square distributions with 2 degrees of freedom. However, for small samples the chi-square approximation in many cases does not agree well with the actual distribution. Several studies ([5],[10],[16],[19]) have given conflicting points of view as to at what point the approximation is "reasonable" for the Pearson chi-square statistic. Standard rules specify that the minimum cell expectation should be 5, with possibly a few smaller. The emphasis here will not be on finding such a rule, but in comparing the likelihood-ratio and Pearson statistics with regard to the approximation. In other words we ask, for small samples, which of the two statistics is better approximated by the asymptotic chi-square distribution?

The model presented in Section 2 assumes that the data consist of two independent trinomials: the individuals have probability vector (p_1,p_2,p_3) and the pairs have probability vector (g_1,g_2,g_3), where the g's are given by (2.1). The maximum likelihood estimates for sample sizes N_I and N_G are given by equations (2.2) through (2.5).

The initial task in this study of the small sample properties is to determine the distribution of the statistics G^2 and X^2 when the null hypothesis holds. Several methods are available to handle such a problem. One method is to determine the distribution of the maximum likelihood estimator $(\hat{p}_1,\hat{p}_2,\hat{p}_3)$ and then find the distributions of the statistics. But the exact distribution of $(\hat{p}_1,\hat{p}_2,\hat{p}_3)$ is not easily derived by theoretical methods. A second method is to use Monte Carlo techniques [12]. This would not derive the exact distribution, but rather an approximation to it. (In fact, Monte Carlo was used selectively for large

N_I and N_G.) The principal method used was that of enumeration. The number of possible outcomes of two trinomials with sample sizes N_I and N_G is given (see [8]) by

$$\text{Outcomes} = \binom{N_I + 2}{2}\binom{N_G + 2}{2} \tag{3.1}$$

For $N_I = N_G = 8$, the number of possible outcomes is 2025. Thus, for a given value of (p_1, p_2, p_3), N_I, and N_G, the distributions of G^2 and X^2 were determined by computer.

One question that arises in the use of this method is how to deal with zero cell counts and zero expected values. As indicated previously, the maximum likelihood estimates were extended by continuity to provided well-defined procedures. In the same manner, when a cell had a zero expected value, it contributed zero to the chi-square statistic.

<u>Figures A and B about here</u>

Figure A gives a contour plot of the mean of G^2 for $N_I = N_G = 8$. Barycentric coordinates were chosen to represent the 3 probabilities (see [14]). Each corner of the triangle represents one of the probability vectors (1, 0, 0), (0, 1, 0), and (0, 0, 1), while a general point in the triangle corresponds to the probability vector (p_1, p_2, p_3). Figure B gives a similar plot for X^2. The asymptotic mean for both statistics is, of course, 2.0. The mean of G^2 overshoots that value for a large set of (p_1, p_2, p_3). The peak value is approximately 2.51. In viewing

Figure B, it can be seen that the mean of the Pearson statistic is a smoother function of (p_1, p_2, p_3) than the mean of G^2. For a large set of (p_1, p_2, p_3), the mean of X^2 is close to 2.0. The peak value is approximately 2.12. Thus, considering the mean only, for $N_I = N_G = 8$ the small sample distribution of the Pearson statistic is better approximated by the asymptotic theory. Similar results hold for $N_I = N_G = 4$ and $N_I = N_G = 16$.

<u>Figures C and D about here</u>

Another method of comparison is to check the agreement of the actual small sample percentage points with the corresponding large sample values. For $N_I = N_G = 8$, Figures C and D present for the likelihood ratio and Pearson statistics, respectively,

$$\alpha_{.05}(p_1, p_2, p_3, N_I, N_G) = P[\text{Statistic} > 5.991], \qquad (3.2)$$

the true probability of rejecting the null hypothesis when the nominal .05 level test is used. Table 2 gives the corresponding probabilities for $N_I = N_G = 4$ and $N_I = N_G = 16$. Results for the .10 and .01 levels give the same general impression.

<u>Table 2 about here</u>

Figure C shows that for small sample sizes the likelihood ratio statistic rejects the null hypothesis more often than the nominal .05 level. For $(p_1, p_2, p_3) = (.4, .3, .3)$, $P(G^2 > 5.991) = .080$ for $N_I = N_G = 4$.

At $N_I = N_G = 8$ with $(p_1, p_2, p_3) = (.5, .2, .3)$, $P(G^2 > 5.991) = .091$, and the maximum probability for $N_I = N_G = 16$ is .073. The corresponding maximum probabilities for the Pearson statistic are .033, .048, and .049. Thus it appears that the Pearson chi-square is better than the likelihood-ratio in terms of not rejecting the null hypothesis more frequently than the nominal value. For $N_I = N_G = 16$, the Pearson is very close to the asymptotic value of 0.05 for a wide range of (p_1, p_2, p_3).

Several questions concerning the likelihood-ratio statistic arise from these results. First, is it still possible that the optimality properties of the likelihood-ratio statistic carry over in spite of the poor characteristics of its null distribution? This will be the subject of another paper comparing the powers of the statistics. Second, can the statistic be easily adjusted to remove some of its poor behavior? And third, exactly how does the likelihood-ratio behave as the "small" sample size increases? An attempt at answering the last question will be given in the next section.

The question of adjusting the likelihood-ratio statistic poses large difficulties. A simple-minded correction for the mean yielded mixed results, partly due to a problem of overcorrection with respect to size. Other corrections involving more moments or quantiles may be possible, but practical use would require a simple multiplicative or additive correction, such as those given in [2] and [4].

4. PROPERTIES OF THE LIKELIHOOD-RATIO CHI-SQUARE STATISTIC.

The asymptotic distribution of G^2 for the model considered here is that of a chi-square variate with 2 degrees of freedom. By examining Figure C and Table 2, it can be seen that for those values of (p_1, p_2, p_3) where the true sizes are very high (> .08) at $N_I = N_G = 8$, the true sizes are near .07 for $N_I = N_G = 16$. Also, although not given above, the mean values of those points with mean above 2.50 in Figure A fell into the range 2.20 - 2.25 for samples of size 16. In these cases, the small sample distribution of G^2 got closer (judged by these criteria) to chi-square as the sample size increased; however, the opposite can also occur. For example, the point (.8, .1, .1) had true size .048 for $N_I = N_G = 8$ and true size .072 for $N_I = N_G = 16$, and the point (.1, .5, .4) had respective true sizes .041 and .038. The larger sample size did not improve the chi-square fit of G^2 at these points.

<u>Table 3 goes about here</u>

For selected values of (p_1, p_2, p_3), Table 3 gives the means and true sizes of the nominal .05 tests for sample sizes $N_I = N_G = 1\ (1)\ 16$. Figure E gives a graph of the mean values of G^2 for (.6, .2, .2). Figure E is indicative of what happens to the mean as the sample size changes. It begins below its asymptotic value, rises to a peak, and descends to the correct value. The true sizes follow a similar pattern. Because of the discreteness of the distribution, the rise and descent may be slightly rocky, but the general pattern ramins the same. Graphs of the mean of G^2 for changing sizes for three other cases are given in Figures F, G, and H. The irregular behavior in Figures F and H is discussed below.

Figures F, F, G, and H go about here

The sample size at which the peak is reached varies considerably depending on the probability vector (p_1, p_2, p_3). Some evidence has been given that the minimum cell expectation governs the closeness of the small sample distribution to asymptotic theory for several chi-square problems (see for example [5], [6], [15], [19]). In the problem at hand, small expected values are found for small values of p_1 (since the first cell for pairs has probability p_1^2) and for very small values of p_2 and p_3. For the points in Table 3, the minimum cell probabilities (minimum cell probability is minimum cell expectation divided by sample size) are given in Table 4, which also summarizes how the mean and true size change as $N_I = N_G$ goes from 1 to 16. Certainly the larger minimum cell expectation cases are closer to the behavior predicted by the asymptotic theory.

Table 4 goes about here

To determine if the irregularities mentioned above are due to the small cell expectations, two cases were selected for further study. The case with $(p_1,p_2,p_3) = (.1,.5,.4)$ was selected because it appeared to be the worst one of the ten studied. $(p_1,p_2,p_3) = (.2,.5,.3)$ was also selected since, although not as bad as (.1,.5,.4), no tailing off of the mean was observed.

For large N_I (or N_G), exact computations become impossible (or at best too expensive) to carry out on the computer. Therefore, the cases were simulated for a large number of values of $N_I = N_G$. Table 5 summarizes the results, giving estimated means and levels of significance (asymptotic value of .05) for 5000 trials of each case.

Table 5 goes about here

Figures I and J go about here

Figures I and J graph the estimated expected values of the likelihood ratio statistic for the two special cases considered here. Note that both mean functions appear to rise to a peak and then tail off just as the exact small sample results showed for more regular cases. The main difference is that the peaks are reached for larger sample sizes. For the case (.2,.5,.3) with minimum cell probability .04, the peak is probably about 40; while for (.1,.5,.4) with minimum cell probability .01, the peak appears to be near to 150. Thus, the general pattern of a rise and then a decline in the expected value as a function of sample size is seen to hold even for cases with small cell expectations. A similar pattern is also evident for the level of significance, although the expected value provides a slightly smoother function.

5. CONCLUSIONS

For one special model with a composite null hypothesis, the small sample distributions of two chi-square statistics were examined. Using as criterion the closeness of small sample distribution to the asymptotic chi-square approximation, the Pearson chi-square statistic is by far the more desirable. The likelihood-ratio statistic has an expected value in excess of the nominal and yields far too many rejections under the null distribution.

It was also noted that the expected value and level of significance for the likelihood-ratio statistic displayed a consistent regularity in which the mean and level rose to a peak and then declined toward the asymptotic value as the sample size increased. This property was shown in the exact computations for relatively large minimum cell expectations and was demonstrated in Monte Carlo simulations for two cases with small minimum cell expectations.

Further investigations are under way comparing the power of these statistics in small sample cases.

REFERENCES

[1] Bahadur, R. R., "An Optimal Property of the Likelihood Ratio Statistic," Proceedings of Fifth Berkeley Symposium on Mathematical Statistical Problems, Berkeley: University of California Press, 1967, 13 - 26.

[2] Bartlett, M. S., "Multivariate Analysis," Journal of the Royal Statistical Society, Series B, 9 (1947), 176 - 197.

[3] Bliss, C. I., Statistics in Biology, Volume 1, New York: McGraw-Hill Book Co., 1967.

[4] Box, G. E. P., "A General Distribution Theory for a Class of Likelihood Criteria," Biometrika, 36 (1949), 317 - 346.

[5] Cochran, W. G., "The χ^2 Test of Goodness of Fit," Annals of Mathematical Statistics, 23 (1952), 315 - 346.

[6] Cramér, H., Mathematical Methods of Statistics, Princeton, N. J.: Princeton University Press, 1945.

[7] Easterling, R. G. and Prairie, R. R., "Combining Component and System Information," Technometrics, 13 (1971), 271 - 280.

[8] Feller, W., An Introduction to Probability Theory and Its Applications, Volume 1, 2nd ed., New York: John Wiley & Sons, Inc., 1957.

[9] Fienberg, S. E. and Larntz, K., "Some Models for Individual-Group Comparisons and Group Behavior," Psychometrika, 36 (1971), 349 - 367.

[10] Fisher, R. A., Statistical Methods for Research Workers, 13th ed., New York: Hafner Publishing Company, 1958.

[11] Goodman, L. A., "Guided and Unguided Methods for Selecting Models for a Set of T Multidimensional Contingency Tables," Journal of the American Statistical Association, 68 (1973), 165 - 175.

[12] Hammersley, J. M. and Handscomb, D. C., Monte Carlo Methods, London: Methuen & Co., 1964.

[13] Hoeffding, W., "Asymptotically Optimal Tests for the Multinomial Distribution," Annals of Mathematical Statistics, 36 (1965), 369 - 401.

[14] Mosteller, F. and Tukey, J. W., "Data Analysis, Including Statistics," in G. Lindsey and E. Aronson, ed., Handbook of Social Psychology, Volume 2, Reading, Mass.: Addison-Wesley, 1968.

[15] Odoroff, C. L., "A Comparison of Minimum Logit Chi-Square Estimation and Maximum Likelihood Estimation in 2 x 2 x 2 and 3 x 2 x 2 Contingency Tables: Tests for Interactions," Journal of the American Statistical Association, 65 (1970), 1617 - 1631.

[16] Roscoe, J. T. and Byars, J. A., "Sample Size Restraints Commonly Imposed on the Use of the Chi-Square Statistic," Journal of the American Statistical Association, 66 (1971), 755 - 759.

[17] Staub, E., "A Child in Distress: The Influence of Age and Number of Witnesses on Children's Attempts to Help," Journal of Personality and Social Psychology, 14 (1970), 130 - 140.

[18] Tate, M. W. and Hyer, L. A., Significance Values for an Exact Multinomial Test and Accuracy of the Chi-Square Approximation, Final Project Report No. 8-B-023, Office of Education, Bureau of Research, U. S. Department of Health, Education, and Welfare, 1969.

[19] Yarnold, J. K., "The Minimum Expectation in X^2 Goodness of Fit Tests and the Accuracy of Approximations for the Null Distribution," Journal of the American Statistical Association, 65 (1970), 864 - 886.

1. OBSERVED TABLE FOR N_1 INDIVIDUALS AND N_2 GROUPS.

Help Grade	Individuals	Groups	Total
I	n_{11}	n_{12}	R_1
II	n_{21}	n_{22}	R_2
III	n_{31}	n_{32}	R_3
Total	N_1	N_2	

2. TRUE PROBABILITIES OF REJECTION USING THE NOMINAL .05 TEST FOR SAMPLE SIZES 4 AND 16.

	$N_I = N_G = 4$		$N_I = N_G = 16$	
(p_1, p_2, p_3)	G^2	X^2	G^2	X^2
.1 .1 .8	.005	.001	.030	.045
.1 .2 .7	.017	.002	.039	.045
.1 .3 .6	.033	.005	.043	.048
.1 .4 .5	.044	.010	.041	.049
.1 .5 .4	.047	.014	.038	.049
.1 .6 .3	.039	.015	.039	.048
.1 .7 .2	.026	.012	.043	.045
.1 .8 .1	.010	.005	.040	.043
.2 .1 .7	.018	.003	.044	.034
.2 .2 .6	.036	.008	.057	.044
.2 .3 .5	.053	.015	.055	.047
.2 .4 .4	.060	.021	.052	.047
.2 .5 .3	.057	.025	.052	.046
.2 .6 .2	.044	.022	.056	.045
.2 .7 .1	.023	.012	.057	.042
.3 .1 .6	.037	.008	.061	.034
.3 .2 .5	.056	.016	.070	.044
.3 .3 .4	.070	.025	.066	.046
.3 .4 .3	.073	.031	.064	.046
.3 .5 .2	.062	.031	.067	.045
.3 .6 .1	.040	.020	.070	.041
.4 .1 .5	.053	.013	.069	.038
.4 .2 .4	.071	.024	.073	.048
.4 .3 .3	.080	.032	.068	.049
.4 .4 .2	.076	.035	.071	.048
.4 .5 .1	.055	.027	.072	.045
.5 .1 .4	.060	.017	.069	.040
.5 .2 .3	.075	.026	.070	.049
.5 .3 .2	.078	.033	.068	.049
.5 .4 .1	.063	.029	.068	.046
.6 .1 .3	.055	.017	.069	.041
.6 .2 .2	.066	.024	.071	.048
.6 .3 .1	.059	.025	.066	.046
.7 .1 .2	.041	.014	.073	.041
.7 .2 .1	.043	.017	.070	.044
.8 .1 .1	.020	.008	.072	.038

NOTE: The values in this table should be compared to the asymptotic value of 0.05.

3. TRUE SIZE AND MEAN OF SELECTED CASES AS SAMPLE SIZE CHANGES.

	(.5,.1,.4)		(.6,.2,.2)		(.8,.1,.1)		(.1,.4,.5)		(.5,.4,.1)		(.2,.6,.2)		(.2,.5,.3)		(.1,.5,.4)		(.5,.25,.25)		(1/3,1/3,1/3)	
$N_I = N_G$	Mean	Size	Mean	Size	Mean	Size	Mean	Size	Mean	Size	Mean	Size	Mean	Size	Mean	Size	Mean	Size	Mean	Size
1	1.48	.000	1.63	.000	1.14	.000	1.25	.000	1.55	.000	1.36	.000	1.49	.000	1.37	.000	1.70	.000	1.63	.000
2	1.98	.030	2.29	.042	1.52	.021	1.54	.018	2.10	.037	1.75	.028	1.94	.033	1.68	.024	2.41	.045	2.25	.040
3	2.17	.037	2.55	.046	1.79	.013	1.62	.023	2.31	.046	1.91	.026	2.06	.032	1.71	.024	2.66	.053	2.47	.046
4	2.26	.060	2.63	.066	2.00	.020	1.64	.044	2.39	.063	2.00	.044	2.09	.057	1.68	.047	2.71	.079	2.54	.076
5	2.31	.058	2.63	.083	2.16	.025	1.64	.032	2.42	.066	2.06	.040	2.09	.050	1.64	.032	2.69	.099	2.53	.086
6	2.34	.070	2.60	.087	2.27	.037	1.64	.050	2.43	.073	2.10	.047	2.09	.062	1.62	.051	2.63	.096	2.50	.091
7	2.36	.060	2.55	.090	2.35	.051	1.64	.038	2.43	.073	2.13	.057	2.10	.066	1.61	.045	2.56	.094	2.47	.088
8	2.37	.067	2.50	.085	2.41	.048	1.65	.041	2.42	.068	2.15	.054	2.11	.059	1.61	.041	2.49	.092	2.43	.080
9	2.37	.065	2.45	.082	2.44	.055	1.65	.034	2.40	.064	2.16	.053	2.12	.054	1.61	.036	2.43	.084	2.40	.072
10	2.37	.067	2.41	.082	2.46	.061	1.66	.036	2.39	.067	2.17	.057	2.13	.055	1.62	.035	2.38	.082	2.37	.073
11	2.37	.067	2.37	.082	2.46	.065	1.67	.044	2.37	.066	2.18	.057	2.14	.053	1.63	.040	2.34	.083	2.35	.073
12	2.36	.066	2.33	.079	2.46	.069	1.68	.038	2.35	.066	2.19	.062	2.15	.054	1.65	.036	2.30	.078	2.32	.073
13	2.35	.067	2.30	.079	2.45	.074	1.69	.041	2.34	.067	2.20	.059	2.16	.052	1.66	.038	2.27	.074	2.32	.066
14	2.34	.069	2.27	.075	2.44	.077	1.70	.037	2.32	.068	2.20	.058	2.16	.051	1.68	.034	2.24	.071	2.28	.068
15	2.33	.069	2.25	.073	2.43	.070	1.72	.042	2.30	.068	2.21	.056	2.17	.052	1.69	.040	2.22	.070	2.27	.070
16	2.32	.069	2.22	.071	2.41	.072	1.73	.041	2.29	.068	2.21	.056	2.18	.052	1.71	.038	2.20	.067	2.25	.067

4. MINIMUM CELL PROBABILITIES FOR SELECTED CASES.

(p_1, p_2, p_3)	M.C.P.	Change for $N_I = N_G = 1$ to 16
(.5,.1,.4)	.10	Rise, Peak, Decline
(.6,.2,.2)	.20	Rise, Peak, Decline
(.8,.1,.1)	.10	Rise, Peak, Decline
(.1,.4,.5)	.01	Irregular Rise
(.5,.4,.1)	.10	Rise, Peak, Decline
(.2,.6,.2)	.04	Rise
(.2,.5,.3)	.04	Rise
(.1,.5,.4)	.01	Irregular Rise
(.5,.25,.25)	.25	Rise, Peak, Decline
(1/3,1/3,1/3)	.111	Rise, Peak, Decline

M.C.P. = Minimum Cell Probability

Minimum Cell Expectation = N_I (or N_G) x M.C.P.

5. MONTE CARLO MEANS AND SIZES FOR TWO SELECTED CASES AS SAMPLE SIZE INCREASES.

	(.2,.5,.3)		(.1,.5,.4)			(.2,.5,.3)		(.1,.5,.4)	
$N_I = N_G$	Mean	Level	Mean	Level	$N_I = N_G$	Mean	Level	Mean	Level
20	2.18	.046	1.81	.041	80	2.09	.060	2.08	.041
24	2.21	.055	1.91	.046	84	2.04	.056	--	--
28	2.19	.054	--	--	88	2.09	.058	--	--
30	--	--	1.99	.048	100	2.04	.053	2.17	.046
32	2.17	.055	--	--	110	2.10	.057	--	--
36	2.21	.059	--	--	120	2.05	.052	2.16	.049
40	2.24	.063	1.97	.041	130	2.01	.053	--	--
44	2.14	.056	--	--	140	2.05	.055	2.20	.054
48	2.22	.063	--	--	150	2.09	.057	--	--
50	--	--	2.04	.038	160	--	--	2.21	.064
52	2.12	.058	--	--	200	--	--	2.11	.056
56	2.14	.059	--	--	240	--	--	2.12	.060
60	2.16	.060	2.12	.044	280	--	--	2.18	.068
64	2.15	.064	--	--	320	--	--	2.08	.061
68	2.13	.064	--	--	360	--	--	2.08	.060
70	--	--	2.15	.043	400	--	--	2.11	.061
72	2.13	.064	--	--	450	--	--	2.03	.058
76	2.13	.060	--	--	500	--	--	2.07	.058

NOTE: All numbers are based on 5000 simulated trials. Approximate standard deviations are .028 and .0031 for the means and levels given above.

FIGURES

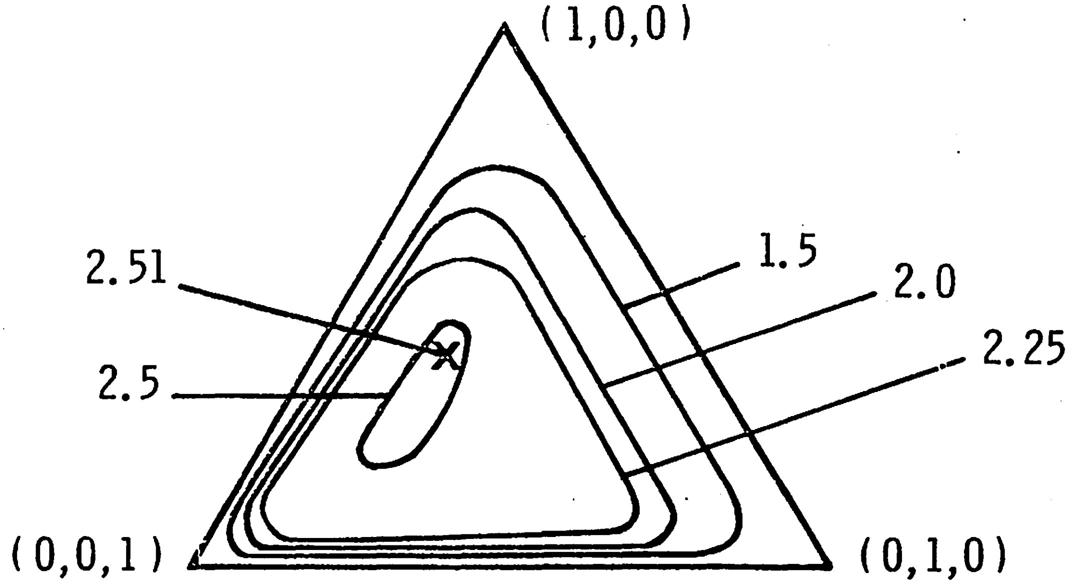

Figure A. Mean of Likelihood Ratio Chi-Square Statistic for $N_I = N_G = 8$.

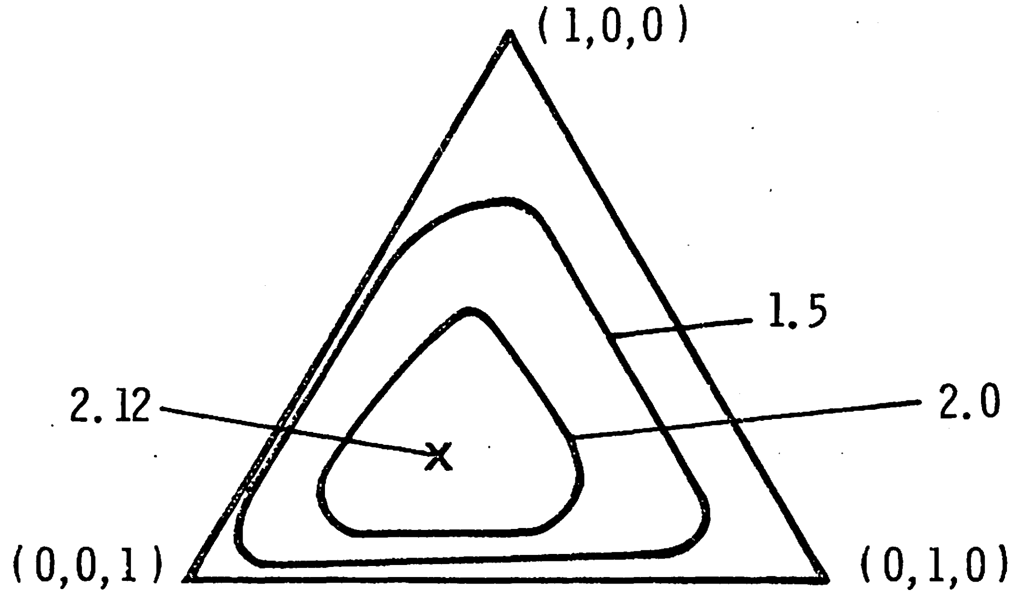

Figure B. Mean of Pearson Chi-Square Statistic for $N_I = N_G = 8$.

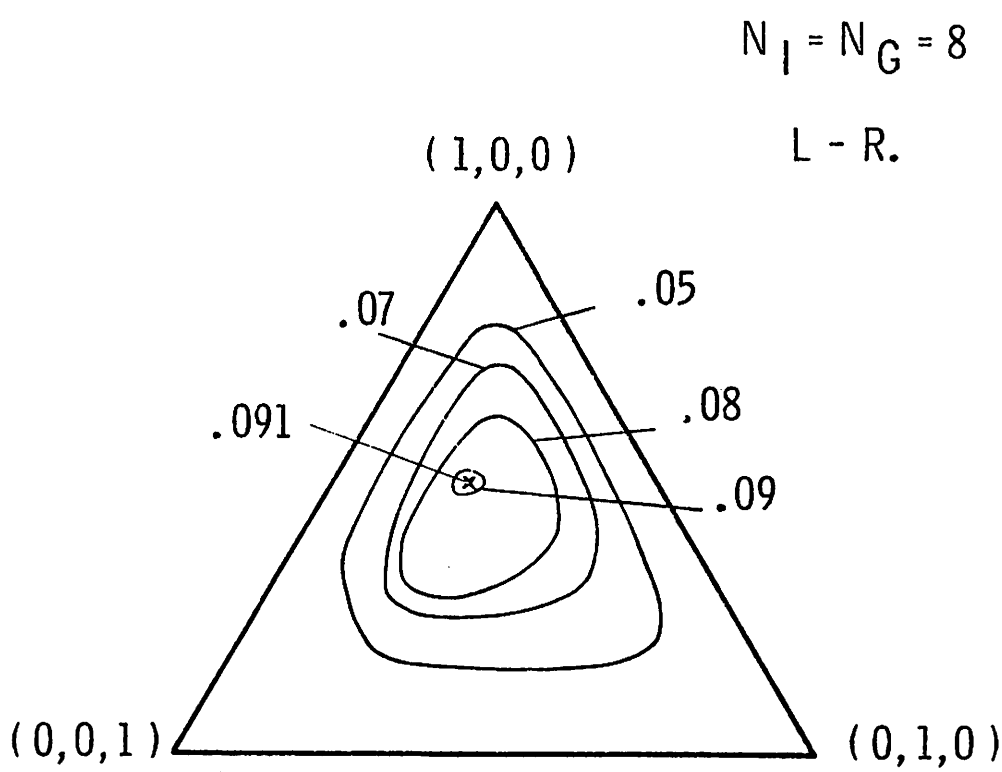

Figure C. Level of Significance of Likelihood Ratio Chi-Square .05 Test for $N_I = N_G = 8$.

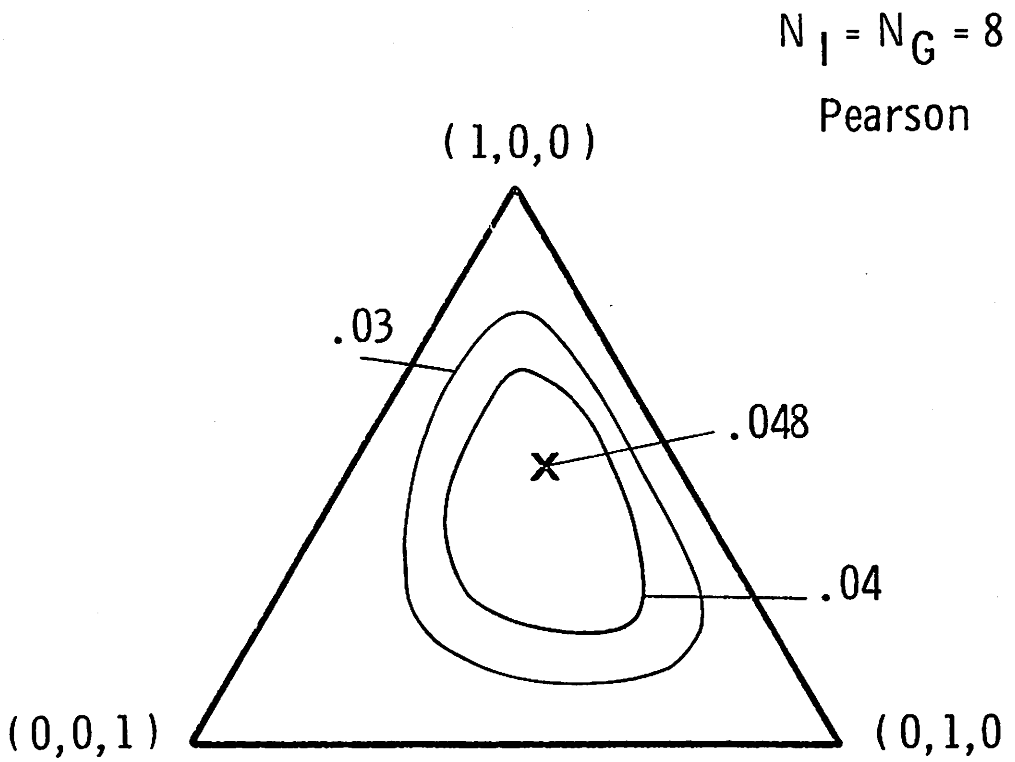

Figure D. Level of Significance of Pearson Chi-Square .05 Test for $N_I = N_G = 8$.

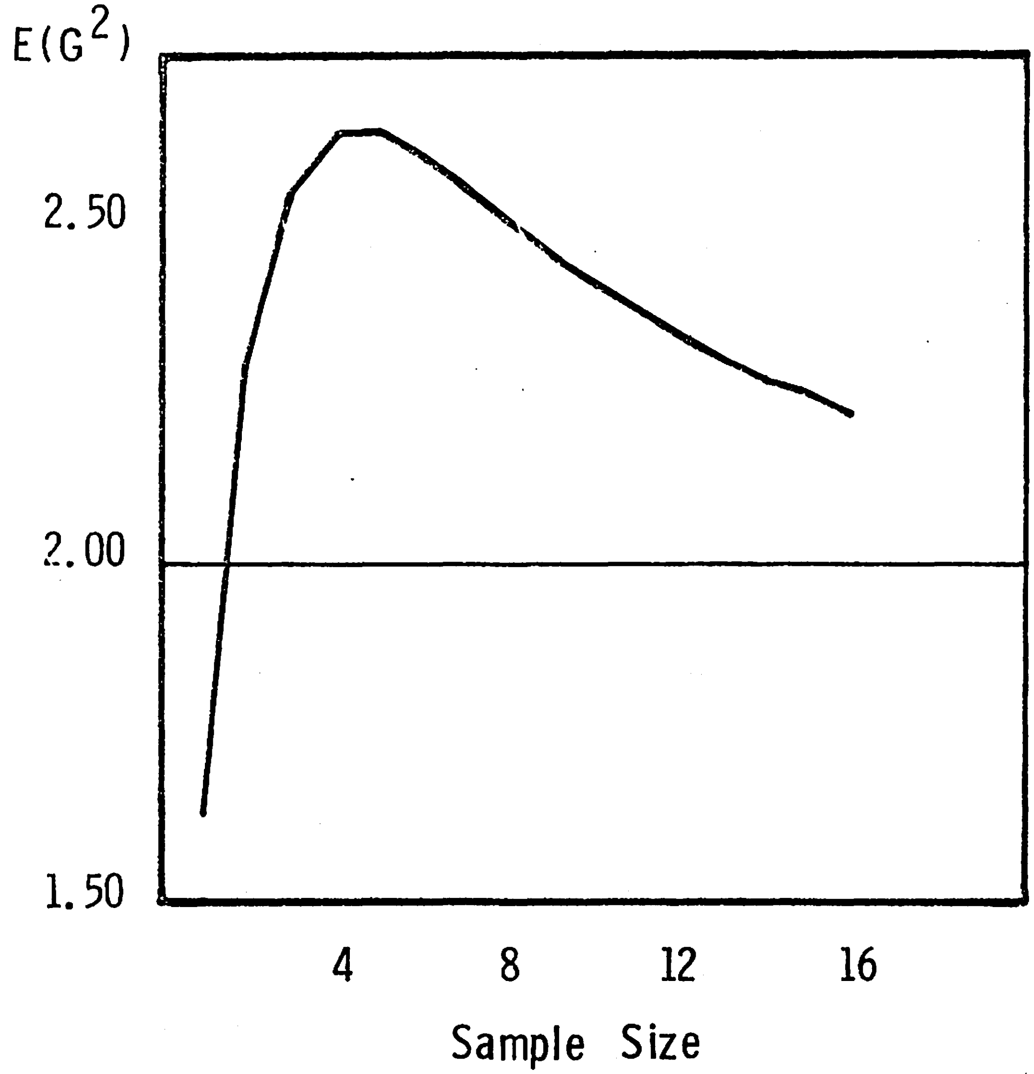

Figure E. Mean of G^2 for (.6,.2,.2) as Sample Size Changes.

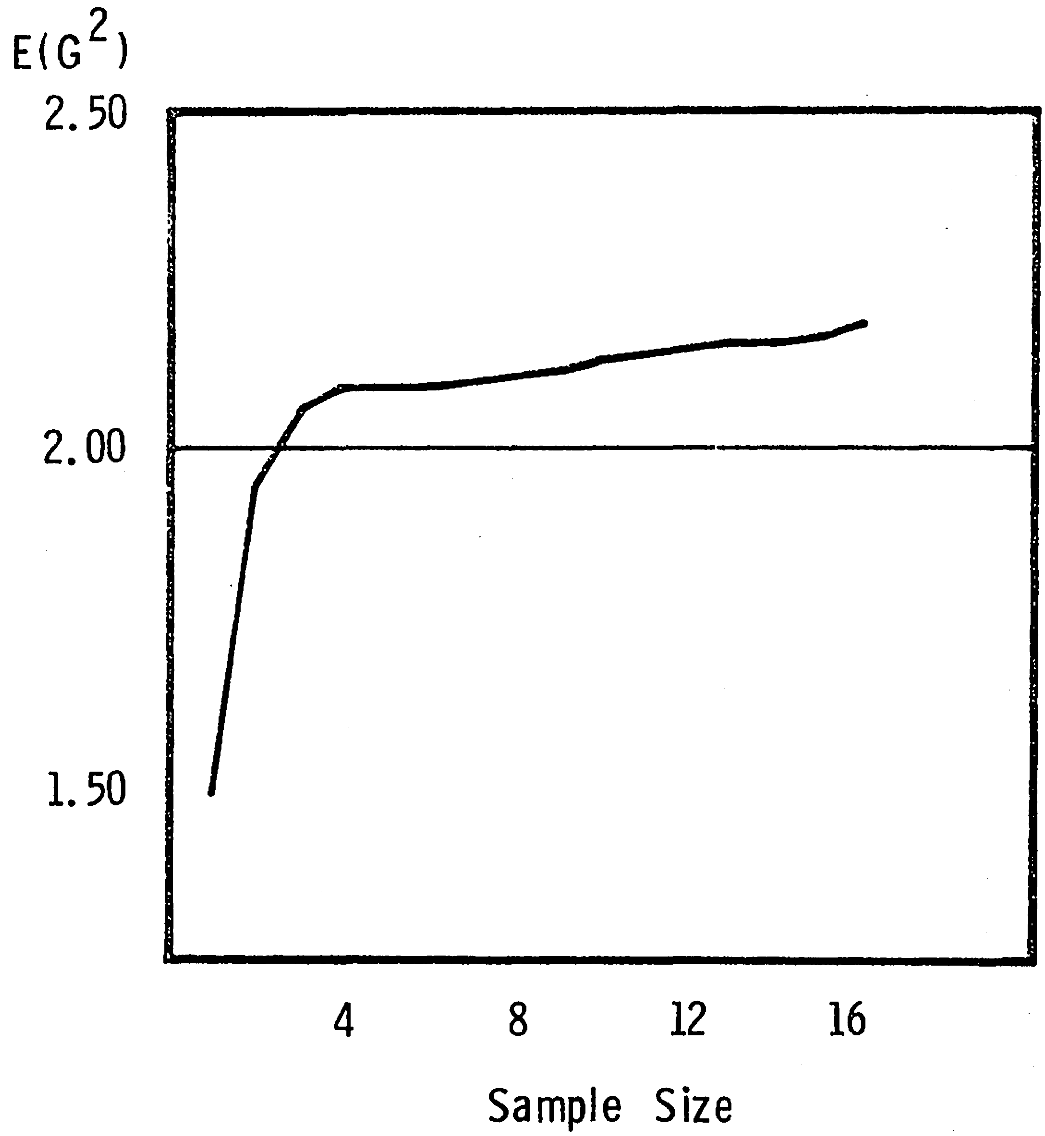

Figure F. Mean of G^2 for (.2,.5,.3) as Sample Size Changes.

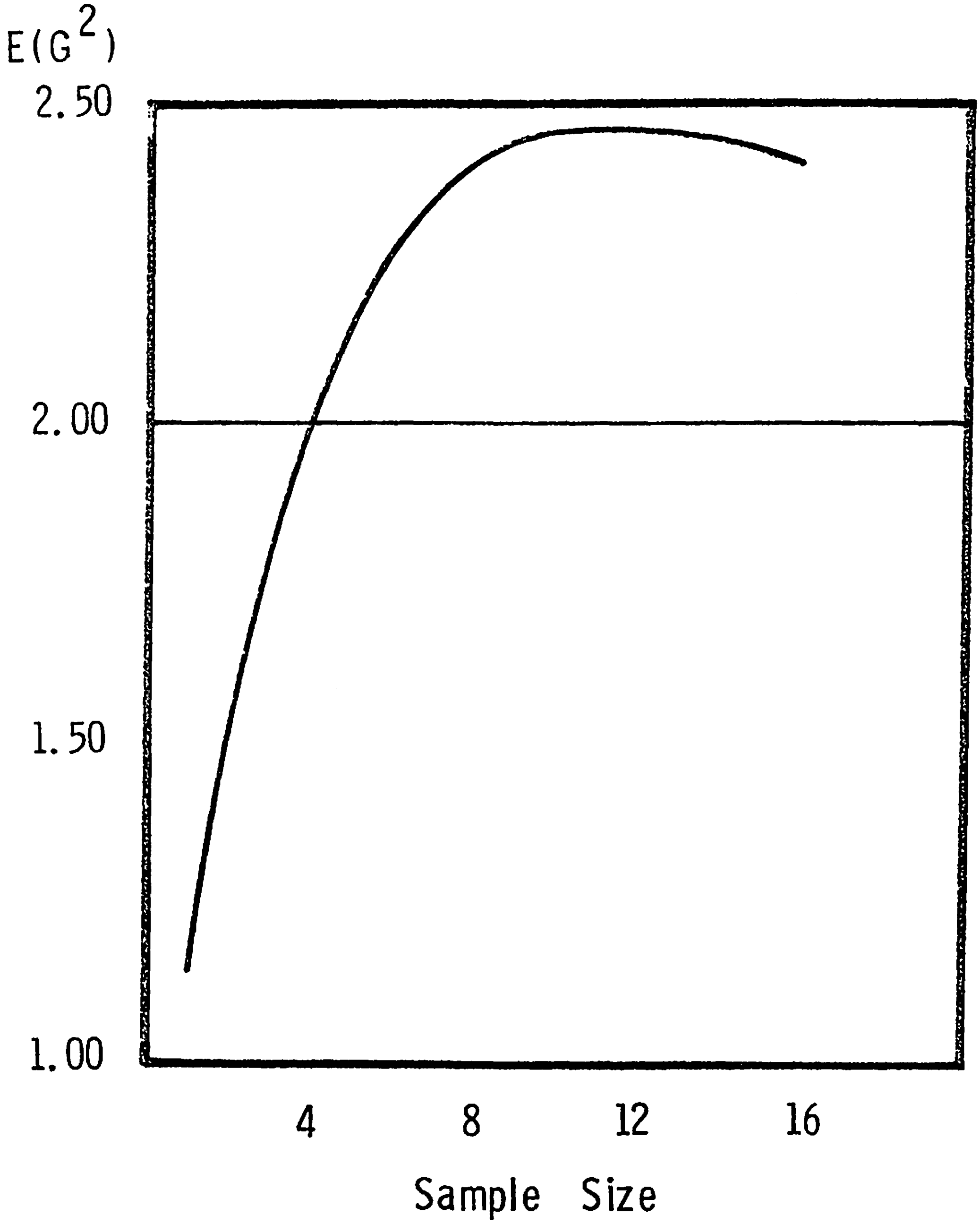

Figure G. Mean of G^2 for (.8,.1,.1) as Sample Size Changes.

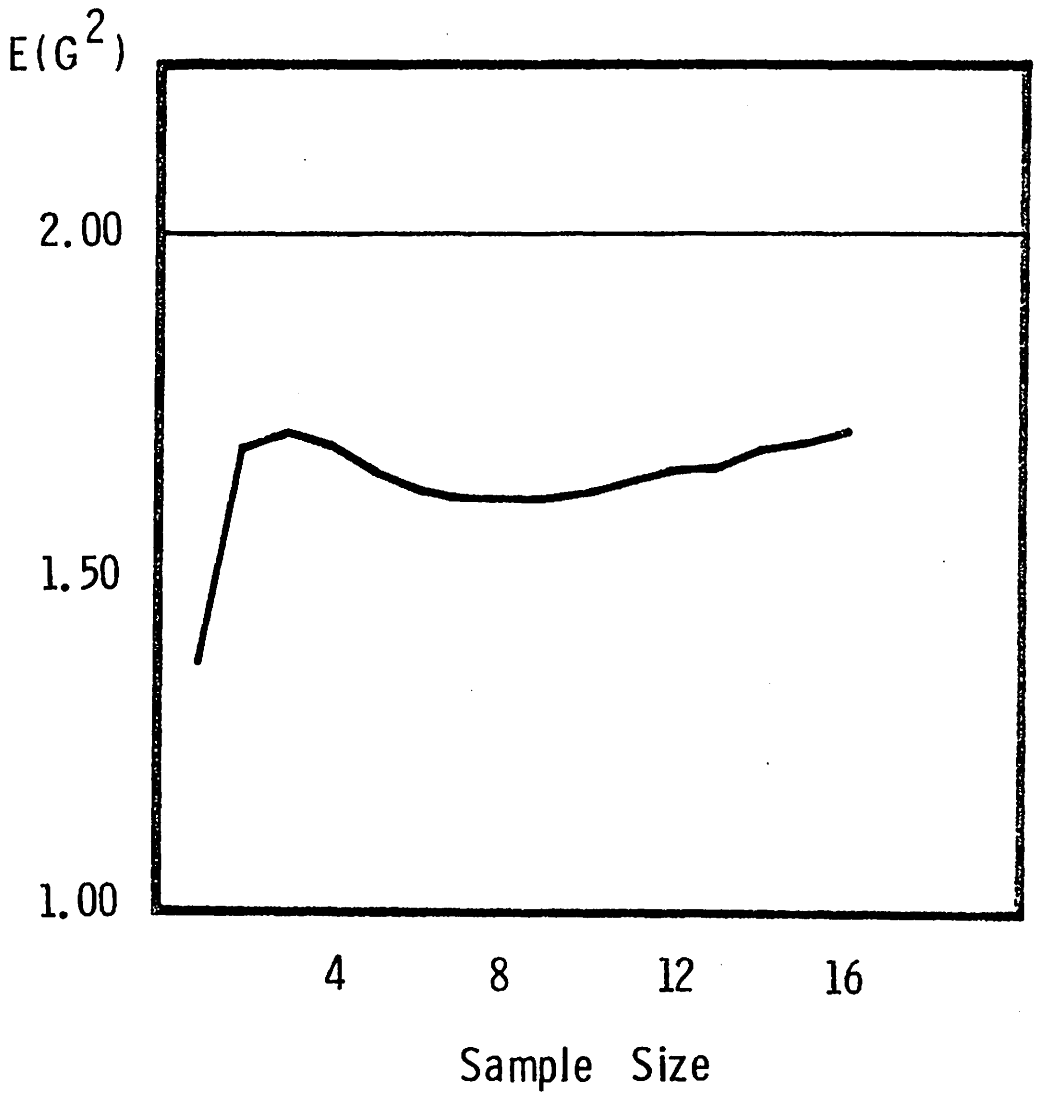

Figure H. Mean of G^2 for (.1,.5,.4) as Sample Size Changes.

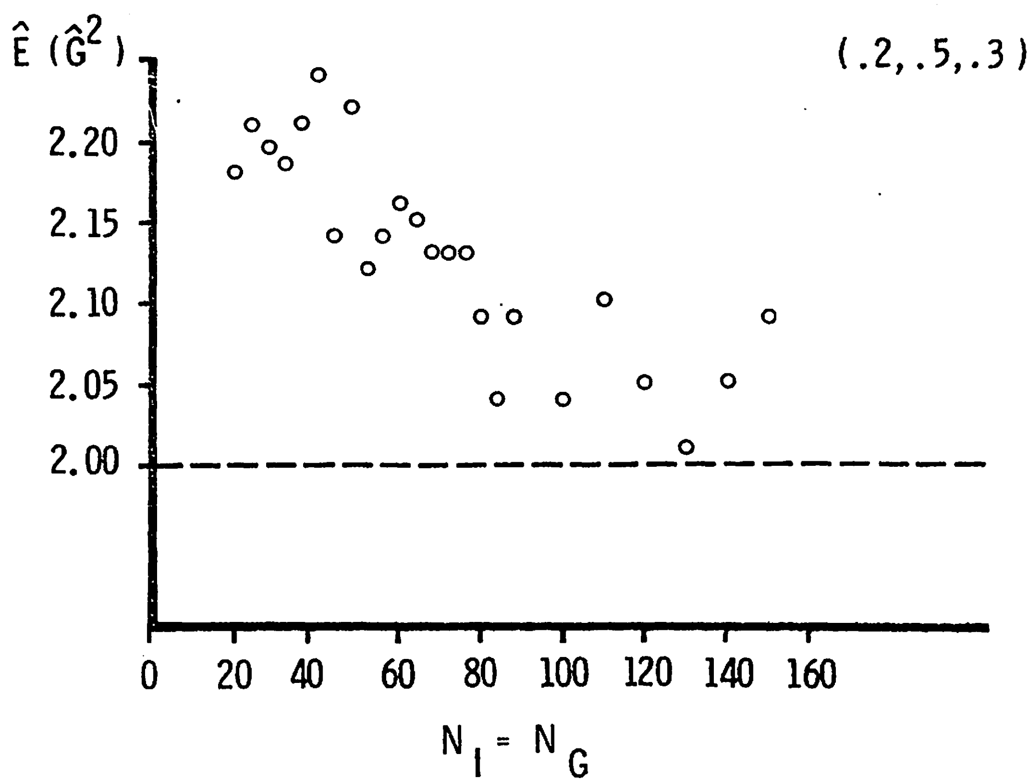

Figure I. Monte Carlo Means of G^2 for (.2,.5,.3) as Sample Size Changes (5000 Trials per Point).

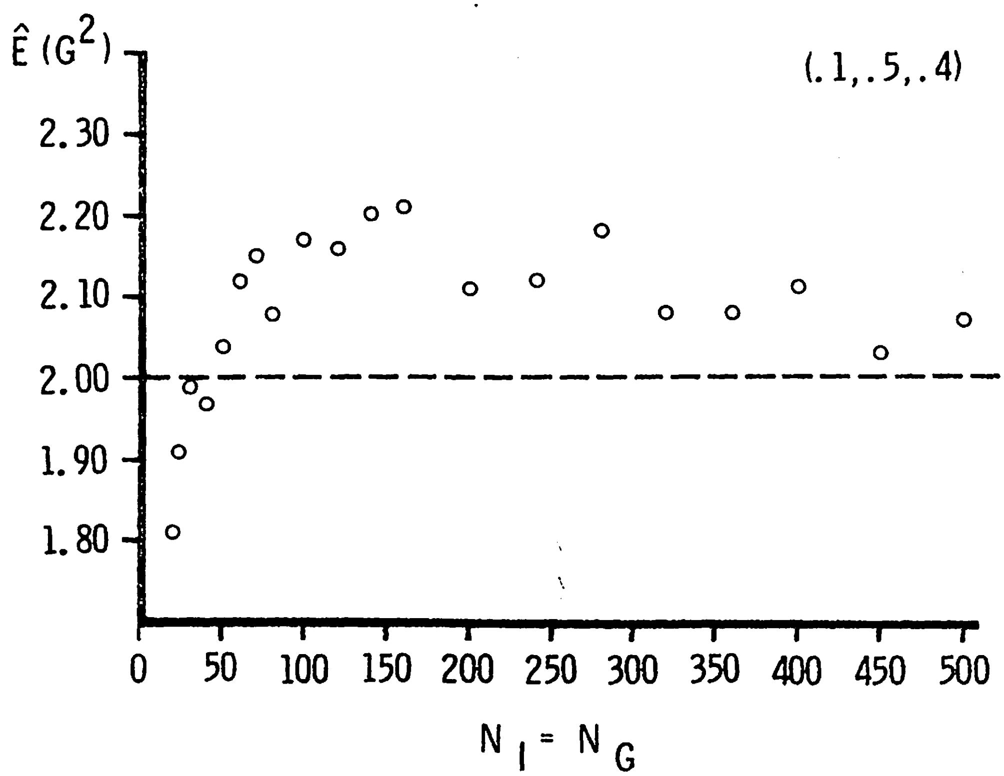

Figure J. Monte Carlo Means of G^2 for (.1,.5,.4) as Sample Size Changes (5000 Trials per Point).