Florida Journal of Educational
Research, Volume XVII, 1975

THE APPLICATION OF SIMULTANEOUS CONFIDENCE INTERVALS FOR MULTINOMIAL POPULATIONS TO AN INVENTORY OF PLANS AND GOALS

John J. Convey

Catholic University of America

The purpose of this paper is to review several procedures for constructing simultaneous confidence intervals for judging contrasts among multinomial populations. The procedures are illustrated by applying them to the responses of a systematic sample of eighth graders in Florida who responded to a section of the Florida State-Wide Eighth-Grade Testing Program entitled "Your Plans and Goals." This section provides the student with the opportunity to report on his occupational and educational aspirations and expectations. The populations compared are white males, black males, white females, and black females.

Theoretical Framework

A multinomial population is one which is partitioned into more than two mutually exclusive and exhaustive subclasses. Let Π_{ij} be the probability that an observation in the jth multinomial population ($j = 1, 2, \ldots, J$) will fall in the ith class ($i = 1, 2, \ldots, I$). Let $\Sigma_i^I = 1 \ \Pi_{ij} = 1$ for all j. A contrast among the J multinomial populations is defined as a linear combination of the Π_{ij}, $\Sigma_{i,j} c_{ij}\Pi_{ij}$, where the c_{ij} are known constants subject to the condition that $\Sigma_{j=1}^J \ c_{ij} = 0$, for all i. Goodman (1964) shows that the constraint that $c_{ij} \neq c_{i1}$ for some i, j is also necessary.

Let n_j be the number of observations in the jth multinomial population, and let n_{ij} be the number of observations in the ith class of the jth population. Maximum likelihood estimators, P_{ij}, of the Π_{ij} are obtained in the usual fashion, $P_{ij} = n_{ij}/n_j$. The variance of P_{ij} is estimated by $P_{ij}(1 - p_{ij})/n_j$.

Confidence Intervals

Usually more than one contrast among the multinomial populations would be of interest in any investigation. If each contrast separately is tested for significance at a particular choice of α, then the probability of a type I error for all contrasts among the populations is greater than the specified α. In order to maintain the probability of a type I error below a specified value, it is necessary to make use of some simultaneous confidence interval procedure. Two such procedures are presented below.

The first procedure is based on a Scheffé strategy (see Miller, 1966). Let Ψ represent a contrast among the multinomial populations.

$\hat{\Psi}$ is the maximum likelihood estimator of Ψ given by $\hat{\Psi} = \Sigma_{i,j}\, c_{i,j} P_{ij}$.

The variance of $\hat{\Psi}$ is estimated by:

$$SE^2\,(\hat{\Psi}) = \Sigma_{j=1}^{J} \frac{1}{n_j} \left[\Sigma_{i=1}^{I}\, c_{ij}^2\, P_{ij}\,(1 - P_{ij})\right] \tag{1}$$

Goodman (1964) gives the simultaneous confidence intervals as:

$$\hat{\Psi} - SE\,(\hat{\Psi})\,L \leq \Psi \leq \hat{\Psi} + SE\,(\hat{\Psi})\,L, \tag{2}$$

where L is the positive square root of the appropriate value from the χ^2 distribution with (I-1) (J-1) degrees of freedom at a significance level of $(1-\alpha)$, and α represents the probability of a type I error for all intervals together. Light (1973) presents a simple example of this procedure for pairwise contrasts among binomial populations.

If not all possible contrasts among the populations are of interest, a Bonferroni strategy may be employed to calculate the intervals (see Miller, 1966). Goodman (1964) gives these intervals as:

$$\text{Lim Pr} \left[\hat{\Psi}_k - SE\,(\hat{\Psi}_k)\,Z_k \leq \Psi \leq \hat{\Psi}_k + SE\,(\hat{\Psi}_k)\,Z_k, \text{ for } k=1,\ldots,G\right] \geq 1-\alpha,\ n\to\infty \tag{3}$$

where G is the number of contrasts of interest and Z_k is the 100 $(1-\beta_k)$ percentile point on the unit normal distribution such that $\Sigma_{k=1}^{G}\,\beta_k = \alpha/2$. If one wishes the Z_k to be all equal, then the level of significance for each contrast is $\alpha/2G$. For most values of G and for the usual values of α, the intervals given by (3) will be shorter than those given by (2).

Homogeneity Hypothesis

Often an hypothesis of interest is that the J multinomial populations are homogeneous:

$$H_o \ : \ \Pi_{i1} = \Pi_{i2} = \ldots = \Pi_{ij}, \quad \text{for } i = 1, \ldots, I. \tag{4}$$

The statistic most commonly used to test this hypothesis is the familiar:

$$X^2 = \Sigma_{i,j} (n_{ij} - n_j p_i)^2 / n_j p_i , \tag{5}$$

where $n_j p_i$ represents the expected frequency for the ijth cell, and p_i is the weighted arithmetic mean. When (4) is true, X^2 will have an approximate X^2 distribution with (I-1)(J-1) degrees of freedom (Marascuilo, 1971).

Goodman (1964) shows that the rejection of (4) by the statistic X^2 will not guarantee the presence of a significant contrast produced by the intervals (2). If one wishes to test that at least one contrast is significant, Goodman proposes a related statistic, Y^2, given by:

$$Y^2 = \Sigma_{i,j} (n_{ij} - n_j p_i^*)^2 / n_{ij}, \tag{6}$$

where $p_i^* = \overline{p}_i / \{\Sigma_{k=1}^{I} \overline{p}_k\}$ and $\overline{p}_i$ is the weighted harmonic mean given by $n / \{\Sigma_{j=1}^{J} (n_i / p_{ij})\}$. Y^2 has an X^2 distribution with (I-1) (J-1) degrees of freedom.

Goodman shows that the rejection of the hypothesis (4) by Y^2 will occur if and only if at least one estimated contrast given by the intervals (2) is significantly different than zero. He also points out that this property does not extend to the intervals given by (3) since not all contrasts are considered by that procedure.

For ease of calculation, Y^2 can be expressed as:

$$Y^2 = \left\{\Sigma_{i=1}^{I} \left(\frac{1}{\Sigma_{j=1}^{J} n_i / p_{ij}}\right)\right\}^{-1} - n \tag{7}$$

where $n = \Sigma_{j=1}^{J} n_j$. When (4) is true, the statistics X^2 and Y^2 are asymtotically equal.

METHOD

Instrument

The "Your Plans and Goals" section of the Florida Eighth Grade battery consists of four items. The first two items, measuring the student's occupational aspirations and expectations respectively, provide the examinee with five lists of job titles. These lists are shown in Table 1. The examinee's task is to choose the list containing the occupations most similar to the kind of work he would most like to do when his education is completed (item 1), and the list of occupations most similar to the kind of work he expects to be doing when he finished his schooling (item 2). Each of the lists contains six occupations requiring a similar level of educational development.

Insert Table 1 about here

The third and fourth items, measuring the student's educational aspirations and expectations respectively, ask the examinee to choose a level of formal education from a list of five levels. The levels, shown in Table 2, range from high school dropout through advanced training beyond college.

Insert Table 2 about here

Subjects

A systematic sample was selected from the population of Florida eighth graders who took the battery in February, 1974. The scores of every fiftieth student were selected from a tape arranged by public, parochial, and private schools. The race and sex of each student were ascertained by demographic information provided by each examinee. The sample contains 930 white males, 255 black males, 902 white females, and 279 black females. Only those students who indicated that they belonged to one of the preceding populations were used in the study.

Procedure

The frequency distribution and the proportion distribution of the responses to each item were calculated for the four populations. Some examinees did not respond to some items; therefore, an additional response classification of blank, indicated by a zero, was included in the distributions.

For each item, the Y^2 statistic was calculated for the 6 x 4 contingency table and compared with the X^2 distribution with 15 degrees of freedom and $\alpha = .05$. All contrasts which compared the same response category for two populations at a time were formed. An example of one of the above would be the contrast which compares the white males with the black males on response category 2. Thirty-six contrasts are possible for each item in the instrument. The critical value for the Bonferroni intervals is $Z = 3.20$ ($\alpha/2G = .05/72 = .00069$), while the value for the Scheffé-like intervals is $L = 5.00$ ($\sqrt{\chi^2} = \sqrt{25}$ at df = 15 and $\alpha = .05$). Since the Bonferroni intervals are shorter, these are used in the analysis.

For each item, the six possible 6 x 2 contingency tables were examined separately. X^2 and Y^2 statistics were calculated for each table. Again, the pairwise contrasts comparing the same response categories were formed, resulting in six comparisons for each pair of populations. Once again, the Bonferroni value ($Z = 2.64$, 6 comparisons) is smaller than the Scheffé value ($L = \sqrt{\chi^2_5} = 3.33$); therefore, the former is used to construct the intervals.

RESULTS

The frequency distributions and proportions distributions are shown in Table 3. An examination of the distributions for items 1 and 2 reveal that the proportions for the more menial occupation listings (2, 3, and 5) increased from the aspiration response to the expectation response, while the proportions for the more prestigious occupations decreased for the same responses. This suggests that the students are aspiring to more prestigious occupations than they realistically expect to achieve. A similar trend is present in the responses to the educational items (3 and 4).

Insert Table 3 about here

Each of the Y^2 statistics calculated for the 6 x 4 tables is significant at $\alpha = .05$. These values are shown in Table 4. The results of the Bonferroni intervals for the pairwise contrasts for each response category are shown also in Table 4. Fifteen of the 36 contrasts for item 1 and 2 are significant; whereas, only one is significant for item 3 and none for item 4. (The reader should recall that a significant Y^2 guarantees the existence of a significant contrast; however, the contrast may not be a simple pairwise one between similar response categories).

Insert Table 4 about here

Several patterns are evident for the two occupational items. Females of both races overwhelmingly favor response category 1 over their male counterparts, while the males favor response category 5. In examining the list of occupations, it appears that some sex bias may be present. In general, there appears to be a dearth of specific female occupations. In their responses, females could have been enticed by the presence of "registered nurse" in category 1 and repelled by the generally masculine occupations listed in category 5.

Some significant race-specific patterns are also apparent. In response category 2, where the most menial occupations are listed, significant differences are noted between the responses of black males and white males on both the aspiration item and the expectation item. In each case, the black males have a higher proportion of responses to this category. However, in no other response category is the proportion of responses of black males significantly different than the proportion of responses of white males. Black females have significantly fewer aspiration responses to category 4, the most prestigious jobs, than each of the other populations. In the expectation responses in this same category, black females continue to have significantly fewer responses than white males and white females, but not black males.

The response patterns of the four groups to the two educational items are rather similar. The only significant result recorded is between black males and black females on the aspiration item in response category 3. The black males show a higher proportion of responses here than do the black females. The black females seem to have compensated for this by choosing category 5 more often than the black males do, although the difference is not significant. The significance of the Y^2 statistic for each educational item demonstrates the existence of some contrast which would differentiate the populations. For item 3, a simple pairwise one is present. Undoubtedly, more complex contrasts are significant for this item and also for item 4. No attempt is made to formulate more complex contrasts in this study. Interested investigators would do well to explore more complex contrasts among the populations in an attempt to examine differences between them.

Insert Table 5 about here

Table 5 contains the X^2 and Y^2 statistics for each of the six 6 x 2 tables with each item. For the most part, both values agree closely. All of the Y^2 values for items 1 and 2 are significant at $\alpha = .05$(critical X^2=11.1,df= However, only five of the 12 Y^2 values for items 3 and 4 are significant.

Insert Table 6 about here

The Bonferroni confidence interval results for the pairwise comparisons for each response category considering just six contrasts simultaneously are shown in Table 6. These intervals are slightly smaller than the ones which were based on 36 contrasts, thus producing less conservative results than those shown in Table 4. The contrasts here are identical to those reported in Table 4 and have the same standard error, only the lengths of the intervals differ since a different critical value is used. The contrasts which become significant under this less conservative procedure are circled. It is noteworthy that in this application, there exists a simple pairwise contrast which is significant for each of the significant Y^2 statistics. Especially noteworthy is that the use of X^2 to compare the responses between black males and black females on item 4 would not have revealed the existence of a significant contrast; whereas, the use of Y^2 does.

The use of the less conservative procedure does not alter the interpretation for items 1 and 2. However, the pattern of significant contrasts is more similar under this procedure. For items 3 and 4, there are a few more contrasts which are significant. The differences in the aspiration item (item 3) again occur in response category 3 and involve all of the black male comparisons. In each case, the black male proportion is higher than each of the other populations. The trade off again exists in the upper two choices (4 and 5) where the black males proportion (.463) is lower than the total proportions for the three other groups (.574, .569, and .552, respectively).

The two significant contrasts relating to item 4 also involve black males. The proportion of black males expecting to be high school dropouts is significantly higher than the white females. It is also higher than the other two populations, but the difference is not significant. On the other hand, a significantly higher number of black females expects to finish high school than the number of their black male counterparts.

SUMMARY AND CONCLUSIONS

Two procedures for constructing simultaneous confidence intervals on contrasts comparing multinomial populations were illustrated. The first procedure, based on a Scheffé strategy, makes use of the X^2 distribution to construct the intervals. The second procedure, based on a Bonferroni strategy, uses the unit normal distribution at significant levels determined by the number of contrasts of interest. The Bonferroni strategy will ordinarily yield shorter intervals than the Scheffé strategy, if not all possible contrasts are of interest. Both strategies maintain the probability of a type I error below a specified value when the contrasts of interest are considered simultaneously.

If the hypothesis of homogeneity is of interest, the Y^2 statistic introduced by Goodman (1964) provides a necessary and sufficient condition for the significance of at least one of the Scheffé-type intervals. The traditional X^2 statistic based on observed and expected frequencies does not have this property. If one wishes to take advantage of the generally shorter Bonferroni intervals, no overall statistic provides a necessary and sufficient condition for the significance of at least one of these intervals.

Since the primary purpose of the paper was to illustrate the methodology of the simultaneous interval procedures, only certain contrasts were examined. As a result of testing simple pairwise contrasts between the same response categories over the four items of the "Your Plans and Goals" instrument, the following conclusions appear to be warranted. The educational aspirations and expectations of the four populations are very similar. Black males appear to have a tendency to aspire to a lower educational level than the other groups. In addition, the black males expect to drop out before finishing high school to a greater extent than their counterparts.

Some sex bias, evidenced by an apparent lack of specific female occupations, appears to be present in the lists of occupations. Because of this apparent bias, it is difficult to interpret and contrast the occupational aspirations and expectations of the four groups. However, it seems that black males and black females do aspire to and expect to achieve more menial occupations than their white counterparts. In an attempt to correct this apparent bias, the lists of occupations have been revised and expanded to include more specific female occupations for the 1975 administration of the instrument.

The confidence interval procedures reviewed here are based on asymptotic theory, so some care should be exercised in using them on rather small populations. Sociologists, school personnel, and other interested researchers are encouraged to investigate more complex contrasts among the populations than have been explored in this paper to attempt to uncover other interesting and potentially important relationships which may be present.

REFERENCES

Goodman, L. A. "Simultaneous confidence intervals for contrasts among multinomial populations." Annals of Mathematical Statistics, 1964, 35, 716-725.

Light, R. L. "Issues in the analysis of qualitative data." In R. M. Travers (Ed.) Second handbook of research on teaching. Chicago: Rand McNally, 1973.

Marascuilo, L. A. Statistical methods for behavioral science research. New York: McGraw-Hill, 1971.

Miller, R. G. Simultaneous statistical inference. New York: McGraw-Hill, 1966.

Table 1

List of Occupations

(Group 1)	(Group 2)	(Group 3)
registered nurse	fruit picker	medical secretary
television cameraman	deliveryman	welder
food inspector	window cleaner	chief telephone operator
airplane navigator	furniture mover	detective
hospital insurance agent	clock assembler	linen-room supervisor
probation officer	sandwich maker	radio repairman

(Group 4)	(Group 5)
college professor	driver of a large truck
vice-president of a large company	apartment house manager
biologist	automobile body repairman
astronomer	mimeograph machine operator
newspaper editor	ambulance attendant
chemist	cashier in a restaurant

Table 2

Educational Levels

1	attend school beyond the eighth grade, but not graduate from high school
2	graduate from high school
3	graduate from a two-year college or technical school
4	graduate from a four-year college
5	graduate from a four-year college and take further advanced training

Table 3

Frequency and Proportion Distributions

Item 1 (Occupational Aspirations)

	WM	BM	WF	BF	WM	BM	WF	BF
0	77	18	65	15	083	071	072	054
1	156	50	371	123	168	196	411	441
2	40	33	23	28	043	129	025	100
3	150	39	165	54	161	153	183	194
4	300	62	215	37	323	243	238	133
5	207	53	63	22	223	208	070	079

Item 2 (Occupational Expectation)

	WM	BM	WF	BF	WM	BM	WF	BF
0	79	19	69	15	085	075	076	054
1	140	30	301	87	151	118	334	312
2	49	34	60	36	053	133	067	129
3	165	44	169	80	177	172	187	287
4	260	54	210	32	280	212	233	115
5	237	74	93	29	255	290	103	104

Item 3 (Educational Aspirations)

	WM	BM	WF	BF	WM	BM	WF	BF
0	71	20	57	16	076	078	063	057
1	28	15	22	11	030	059	024	039
2	137	35	153	56	147	137	170	201
3	161	67	157	42	173	263	174	151
4	206	43	202	54	222	169	224	194
5	327	75	311	100	352	294	345	358

Item 4 (Educational Expectation)

	WM	BM	WF	BF	WM	BM	WF	BF
0	74	22	60	17	080	086	067	061
1	36	18	23	12	039	071	025	043
2	188	46	218	78	202	180	242	280
3	177	49	169	41	184	192	187	147
4	253	51	239	63	272	200	265	226
5	208	69	193	68	224	271	214	244

Table 4

Y^2 for Each Item and Results of Pairwise

Contrasts Based on 36 Comparisons

	1 (Y^2=335.97)						2 (Y^2=276.01)						3 (Y^2=29.21)						4 (Y^2=33.81)					
	0	1	2	3	4	5	0	1	2	3	4	5	0	1	2	3	4	5	0	1	2	3	4	5
WM-BM			x						x															
WM-WF		x			x	x		x				x												
WM-BF		x			x	x		x	x	x	x	x												
BM-WF		x	x			x		x				x												
BM-BF		x			x	x		x		x		x				x								
WF-BF			x		x					x	x													

x indicates a significant contrast

Table 5

X^2 and Y^2 Values for Each 6x2 Table Within Each Item

	Item 1		Item 2		Item 3		Item 4	
	X^2	Y^2	X^2	Y^2	X^2	Y^2	X^2	Y^2
WM-BM	29.97	20.69	24.81	18.80	17.67	15.91	11.06	11.03
WM-WF	184.47	206.32	128.37	139.22	3.20	4.55	7.09	7.93
WM-BF	136.81	156.37	109.23	131.66	6.98	6.81	10.74	10.96
BM-WF	108.46	86.94	92.21	100.54	21.83	18.64	21.96	18.50
BM-BF	52.06	57.79	63.09	71.51	15.41	15.78	11.02	11.35
WF-BF	40.69	32.89	36.59	40.15	4.73	4.96	7.62	7.78

(critical X_5^2 = 11.1, α = 0.05)

Table 6

Results of Pairwise Contrasts Based on 6 Comparisons

	Item 1						Item 2						Item 3						Item 4					
	0	1	2	3	4	5	0	1	2	3	4	5	0	1	2	3	4	5	0	1	2	3	4	5
WM-BM			x	ⓧ					x							ⓧ								
WM-WF		x			x	x		x				x												
WM-BF		x	ⓧ		x	x		x	x	x	x	x												
BM-WF		x	x			x		x	ⓧ			x				ⓧ				ⓧ				
BM-BF		x			x	x		x		x	ⓧ	x				x					ⓧ			
WF-BF			x		x				ⓧ	x	x													

x indicates a significant contrast

ⓧ indicates a contrast which is significant when intervals are based on 6 comparisons, but not significant when intervals are based on 36 comparisons