Behavior Science Research
Vol. 18, No. 2, Summer 1983

An Exact Significance Test for Three-Way Interaction Effects

Douglas R. White, Robert Pesner, and Karl P. Reitz

ABSTRACT: A modification of Fisher's exact test for the 2×2×2 contingency table is proposed as a test of the null hypothesis of no three-way statistical interaction among variables, controlling for the two-way or first-order correlations. The test uses a truncated hypergeometric distribution, limited by the bivariate marginal totals of the variables. Possible generalizations to L×M×N *tables are discussed. The test is also applicable to the null hypothesis of no difference in the magnitude of correlation in a comparison of two bivariate distributions. Illustrations of each application are provided. One obvious use in cross-cultural or survey research is as a test of the replication of a correlation in different subsamples of a population.*

Much attention in recent years is paid to developing statistical techniques for applications in social scientific research, particularly in developing models for discrete or ordinal variables. Perhaps the most important breakthrough has been in the development of log-linear techniques based on the cross-product ratio. Goodman's approach has made it possible to decompose complex patterns in *N*-way tables into first-, second-, or higher-order interactions. The strength of the cross-product (log linear) model is that since most correlation coefficients are also based on cross-product ratio measures, the significance tests are appropriate for testing correlational inferences. This also limits their applicability. Recent work on entailment structures (White, Burton, and Brudner 1977) represents an instance where the correlational model is rejected in favor of discrete if-then statements about contingent relationships, with no assumption of invariant cross-products. Such discrete relationships can best be analyzed using statistical tests based on the hypergeometric probability distribution, involving discrete sampling without replacement. This forms the starting point for our approach, particularly the application of the hypergeometric distribution known as Fisher's exact test.[1]

Fisher's exact test can be used to measure the statistical significance of differences between two discrete univariate distributions (or the probability of the observed or more extreme differences) under the null hypothesis of

the equivalence of the distributions.[2] For cross-classification or contingency tables, Fisher's test also provides a test of significance for the correlation between two discrete variables.[3] For dichotomous variables, producing 2×2 tables, Fisher's test is quite simply applied. The test is generalizable to cases involving $M \times N$ tables, but with some difficulty since an explicit ordering of possible distributions by degree of departure from independence must be defined.[4]

Bartlett (1935) was the first to propose a generalization of Fisher's exact test to find significant interaction in $2 \times 2 \times 2$ contingency tables. He developed both an exact and an asymptotic test. His tests have been generalized by a number of authors. Zelen (1971) extended his tests to $2 \times 2 \times k$ tables, Gart (1972) to $2 \times j \times k$ tables, and Patil (1974) to $i \times j \times k$ tables. The basic distribution formulas for specific models can be found in a variety of sources (see for example Bishop et al. (1975) for the $2 \times 2 \times k$ case and Halperin et al. (1977) for the $i \times j \times k$ case). These authors have mainly been concerned with the asymptotic version of these tests, with little attention given to the implementation of the exact tests under different sampling models.

In this paper we redevelop Bartlett's exact test to measure the statistical significance of a difference between (a) two dichotomous bivariate distributions and (b) second-order interaction effects in systems of three variables, with the three bivariate distributions between the three variables held constant. This three-way interaction is directly analogous to interaction between two variables and is our main interest in developing the statistical model used in this paper. We also distinguish cases where only one of the three possible bivariate distributions between three variables is held constant. We conclude with a brief discussion about generalizing these results to polychotomous variables ($L \times M \times N$ tables).

Differences Between Two Discrete Bivariate Distributions

As in the case of the difference between two discrete univariate distributions, the null hypothesis asserts the *equivalence* of the two distributions. Obviously, this is only possible if the two distributions are of the same dimensions, say $M \times N$. Provisional acceptance of the null hypothesis generates the corollary that the two populations are structurally equivalent as far as the categories involved in the distribution are concerned. This in turn allows the two distributions to be combined and considered as complementary exhaustive samples of a larger population formed by combining the two distributions, with all three bivariate distributions

held constant. This takes the tabular form of a $2 \times M \times N$ contingency table, where the value of the first variable indicates one or the other of the original bivariate distributions. We start below with the simple case of a $2 \times 2 \times 2$ table (dichotomous bivariate distributions).

Three-way Interaction

First-order interaction refers to relationships between pairs of variables. However, this is only the simplest level on which interaction between variables takes place. Given a system of three variables, say X_1, X_2, and X_3, it may be of interest to see if, for example, X_1 affects the relationships between X_2 and X_3. This is second-order interaction. Given larger systems of variables this concept can be extended as far as is desirable.

In tests of statistical significance of three-way interaction, the null hypothesis asserts the independence of a univariate distribution (the control variable) from all the various bivariate distributions present in the system of variables. For the purposes of this test, therefore, all bivariate distributions present in the system are held constant. We start with a system of three dichotomous variables, which takes the form of a $2 \times 2 \times 2$ contingency table.

Table 1: Notation

X_1, X_2, X_3	= A set of three variables taking values {1, 2} defined over N observations.
$A_{i..}$	= The number of observations for variable X_1 having value i, where $i \leq \{1,2\}$
$A_{.i.}$	= The number of observations for variable X_2 having value i, where $i \leq \{1,2\}$
$A_{..i}$	= The number of observations for variable X_3 having value i, where $i \leq \{1,2\}$
A_{ijk}	= The number of observations in a cell of the trivariate distribution taking values i, j, k, where $i, j, k \leq \{1,2\}$
a_{ijk}	= A possible number of observations in a cell of the trivariate distribution, given the marginal constraints $A_{.jk}$, $A_{i.k}$, $A_{ij.}$
$\bar{a}_{ijk}$	= The largest possible a_{ijk} given the above marginal constraints
$\underline{a}_{ijk}$	= The largest possible a_{ijk} given the above marginal constraints.

The Model

The three-way interaction problem as well as comparison of two bivariate distributions begin with the statistical analysis of a $2\times2\times2$ contingency table, with one variable acting as control. In interaction, of course, each variable takes a turn as the control variable. This involves a system of three dichotomous variables, X_1, X_2, and X_3. A notation for the statistical analysis of a $2\times2\times2$ table is defined in Table 1 and illustrated in Figure 1. A_{ijk} is the actual cell value in a trivariate distribution: the subscripts stand for the values of the first, second, and third variables (X_1, X_2, and X_3), respectively. Marginal totals for the first variable are designated $A_{1..}$ and $A_{2..}$, for the second variable $A_{.1.}$ and $A_{.2.}$, and so forth. Bivariate distributions, while not shown in the table, can be designated as follows: for the bivariate distribution (cell values in the 2×2 table) between X_2 and X_3, for example, the cell values are $A_{.11}$, $A_{.12}$, $A_{.21}$, and $A_{.22}$ (see Figure 1). The total number of observations is $A_{...}(=N)$.

Both cases dealt with in this paper involve examining the possible cell values in the trivariate distribution while holding all three bivariate distributions constant. Small letter-subscripted values, such as a_{ijk}, designate possible cell values in the trivariate distribution so constrained. Minimal and maximal cell values under these constraints are designated $\underline{a}_{ijk}$ and $\bar{a}_{ijk}$, respectively.

Figure 1 shows a schematic representation of $2\times2\times2$ distributions with bivariate distributions projected as 2×2 marginal total matrices up, out, and to the right, and univariate distributions as marginal totals of these matrices.

The computation of the probability, $P(A_{111})$, of an observed distribution, is derived by consideration of the number of ways of drawing A_{111}, A_{112}, A_{121}, and A_{122} observations, respectively, out of the bivariate marginal totals in the outward matrix of Figure 1: $A_{.11}$, $A_{.12}$, $A_{.21}$, and $A_{.22}$. The number of ways of drawing such samples, without replacement, is [5]

$$\binom{A_{.11}}{A_{111}}\binom{A_{.12}}{A_{112}}\binom{A_{.21}}{A_{121}}\binom{A_{.22}}{A_{122}} \tag{1}$$

The total sample space for all possible ways of drawing samples is

$$\sum_{i=\underline{a}_{111}}^{\bar{a}_{111}}\binom{A_{.11}}{i}\binom{A_{.12}}{A_{11.}-i}\binom{A_{.21}}{A_{1.1}-i}\binom{A_{.22}}{A_{1..}-A_{11.}-A_{1.1}+i} \tag{2}$$

Table 1 gives the definitions of a_{111} and a_{111}.

The point probability of drawing A_{111}, A_{112}, A_{121}, and A_{122} observations is therefore (identifying this probability by reference to the 111 cell)

$$P(a_{111}=A_{111}) = \frac{\binom{A_{.11}}{A_{111}}\binom{A_{.12}}{A_{112}}\binom{A_{.21}}{A_{121}}\binom{A_{.22}}{A_{122}}}{\sum_{i=\underline{a}_{111}}^{\bar{a}_{111}} \binom{A_{.11}}{i}\binom{A_{.12}}{A_{11.}-i}\binom{A_{.21}}{A_{1.1}-i}\binom{A_{.22}}{A_{1..}-A_{11.}-A_{1.1}+i}} \quad (3)$$

[The same probability will be obtained for $P(a_{211} = A_{211})$.]

Calculating the point probabilities for every value of i between $\underline{a}_{111}$ and $\bar{a}_{111}$ will produce a point probability distribution for the given trivariate distribution under the above-mentioned constraints.

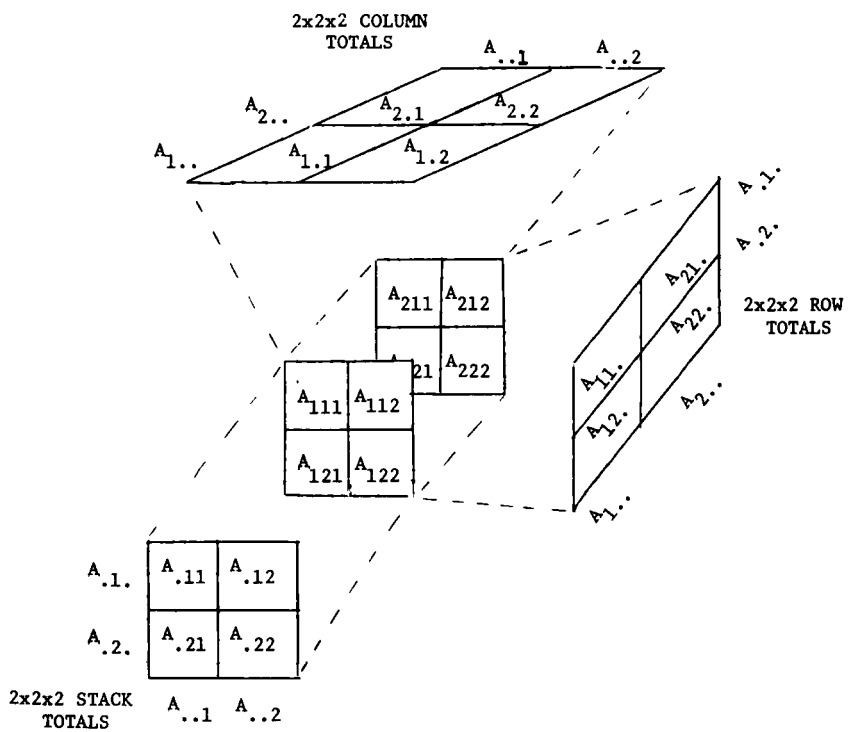
In much statistical hypothesis testing, however, what is compared to a previously established level of significance is not a point probability, but a cumulative probability. This cumulative probability is calculated by adding all the point probabilities of possible outcomes equally or more extreme than the actual outcome, thus forming one tail of the probability distribution. This concept of a tail depends on a unilinear directed ordering of the point probabilities based on a concept of degree of departure from statistical independence. In contingency tables with one degree of freedom, where a given value in any single cell completely determines the other cells, this order is always based on either increasing or decreasing values in (any) single cell. As $2\times2\times2$ tables have one degree of freedom, this procedure is fully applicable.

For any given outcome, except the two extremes of the distribution, the question arises as to which direction is appropriate for summing the cumulative probability. In one-tailed tests involving alternate hypotheses of the form $a \leq c$ or $a \geq c$, the alternate hypothesis establishes the direction: summing up to the extreme in the former case and down to the extreme in the latter.

In two-tailed tests, with alternate hypotheses of the form $a \neq c$, this method is unavailable. When using symmetrical probability distributions,

such as normal or *t*-distributions, cumulative probabilities are summed away from the mean. But hypergeometric distributions are not normally symmetrical, and using this method may involve adding point probabilities greater than that of the observed outcome. Therefore, we have substituted the method of summing along the generally descending direction of the distribution. While this may be immediately determined by examining a list of the point probabilities, a simpler method suitable for computer calculation is utilized in the following equations for the cumulative probability:

Figure 1: Schematic of Marginal Totals for a 2x2x2 Trivariate Distribution



$$P(a_{111} = A_{111} \text{ or more extreme})$$

$$= P(a_{111} \leq A_{111})$$

$$= \frac{\displaystyle\sum_{i=\underline{a}_{111}}^{A_{111}} \binom{A_{.11}}{i}\binom{A_{.12}}{A_{11.}-i}\binom{A_{.21}}{A_{1.1}-i}\binom{A_{.22}}{A_{1..}-A_{11.}-A_{1.1}+i}}{\displaystyle\sum_{i=\underline{a}_{111}}^{\bar{a}_{111}} \binom{A_{.11}}{i}\binom{A_{.12}}{A_{11.}-i}\binom{A_{.21}}{A_{1.1}-i}\binom{A_{.22}}{A_{1..}-A_{11.}-A_{1.1}+i}}$$

or

$$P(a_{111} = A_{111} \text{ or more extreme})$$

$$= P(a_{111} A_{111})$$

$$= \frac{\displaystyle\sum_{i=\underline{a}_{111}}^{A_{111}} \binom{A_{.11}}{i}\binom{A_{.12}}{A_{11.}-i}\binom{A_{.21}}{A_{1.1}-i}\binom{A_{.22}}{A_{1..}-A_{11.}-A_{1.1}+i}}{\displaystyle\sum_{i=\underline{a}_{111}}^{\bar{a}_{111}} \binom{A_{.11}}{i}\binom{A_{.12}}{A_{11.}-i}\binom{A_{.21}}{A_{1.1}-i}\binom{A_{.22}}{A_{1..}-A_{11.}-A_{1.1}+i}} \tag{4}$$

whichever is *smaller.*

Or to simplify the notation further:

$$P(a_{111} = A_{111} \text{ or more extreme})$$

$$= P(a_{111} \leq A_{111}) \text{ or } 1 - P(a_{111} \leq A_{111}) + P(a_{111} = A_{111})$$

whichever is smaller.[6]

[The same probabilities would be obtained using formulas for $P(a_{211} = A_{211}$ or more extreme).]

It should be noted that the resulting probability distribution is *not* a full hypergeometric distribution. A full distribution with four combinatorials in the numerator would have the form:

$$\frac{(A)(B)(C)(\quad D \quad)}{(a)(b)(c)(t-a-b-c)}$$
$$\overline{(A+B+C+D)}$$
$$(\quad t \quad)$$

and would require the variables *a, b,* and c to take on any values greater than or equal to 0 such that $a+b+c \leq t$. It would also be subject to the ordering problems discussed below. However, in our model, the variables analogous to *b* and *c* (namely A_{112} and A_{121}) are determined by the variable analogous to *a* (namely A_{111}). This, of course, derives from the fact that given all marginals, a 2×2 (or a $2 \times 2 \times 2$) contingency table, having but one degree of freedom, is totally determined once any one cell is known. Thus the problem of explicitly ordering possible distributions by degree of departure from independence is avoided.[7]

This reduction of the sample space from the theoretical full hypergeometric distribution is a result of the dual constraints placed on the $2 \times 2 \times 2$ tables under investigation, as discussed above: (a) it is constrained by the bivariate distribution X_2 versus X_3 and (b) it is constrained by the marginals in the table itself, which are, in fact, the values of the other two possible bivariate distributions, X_1 versus X_2 and X_1 versus X_3.

By way of contrast, significance testing involving a relation between a univariate and a bivariate distribution is an example of a full hypergeometric distribution involving three variables. In one form this is the classic three-variable case, central to much quantitative social science research. For example, in such a standard work as Rosenberg's *Logic of Survey Analysis,* [8] the intervention of a third "test" variable is the act which enables the researcher to determine the nature of a relationship between two variables found in the data. Of course, Rosenberg's treatment is purposefully unsophisticated statistically. He restricts himself to a simple comparison of percentages, thus subjecting his data merely to the first stage of refinement. He makes use of no probability model and thus does no real statistical hypothesis testing. He uses no criteria to measure the significance of the differences he finds among the percentages. One such criterion can be based on the hypergeometric model.

Where the control variable is dichotomous, the three variable case takes the form of a $2 \times M \times N$ contingency table. The null hypothesis asserts the

independence of the univariate and bivariate distributions. In this case only the given bivariate distribution is held constant and which variables are involved will affect the resulting probabilities. Using the notation developed above, in the case where X_1 is the control variable, the equations become

$$P(a_{111}=A_{111}) = \frac{\binom{A_{.11}}{A_{111}}\binom{A_{.12}}{A_{112}}\binom{A_{.21}}{A_{121}}\binom{A_{.22}}{A_{122}}}{\binom{A_{...}}{A_{1..}}}$$

and

$$P(a_{111} = A_{111} \text{ or more extreme})$$

$$= P(a_{111} \geq A_{111})$$

$$= \frac{\sum_{i=\underline{a}_{111}}^{A_{111}} \binom{A_{.11}}{i}\binom{A_{.12}}{A_{11.}-i}\binom{A_{.21}}{A_{1.1}-i}\binom{A_{.22}}{A_{1..}-A_{11.}-A_{1.1}+i}}{\binom{A_{...}}{A_{1..}}}$$

or

$$= P(a_{111} \geq A_{111}) = 1 - P(a_{111} \leq A_{111}) + P(a_{111} = A_{111}) \qquad (5)$$

whichever is *smaller.*

[The same probabilities would be obtained using formulas for $P(a_{211} = A_{211}$ or more extreme).] The difference in the denominator arises from the fact that the relevant marginals of the bivariate distributions of X_1 versus X_2 and X_1 versus X_3 ($A_{1.1}$, $A_{1.2}$, $A_{11.}$, and $A_{12.}$) are not prior constraints on the controlled table values.

Whichever formulas are applicable to the data at hand, once the probability has been computed it can be compared with whatever level of significance has been previously established to determine whether the null hypothesis should be accepted or rejected. We must emphasize that this is a test of significance only, and says nothing about the nature or direction of any differences between the two bivariate distributions (if that is what is

being tested), about the type of interaction involved (if that is what is being tested), or about the relation between a univariate and a bivariate distribution (if that is what is being tested).

Proof of Symmetry. One of the important features of this significance test is that it is symmetric for any of the three possible control variables. Here we offer a proof of symmetry. Such proof is necessary since the denominator [Equation (2)] of our point probability is not equal for permutations of the variables.

For any given value i in cell 111, the number of ways of drawing samples holding variable X_1 constant is given by Equation (1), substituting i for value A_{111}, as in Equation (2). Now consider the ratio of ways of drawing samples with value i in cell 111 to the ways of drawing samples with value $i+1$:

$$\frac{\binom{A_{.11}}{i}\binom{A_{.12}}{A_{11.}-i}\binom{A_{.21}}{A_{1.1}-i}\binom{A_{.22}}{A_{1..}-A_{11.}-A_{1.1}+i}}{\binom{A_{.11}}{i+1}\binom{A_{.12}}{A_{11.}-i-1}\binom{A_{.21}}{A_{1.1}-i-1}\binom{A_{.22}}{A_{1..}-A_{11.}-A_{1.1}+i+1}} \tag{6}$$

Simplifying by cancelling factorials, Equation (6) becomes

$$= \frac{i+1}{A_{1.1}-1}\,\frac{A_{1.2}-A_{11.}+i+1}{A_{11.}-i}\,\frac{A_{2.1}-A_{.11}+i+1}{A_{.11}-i}\,\frac{A_{.1.}-A_{11.}-A_{.11}+i+1}{A_{2.2}-A_{.1.}-A_{11.}-A_{.11}-i} \tag{7}$$

$$= \frac{i+1}{A_{1.1}-i}\,\frac{a_{122}+1}{A_{11.}-i}\,\frac{a_{221}+1}{A_{.11}-i}\,\frac{a_{212}+1}{a_{222}} \tag{8}$$

This result, however, is perfectly symmetric whether we begin with variable X_1, X_2, or X_3 as the control. Hence the point probabilities for each value of i must be the same for each control variable. Q.E.D.

An example of symmetric results taking each variable as a control is shown in Table 2. Given the marginal constraints of the observed distribution, there are three possible theoretical distributions where in Equation (2), $i = 1, 2, 3$. Computations are shown for X_1 as control, X_2 as control, and X_3 as control, leading to identical point probability estimates for each value of i, as predicted.

The Sampling Model

Every test of significance corresponds to the probability of some event under a random sampling model. For the truncated hypergeometric distribution the underlying sampling model is theoretically simple but practically cumbersome. A sample of the observed size N is picked at random, with replacement, from a population in which all three-way combinations of values of the variables occur with equal frequency. Each sample is tabulated in a $2\times2\times2$ table. Samples that do not correspond to the observed marginals of this table are rejected. The sampling distribution of the remaining $2\times2\times2$ tables is described by the truncated hypergeometric model. There are equivalent designs that are less cumbersome, but all constrain the sample according to marginal constraints, which is the key feature of the hypergeometric distribution.

Comparison with the Log-Linear Model

Goodman's log-linear analysis defines the null hypothesis of no N-way interaction in terms of equality of cross-product ratios. In the 2×2 table this is expressed by $ad/bc = 1$, where a, b, c, d are cell values. In the

Table 2: Interaction Tests for a 2x2x2 Table Showing Symmetry

	$X_1 = 1$		$X_1 = 2$	
	X_3=1	2	1	2
$X_2 = 1$	2	1	2	2
$X_2 = 2$	3	4	1	3

Observed Distribution

Theoretical Samples		Expected Distributions: X_1 as control	X_2 as control	X_3 as control
i = 1				
1 2	3 1	$\binom{4}{1}\binom{3}{2}\binom{4}{4}\binom{7}{3} = 420$	$\binom{5}{1}\binom{5}{2}\binom{3}{3}\binom{5}{1} = 250$	$\binom{3}{1}\binom{7}{3}\binom{4}{2}\binom{4}{0} = 420$
4 3	0 4	p = .12	p = .12	p = .12
i = 2				
2 1	2 2	$\binom{4}{2}\binom{3}{1}\binom{4}{3}\binom{7}{4} = 2520$	$\binom{5}{2}\binom{5}{1}\binom{3}{2}\binom{5}{2} = 1500$	$\binom{3}{2}\binom{7}{3}\binom{4}{2}\binom{4}{1} = 2520$
3 4	1 3	p = .73	p = .73	p = .73
i = 3				
3 0	1 3	$\binom{4}{3}\binom{3}{0}\binom{4}{2}\binom{7}{5} = 504$	$\binom{5}{3}\binom{5}{0}\binom{3}{1}\binom{5}{3} = 300$	$\binom{3}{3}\binom{7}{2}\binom{4}{1}\binom{4}{2} = 504$
2 5	2 2	p = .15	p = .15	p = .15

$2 \times 2 \times 2$ table the no three-way interaction hypothesis is specified by $(ad/bc)\ (eh/fg) = 1$. Log-linear analysis is based on decomposition of the contributions to the various cross-product ratios: for three variables there is one three-way cross product and three two-way cross products. Since the cross-product ratio underlies many of the correlational models such as Phi (Pearson's product-moment coefficient for the 2×2) and Gamma, log-linear is an appropriate model for testing statistical hypotheses about correlational models.

There are, however, statistical models that are not correlational, and for which lack of interaction is not necessarily defined by equality of cross-product ratios. Entailment analysis (White, Burton, and Brudner 1977) is one such model, where zero cells or near-zero cells are hypothesized in a series of 2×2 tables in a given data set, or a lower rate of exceptions is hypothesized for entailments of the form "If X then Y" than for their converses. For the 2×2 case log-linear and Fisher's exact test happen to converge in that the hypothesis of noninteraction is identical: the hypergeometric expectation of no interaction is one where the cross-products are equal.

Log-linear and the hypergeometric methods diverge, however, in the case of three-way interaction. In the case of entailment analysis, we may be trying to test, in this case, whether an entailment "If X then Y" is replicated under control conditions for the presence or absence of Z. *The log-linear model is inapplicable in that no equality of cross-product ratios can be logically derived* from the entailment model for the case of no interaction. The appropriate model is that developed here: given that the bivariate frequencies of the variables are fixed (all two-way interactions held constant), which is the probability of getting the observed distribution by chance? That this involves a random procedure for filling the cells of the $2 \times 2 \times 2$ table under constraint should be no surprise. It should also come as no surprise that the three-way exact test gives a different solution for the expected values of cells in the $2 \times 2 \times 2$ table under the hypothesis of no interaction than the log-linear method.

One of the attractive features of the three-way exact test is that it gives an exact computational formula for the expected values of cells, with no interaction, in the three-way table. A computational solution for these values in the log-linear approach has been proven to be impossible, and they are obtained by iterative methods. This in itself entails that the two methods give divergent results, or that different statistical models are involved. Practically, however, we have found that the methods give expected values that are extremely close under most marginal constraints.

Evaluation

Besides the log-linear model, the only other statistical measures known to us that can be used to measure statistical significance in the cases dealt with here are χ^2 and w^2.[9] However these statistical measures suffer from several shortcomings if applied to discrete data. In the first place, if the total number of cases in each controlled contingency table is less than 50 χ^2 or w^2 loses accuracy. Our statistic is not subject to any such limitation. Secondly, w^2 requires the identification of an independent variable, as it is based on the asymmetric statistic, Somer's $d_{y/x}$. As we have demonstrated above, independence or dependence is irrelevant with our model. Thirdly, and most importantly, w^2 is based on χ^2, both of which involve the assumption that the variables being analyzed are continuous. Our model, of course, makes the opposite assumption of the discreteness of the variables. Finally, both involve the assumption that sampling is done *with* replacement, while our model assumes sampling without replacement. However, our statistic as developed so far suffers from the serious drawback of being limited to dichotomous variables. It is possible to generalize for polychotomous variables, but this involves theoretical assumptions that may not be justified in specific research situations.

Taking first the $2\times2\times N$ case, where the first variable is the control variable, each controlled table will be of type $2\times N$. In the equation for P, the numerator will now have $2\times N$ instead of $2\times2=4$ terms, and of these $(2-1)\times(N-1)=N-1$ must be known before the rest are determined. Thus there is no single given ordering to the various possibilities, as these are made up of ascending and descending values for $N-1$ cells. The problem that then arises is to order these possible combinations by degree of departure from statistical independence. If the measuring scale involved is ordinal there may be a sound approach to doing this; if the scale is only nominal any approach will be arbitrary.

The problem is similar but compounded in the $2\times M\times N$ case. Even if both polychotomous variables are measured with ordinal scales, it will probably be rare that an ordering of the $M\times N$ combination of measurements possible will be given unambiguously.

Finally, the $L\times M\times N$ case. This, in fact, is merely a logical extension or generalization from the preceding cases, derived by allowing each variable in turn to be the control variable. Clearly the ordering problem here reaches a third level of complexity. In this case $(L-1)\times(M-1)\times(N-1)$ values must be given before the rest are determined, and $L\times M\times N$ possible combinations of values must be ordered. Obviously the solution to this problem will rarely be unambiguous.

Following a suggestion of Pierce's (1970) we can offer an overall solution to the ordering problem.[10] An unambiguous order may be obtained by merely listing all the point probabilities in numerical order and cumulate all point probabilities less than or equal to the point probability of the actual outcome. As Pierce says, the resulting test is "nondirectional"[11] and will work unless a "one-tailed" test is insisted upon, which in our opinion has no analogue even the the $2 \times 2 \times L$ case unless a particular ordering coefficient (e.g., Gamma) is specified, as illustrated under "Applications" (Table 6).

It should be understood, however, that the calculation of exact *point* probabilities for a particular table of any dimension is theoretically unambiguous (if practically tedious). Thus once the ordering problem is solved, calculation of cumulative probabilities can proceed in a relatively straightforward manner.

This suggestion allows unlimited extension of our model to more complex research problems. Some such possible extensions are interaction effects between univariate and multivariate distributions or between bivariate or higher-order and multivariate distributions (full hypergeometric distributions), comparisons between two multivariate distributions and higher-order interaction third-order analysis (both reduced hypergeometric distributions). We have applied interaction analysis of the $2 \times 2 \times 2$ case to large systems of variables, which we hope to discuss in a future paper.

Applications

Three applications are illustrat·

1. significance testing for difference between two bivariate distributions of dimensions 2×2;
2. significance testing for interaction effects among three dichotomous variables; and
3. significance testing for interaction effects in the $2 \times 3 \times 3$ case.

The first example is from a study by Brudner-White (1978) on the concomitants of language variability in an Austrian village near the Yugoslavian border. Her contention is that occupational endogamy is stronger than language identity as an occupational marker for the farmer populations that control access to local agrarian resources, and that language endogamy is consequently an epiphenomenon of occupational endogamy. Tables 3 and 4 present the data for occupational and language endogamy, respectively.

Table 3: Occupational Endogamy

		WIFE: daughter of Farmer	WIFE: daughter of Non-Farmer
HUSBAND	Farmer	80	7
	Non-Farmer	15	34

Table 4: Language Endogamy

		WIFE Bilingual	WIFE Monolingual
HUSBAND	Bilingual	98	16
	Monolingual	7	15

Using Equation (4) the statistical significance of the difference between these two bivariate distributions is $P = .47$. Thus, although the occupational endogamy is greater than the language endogamy (Romney's normalized measures of endogamy are .67 and .57, respectively),[12] the difference is not statistically significant at $P = .05$.

The second example is from Murdock and White's (1969) cross-cultural sample. When the worldwide association between patrilineality and bridewealth is broken down by region, as shown in Table 5, there are significant differences [$P = .0003$ using Equation (3)] between societies in the insular Pacific region and those outside of this region. Within this region, in fact, the direction of the relationship is reversed, as shown by the Gamma coefficients in Table 5.

The third example is a hypothetical illustration of use of the significance test with a table of higher dimensionality, in this case $2 \times 3 \times 3$. Assuming that for a population of 19 cases that the cross-classification of two nominal three-category variables is as follows:

1	4	3
3	2	0
3	2	1

Table 5: Regional Replication of the Association between Patrilineality and Bridewealth

INSULAR PACIFIC SOCIETIES

Bridewealth	Patrilineal Groups: Absent	Patrilineal Groups: Present
Present	8	5
Absent	9	9

Gamma = - .23

Significance Test for Difference: P = .0003

SOCIETIES OUTSIDE THE INSULAR PACIFIC

Bridewealth	Patrilineal Groups: Absent	Patrilineal Groups: Present
Present	11	46
Absent	76	22

Gamma = + .87

Nine cases are drawn from this population as possessing a certain characteristic. Does the bivariate distribution in the new subsample resemble that of the old? For illustration, we assume that the new distribution is as follows:

0	1	2
3	0	0
0	2	1

What is the probability of sampling this distribution randomly from the larger population, with the qualification that the only valid or comparable samples are ones with the same row and column totals? Table 6 shows the seven valid samples by this criterion, and the number of ways of drawing such samples at random from the total population. The point probability of each valid sample is computed as the proportion of ways of drawing each given valid sample over the total ways of drawing any valid sample. The probability of drawing the observed sample is $P = .03$, using a generalization of Equation (4). Using the nondirectional method for obtaining cumulative probabilities (the sum of all point probabilities equal to or less than the observed), the cumulative probability is also $P = .03$, as this is the smallest of all the point probabilities.

Table 6 also shows how the cumulative probability would differ if ordinal assumptions can be made about the variables, and if the tables are ordered by a correlation coefficient such as Gamma. The observed distribution, while it is the least likely to occur by chance, is not the one with the most positive Gamma coefficient. The cumulative (directional) probability in this case changes to $P = .07$, which is not as significant as the (unordered) significance test at the nominal level of association.

Conclusion

In this paper we have presented a versatile model for measuring statistical significance for use in testing hypotheses involving discrete polychotomous variables. It is most easily applied to dichotomous variables but is generalizable, with some difficulty, to those of higher order. We concentrated on two three-way applications of the model: comparisons of two bivariate distributions and second-order interaction effects in systems of three variables. Finally, we suggested several directions in which this model can be fruitfully extended.

We would be the first to assert the extreme simplicity of our model com-

Table 6: Interaction Tests for a 2x3x3 Table Formed by Sampling from a 3x3 Distribution

Population:

1	4	3
3	2	0
3	2	1

Samples:	Ways of Drawing each Sample:	Point p	ORDINAL MEASURES Cumulative P*	Gamma	Cumulative P**
0 0 3 1 2 0 2 1 0	$\binom{1}{0}\binom{4}{0}\binom{3}{3}\binom{3}{1}\binom{2}{2}\binom{0}{0}\binom{3}{2}\binom{2}{1}\binom{1}{1} = 18$	.04	.15	-.91	.04
0 0 3 2 1 0 1 2 0	$\binom{1}{0}\binom{4}{0}\binom{3}{3}\binom{3}{2}\binom{2}{1}\binom{0}{0}\binom{3}{1}\binom{2}{2}\binom{1}{0} = 18$	.04	.15	-.65	.08
0 1 2 1 2 0 2 0 1	$\binom{1}{0}\binom{4}{1}\binom{3}{2}\binom{3}{1}\binom{2}{2}\binom{0}{0}\binom{3}{2}\binom{2}{0}\binom{1}{1} = 108$	.22	.48	-.62	.30
0 1 2 2 1 0 1 1 1	$\binom{1}{0}\binom{4}{1}\binom{3}{2}\binom{3}{2}\binom{2}{1}\binom{0}{0}\binom{3}{1}\binom{2}{1}\binom{1}{1} = 252$	.52	1.00	-.40	.70***
1 0 2 1 2 0 1 1 1	$\binom{1}{1}\binom{4}{0}\binom{3}{2}\binom{3}{1}\binom{2}{2}\binom{0}{0}\binom{3}{1}\binom{2}{1}\binom{1}{1} = 54$	.11	.26	-.20	.18
0 1 2 3 0 0 0 2 1	$\binom{1}{0}\binom{4}{1}\binom{3}{2}\binom{3}{3}\binom{2}{0}\binom{0}{0}\binom{3}{0}\binom{2}{2}\binom{1}{1} = 12$	.03	.03	-.13	.07
1 0 2 2 1 0 0 2 1	$\binom{1}{1}\binom{4}{0}\binom{3}{2}\binom{3}{2}\binom{2}{1}\binom{0}{0}\binom{3}{0}\binom{2}{1}\binom{1}{1} = 18$	.04	.15	+.05	.04

* Nondirectional

** Directional

*** Lesser of two directional values, the other being .82.

pared to much of the statistical work being used in contemporary social scientific research. Yet its simplicity is by no means an indication of its limited applicability—quite the contrary. In fact, we feel it is very widely applicable, including in some situations where much more complicated statistical techniques have been applied without sufficient regard for the theoretical assumptions implicit in the techniques. Hopefully, the availability of a versatile model that is not subject to the limitations of assumptions about the number of observations and the continuous quality of variables will encourage the use of more appropriate models in social scientific research.

Notes

1. See Fisher (1954:96-97), Siegel (1956), Lieberman and Owen (1961), and Pierce (1970).
2. In a statistical test of the difference between two discrete univariate distributions, the null hypothesis would be that the two distributions are *equivalent*. Under the assumption of the null hypothesis, therefore, the populations from which these distributions are taken must also be considered as structurally equivalent, *as far as the categories used to construct the distributions are concerned.* Thus the univariate distributions can be combined into a single contingency table, which is then analyzable with Fisher's exact test. For computational details, see Pierce (1970:110-113).
3. In contrast to the preceding case, here the null hypothesis asserts the *independence* of the two variables. Here the contingency table is given and Fisher's test can be applied directly. For complete details, see Pierce (1970:116-120).
4. See Pierce (1970:127-130). This problem also arises with our model; we discuss a solution below.
5. This, of course, is only one of the three possible ways of calculating this probability, all of which will yield equivalent results. In this case the control variable is X_1. See the following section for proof of symmetry.
6. Copies of a FORTRAN program that calculates this probability are available from the authors upon request.
7. Even in the 2×2 and $2\times2\times2$ case there are actually two possible orderings, ascending or descending for any given cell. We let the data itself choose between these with our alternative definitions for $P(a_{111} = A_{111}$ or more extreme).
8. See Rosenberg (1968).
9. See Ploch (1974).
10. cf. Pierce (1970:129).
11. *Ibid.*
12. For a 2×2 table, Romney's (1971) coefficient of endogamy is achieved by normalizing the rows and columns of the table, and then computing $(d\text{-}o)/(d+o)$, where d is the number of cases in the diagonal, and o is the number of cases in the off-diagonal.

References

Bartlett, M.S.
1935 Contingency Table Interactions. Journal of the Royal Statistical Society Supp. 2:248-52.

Bishop, Yvonne M., Stephen E. Feinberg, and Paul W. Holland.
1975 Discrete Multivariate Analysis: Theory and Practice. Cambridge, MA: The MIT Press.

Brudner-White, Lilyan A.
1978 Occupational Concomitants of Language Variability in Southern Austrian Bilingual Communities. *In* Advances in the Study of Societal Multilingualism. J.A. Fishman, ed. The Hague: Mouton.

Fisher, R.A.
1954 Statistical Methods for Research Workers. (12th Ed.). New York: Hafner Publishing Co.

Gart, John J.
1972 Interaction Tests for $2 \times s \times t$ Contingency Tables. Biometrica 59:309-316.

Halperin, Max, James H. Ware, David P. Byar, Naltran Mantil, Charles C. Brown, James Koziol, Mitchell Gail, and Sylvan B. Green.
1977 Testing for Interaction in an $I \times J \times K$ Contingency Table. Biometrica 64:271-75.

Lieberman, G.J. and D.B. Owen.
1961 Tables of the Hypergeometric Probability Distribution. Stanford, CA: Stanford University Press.

Murdock, G.P. and D.R. White.
1969 Standard Cross-Cultural Sample. Ethnology 8:329-369.

Patil, Kashinath D.
1974 Interaction Test for Three-Dimensional Contingency Tables. Journal of the American Statistical Association 69:164-68.

Pierce, Albert.
1970 Fundamentals of Nonparametric Statistics. Belmont, California: Dickenson.

Ploch, D.
1974 Ordinal Measures of Association: The General Linear Model. In Measurement in the Social Sciences. H. Blalock, ed. Chicago: Aldine.

Romney, A. Kimball.
1971 Measuring Endogamy. In Explorations in Mathematical Anthropology. Paul Kay, ed. Cambridge, MA: MIT Press.

Rosenberg, Morris
1968 The Logic of Survey Analysis. New York: Basic Books.

Siegel, Sidney.
1956 Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill.

White, Douglas R., Michael L. Burton, and Lilyan A. Brudner.
1977 Entailment Theory and Method: A Cross-Cultural Analysis of the Sexual Division of Labor. Behavioral Science Research 12:1-24.

Zelen, M.
1971 The Analysis of Reversal 2×2 Contingency Tables. Biometrica 58:129-37.