

TECHNICAL INFORMATION**EXACT NONPARAMETRIC STATISTICAL METHODS IN BEHAVIOR ANALYSIS**

Curtis K. Deutsch, Eric Lauer,

UNIVERSITY OF MASSACHUSETTS MEDICAL SCHOOL – SHRIVER CENTER

Nitin Patel and Cyrus Mehta

MASSACHUSETTS INSTITUTE OF TECHNOLOGY, HARVARD SCHOOL OF PUBLIC HEALTH,
AND CYTEL SOFTWARE CORPORATION

In psychology in general and the Experimental Analysis of Human Behavior in particular, we routinely make inferences about behavior that are drawn from a limited number of observations. Statistical methods are designed to aid us in drawing inferences, and sometimes experimental effects are so dramatic that these tests are not even required. Such is the case when behavior analytic methods are so powerful as to create non-overlapping distributions of the behavior under study. When this experimental outcome is not cost-effective, the behavior analyst turns to statistics that can accommodate small-N studies. These are usually nonparametric tests, which make no assumptions about underlying distributions.

Exact Methods

Since the mid-1950s, the classic manual for these methods has been Siegel's *Nonparametric Statistics for the Behavioral Sciences*. This text provided a small number of inferential tests, many of them of limited statistical power. One of these methods that investigators have historically relied on has been the computation of exact probability values. Here, all possible experimental outcomes (permutations) are taken into consideration so the likelihood of the given outcome can be computed. This approach was pioneered by R.A. Fisher (1925), and his Exact Test quickly became a statistical staple in small-N studies.

Fisher's Exact Test was designed for simple 2 X 2 tables. But often an analysis required a higher-dimension table, and the calculations became correspondingly more compute-intensive. Until recently, this limited their use. But with the advent of more computing power --- and importantly, more efficient algorithms for permutational inference --- it was possible to generalize Fisher's approach.

In recent years, a new tool has been developed which may find broad application in behavior analysis. In an effort to develop specialized software to work with smaller sample sizes, a group of Cambridge, MA, statisticians developed *StatXact* (1995). This package accommodates data sets with a modest number of observations, unbalanced cell-counts, and single-subject designs. It deals with the complex issue of data distribution by allowing the user to calculate, usually with near certainty, the exact distribution of the test statistic (Mehta & Patel, 1998).

StatXact's algorithms can ascertain exact p-values for most data under these nonparametric circumstances. But sometimes when the sample sizes are substantial, the limitation of computability is encountered once again. Then it is possible to use an *estimation* of the exact outcome, often by using a Monte Carlo enumeration. This estimates a random subset of all the possible outcomes, very much like "rolling the dice" hundreds of times, as the gambling metaphor implies. Normally, this provides a 99% accurate answer.

Organization

StatXact's documentation provides a summary "road map" that outlines when to use each test, with theory and examples provided. The choices of tests are conditioned on the use of data type (binomial, nominal, ordinal, and continuous data sets), use of both related vs. independent samples (those with paired and random values), and the dimensions of the data table (determined by the numbers of rows and columns in the data table).

The behavior analyst will recognize a number of familiar tests in the menu. For example, StatXact provides Binomial and Chi-Squared tests for one-sample cases, Sign tests for related samples, and Fisher and Mann-Whitney tests for independent samples. Also, there are many newer tests that are tailor-made for small-N studies.

Further, there is a rich assortment of correlational measures. For instance, in the case of measuring associations for ordinal data, it

We appreciate the technical assistance of Emily Saunders. This research was supported by NIH grants P01 HD25995 and P20 HD04147 to W.J. McIlvane (PI).

provides Spearman correlates when assumptions of normal distributions are unknown. In addition, StatXact has contingency coefficients for nominal data, when the user wants to examine the magnitude of an association between variables with differing row and column dimensions. Ordered and nominal data sets can be measured for agreement using Weighted and Cohen's Kappa, respectively.

Worked Example

Consider the comparison of the efficacy of four randomly assigned teaching methods for a given task, each administered to a small sample of subjects ($N = 6$ or 7). Responses (Yes or No) signify whether or not a subject met criterion for learning a task. Using StatXact, the following table was produced:

Outcome	Teaching Method				TOTAL
	A	B	C	D	
Yes	7	2	3	2	14
No	0	4	4	4	12
TOTAL	7	6	7	6	26

The data are entered directly into an interactive online environment, without the need for a pre-existing database. By running an exact Pearson's Chi-Square test, the above findings were found to be significant using the nonparametric exact method. Below, is the output of the program for a Chi-Square Test for Independence:

Statistic based on the observed 2 by 4 table(x):

$CH(X)$: Pearson Chi-Square Statistic = 8.372

Exact p-value and point probability:

$Pr \{ CH(X) .GE. 8.372 \} = 0.0365$

$Pr \{ CH(X) .EQ. 8.372 \} = 0.0016$

This means that the precise experimental outcome was ($p=.0016$). It is then necessary to estimate this probably and the likelihood of all the *less likely* outcomes; these turn out to be $p=.0365$, well within the cut-off for statistical significance.

This table yielded a significant difference for the whole outcome of controlled methods. How about a particular contrast, such as Method A versus Method B? An additional comparison of the first two methods, noted in the two-by-two table below, yielded the following findings using a Fisher's Exact Test:

Outcome	Method A	Method B	TOTAL
Yes	7	2	9
No	0	4	4
TOTAL	7	6	13

Here is the output for the StatXact software for a Fisher's Exact Test (2 X 2 table):

Exact p-value and point probabilities:

Two-sided:

$Pr \{ FI(X) .GE. 6.264 \} = Pr \{ P(X) .LE. 0.0210 \} = 0.0210$

$Pr \{ FI(X) .EQ. 6.264 \} = Pr \{ P(X) .EQ. 0.0210 \} = 0.0210$

One-sided: Let y be the value in Row 1 and Column 1:

$y = 7 \min(Y) = 3 \max(Y) = 7 \text{ mean}(Y) = 4.846 \text{ std}(Y) = 0.8635$

$Pr \{ Y .GE. 7 \} = 0.0210$

$Pr \{ Y .EQ. 7 \} = 0.0210$

As seen above, the two methods were found to have significantly different outcomes at the $p < .02$ level. Moreover, it shows that there is a 21-in-1000 chance of obtaining cell entries like this if there were no row-by-column interactions (no relationship between independent and dependent variables).

Summary

The contingency tables described above are examples of cases in which exact inference is necessary for statistical analysis. There is only a small N in this sample, with a limited number of observations. In these cases, nonparametric methods should be applied.

Many larger software programs are unable to compute more than single two-by-two exact nonparametric tables. For example, SAS (1995) is able to compute Fischer's exact test; however, there are many scenarios in which it cannot find exact p-values; StatXact can, and it makes both exact and Monte Carlo calculations available in the same package. In a more limited set of exact tests, SPSS (1995) also provides software that supports these methods.

Information regarding StatXact can be found at www.cytel.com, which discusses the recent release of the newest version, StatXact 4. This site includes reviews, documentation, workshops, and technical papers. Also, it is expected that demos of StatXact 5 will be available at www.cytel.com within the Fall of 2001. An additional website devoted to the discussion of exact methodology can be found at Exact-Stats: www.mailbase.ac.uk/lists-a-e/exact-stats.

For other related inquiries, the SAS and SPSS websites are located at: www.sas.com and www.spss.com, respectively.

References

- Fischer, R.A. (1925). *Statistical Methods for Research Workers*, Edinburgh: Oliver & Boyd.
- Mehta, C. & Patel, N. (1998). Exact Inference for Categorical Data. In: Armitage, P & Coltin, T (Eds.). *Encyclopedia of Biostatistics* (Vols.1-6). NY: John Wiley.
- SAS Institute Inc. (1995). *SAS/Stat User's Guide, Version 6.11*. Cary, NC: The SAS Institute.

- Siegel, S (1956). *Nonparametric Statistics for the Behavioral Sciences*, NY: McGraw-Hill.
- SPSS *Exact Tests for Windows* (1995). Chicago: SPSS Inc.
- StatXact-3 for Windows (1995). *Software for Exact Nonparametric Inference*. Cambridge, MA: Cytel Software Corporation.