A program of a recursive algorithm for a summed multinomial density function

RAYMOND K. FINK

School of Aeronautics and Astronautics
Purdue University, West Lafayette, Indiana 47907

and

HERBERT MOSKOWITZ

Krannert School of Industrial Administration
Purdue University, West Lafayette, Indiana 47907

In many statistical decision problems, one must account for the inaccurate reporting of experimental outcomes. This requires that the matrix $p_n(Y|X)$ (the probabilities of the adjudged or reported experimental outcomes given the true experimental outcomes of an experiment) be determined for each of the n possible experiments considered in the problem. For example, in inspector error models in quality control (auditing) having Cartesian product structures, it is necessary to compute $p_n(Y|X)$ where Y and X are the reported and the true number of defective (noncompliance) items respectively in the sample (audit), in order to determine an optimal economic sampling (auditing) plan (Collins, Note 1).

Let us view an experiment as consisting of n independent trials, each trial having four mutually exclusive and exhaustive outcomes: (1) report success when a true success; (2) report failure when a true success; (3) report success when a true failure; (4) report failure when a true failure. If we can also assume that the probability of a given element remains constant for each trial, each element of the matrix $p_n(Y|X)$ can be computed by means of a summed multinomial density function (Fink & Moskowitz, Note 2).

However, the computational complexity of this density function and the fact that most applications (e.g., algorithms for determining optimal quality control inspection [auditing] plans) examine successively increasing sample sizes, highlight the value of a recursive formulation of $p_n(Y|X)$ (Moskowitz & Fink, Note 3). The program described below is a recursive algorithm for computing the $p_n(Y|X)$ Markov array, thus circumventing the use of the computationally onerous summed multinomial density function. The efficiency of the algorithm is exemplified by the fact that it takes about 60 sec to generate the arrays for all sample sizes from $n = 1$ to 150 with the algorithm, while a summed multinomial solution requires 348 sec just to compute the $p_n(Y|X)$ matrix for $n = 150$ only.

Description. The basis for the recursive algorithm is as follows. If the experiment is composed of n independent trials, then the probabilities of outcomes for n trials can be determined by the probabilities of outcomes for $(n - 1)$ trials and the probabilities of outcomes on the n -th trial. Since we assume that trial probabilities are identically distributed $p_n(Y|X)$ can be calculated from the transition equation

$$\begin{aligned} p_n(Y = j|X = k) = & p_{n-1}(Y = j|X = k) \cdot p_1(Y = 0|X = 0) \\ & + p_{n-1}(Y = j|X = k - 1) \cdot p_1(Y = 0|X = 1) \\ & + p_{n-1}(Y = j - 1|X = k) \cdot p_1(Y = 1|X = 0) \\ & + p_{n-1}(Y = j - 1|X = k - 1) \cdot p_1(Y = 1|X = 1) \end{aligned} \quad (1)$$

$$0 \leq j \leq n \quad 0 \leq k \leq n$$

This work was completed while the second author was a Guest Professor at the Sonderforschungsbereich, Sozial-und Wirtschaftspychologische Entscheidungsforschung, Der Universität Mannheim (WH), Mannheim, West Germany.

with appropriate boundary conditions for $j = 0, j = n, k = 0$, or $k = n$.

The algorithm can also be used as a subroutine in conjunction with other computational algorithms for determining optimal economic sampling (compliance) plans with inspector (auditor) error in quality control (auditing) problems (Moskowitz & Fink, Note 3). While the algorithm, as a subroutine in such applications, would principally be called to compute $p_n(Y|X)$ for successively increasing n 's, it could also be used to compute $p_n(Y|X)$ for any specified n simply by calling it n times.

Input. Inputs to the program are: (1) n , the sample size, and (2) the $p_1(Y|X)$ array ("ONE") — this is a 2×2 array for $n = 1$, whose rows must sum to one and whose columns are the reports and rows are the true experimental outcomes. If used as a subroutine for computing an optimal sampling plan, then the $n \times n$ $p_{n-1}(Y|X)$ array ("PYX") serves as an input to obtain $p_n(Y|X)$.

Output. The output of the program is either $p_n(Y|X)$ or $p_n(Y|X) \forall n = 1, \dots, n$.

Computer and language. The program is written in FORTRAN IV for the CDC 6500 computer. Conversion to other computing systems is straightforward.

Restrictions. In its present form the recursive algorithm assumes constant, identically distributed trial probabilities. In the context of inspector error models, this implies constant inspector error independent of sample size. However, more sophisticated models of inspector error are easily accommodated by the algorithm.

For example, if trial probabilities remain constant in each trial (i.e., inspector error remains constant for each item inspected in the sample) for a given sample size n , but changes with increasing n , we can reflect this by exogeneously adjusting $p_1(Y|X)$ for each sample size n , and setting it to $p_1^n(Y|X)$, the inspector error probabilities for a single sample out of a sample size n . For each n , we must then use the recursive algorithm to regenerate $p_1^n(Y|X)$, $p_2^n(Y|X), \dots, p_n^n(Y|X)$; while this slows things down considerably, it is still much faster than the summed multinomial approach.

A computationally simpler and more realistic inspector error model is one in which the trials are not identically distributed, but change with sample size such that $p_k^n(Y|X)$, the error probabilities at the k -th trial, are constant for any sample size n , but $p_k^n(Y|X)$ and $p_{k+1}^n(Y|X)$ are not identical. This may reflect the effects of fatigue on the inspector's performance; in the transition equation (1), this implies variable step-transition probabilities. For this case, we can simply adjust $p_1^n(Y|X)$ as necessary, and calculate $p_{k+1}^n(Y|X)$ from $p_k^n(Y|X)$ with virtually no increase in computation time. The program is limited to the dimension $PYX + 1 (= n + 1)$.

Availability. A more detailed description of the program and algorithm as well as a program listing may be obtained free by writing Herbert Moskowitz, Krannert Graduate School of Industrial Administration, Purdue University, West Lafayette, Indiana 47907.

REFERENCE NOTES

1. Collins, R. D. Statistically and economically based attribute acceptance sampling models with inspector errors. PhD dissertation in industrial engineering and operations research, Virginia Polytechnic Institute and State University, May, 1974.
2. Fink, R. K., & Moskowitz, H. A recursive algorithm for a summed multinomial density function. Institute for Research in the Behavioral, Economic, and Management Sciences, Krannert Graduate School of Industrial Administration, Purdue University, West Lafayette, Indiana, No. 510, May, 1975.
3. Moskowitz, H., & Fink, R. K. A Bayesian algorithm incorporating inspector errors for quality control and auditing. Institute for Research in the Behavioral, Economic and Management Sciences, Krannert Graduate School of Industrial Administration, Purdue University, West Lafayette, Indiana, No. 524, August, 1975.