The Annals of Statistics
1977, Vol. 5, No. 1, 235–236

A NOTE ON INVERSE SAMPLING PROCEDURES FOR SELECTING THE BEST BINOMIAL POPULATION

BY DONALD A. BERRY AND D. H. YOUNG

University of Minnesota and Brunel University

Two inverse sampling procedures, one that uses the classical vector-at-a-time observation rule and another that uses the play-the-winner observation rule, are shown to select the best of k binomial populations with the same probability, independent of the probabilities of success. This shows that the play-the-winner rule is better from the point of view that both the sample size and number of failures of each population are stochastically smaller using play-the-winner than vector-at-a-time.

1. The problem. Given k independent binomial populations $\pi_1, \cdots, \pi_k$ we wish to select the "best" population, the one with the largest probability of success. Denote the ordered success probabilities by $p_{[1]} \leqq \cdots \leqq p_{[k]}$. Sobel and Weiss (1972) considered this problem using the so-called "indifference zone" approach; they compared procedures which have a probability of correct selection (PCS) greater than or equal to P^* whenever $p_{[k]} - p_{[k-1]} \geqq \Delta^*$, where P^* and Δ^* are preassigned constants. The two procedures which most interested them are called R_{VT} and R_{PW}.

Under the play-the-winner procedure R_{PW} the populations are placed in random order at the outset; relabel the populations so that the order is $\pi_1, \cdots, \pi_k$. Observations are taken on π_1 until a failure is obtained, then observations are taken on π_2 until another failure is obtained, and so on, always setting aside a population whenever it yields a failure. Observation ceases whenever one of the populations yields r successes and that population is selected as being best. The value of r depends on (P^*, Δ^*) and, quite naturally, is increasing in P^* and decreasing in Δ^*. Should the cycle $(\pi_1, \cdots, \pi_k)$ be completed with each population yielding a failure, the cycle is repeated until one population yields a total of r successes.

Under the vector-at-a-time procedure R_{VT} one observation per population is taken until at least one of the populations yields s successes, so that the sample size is the same for every population. Again, s depends on (P^*, Δ^*). If more than one population yields s successes at the same time, one of them is selected at random.

Sobel and Weiss show that for the least favorable configuration $p_{[1]} = \cdots = p_{[k-1]} = p_{[k]} - \Delta^*$, the two procedures have the same PCS when $r = s$ and hence that they require exactly the same selection constant defining the stopping rule. In Section 2 of this note we generalize this result, showing that when $r = s$ the two procedures have the same PCS for all $p_1, \cdots, p_k$. This result is important

Received January 1975; revised June 1976.

AMS 1970 *subject classification.* Primary 62F07.

Key words and phrases. Binomial selection, play-the-winner sampling, vector-at-a-time sampling, inverse sampling.

as it shows that the selection constants for the two procedures will be equal whatever formulation is adopted for the required probability of correct selection. For example, in many situations it would be more meaningful for the experimenter to base his probability requirement for correct selection on the ratios of the single trial probabilities of success rather than their differences.

Our method of proof also shows that for any $p_1, \cdots, p_k$ and r, the sample size on each population is stochastically smaller using R_{PW} than using R_{VT}. This result is stronger than the corresponding one for expected sample sizes or expected total sample size which Sobel and Weiss showed to hold asymptotically ($r \to \infty$).

2. Proof that R_{PW} and R_{VT} have same PCS. Consider a complete experiment accomplished using R_{PW}. Use the experimental results to generate a partial matrix of 0's and 1's in the following fashion. Row i gives the results on π_i in order from first observation to last; "0" means failure, "1" means success. Suppose that population π_{i^*} is selected by R_{PW} and that the numbers of observations and failures on π_i are n_i and f_i for $i = 1, \cdots, k$. Letting $n = n_{i^*}$, it follows that $n - f_{i^*} = r$, $n > n_i$ for $i \neq i^*$, and $f_1 = \cdots = f_{i^*-1} = f_{i^*} + 1 = \cdots = f_k + 1$. Now suppose that R_{VT} had been used instead with $s = r$, so that the partial matrix is filled out to have k rows and n columns. If $i \in \{1, \cdots, i^* - 1\}$ then $n - f_i < r$ and π_i cannot be selected at or before the nth stage using R_{VT} even if π_i fills out its row with 1's. If $i \in \{i^* + 1, \cdots, k\}$ then $n - f_i = r$ and π_i cannot be selected before the nth stage—it can be selected at the nth only if it fills out its row with 1's. Consider the class $\{\pi_{i_1}, \cdots, \pi_{i_m}\}$ of populations which have r successes in the completed experiment where $i_1 < \cdots < i_m$. Since the populations were randomly ordered initially, selecting π_{i_1} is equivalent to randomizing among this class. In view of the previous discussion, $i_1 = i^*$, and the procedures select the same population for every experimental matrix. The probability that a particular population becomes the first to yield the rth success and does so at the nth stage is independent of whether the matrix is filled out (R_{VT}) or not (R_{PW}). Therefore the procedures (with $r = s$) have the same PCS for any $p_1, \cdots, p_k$.

The method of proof used here is not directly generalizable to arbitrary selection problems, but it can apply to a wider class of procedures than the one considered. The procedures dealt with here were considered because of their importance in the literature.

REFERENCE

SOBEL, M. and WEISS, G. H. (1972). Play-the-winner and inverse sampling for selecting the best of $k \geqq 3$ binomial populations. *Ann. Math. Statist.* **43** 1808–1826.

DEPARTMENT OF THEORETICAL STATISTICS
SCHOOL OF STATISTICS
UNIVERSITY OF MINNESOTA
270 VINCENT HALL
MINNEAPOLIS, MINNESOTA 55455