A Sequential Test for Two Binomial Populations

(statistical/clinical trials)

HERBERT ROBBINS

Columbia University, New York, N.Y. 10027; and Brookhaven National Laboratory, Upton, New York 11973

Contributed by Herbert Robbins, September 9, 1974

ABSTRACT A sequential procedure is proposed for deciding which of two sequences of Bernoulli trials has the greater success probability. The procedure covers cases in which the two sequences are being observed at different rates, whether by design or otherwise, as well as the case of pairwise observation. The error probability of the procedure depends approximately only on the odds ratio of the two probabilities, a desirable property in deciding which of two medical treatments has the greater probability of success.

Let $x_1, \dots$ be independent observations, with $P(x_i = 1) = p_1$, $P(x_i = 0) = q_1 = 1 - p_1$, and let $y_1, \dots$ be independent and independent of the x 's, with $P(y_i = 1) = p_2$, $P(y_i = 0) = q_2 = 1 - p_2$. We wish at some stage to terminate our observation of the two sequences and assert either that $p_1 > p_2$ or that $p_2 > p_1$. (The possibility of stopping without a decision is excluded.) The procedure that we propose is the following. For any positive integers m and n let

$$u_m = x_1 + \dots + x_m, v_n = y_1 + \dots + y_n, z_{m,n} = 2(nu_m - mv_n)/(m + n).$$

Stop with $(M, N) = \text{first } (m, n) \text{ such that } |z_{m,n}| \geq B$, where B is some preassigned positive constant, and assert that $p_1 > p_2$ or $p_2 > p_1$ according as $z_{M,N} \geq B$ or $z_{M,N} \leq -B$.

If $\lambda = p_1q_2/p_2q_1$ is the odds ratio, then when $p_1 > p_2$, so that $\lambda > 1$, we shall show that the error probability, $P(\text{assert } p_2 > p_1)$, is

$$P(\text{error}) \lesssim 1/(\lambda^B + 1) \quad [1]$$

while the terminal sample sizes M, N are such that

$$E\left(\frac{2MN}{M+N}\right) \gtrsim \frac{B(\lambda^B - 1)}{|p_1 - p_2|(\lambda^B + 1)} \left(\simeq \frac{B^2}{2pq} \text{ if } p_1 = p_2 = p\right). \quad [2]$$

(When $p_2 > p_1$, [1] and [2] hold with λ replaced by $1/\lambda$.)

In the case of pairwise observation, with $m = n$ at each stage, and hence $M = N$, the formulas [1] and [2] are exact when B is a positive integer; here $z_{n,n} = u_n - v_n$ and $2MN/(M + N) = N = \text{number of pairs } (x_i, y_i) \text{ observed before termination}$. This case was treated by A. Wald (1), and the present procedure may be regarded as a generalization of his, although even in the pairwise case it is conceptually somewhat different.

In what follows we shall first prove that formulas [1] and [2] are exact in the pairwise case when B is a positive integer, and then give a heuristic argument for their approximate validity when $m \neq n$, but the sampling rule that determines

whether the next observation is to be an x or a y is symmetric in the two sequences.

The pairwise case

The probability of observing any particular sequence $x_1, \dots, x_n; y_1, \dots, y_n$ of $2n$ 0's and 1's is

$$f_{p_1, p_2}(x_1, \dots, x_n; y_1, \dots, y_n) = p_1^{u_n} q_1^{n-u_n} p_2^{v_n} q_2^{n-v_n};$$

the likelihood ratio obtained by interchanging p_1 and p_2 is

$$f_{p_1, p_2}/f_{p_2, p_1} = \lambda^{u_n - v_n}.$$

Hence, when B is a positive integer and $\lambda > 1$, the error probability is

$$\begin{aligned} P(\text{error}) &= P_{p_1, p_2}(\text{assert } p_2 > p_1) = P_{p_1, p_2}(u_N - v_N = -B) \\ &= \sum_{n=1}^{\infty} \sum_{\{N=n, u_n - v_n = -B\}} f_{p_1, p_2} \\ &= \sum_{n=1}^{\infty} \sum_{\{N=n, u_n - v_n = -B\}} \lambda^{u_n - v_n} f_{p_2, p_1} \\ &= \lambda^{-B} P_{p_2, p_1}(\text{assert } p_2 > p_1) \\ &= \lambda^{-B} P_{p_1, p_2}(\text{assert } p_1 > p_2) \\ &= \lambda^{-B}[1 - P(\text{error})], \end{aligned}$$

so that [1] holds exactly. Now by Wald's lemma and [1]

$$\begin{aligned} E(u_N - v_N) &= (p_1 - p_2)E(N) \\ &= B[(1 - P(\text{error})) - BP(\text{error})], \end{aligned}$$

so that [2] holds exactly when $p_1 > p_2$. The case $p_1 = p_2$ follows from the fact that then $E(u_N - v_N)^2 = B^2 = 2pqE(N)$ by Wald's lemma for second moments; alternatively, we can let $p_2 \rightarrow p_1$ and use L'Hospital's rule.

The non-pairwise case

The probability of observing any particular sequence $x_1, \dots, x_m; y_1, \dots, y_n$ of $m + n$ 0's and 1's is

$$f_{p_1, p_2}(x_1, \dots, x_m; y_1, \dots, y_n) = p_1^{u_m} q_1^{m-u_m} p_2^{v_n} q_2^{n-v_n};$$

the likelihood ratio obtained by interchanging p_1 and p_2 is

$$f_{p_1, p_2}/f_{p_2, p_1} = \lambda^{u_m - v_n} (q_2/q_1)^{n-m}. \quad [3]$$

Putting $p_1 = p + \epsilon$, $p_2 = p - \epsilon$, if $\epsilon > 0$ is small compared to p and q then

$$\begin{aligned} \lambda &= \left(1 + \frac{\epsilon}{p}\right) \left(1 + \frac{\epsilon}{q}\right) / \left(1 - \frac{\epsilon}{p}\right) \left(1 - \frac{\epsilon}{q}\right) \\ &\simeq e^{2\epsilon[(1/p) + (1/q)]} = e^{2\epsilon/pq} \end{aligned}$$

$$q_2/q_1 = \left(1 + \frac{\epsilon}{q}\right) / \left(1 - \frac{\epsilon}{q}\right) \simeq e^{2\epsilon/q},$$

so that by Eq. [3]

$$f_{p_1, p_2} / f_{p_1, p_1} \simeq e^{(2\epsilon/pq)[u_m - v_n + p(n-m)]} \simeq \lambda^{u_m - v_n + p(n-m)}. \quad [4]$$

Now

$$E\left(\frac{u_m + v_n}{m + n}\right) = \frac{mp_1 + np_2}{m + n} = p + \frac{(m - n)\epsilon}{m + n}$$

$$\text{Var}\left(\frac{u_m + v_n}{m + n}\right) = \frac{mp_1q_1 + np_2q_2}{(m + n)^2} \leq \frac{1}{4(m + n)},$$

so that with high probability when $m + n$ is large

$$\frac{u_m + v_n}{m + n} \simeq p + \frac{(m - n)\epsilon}{m + n},$$

and hence

$$p \simeq \frac{u_m + v_n + (n - m)\epsilon}{m + n}. \quad [5]$$

Substituting this estimate of p into the exponent of λ into [4] gives

$$u_m - v_n + p(n - m) \simeq \frac{(m + n)(u_m - v_n) + (n - m)[u_m + v_n + (n - m)\epsilon]}{m + n}$$

$$= \frac{2(nu_m - mv_n)}{m + n} + \frac{(n - m)^2\epsilon}{m + n} = z_{m,n} + \frac{(n - m)^2\epsilon}{m + n},$$

so that by [4] with $\epsilon > 0$ and $\lambda > 1$

$$\frac{f_{p_1, p_2}}{f_{p_1, p_1}} \simeq \lambda^{z_{m,n} + (n - m)^2\epsilon/(m + n)} \geq \lambda^{z_{m,n}}.$$

Hence when $\lambda > 1$ the probability of correct decision is

$$P_{p_1, p_2}(\text{assert } p_1 > p_2) = \sum_{m, n=1}^{\infty} \left\{ (M, N) = (m, n); z_{m,n} \geq B \right\} f_{p_1, p_2}$$

$$\gtrsim \sum_{m, n=1}^{\infty} \left\{ (M, N) = (m, n), z_{m,n} \geq B \right\} \lambda^{z_{m,n}} f_{p_1, p_1}$$

$$\geq \lambda^B P_{p_2, p_1}(\text{assert } p_1 > p_2) = \lambda^B P_{p_1, p_2}(\text{assert } p_2 > p_1)$$

$$= \lambda^B [1 - P_{p_1, p_2}(\text{assert } p_1 > p_2)],$$

so that

$$P_{p_1, p_2}(\text{assert } p_1 > p_2) \gtrsim \frac{\lambda^B}{\lambda^B + 1},$$

and hence

$$P(\text{error}) = 1 - P_{p_1, p_2}(\text{assert } p_1 > p_2) \lesssim \frac{1}{\lambda^B + 1}.$$

This provides a heuristic justification of [1] when p_1, p_2 and B are such that the approximations made above are valid. Since we have made no assumptions about the relation between m and n as sampling proceeds, other than that both become infinite (to assure termination with probability 1), it would be difficult to be more precise about the accuracy of the approximation [1] in all cases. We may hope, at least, that when it is not a good approximation it is because $P(\text{error})$ is too small to be of any practical significance.

To justify [2], we have by neglecting overshoot and using [1]

$$E(z_{M,N}) \simeq B[1 - P(\text{error})] - BP(\text{error}) \simeq \frac{B(\lambda^B - 1)}{\lambda^B + 1}. \quad [6]$$

But since $u_M \simeq M p_1$ and $v_N \simeq N p_2$ with high probability, we have by a nontrivial act of faith the approximation

$$E(z_{M,N}) \simeq 2E\left(\frac{NMp_1 - MNp_2}{M + N}\right) = (p_1 - p_2)E\left(\frac{2MN}{M + N}\right),$$

which together with [6] gives [2] for $p_1 > p_2$.

In the nonpairwise case many questions remain to be answered by Monte Carlo simulation for various p_1, p_2 pairs and by purely analytic means, e.g., in the form of asymptotic results as $\lambda \rightarrow 1$ or $B \rightarrow \infty$. Even more interesting are the allocation problems in the case of clinical trials where the objective is to minimize the expected value of, say, the number of patients given the inferior treatment before a decision is reached. The method of estimating a nuisance parameter p as in [5] works also in other two-population decision problems, including the normal case that was treated by more exact methods in ref. 2.

This research was supported by PHS Grant 5-RO1-GM-16895-06.

1. Wald, A. (1947) in *Sequential Analysis* (John Wiley and Sons, New York), pp. 106-116.
2. Robbins, H. & Siegmund, D. O. (1974) "Sequential tests involving two populations," *J. Amer. Stat. Ass.* **69**, 132-139.