8250 SP24

Appl. sci. Res. Section B, Vol. 3

SEQUENTIAL TEST WITH THREE POSSIBLE DECISIONS FOR TESTING AN UNKNOWN PROBABILITY *)

by J. DE BOER

Mathematical Centre, Amsterdam, Netherlands

Summary

A three-decisions sequential test developed by Sobel and Wald, is applied for testing an unknown probability. Some simplifications for this case are introduced and the test is compared with another test for the same hypothesis, having fixed sample size.

§ 1. *Introduction.* M. Sobel and A. Wald [1]) developed a sequential test for the following problem: Dividing the real axis by two points into three intervals, we may set up three hypotheses, each stating that the unknown mean of a normal distribution belongs to one of the intervals.

The method of Sobel and Wald enables one to test these hypotheses, i.e. to decide, with a known confidence level, in which one of the three intervals the mean is lying. Its use is not restricted to the case of testing the mean of a normal distribution (although Sobel and Wald especially considered this problem), but it may be modified for testing any unkown parameter of a distribution function of given form, as is done in this paper for the binomial case, i.e. for testing an unknown probability. In that case some simplifications of the procedure are introduced and the test, truncated or not, is compared, by means of numerical examples, with a classical test for the same hypotheses, viz. a test with fixed sample size.

§ 2. *Mathematical model.* An experiment will be considered with two possible outcomes A and B, their unknown probabilities being

*) Report SP 24 of the Statistical Department of the Mathematical Centre, Amsterdam.

p and $q = 1 - p$ respectively. The test is based upon a number of mutually independent experiments, this number being random.

Two numbers a_{12} and a_{32} between 0 and 1 are given (or chosen in connection with the particular problem) and the hypotheses H_1, H_2 and H_3 are formulated as follows:

$$H_1 : p < a_{12}, \quad H_2 : a_{12} \leq p \leq a_{32}, \quad \mathrm{H}_3 : p > a_{32}.$$

Further, for the performance of the test, two points p_1 and p_2 must be chosen on both sides of a_{12}, and two points p_2'' and p_3 on both sides of a_{32}, satisfying the relations

$$0 < p_1 < a_{12} < p_2' < p_2'' < a_{32} < p_3 < 1.$$

The situation is sketched in fig. 1.

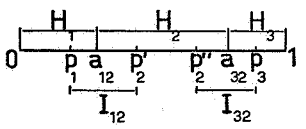

Fig. 1. Partition of the parameter space.

The following definition of the concepts "correct" and "incorrect decision" is introduced:

TABLE I

Correct and incorrect decisions		
true value of p	correct decision	incorrect decisions
$p \leq p_1$	Accept H_1	Accept H_2 or H_3
$p_1 < p < p_2'$	,, H_1 or H_2	,, H_3
$p_2' \leq p \leq p_2''$	,, H_2	,, H_1 or H_3
$p_2'' < p < p_3$	,, H_2 or H_3	,, H_1
$p \geq p_3$	,, H_3	,, H_1 or H_2

The intervals denoted by I_{12} and I_{32} in fig. 1, are called zones of indifference, in connection with table I. The numbers a_{12} and a_{32} are not further of importance for the performance of the test.

§ 3. *Description of the test.* The path, corresponding to the consecutive observations, may be traced out in a rectangular point-lattice with a rectangular coordinate system. This is done, beginning from the origin O, by making one step in vertical direction at each observation with result A and one step in horizontal direction at each observation with result B. In this plane, acceptance regions G_1, G_2 and G_3 are constructed for H_1, H_2 and H_3 respectively, with the two decisions-sequential test of W a l d [2]) (cf. fig. 2).

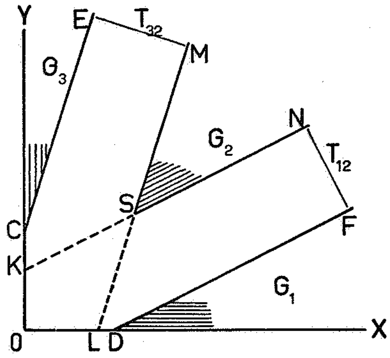

Fig. 2. Scheme of a sequential test T with three possible decisions.

In fig. 2, DF and KN are the two lines of a common sequential test T_{12} for testing the hypothesis $p \leq p_1$ against $p \geq p_2'$. In the same way, CE and LM are the lines of the Wald-test T_{32} for testing $p \leq p_2''$ against $p \geq p_3$. Thus passing *) these lines has the following meaning:

Passing DF leads to acceptance of $p \leq p_1$ according to T_{12},
,, KN ,, ,, ,, ,, $p \geq p_2'$,, ,, T_{12},
,, LM ,, ,, ,, ,, $p \leq p_2''$,, ,, T_{32},
,, CE ,, ,, ,, ,, $p \geq p_3$,, ,, T_{32}.

It is useful to adopt the following notations:

α_{12} = prob. of acceptance of $p \geq p_2'$ in T_{12}, if $p = p_1$,
α_{21} = ,, ,, ,, ,, $p \leq p_1$,, T_{12}, ,, $p = p_2'$,
α_{23} = ,, ,, ,, ,, $p \geq p_3$,, T_{32}, ,, $p = p_2''$,
α_{32} = ,, ,, ,, ,, $p \leq p_2''$,, T_{32}, ,, $p = p_3$.

Now consider the relations

$$x_L \leq x_D \text{ and } y_K \leq y_C. \tag{1}$$

If these inequalities are not satisfied, there exists e.g. the possibility of passing DF and then CE, without passing LS, which would lead to

*) By passing we mean crossing or reaching.

acceptance of $p \leq p_1$ in T_{12} and of $p \geq p_3$ in T_{32}. It is clear that it is desirable to avoid such possibilities and thus to satisfy the inequalities (1). From the equations of the four lines in fig. 2 (cf. [2]), p. 90 ff.), it follows that, in order that the relations (1) be satisfied, the following inequalities are sufficient:

$$\frac{\log \dfrac{\alpha_{21}}{1-\alpha_{12}}}{\log \dfrac{q_2'}{q_1}} \geq \frac{\log \dfrac{1-\alpha_{23}}{\alpha_{32}}}{\log \dfrac{q_2''}{q_3}} \quad \text{and} \quad \frac{\log \dfrac{\alpha_{23}}{1-\alpha_{32}}}{\log \dfrac{p_2''}{p_3}} \geq \frac{\log \dfrac{1-\alpha_{21}}{\alpha_{12}}}{\log \dfrac{p_2'}{p_1}}, \qquad (2)$$

So these are the restrictions to be imposed. Then there are the following possibilities of the path of observations, with corresponding decisions:

TABLE II

Possible paths and corresponding decisions of the test T

If we pass	the decision is
LS and then DF	accept H_1
LS ,, ,, SN	,, H_2
LS ,, ,, KS	,, H_2
KS ,, ,, LS	,, H_2
KS ,, ,, SM	,, H_2
KS ,, ,, CE	,, H_3

§ 4. *Properties of the test.* According to table I, the probability of an incorrect decision (the operating characteristic of the test T) is a function of p, which we shall denote by $\gamma(p)$. Its maximum α over the interval [0.1] is called the true level of significance of test T. Denoting the probability of acceptance of the hypothesis H_i in test T_{i2} ($i = 1,3$), when the parameter value is p, by $\beta(p)$ we may set up the following table, derived from table I and the definition of α_{ij}:

TABLE III

Relations between γ and β_i ($i = 1,3$)

for	relation
$p \leq p_1$	$\gamma = 1 - \beta_1$
$p = p_1$	$\gamma = \alpha_{12}$
$p_1 < p < p_2'$	$\gamma = \beta_3$
$p_2' \leq p \leq p_2''$	$\gamma = \beta_1 + \beta_3$
$p_2'' < p < p_3$	$\gamma = \beta_1$
$p = p_3$	$\gamma = \alpha_{32}$
$p \geq p_3$	$\gamma = 1 - \beta_3$

From the definition of α and the shape of $\beta_i(p)$ $(i = 1,3)$ (cf. [2]), p. 51), it thus follows that

$$\alpha = \operatorname{Max} \{\alpha_{12}, \beta_1 (p_2') + \beta_3(p_2'), \beta_1(p_2'') + \beta_3(p_2''), \alpha_{32}\}.$$

Moreover we have the inequality

$$\operatorname{Max} \{\beta_1(p_2') + \beta_3(p_2'), \beta_1(p_2'') + \beta_3(p_2'')\} < \alpha_{21} + \alpha_{23}$$

for $\quad \beta_3(p_2') < \beta_3(p_2'') = \alpha_{23} \quad$ and $\quad \beta_1(p_2'') < \beta_1(p_2') = \alpha_{21}.$

Now for fixed α, an obvious choice for the α_{ij} is:

$$\alpha_{12} = \alpha_{32} = \alpha_{21} + \alpha_{23} = \alpha,$$

because then it is certain that α is reached, but not exceeded. In the symmetric case $(p_1 + p_3 = p_2' + p_2'' = 1)$, which is used later for comparison of T with a classical test, one takes moreover

$$\alpha_{21} = \alpha_{23} = \tfrac{1}{2}\alpha.$$

It is easy to prove that then the relations (2) are satisfied.

The $\beta_i(p)$ $(i = 1,3)$ are computed with the formulae given in [2]), p. 50 ff. Then, with table III, γ can be determined and thus α too.

The mathematical expectation of the number $\boldsymbol{n}$ *) of observations, needed until a decision in test T is reached, when the parameter value is p, is denoted by $\mathcal{E}(\boldsymbol{n}/p)$. Analogously $\mathcal{E}(\boldsymbol{n}_{12}/p)$ for T_{12}, and $\mathcal{E}(\boldsymbol{n}_{32}/p)$ for T_{32}. To get a rough sketch of $\mathcal{E}(\boldsymbol{n}/p)$, the following approximations are used:

a) For $p < p_2'$: $\mathcal{E}(\boldsymbol{n}/p) \approx \mathcal{E}(\boldsymbol{n}_{12}/p)$,

b) For $p > p_2''$: $\mathcal{E}(\boldsymbol{n}/p) \approx \mathcal{E}(\boldsymbol{n}/_{32}p)$.

It is evident that $\mathcal{E}(\boldsymbol{n}/p)$ possesses in (p_2', p_2'') a local minimum. Thus we get a function of the type of fig. 3.

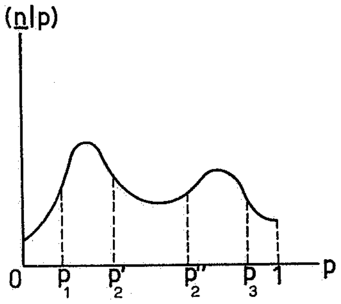

Fig. 3. Sketch of $\mathcal{E}(\boldsymbol{n}/p)$.

*) Random variables are indicated by bold type symbols; the same symbols, not boldly printed, being used to denote values assumed by these random variables.

For an exact formula for $\mathcal{E}(\boldsymbol{n}/p)$, one may proceed as follows: Let $\boldsymbol{n}_{12}$ and $\boldsymbol{n}_{32}$ be the number of observations needed to arrive at a decision in T_{12} and T_{32} respectively, whilst $\boldsymbol{n}^*$ is the number of steps needed to pass the boundary KSL (cf. fig. 2) for the first time. Two types of paths are distinguished:

1°. Paths terminating at one of the boundaries of T_{12}, denoted by $w_{12}^{(j)}$, where $j = 1, 2, \ldots.$ (the number of these paths is enumerable). All these paths pass LS. Let $n_{12}^{(j)}$ be the total number of steps of the path $w_{12}^{(j)}$ and $m_{12}^{(j)}$ the number of steps until the first point of intersection with LS.

2°. Paths terminating at one of the boundaries of T_{32}, denoted by $w_{32}^{(j)}$ $(j = 1, 2, \ldots)$. All these paths pass KS. Let $n_{32}^{(j)}$ be the total number of steps, and $m_{32}^{(j)}$ the number until the first point of intersection with KS.

Indicating the probability of a path w, at given p, by $\mathcal{P}(w/p)$, the following relations hold:

$$\mathcal{E}(\boldsymbol{n}/p) = \sum_i P(w_{12}^{(i)}/p) \,.\, n_{12}^{(i)} + \sum_j P(w_{32}^{(j)}/p) \,.\, n_{32}^{(j)},$$

$$\mathcal{E}(\boldsymbol{n}_{12}/p) = \sum_i P(w_{12}^{(i)}/p) \,.\, n_{12}^{(i)} + \sum_j P(w_{32}^{(j)}/p) \,.\, m_{32}^{(j)},$$

$$\mathcal{E}(\boldsymbol{n}_{32}/p) = \sum_i P(w_{12}^{(i)}/p) \,.\, m_{12}^{(i)} + \sum_j P(w_{32}^{(j)}/p) \,.\, n_{32}^{(j)},$$

$$\mathcal{E}(\boldsymbol{n}^*/p) = \sum_i P(w_{12}^{(i)}/p) \,.\, m_{12}^{(i)} + \sum_j P(w_{32}^{(j)}/p) \,.\, m_{32}^{(j)}.$$

Hence it immediately follows that

$$\mathcal{E}(\boldsymbol{n}/p) = \mathcal{E}(\boldsymbol{n}_{12}/p) + \mathcal{E}(\boldsymbol{n}_{32}/p) - \mathcal{E}(\boldsymbol{n}^*/p).$$

The first two terms of the right hand member of this equation can be calculated just as W a l d does ([2]), p. 53), and $\mathcal{E}(\boldsymbol{n}^*/p)$ is a series with a finite number of terms, each term containing a probability which may be found in a table of the binomial distribution [3]).

R e m a r k: S o b e l and W a l d do not derive an exact formula for $\mathcal{E}(\boldsymbol{n}/p)$, but upper and lower bounds ([1]), p. 513 ff.). Their method, however, is much more complicated and possesses no advantages over the one given above, at least not for the problem treated here.

§ 5. *Truncation of the test.* $\mathcal{E}(\boldsymbol{n}/p)$ can assume great values if p lies in I_{12} or in I_{32}, as is shown in fig. 3. Because it is often desirable to set an upper limit N to the number of observations, a rule is given in order to truncate the test T at $\boldsymbol{n} = N$ i.e. in fig. 2, at the line $x + y = N$.

This rule is as follows: Draw two lines through O in fig. 2, parallel with DF and CE respectively, which divide the line $x + y = N$ into three parts. The part, next to the x-axis (or y-axis respectively) is the region of acceptance for H_1 (or H_3 respectively), whilst the middle part leads to acceptance of H_2. The truncated test thus obtained is denoted by ${}_NT$, and the probabilities of acceptance of H_i $(i = 1,3)$ by $\beta_i(p, {}_NT)$ $(i = 1,3)$. This way of truncation is analogous to that in the ordinary sequential test of Wald.

Evidently the true level of significance of T will become greater by truncation. For its computation, as well as for that of $\mathcal{E}(\boldsymbol{n}/p)$ in the truncated case, we need the probabilities $P(\boldsymbol{n} = n/p)$, with $n = 0, 1, \ldots, N$, because

$$\mathcal{E}(\boldsymbol{n}/p) = \sum_{n=0}^{N} P(\boldsymbol{n} = n/p) . n,$$

where $P(\boldsymbol{n} = n/p)$ is the probability that ${}_NT$ terminates at the n^{th} observation. These probabilities are found with the help of a table of the binomial distribution.

§ 6. *The classical test.* Now consider a test T^* for H_1, H_2 and H_3 (cf. fig. 1), based upon a *constant number* (N) *of observations.* Let m be the number of observations A among these N, then T^* is defined as follows:

$$T^* \begin{cases} \text{accept } H_1, \text{ if } m < a_{12} N, \\ \quad ,, \quad H_2, \;,, \; a_{12} N \leq m \leq a_{32} N, \\ \quad ,, \quad H_3, \;,, \; m > a_{32} N. \end{cases}$$

Of course we need not always perform all the observations, e.g. H_3 may be accepted as soon as more than $a_{32} N$ observations A are found. But this does not affect the process of sampling which consists in taking a sample of size N in each case. Therefore only the number N is considered for comparison with the $\mathcal{E}\boldsymbol{n}$ of a sequential test.

The correctness and incorrectness of decisions is defined again as in table I of § 2. The quantities $\gamma^*(p)$ and α^* (analogous to $\gamma(p)$ and α) are easily calculated. It may be remarked here that, unlike in the case of the sequential test, the numbers a_{12} and a_{32} are very important in this test.

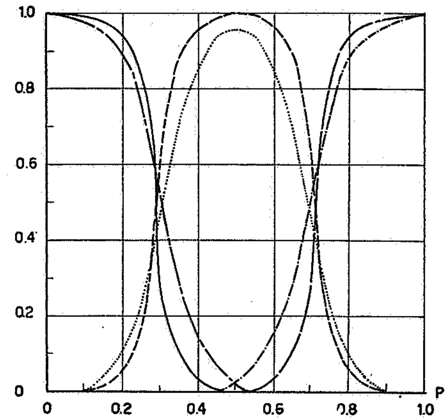

Fig. 4. Characteristics of $T^{(1)}$ and ${}_{25}T^{(1)}$.

$= \beta_1(p, T^{(1)})$

$= \beta_3(p, T^{(1)})$

$= 1 - (\beta_1(p, T^{(1)}) + \beta_3(p, T^{(1)}))$

$= \beta_1(p, {}_{25}T^{(1)})$

$= \beta_3(p, {}_{25}T^{(1)})$

$= 1 - (\beta_1(p, {}_{25}T^{(1)}) + \beta_3(p, {}_{25}T^{(1)}))$

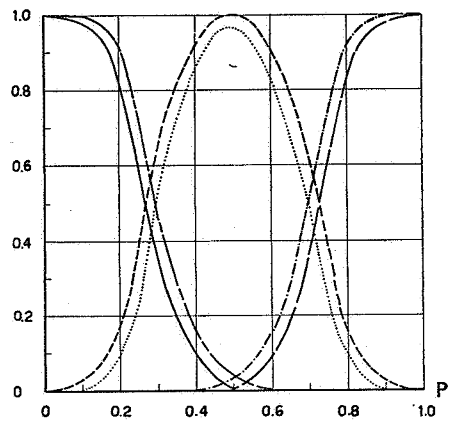

Fig. 5. Characteristics of $T^{(2)}$ and ${}_{25}T^{(2)}$.

$= \beta_1(p, T^{(2)})$

$= \beta_3(p, T^{(2)})$

$= 1 - (\beta_1(p, T^{(2)}) + \beta_3(p, T^{(2)})$

$= \beta_1(p, {}_{25}T^{(2)})$

$= \beta_3(p, {}_{25}T^{(2)})$

$= 1 - (\beta(p, {}_{25}T^{(2)}) + \beta(p, {}_{25}T^{(2)}))$

§ 7. *Comparison of T and T^* by numerical examples.* Since a comparison of T and T^* can practically only be done by numerical examples, some cases, which for convenience have been taken symmetrical, are considered (cf. § 4). Two sequential tests, $T^{(1)}$ and $T^{(2)}$, both with $\alpha=0.05$, are truncated at $N = 25$, yielding the tests ${}_{25}T^{(1)}$ and ${}_{25}T^{(2)}$ with true levels of significance $\alpha^{(1)}$ and $\alpha^{(2)}$ respectively.

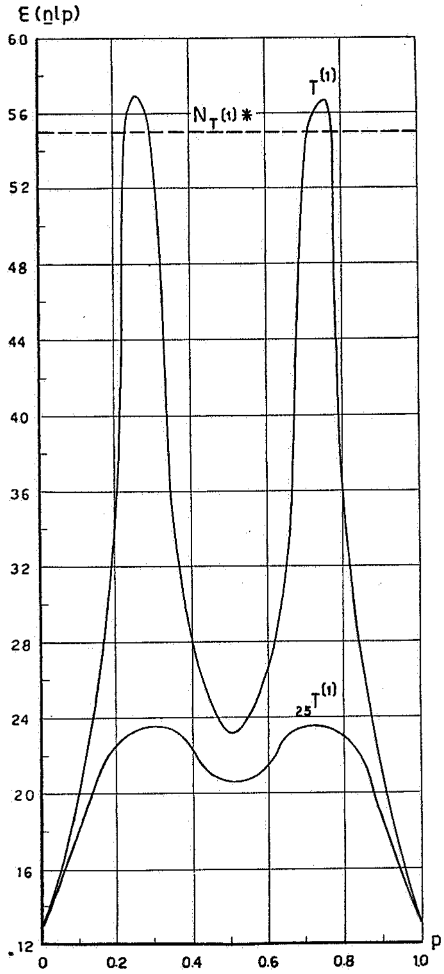

Fig. 6. $\mathcal{E}(\boldsymbol{n}/p)$ of $T^{(1)}$ and ${}_{25}T^{(1)}$ compared with $N_T(1)^*$.

These four tests are compared with the classical tests $T^{(1)*}$, $T^{(2)*}$, ${}_{25}T^{(1)*}$ and ${}_{25}T^{(2)*}$ respectively, having nearly the same operating characteristic and the same true level of significance as the corresponding sequential test. This comparison refers to $\mathcal{E}\boldsymbol{n}$, i.e. the average number of observations for ${}_{25}T^{(j)}$, and N, i.e. the number of observations for $T^{(j)*}$, needed to obtain a true level of significance $\alpha^{(j)}$ $(j = 1, 2)$.

By the way, we note the increase of the true level of significance caused by truncation of a sequential test (cf. § 5).

$$T^{(1)} \text{ and } T^{(2)} \text{ are defined by}$$
$$T^{(1)}: \ p_1 = 0.20, \ p_2' = 0.40,$$
$$T^{(2)}: \ p_1 = 0.15, \ p_2' = 0.45.$$

For the classical tests, a_{12} is taken as the midpoint of the interval (p_1, p_2') and, because of the symmetry, a_{32} is the midpoint of (p_2'', p_3). Thus we have

$$T^{(1)*} \text{ and } {}_{25}T^{(1)*}: \ p_1 = 0.20, \quad a_{12} = 0.30, \quad p_2' = 0.40,$$
$$T^{(2)*} \text{ and } {}_{25}T^{(2)*}: \ p_1 = 0.15, \quad a_{12} = 0.30, \quad p_2' = 0.45.$$

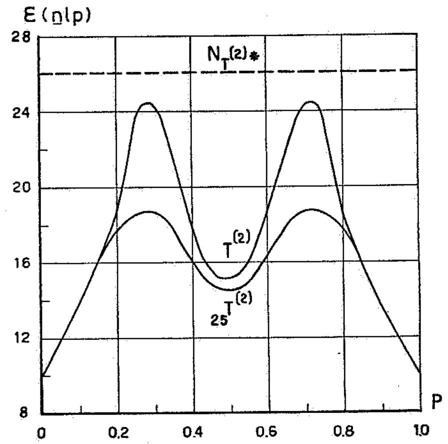

Fig. 7. $\mathcal{E}(\boldsymbol{n}/p)$ of $T^{(2)}$ and ${}_{25}T^{(2)}$ compared with $N_T(2)*$.

The operating characteristics of $T^{(j)}$ and the corresponding ${}_NT^{(j)}$ $(j = 1, 2)$ have been drawn in figs. 4 and 5. We find

$$\alpha^{(1)} = 0.15, \quad \alpha^{(2)} = 0.07.$$

For N we get the following values:

$$T^{(1)*}: \ N = 55, \qquad {}_{25}T^{(1)*}: \ N = 25,$$
$$T^{(2)*}: \ N = 26, \qquad {}_{25}T^{(2)*}: \ N = 20.$$

These values of N and the functions $\mathcal{E}(\boldsymbol{n}/p)$ for the corresponding sequential tests have been drawn in figs. 6 and 7. We see that

$$\mathcal{E}(\boldsymbol{n}/p) < N$$

for the four pairs of tests and all values op p, except for the case of $T^{(1)}$ and $T^{(1)*}$ with p in the neighbourhood of 0.25 and 0.75.

Thus, for each value of p, the truncated sequential tests need on the average less observations than the corresponding classical tests, when the true level of significance is the same in both cases. The difference, moreover, is considerable.

§ 8. *Conclusions.* A comparison of the sequential and the classical tests by means of some numerical examples, shows that as a rule the non-truncated test needs on the average less observations than the corresponding classical test with the same true level of significance, except for tests where the zones of indifference are relatively small and moreover the real value of p lies about the middle of a zone of indifference. For the truncated sequential test the same rule holds with respect to the classical test, in this case without the exception mentioned above. This means that, in these examples, truncated sequential tests require on the average, for every value of p, less observations than classical tests with the same true level of significance.

Received 2nd September, 1953.

REFERENCES

1) Sobel, M. and A. Wald, A sequential decision procedure for choosing one of three hypotheses concerning the unknown mean of a normal distribution. The annals of mathematical statistics, **20** (1949) 502.
2) Wald, A., Sequential analysis, John Wiley and Sons, Inc. N.Y. (1947).
3) Tables of the binomial probability distribution, National Bureau of Standards, Applied Mathematics Series, **6** (1949).