Selecting the Best One of Several Binomial Populations

By MILTON SOBEL and MARILYN J. HUYETT

(Manuscript received July 17, 1956)

Tables have been prepared for use in any experiment designed to selec that particular one of k binomial processes or populations with the highest (long time) yield or the highest probability of success. Before experimentation the experimenter chooses two constants d^* and P^* ($0 < d^* \leqq 1$; $0 \leqq P^* < 1$) and specifies that he would like to guarantee a probability of at least P^* of a correct selection whenever the true difference between the long-time yields associated with the best and the second best processes is at least d^*. The tables show the smallest number of units required per process to be put on test to satisfy this specification. Separate tables are given for $k = 2$, 3, 4 and 10. Each table gives the result for $d^* = 0.05$ (0.05) 0.50 and for $P^* = 0.50, 0.60, 0.70, 0.75, 0.80, 0.85, 0.90, 0.95$, and 0.99. For values of d^* and P^* not considered in the tables, graphs are given on which interpolation can be carried out. Graphs have also been constructed to make possible an interpolation or extrapolation for other values of k. An alternative specification is given for use when the experimenter has some a priori knowledge of the processes and their probabilities of success. This specification is then compared with the original specification. Applications of these tables to different types of problems are considered.

INTRODUCTION AND SUMMARY

A frequently encountered problem is that of selecting the "best" one of k ($k \geqq 2$) processes or populations on the basis of the same number n of observations from each process. We shall assume that the given processes are all binomial or "go — no go" processes and that the best process is the one with the highest probability of obtaining a "success" on a single observation. We shall consider a single sample or nonsequential procedure which means that the common number n of observations from each process is to be determined before experimentation starts. The corresponding sequential problem is being investigated.[1]

Briefly, the technique employed here is to let the experimenter decide how "close" the best and second best processes can be before he is willing to relax his control on the probability of a correct selection. The selection of a best process will, of course, be made on the basis of the largest observed frequency of "success"; the only remaining problem is to determine the value of n. Tables and graphs which cover almost all practical problems in this framework are given for determining the required value of n. In particular, tables and graphs are given for $k = 2$, 3, 4 and 10. Graphs are also given to approximate the result for any value of k up to 100.

This problem arises in many widely different fields of endeavor; we shall briefly consider two industrial applications. One application of the binomial problem is to comparative yield studies. Here success corresponds to the making of a good unit, and the goal is to select the process with the highest (long-time) yield. Another application of the binomial problem is to comparative life testing studies. In this case the experimenter selects a fixed time T and defines the best process as the one for which the probability of any one unit surviving this time T is highest. Then, of course, a successful unit is one which survives the time T. In treating this as a binomial we are discarding the information contained in the exact times of failure. In many cases the times of failure are either unknown or very inexact; in other cases it is not known how to utilize the knowledge of the exact times of failure. Hence, it would be valuable to know the results for the more basic binomial problem. The time T is considered fixed throughout; its value is determined by non-statistical considerations. The specification and the final decision of the experimenter all refer to this predetermined time T. It should be noted that the experimenter cannot use information obtained from the continuation of the test beyond time T since the best process for T is not necessarily the best process for a longer time, say $10T$. This binomial type of analysis has the advantage that it does not assume any particular form of the life distribution. In particular, the assumption of exponential life is avoided.

The presence of à priori information changes the number of observations required. An alternative specification is given which is justified by certain à priori information based on past experimentation. The amount of saving is briefly examined. This area of utilizing à priori information to reduce the number of observations required should be investigated further.

The treatment of the problem in this paper is based on the assumption that, after experimentation is carried out, the experimenter must choose one of the k processes and assert that it is best. If he allows himself the possibility of hedging and asserting that he needs further experimentation, then the problem changes and the tables of this paper are not appropriate.

The following additional assumptions will be made:

1. Observations from the same or different processes are independent.
2. Observations from the same process have a common fixed probability of "success".
3. There is no chance of error in determining whether a success or a failure has occurred.

The assumption of a common probability fixed once and for all for each process is one that should be checked carefully in any practical application of the results in this paper. Roughly speaking, this assumption states that each of the processes is in a state of statistical control as far as the probability of success is concerned.

We shall consider only the case in which the same number n of observations are taken from each process. This is certainly reasonable for a single sample procedure if no à priori information is assumed.

STATISTICAL FORMULATION OF THE PROBLEM

Each of k given binomial populations Π_i is associated with a fixed probability of success p_i where $0 \leqq p_i \leqq 1$ $(i = 1, 2, \cdots, k)$. For example, in the yield problem p_i is the long-time yield for process Π_i or the probability that any one unit from Π_i is a good one. Let the ordered values of the p_i be denoted by

$$p_{[1]} \geqq p_{[2]} \geqq \cdots \geqq p_{[k]} \quad (1)$$

No à priori information is assumed about the values of the p_i or about the correspondence between the ordered $p_{[i]}$ and the k identifiable populations Π_i. In particular, we have no idea before experimentation starts whether $p_{[1]}$ is associated with $\Pi_1, \Pi_2, \cdots$, or Π_k.

The problem is to select the population associated with $p_{[1]}$ on the basis of n observations from each population. If there are t ties for first place, say

$$p_{[1]} = p_{[2]} = \cdots = p_{[t]} > p_{[t+1]} \qquad (t < k) \quad (2)$$

then we shall certainly be content with the selection of any one of the associated t populations as the best one.

As an index of the true difference (or distance) between the best and second best populations we introduce the symbol

$$d = p_{[1]} - p_{[2]} \quad (3)$$

It is assumed that if the difference d between the best and second best populations is small enough, then the error involved in wrongly selecting the second best process as the best one is an error of little or no consequence. The experimenter is therefore asked to specify two quantities which will determine the number n of observations he is required to take from each process.

> Specification: He specifies the smallest value d^* $(0 < d^* \leqq 1)$ of d for which it would be economically desirable to make the correct selection. He also specifies a probability P^* $(0 \leqq P^* < 1)$ of making a correct selection that he would like to guarantee whenever the true difference $d \geqq d^*$. (4)

Letting $P_{CS} = P_{CS}(p_{[1]}, \cdots, p_{[k]})$ denote the probability of a correct selection we can now rewrite the specification that the experimenter wants to satisfy in the simple form

$$P_{CS} \geqq P^* \quad \text{for} \quad d \geqq d^* \tag{5}$$

[The word "specification" will be used below to denote the specified pair of constants (d^*, P^*) as well as the condition (5); it will be clear from the text which is meant.] Since the final selection is to be made on the basis of the observed frequency of success, the essential problem is to find the number n of observations required per process to satisfy the specification (5).

The possibility that d may be less than d^* is not being overlooked. The region $d < d^*$ is being regarded as a zone of indifference in the sense that if $d < d^*$, then we do not care which process is selected as best so long as its p-value is within d^* of the highest p-value $p_{[1]}$. For values of $P^* \leqq 1/k$ no tables are needed since a probability of $1/k$ can be attained by chance alone.

Some comments on the above approach and on a possible modification have been placed in Appendix I in order to preserve the continuity of the paper.

CONFIDENCE STATEMENT

After the experiment is completed and the selection of a best process is made, the experimenter can make a confidence statement with confidence level P^*. Let p_s denote the true p-value of the selected population and let p_u denote the maximum true p-value over all unselected populations. Then the confidence statement, consisting of two sets of inequalities

$$\begin{Bmatrix} p_{[1]} - d^* \leqq p_s \leqq p_{[1]} \\ p_{[2]} \leqq p_u \leqq p_{[2]} + d^* \end{Bmatrix} \quad \textit{or} \quad \begin{Bmatrix} 0 \leqq p_{[1]} - p_s \leqq d^* \\ 0 \leqq p_u - p_{[2]} \leqq d^* \end{Bmatrix}$$

has confidence level P^*. It should be noted that the above confidence statement is not a statement about the value of any p but is a statement about the correctness of the selection made.

LEAST FAVORABLE CONFIGURATION

The main idea used in the construction of the tables was that of a least favorable configuration. Before defining this concept we shall define the set of configurations

$$p_{[1]} - d = p_{[2]} = p_{[3]} = \cdots = p_{[k]} \tag{6}$$

obtained by letting d in (6) vary over the closed interval $(d^*, 1)$ as the *Less-Favorable* set of configurations. It is intuitively clear and will be rigorously shown in Appendix II that if our procedure satisfies the specification for any true configuration (6) with $d = d^0$ and $p_{[1]} = \overset{0}{p}_{[1]}$, then it will also satisfy the specification when

$$\overset{0}{p}_{[1]} - d^0 \geqq p_{[2]} \geqq p_{[3]} \geqq \cdots \geqq p_{[k]} \tag{7}$$

Of course, we shall be interested particularly in the case in which d equals the specified value d^*. If $d = d^*$ is fixed in (6), then (6) specifies the differences between the p-values, but the "location" of the set is still not specified. We shall use $p_{[1]}$ to locate the set of p-values. The probability P_{CS} of a correct selection for configurations like (6) with $d = d^*$ depends not only on d^*, n and k but also on the location $p_{[1]}$ of the largest p-value (except for the special case $k = 2$ and $n = 1$). [In the corresponding problem for selecting the largest population mean of k independent *Normal* distributions with unit variance,[2] this probability P_{CS} depends *only on the differences* and, hence, only on d in the configuration corresponding to (6)].

When (6) holds with any *fixed* value of d, the probability P_{CS} (for any fixed n) may be regarded as a function of $p_{[1]}$ where $d \leqq p_{[1]} \leqq 1$). This function is continuous and bounded over a closed interval and therefore assumes its minimum value at some point $p_{[1]}(d) = p_{[1]}(d;n)$ in the closed interval $(d,1)$. Fig. 1(b) gives the value of $p_{[1]}(d)$ as a function of d for $k = 3$ and for $n = 1, 2, 4, 10$ and ∞. For any particular value of n and for $d = d^*$ we shall be particularly interested in the value $p_{[1]}^L = p_{[1]}(d^*;n)$ since this (as shown in Appendix II) gives the smallest probability P_{CS}^L of a correct selection *for all the configurations included in the statement of the experimenter's specification*. This particular configuration (6) with $d = d^*$ and $p_{[1]} = p_{[1]}^L$ (which depends on n) is called the *Least-Favorable Configuration*.

Although the least favorable configuration depends on n, it has been empirically found that for $n \geqq 10$ (and in some cases for $n \geqq 4$) the least favorable configuration is approximately given by $p_{[1]}^L = \frac{1}{2}(1 + d^*)$ in which the two values, $p_{[1]}^L$ and $p_{[2]}^L = p_{[1]}^L - d^*$ are symmetric about $\frac{1}{2}$. This symmetric configuration clearly does not depend on n. Fig. 1(b) shows that as $n \to \infty$ the least favorable configuration approaches this symmetric configuration (i.e., the straight line marked $n = \infty$) *quite rapidly* for any value of d. In Appendix III it is proved that the symmetric configuration is least favorable as $n \to \infty$. Fig. 1(a) shows for $k = 3$, $n = 10$, and any value of d the error in P_{CS} which arises as a result of using the symmetric configuration instead of the true least favorable configuration.

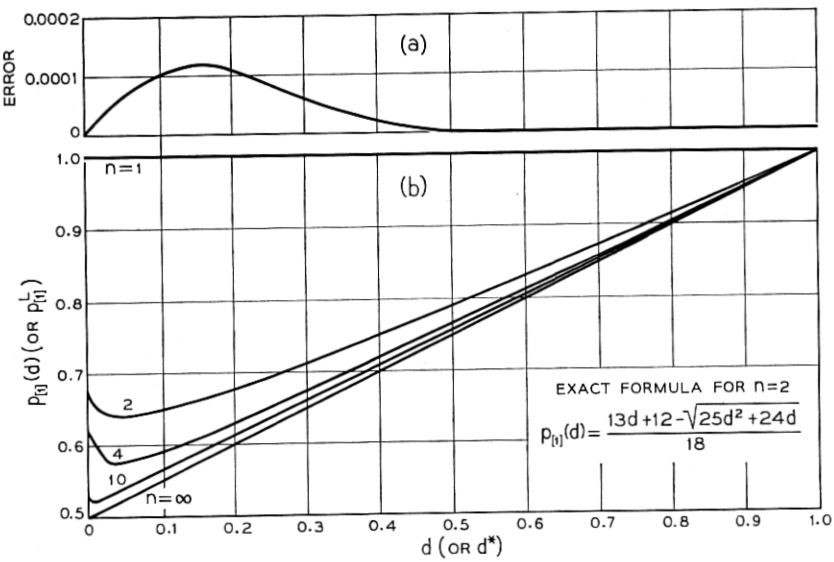

Fig. 1 — (a) Error in P_{CS} as a result of using the symmetric configuration instead of the least favorable configuration for $k = 3$, $n = 10$, and any common true difference d. (b) Least favorable value $p_{[1]}(d)$ of $p_{[1]}$ as a function of the common true difference $d = p_{[1]} - p_{[2]}$, $i \geqq 2$, for $k = 3$ and selected values of n. (for $d = d^*$, $p_{[1]}(d) = p_{[1]}^L$)

CONSTRUCTION OF THE TABLES

Consider any fixed value of d^*. For each of a set of increasing values of n the minimum probability P_{CS}^L of a correct selection for $d \geqq d^*$ (i.e., the probability for the least favorable configuration) was computed. These calculations were then inverted to find the smallest n for which the P_{CS}^L is greater than or equal to the specified value P^*. Tables I through IV give the smallest value of n for $k = 2, 3, 4$, and 10, for $d^* = 0.05\ (0.05)\ 0.50$, and for selected values of P^*. Graphs corresponding to these tables are given in Figs. 2 through 5.

For small values of n (say, $n < 10$) it was necessary to approximate $p_{[1]}^L$ by calculating the P_{CS} *exactly* for several values of $p_{[1]}$ and proceeding in the direction of the minimum probability P_{CS}^L. For the special case $n = 2$ and $k = 3$ an explicit formula for $p_{[1]}^L$ is given on Fig. 1.

For large values of n (say, $n > 10$) the P_{CS}^L was calculated by assuming the symmetric configuration. Here it was necessary to make use of the normal approximation to the binomial. Fortunately the appropriate table needed in this normal approximation is already published.[2] The proof that this table is appropriate is given in Appendix III. The resulting value of n is given by

$$n = \frac{B}{d^{*2}}(1 - d^{*2}) \cong \frac{B}{d^{*2}} \tag{8}$$

where the constant B, depending on P^* and k, is equal to $\frac{1}{4}C^2$ and C is the entry in the appropriate column of Table I of R. E. Bechhofer's paper.[2] A short table of B values, Table V (see page 550), is included in this paper to make it self-contained.

The middle expression in (8) will be referred to as the normal approximation and the right hand expression in (8) will be referred to as the "straight line" approximation. In many cases it has been *empirically* found that these two expressions give close lower and upper bounds to the true value. Thus by noting the curves drawn in Figs. 4 and 6 for $k = 4$, $P^* \geqq 0.75$ it appears that for all values of d^* the true P_{CS}^L is between the normal approximation and the straight line approximation. Assuming this to be so, it follows that for $k = 4$, $P^* \geqq 0.75$ the required value of n satisfies the inequalities

$$\left[\frac{B}{d^{*2}}(1 - d^{*2})\right] \leqq n \leqq \left[\frac{B}{d^{*2}}\right] \tag{9}$$

where $[x]$ denotes the smallest integer greater than or equal to the enclosed quantity x. This result (9) is empirical and not based on any mathematically proven inequalities. It is used here only to estimate the

TABLE I — NUMBER OF UNITS REQUIRED PER PROCESS TO GUARANTEE A PROBABILITY OF P^* OF SELECTING THE BEST OF k BINOMIAL PROCESSES WHEN THE TRUE DIFFERENCE $p_{[1]} - p_{[2]}$ IS AT LEAST d^*. ($k = 2$)

The three values in each group are: (1) Normal approximation, (2) Straight line approximation, and (3) Smallest integer required.

d^*	P^* 0.50	0.60	0.75	0.80	0.85	0.90	0.95	0.99
0.05	0 0 0	12.81 12.84 14	90.77 90.99 92	141.30 141.66 142	214.29 214.83 215	327.66 328.48 329	539.77 541.12 541	1079.70 1082.41 1082
0.10	0 0 0	3.18 3.21 4	22.52 22.75− 23	35.06 35.41 36	53.17 53.71 54	81.30 82.12 83	133.93 135.28 135	267.90 270.60 270
0.15	0 0 0	1.39 1.43 2	9.88 10.11 11	15.39 15.74 16	23.33 23.87 24	35.68 36.50− 37	58.78 60.12 60	117.57 120.27 120
0.20	0 0 0	0.77 0.80 1	5.46 5.69 6	8.50− 8.85+ 9	12.89 13.43 14	19.71 20.53 21	32.47 33.82 34	64.94 67.65+ 67
0.25	0 0 0	0.48 0.51 1	3.41 3.64 4	5.31 5.67 6	8.06 8.59 9	12.32 13.14 14	20.29 21.64 22	40.59 43.30 42
0.30	0 0 0	0.32 0.36 1	2.30 2.53 3	3.58 3.93 4	5.43 5.97 6	8.30 9.12 9	13.68 15.03 15	27.36 30.07 29
0.35	0 0 0	0.23 0.26 1	1.63 1.86 2	2.54 2.89 3	3.85− 4.38 5	5.88 6.70 7	9.69 11.04 11	19.38 22.09 21
0.40	0 0 0	0.17 0.20 1	1.19 1.42 2	1.86 2.21 3	2.82 3.36 4	4.31 5.13 5	7.10 8.46 9	14.21 16.91 16
0.45	0 0 0	0.13 0.16 1	0.90 1.12 2	1.39 1.75− 2	2.11 2.65+ 3	3.23 4.06 4	5.33 6.68 7	10.65+ 13.36 13
0.50	0 0 0	0.10 0.13 1	0.68 0.91 1	1.06 1.42 2	1.61 2.15− 3	2.46 3.28 4	4.06 5.41 5	8.12 10.82 10

TABLE II — NUMBER OF UNITS REQUIRED PER PROCESS TO GUARANTEE A PROBABILITY OF P^* OF SELECTING THE BEST OF k BINOMIAL PROCESSES WHEN THE TRUE DIFFERENCE $p_{[1]} - p_{[2]}$ IS AT LEAST d^*. ($k = 3$)

The three values in each group are: (1) Normal approximation, (2) Straight line approximation, and (3) Smallest integer required.

d^*	P^*							
	0.50	0.60	0.75	0.80	0.85	0.90	0.95	0.99
0.05	30.89 30.97 31	78.16 78.36 79	205.06 205.58 206	272.36 273.04 273	363.06 363.97 364	496.14 497.38 498	732.63 734.46 735	1305.21 1308.49 1308
0.10	7.66 7.74 8	19.39 19.59 20	50.88 51.39 52	67.58 68.26 69	90.08 90.99 91	123.10 124.34 125	181.78 183.62 184	323.85+ 327.12 327
0.15	3.36 3.44 4	8.51 8.71 9	22.33 22.84 23	29.66 30.34 31	39.53 40.44 41	54.02 55.26 55	79.77 81.61 82	142.12 145.39 145
0.20	1.86 1.94 3	4.70 4.90 5	12.33 12.85− 13	16.38 17.07 17	21.84 22.75− 23	29.84 31.09 31	44.07 45.90 46	78.51 81.78 81
0.25	1.16 1.24 2	2.94 3.13 4	7.71 8.22 9	10.24 10.92 11	13.65− 14.56 15	18.65+ 19.90 20	27.54 29.38 29	49.07 52.34 52
0.30	0.78 0.86 2	1.98 2.18 3	5.20 5.71 6	6.90 7.58 8	9.20 10.11 10	12.57 13.82 14	18.57 20.40 20	33.08 36.35− 35
0.35	0.55+ 0.63 2	1.40 1.60 2	3.68 4.20 5	4.89 5.57 6	6.52 7.43 8	8.91 10.15 10	13.15+ 14.99 15	23.43 26.70 26
0.40	0.41 0.48 1	1.03 1.22 2	2.70 3.21 4	3.58 4.27 5	4.78 5.69 6	6.53 7.77 8	9.64 11.48 11	17.17 20.45− 20
0.45	0.30 0.38 1	0.77 0.97 2	2.02 2.54 3	2.69 3.37 4	3.58 4.49 5	4.90 6.14 6	7.23 9.07 9	12.88 16.15+ 15
0.50	0.23 0.31 1	0.59 0.78 2	1.54 2.06 3	2.05− 2.73 3	2.73 3.64 4	3.73 4.97 5	5.51 7.34 7	9.81 13.08 12

TABLE III — NUMBER OF UNITS REQUIRED PER PROCESS TO GUARANTEE A PROBABILITY OF P^* OF SELECTING THE BEST OF k BINOMIAL PROCESSES WHEN THE TRUE DIFFERENCE $p_{[1]} - p_{[2]}$ IS AT LEAST d^*. ($k = 4$)

The three values in each group are : (1) Normal approximation, (2) Straight line approximation, and (3) Smallest integer required.

d^*	P^*							
	0.50	0.60	0.75	0.80	0.85	0.90	0.95	0.99
0.05	69.85− 70.02 71	132.65+ 132.99 134	282.27 282.98 283	357.52 358.42 359	456.82 457.96 458	599.53 601.03 601	848.30 850.42 850	1438.12 1441.72 1442
0.10	17.33 17.51 18	32.91 33.25− 34	70.04 70.74 71	88.71 89.61 90	113.35− 114.49 114	148.76 150.26 150	210.48 212.61 212	356.83 360.43 360
0.15	7.61 7.78 8	14.44 14.78 15	30.74 31.44 32	38.93 39.82 40	49.74 50.88 51	65.29 66.78 67	92.37 94.49 94	156.61 160.19 160
0.20	4.20 4.38 5	7.98 8.31 9	16.98 17.69 18	21.51 22.40 23	27.48 28.62 29	36.06 37.56 38	51.03 53.15+ 53	86.50+ 90.12 89
0.25	2.63 2.80 3	4.99 5.32 6	10.61 11.32 12	13.44 14.34 14	17.17 18.32 18	22.54 24.04 24	31.89 34.02 34	54.06 57.67 57
0.30	1.77 1.95− 3	3.36 3.69 4	7.15+ 7.86 8	9.06 9.96 10	11.58 12.72 13	15.19 16.70 17	21.50− 23.62 23	36.44 40.05− 39
0.35	1.25+ 1.43 2	2.38 2.71 3	5.07 5.77 6	6.42 7.31 7	8.20 9.35− 9	10.76 12.27 12	15.23 17.36 17	25.82 29.42 28
0.40	0.92 1.09 2	1.75− 2.08 3	3.71 4.42 5	4.70 5.60 6	6.01 7.16 7	7.89 9.39 9	11.16 13.29 13	18.92 22.53 21
0.45	0.69 0.86 2	1.31 1.64 2	2.79 3.49 4	3.53 4.42 5	4.51 5.65+ 6	5.92 7.42 7	8.37 10.51 10	14.19 17.80 17
0.50	0.53 0.70 2	1.00 1.33 2	2.12 2.83 3	2.69 3.58 4	3.43 4.58 5	4.51 6.01 6	6.38 8.50+ 8	10.81 14.42 13

TABLE IV — NUMBER OF UNITS REQUIRED PER PROCESS TO GUARANTEE A PROBABILITY OF P^* OF SELECTING THE BEST OF k BINOMIAL PROCESSES WHEN THE TRUE DIFFERENCE $p_{[1]} - p_{[2]}$ IS AT LEAST d^*. ($k = 10$)

The three values in each group are: (1) Normal approximation, (2) Straight line approximation, and (3) Smallest integer required.

d^*	P^*							
	0.50	0.60	0.75	0.80	0.85	0.90	0.95	0.99
0.05	216.96 217.50+ 218	312.51 313.29 314	511.15+ 512.43 513	604.04 605.55+ 606	722.50− 724.31 725	887.54 889.77 890	1165.49 1168.41 1169	1798.01 1802.51 1803
0.10	53.83 54.38 55	77.54 78.32 79	126.83 128.11 128	149.87 151.39 151	179.27 181.08 181	220.22 222.44 222	289.18 292.10 291	446.12 450.63 449
0.15	23.62 24.17 25	34.03 34.81 35	55.66 56.94 57	65.77 67.28 67	78.67 80.48 80	96.64 98.86 98	126.90 129.82 129	195.77 200.28 198
0.20	13.05+ 13.59 14	18.80 19.58 20	30.75− 32.03 32	36.33 37.85− 38	43.46 45.27 45	53.39 55.61 55	70.10 73.03 72	108.15 112.66 111
0.25	8.16 8.70 9	11.75− 12.53 13	19.22 20.50− 20	22.71 24.22 24	27.16 28.97 29	33.37 35.59 35	43.82 46.74 46	67.59 72.10 70
0.30	5.50− 6.04 7	7.92 8.70 9	12.95+ 14.23 14	15.31 16.82 17	18.31 20.12 20	22.49 24.72 24	29.53 32.46 32	45.56 50.07 48
0.35	3.90 4.44 5	5.61 6.39 7	9.18 10.46 11	10.84 12.36 13	12.97 14.78 15	15.93 18.16 18	20.92 23.85− 23	32.28 36.79 35
0.40	2.85+ 3.40 4	4.11 4.90 5	6.73 8.01 8	7.95− 9.46 10	9.51 11.32 11	11.68 13.90 13	15.34 18.26 17	23.66 28.16 26
0.45	2.14 2.69 3	3.08 3.87 4	5.05− 6.33 6	5.96 7.48 8	7.13 8.94 9	8.76 10.98 11	11.50+ 14.42 14	17.75− 22.25+ 20
0.50	1.63 2.18 3	2.35− 3.13 4	3.84 5.12 5	4.54 6.06 6	5.43 7.24 7	6.67 8.90 9	8.76 11.68 11	13.52 18.03 16

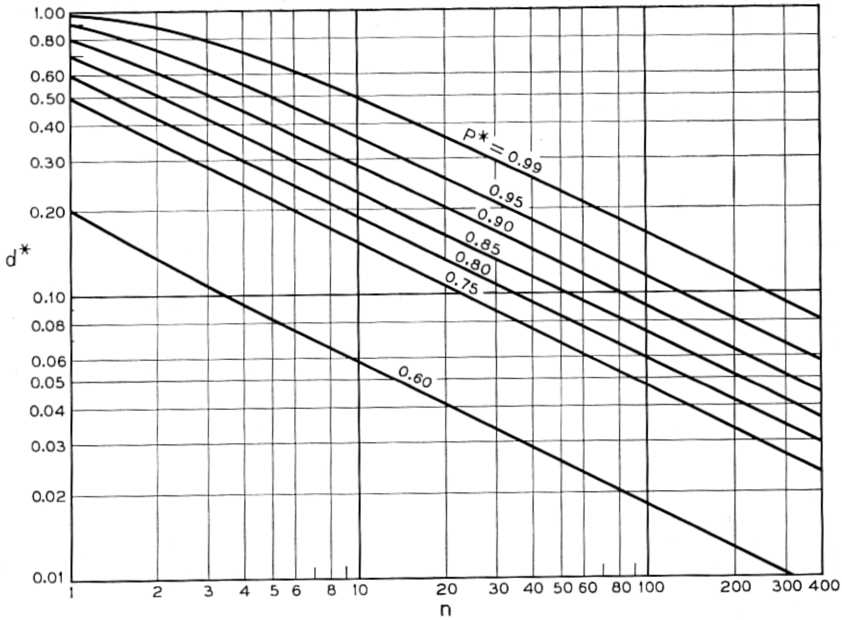

Fig. 2. — Number of units n required per process to guarantee a probability of P^* of selecting the better of two binomial processes when the true difference d is at least d^*.

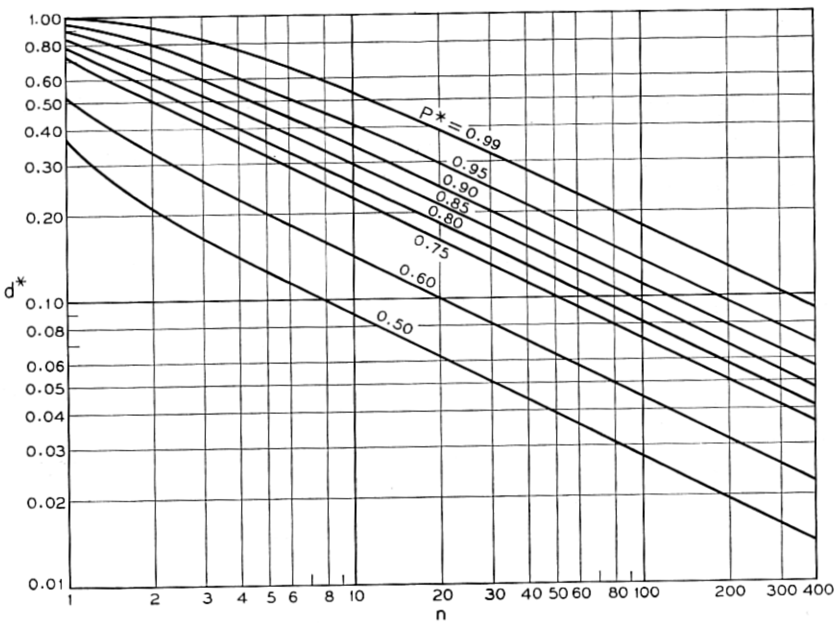

Fig. 3. — Number of units n required per process to guarantee a probability of P^* of selecting the best of three binomial processes when the true difference d is at least d^*.

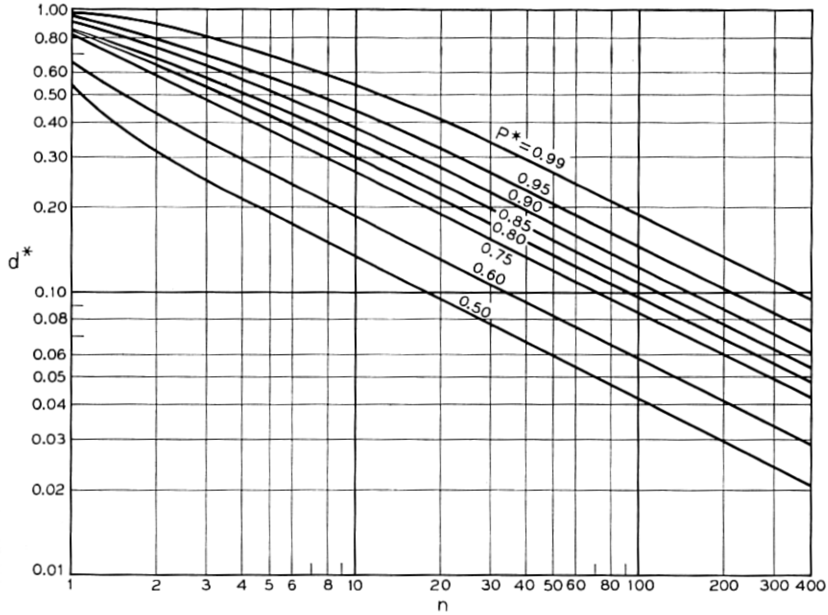

Fig. 4 — Number of units n required per process to guarantee a probability of P^* of selecting the best of four bionimal processes when the true difference d is at least d^*.

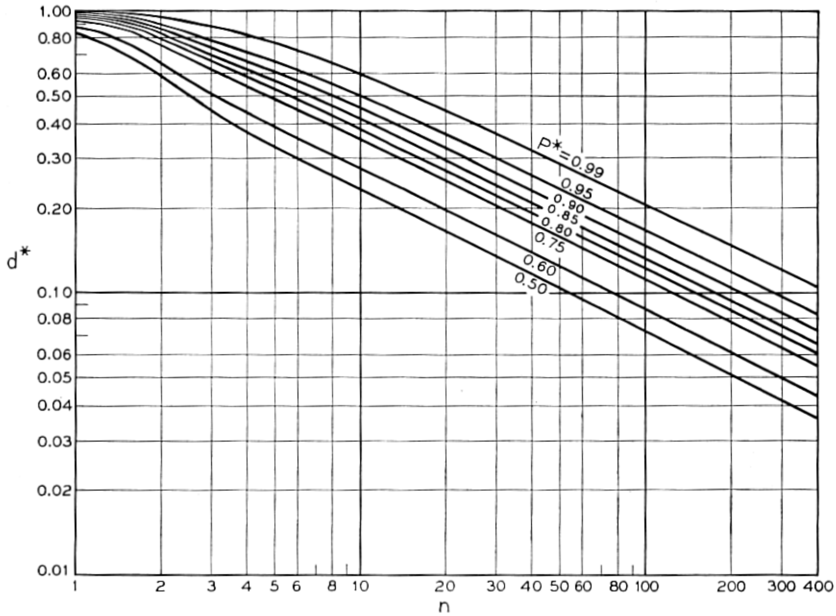

Fig. 5 — Number of units n required per process to guarantee a probability of P^* of selecting the best of ten binomial processes when the true difference d is at least d^*.

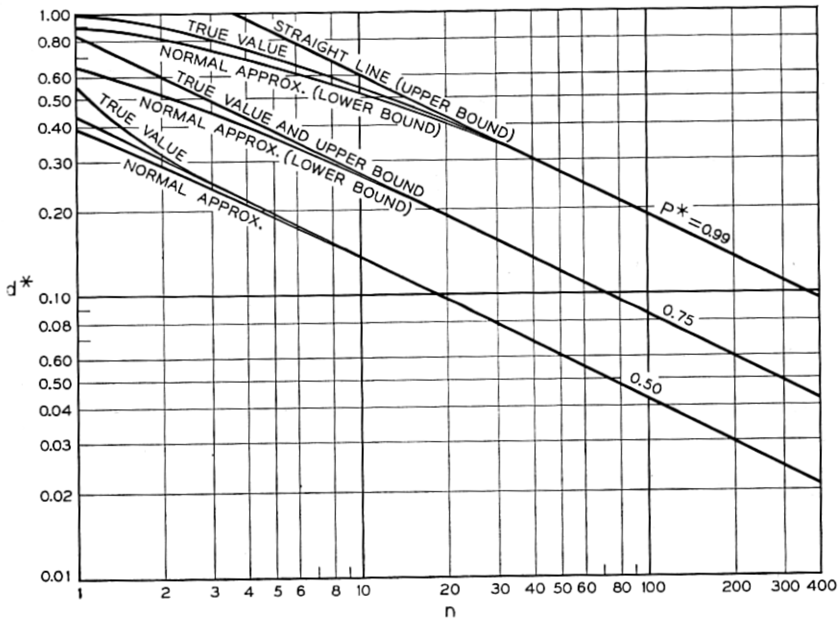

Fig. 6 — Bounds for the number of units n required per process to guarantee a probability of P^* of selecting the best of four binomial processes when the true difference d is at least d^*.

TABLE V — VALUES OF $B = \frac{1}{4}C^2$ TO BE USED WITH THE NORMAL APPROXIMATION (8) WHERE C IS OBTAINED FROM TABLE I OF R. E. BECHHOFER'S PAPER[1]

Prob. of Correct Selection	$k = 2$	$k = 3$	$k = 4$	$k = 10$
0.99	2.7060	3.2712	3.6043	4.5063
0.95	1.3528	1.8362	2.1261	2.9210
0.90	0.8212	1.2434	1.5026	2.2244
0.85	0.5371	0.9099	1.1449	1.7965+
0.80	0.3541	0.6826	0.8961	1.5139
0.75	0.2275−	0.5139	0.7074	1.2811
0.70	0.1375−	0.3832	0.5575−	1.0892
0.65	0.0742	0.2792	0.4347	0.9256
0.60	0.0321	0.1959	0.3325−	0.7832
0.55	0.0079	0.1294	0.2468	0.6569
0.50	0.0000	0.0774	0.1751	0.5438

order of magnitude of the error in our large sample calculations. For example, if $k = 4$, $d^* = 0.05$ and $P^* = 0.90$, then from Table V we find that $B = 1.5026$ and the two expressions in (8) yield 599.54 and 601.04. Hence, it would follow from (9) that n is 600 or 601 or 602. Based on an investigation of the behavior of these two approximations in the case of smaller P^* or larger d^* values, it is estimated that the true value of n is 601. Even if the correct value is 600 or 602 the error would be less than $\frac{1}{6}$ of 1 per cent. Fig. 6 illustrates these bounds on the P_{CS} for $k = 4$, $P^* = 0.50$, 0.75 and 0.99. For $P^* \leqq 0.60$ the straight line approximation is a closer lower bound than the normal approximation.

It is estimated that all integer entries in Tables I through IV have an error of at most 1 per cent and, in particular, that all entries under 100 are exact.

OTHER VALUES OF k

In addition to the tables and graphs for $k = 2$, 3, 4 and 10 there are also graphs (Figs. 7 through 14) on which interpolation can be carried out for $k = 5$ through 9 and on which extrapolation can be carried out for $k = 11$ through 100. By plotting n versus log k (or n versus k on semi-log paper) and drawing a straight (dashed) line through the values of n for $k = 4$ and $k = 10$ we obtain results which are remarkably good approximations for $k > 10$. The solid curve in these figures connects the true values obtained for $k = 2$, 3, 4 and 10.

For large values of k the theoretical justification for a straight line approximation is given in Appendix V. In order to check the accuracy of our procedure of drawing the straight line through the values of n computed for $k = 4$ and $k = 10$, we have chosen two points at $k = 101$ for an independent computation of the probability of a correct selection. For $P^* = 0.90$, $d^* = 0.10$ and $k = 101$ the dashed line in Fig. 12 gives n as approximately 400. To check this we computed the normal approximation to the probability P_{CS}^L of a correct selection for the least favorable configuration in the form

$$P_{CS}^L \cong \int_{-\infty}^{\infty} F^{100}(x + h)f(x)\, dx = \frac{1}{\sqrt{n}} \int_{-\infty}^{\infty} F^{100}(x\sqrt{2} + h)e^{-x^2}\, dx \qquad (10)$$

where

$$h = \frac{2d^* \sqrt{n}}{\sqrt{1 - d^{*2}}} \text{ (= 4.02015 in this example)} \qquad (11)$$

$f(x)$ is the normal density and $F(x)$ is its c.d.f. This was computed by a method suggested by Salzer, Zucker and Capuano[3] and the result was

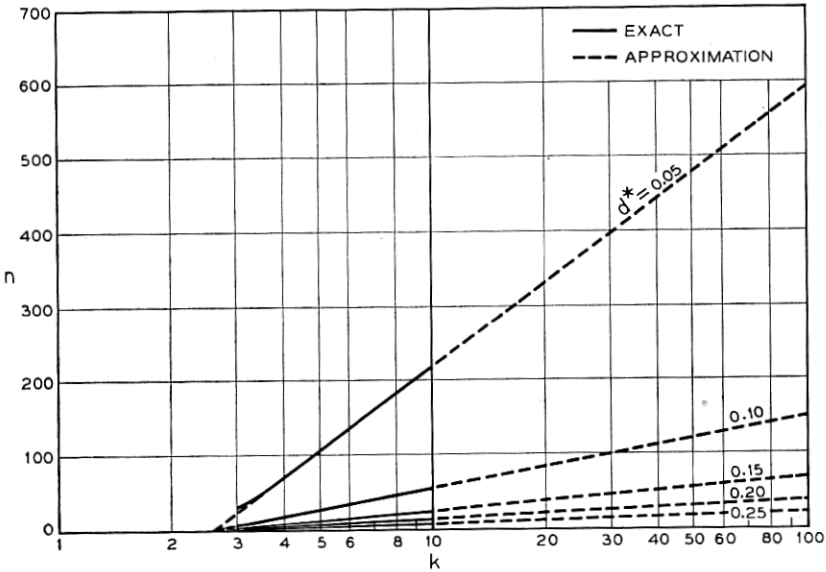

Fig. 7 — Number of units n required per process to guarantee a probability of $P^* = 0.50$ of selecting the best of k binomial processes when the true difference $p_{[1]} - p_{[2]}$ is at least d^*.

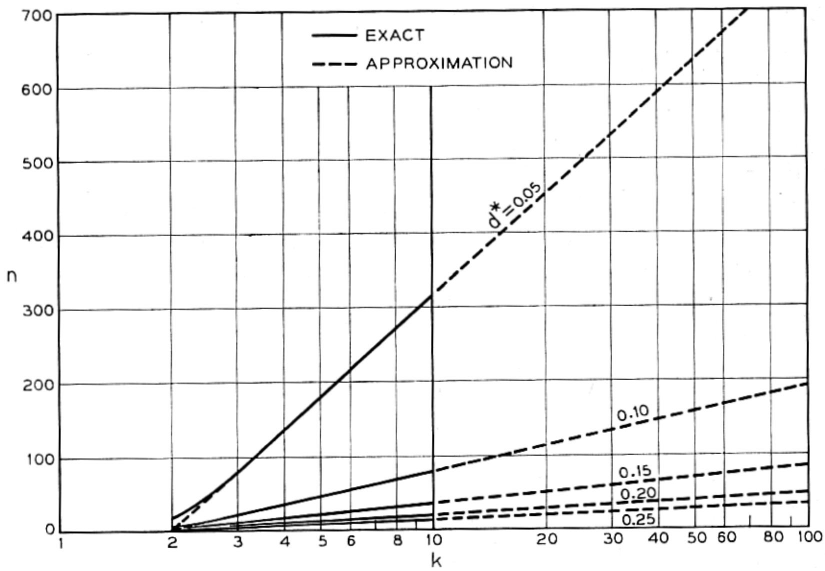

Fig. 8 — Number of units n required per process to guarantee a probability of $P^* = 0.60$ of selecting the best of k binomial processes when the true difference $p_{[1]} - p_{[2]}$ is at least d^*.

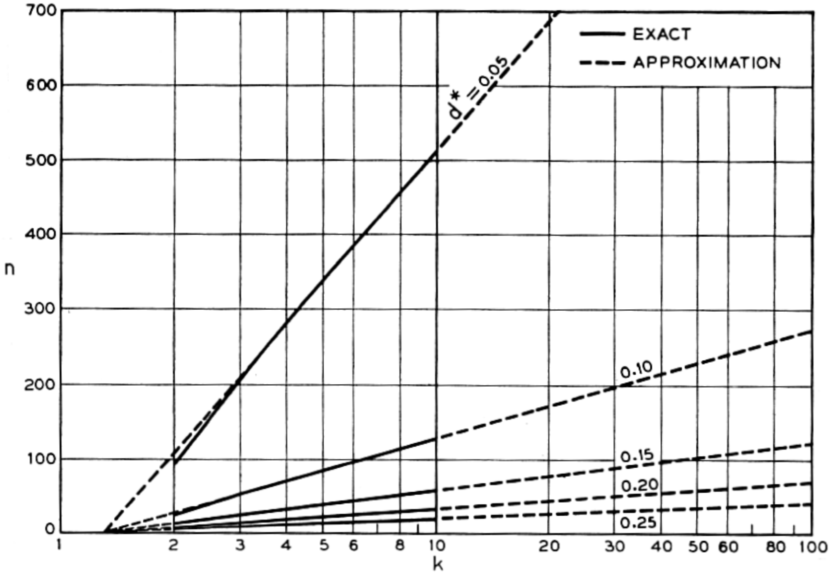

Fig. 9 — Number of units n required per process to guarantee a probability of $P^* = 0.75$ of selecting the best of k binomial processes when the true difference $p_{[1]} - p_{[2]}$ is at least d^*.

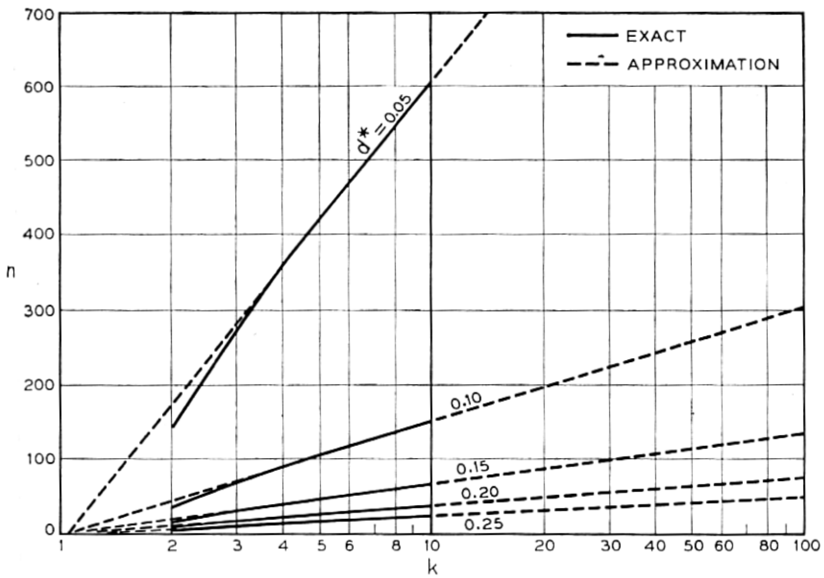

Fig. 10 — Number of units n required per process to guarantee a probability of $P^* = 0.80$ of selecting the best of k binomial processes when the true difference $p_{[1]} - p_{[2]}$ is at least d^*.

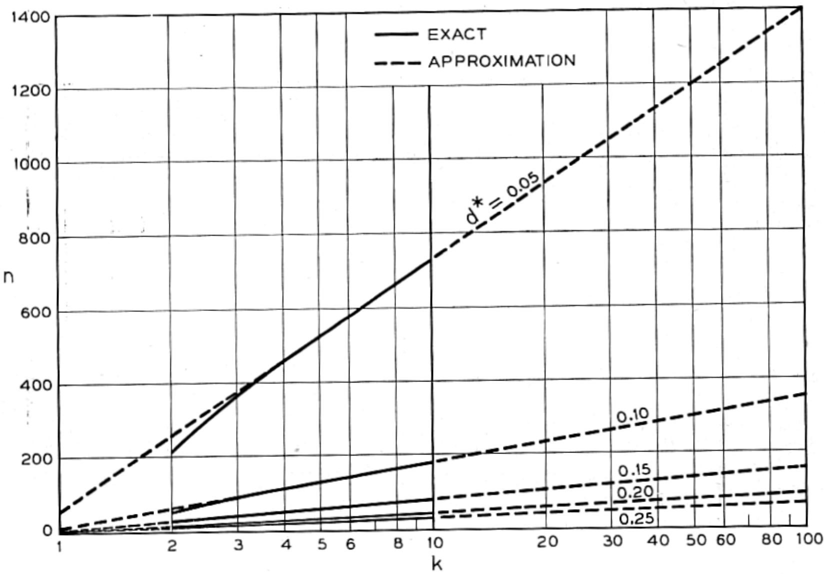

FIG. 11. — Number of units n required per process to guarantee a probability of $P^* = 0.85$ of selecting the best of k binomial processes when the true difference $p_{[1]} - p_{[2]}$ is at least d^*.

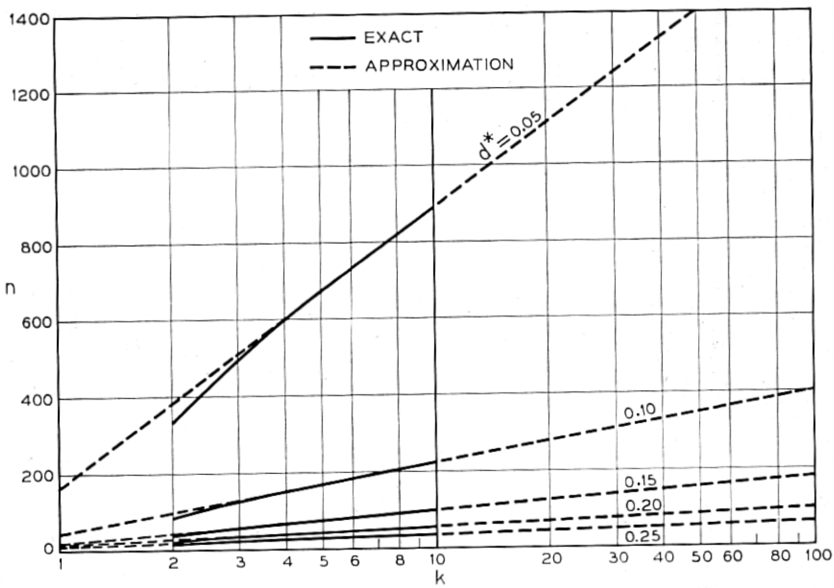

Fig. 12 — Number of units n required per process to guarantee a probability of $P^* = 0.90$ of selecting the best of k binomial processes when the true difference $p_{[1]} - p_{[2]}$ is at least d^*.

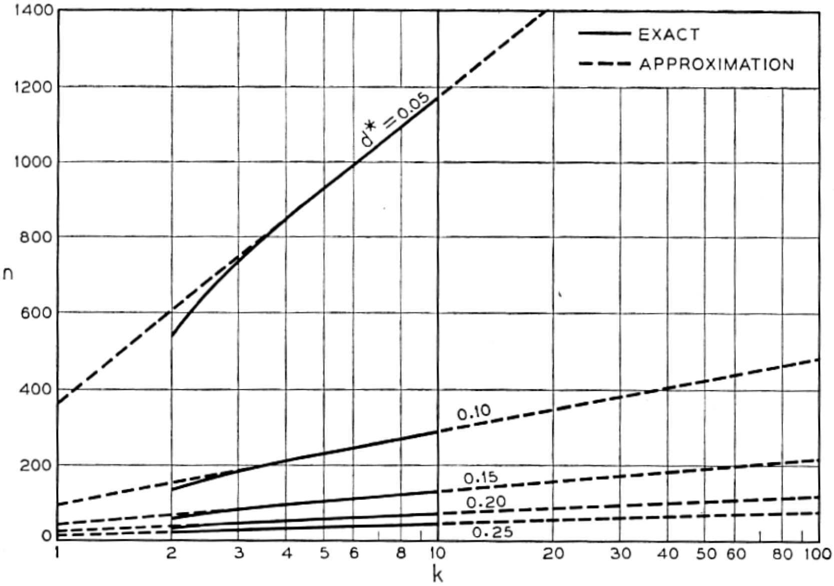

Fig. 13 — Number of units n required per process to guarantee a probability of $P^* = 0.95$ of selecting the best of k binomial processes when the true difference $p_{[1]} - p_{[2]}$ is at least d^*.

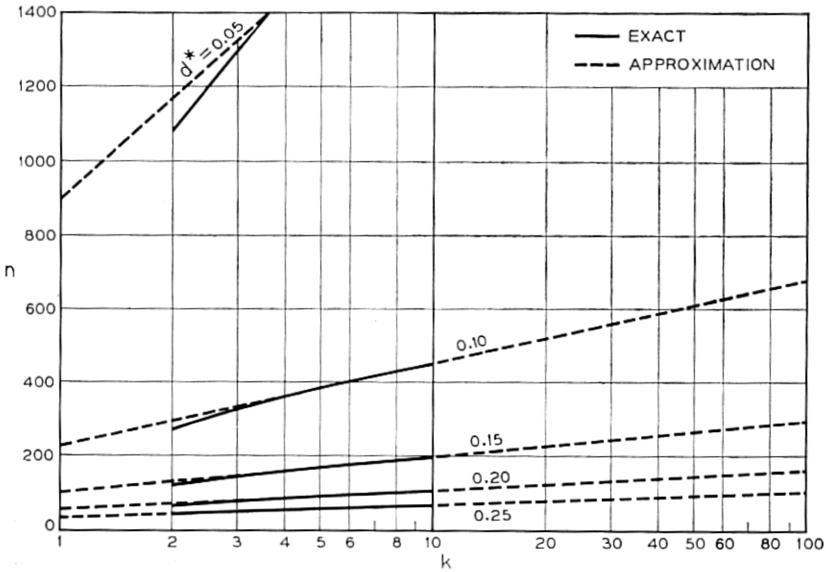

Fig. 14 — Number of units n required per process to guarantee a probability of $P^* = 0.99$ of selecting the best of k binomial processes when the true difference $p_{[1]} - p_{[2]}$ is at least d^*.

$P_{CS}^{L} \cong 0.9168$ as compared to the value 0.90 in Fig. 12. The expression (10) is derived in Appendix III. Another check was made at $P^* = 0.99$, $d^* = 0.20$ and $k = 101$. The value of n from Fig. 14 is 162. The value of the P_{CS}^{L} computed from (10) using Salzer, Zucker and Capuano[3] is 0.9925+.

Further calculation using (10) yielded the more accurate results 378 and 154 instead of 400 and 162, respectively, in the above illustrations. The error in both cases is less than 6 per cent; for smaller values of k the percentage error will, of course, be much less.

For interpolation the results are estimated to be within 1 per cent of the correct value. For example we estimate from Fig. 11 that the required value of n for $k = 5$, $P^* = 0.85$ and $d^* = 0.05$ is 523. This value was computed by the normal approximation and found to be 522.

TIED POPULATIONS

In computing the tables and graphs it was assumed that if two or more populations are tied for first place then one *of these* is selected by a chance device which assigns equal probability to each of them. The experimenter may want to select one of these contenders for first place by economic or other considerations. In most practical problems we may assume that such a selection is at random as far as the probability of a correct selection is concerned. Hence, it appears reasonable to use the tables in this paper without any corrections even when the rule for tied populations is altered in the manner described above.

It is interesting to note that in the yield problem the experimenter may settle the question of ties for first place by taking more observations until the tie is broken. However, in the life-testing problem he may not settle ties by letting the test run beyond time T since the best process for time T is not necessarily the best for a time greater than T.

In some applications when there are two or more populations tied for first place, the experimenter may prefer to recommend all these contenders for first place rather than select one of them by a chance device. In this case we shall agree to call the selection a correct one if the recommended set contains the best population (or, when $p_{[1]} = p_{[2]}$, if the recommended set contains at least one of the best populations). Exact tables for the procedure so altered have not been computed. However, if the value of n is large and this rule for tied populations is used, then the experimenter may reduce the tabled values by an amount equal to the largest integer contained in $1/d^*$. For example, using the above rule for tied populations for the case $k = 2$, $P^* = 0.99$, $d^* = 0.30$, the tabled value 29 can be reduced by 3 giving the result 26.

ALTERNATIVE SPECIFICATION

If the experimenter has some à priori knowledge about the processes, then he will prefer to specify the following *three* quantities in order to determine the number n of observations he is required to take from each process.

Specification: He specifies $p^*_{[1]}$ and $p^*_{[2]}$ $(0 \leqq p^*_{[2]} \leqq p^*_{[1]} \leqq 1)$ in the neighborhood of his estimate of the probabilities associated with his processes. He also specifies a probability P^* $(0 \leqq P^* < 1)$ that he would like to guarantee of making a correct selection whenever the true $p_{[1]} \geqq p^*_{[1]}$ and the true $p_{[2]} \leqq p^*_{[2]}$. (12)

TABLE VI — NUMBER OF UNITS REQUIRED PER PROCESS TO GUARANTEE A PROBABILITY OF P^* OF SELECTING THE BETTER OF TWO BINOMIAL PROCESSES WHEN THE TRUE $p_{[1]} \geqq p^*_{[1]}$ AND THE TRUE $p_{[2]} \leqq p^*_{[2]}$. (ALTERNATIVE SPECIFICATION, $k = 2$)

P^*	$p^*_{[1]} = 0.75$ $p^*_{[2]} = 0.60$	$p^*_{[1]} = 0.95$ $p^*_{[2]} = 0.80$	$p^*_{[1]} = 0.90$ $p^*_{[2]} = 0.80$	$p^*_{[1]} = 0.85$ $p^*_{[2]} = 0.80$	$p^*_{[1]} = 0.95$ $p^*_{[2]} = 0.90$
0.50	1	1	1	1	1
0.60	2	2	3	9	6
0.75	10	6	13	53	27
0.80	14	8	19	83	40
0.85	21	11	28	124	60
0.90	32	16	42	189	91
0.95	53	25	68	312	149
0.99	106	49	135	623	298

TABLE VII — NUMBER OF UNITS REQUIRED PER PROCESS TO GUARANTEE A PROBABILITY OF P^* OF SELECTING THE BEST OF FOUR BINOMIAL PROCESSES WHEN THE TRUE $p_{[1]} \geqq p^*_{[1]}$ AND THE TRUE $p_{[2]} \leqq p^*_{[2]}$. (ALTERNATIVE SPECIFICATION, $k = 4$)

P^*	$p^*_{[1]} = 0.75$ $p^*_{[2]} = 0.60$	$p^*_{[1]} = 0.95$ $p^*_{[2]} = 0.80$	$p^*_{[1]} = 0.90$ $p^*_{[2]} = 0.80$	$p^*_{[1]} = 0.85$ $p^*_{[2]} = 0.80$	$p^*_{[1]} = 0.95$ $p^*_{[2]} = 0.90$
0.50	7	4	10	42	21
0.60	14	8	18	79	39
0.75	28	14	37	168	80
0.80	35	18	46	211	101
0.85	45	22	59	268	128
0.90	59	28	77	350	171
0.95	83	39	107	493	239
0.99	139	65	182	831	399

Again we can rewrite the specification that the experimenter wants to satisfy in the simple form

$$P_{cs} \geqq P^* \quad \text{for} \quad p_{[1]} \geqq p_{[1]}^* \quad \text{and} \quad p_{[2]} \leqq p_{[2]}^* \tag{13}$$

Tables VI and VII give the number of observations required per process for several selected triplets of specified constants $(p_{[1]}^*, p_{[2]}^*, P^*)$ when $k = 2$ and $k = 4$. These results are also given in graphical form in Figs. 15 and 16.

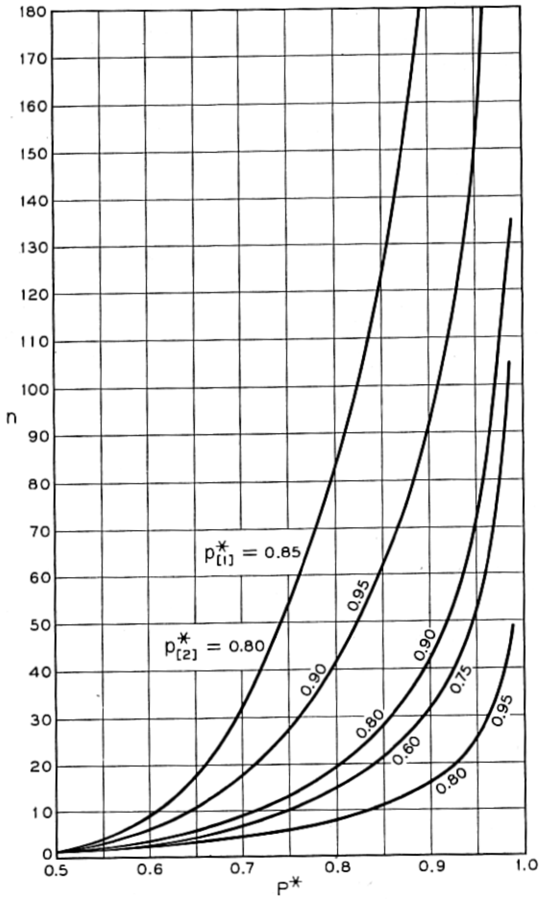

Fig. 15 — Number of units required per process to guarantee a probability of P^* of selecting the better of two binomial processes when the true $p_{[1]} \geqq p_{[1]}^*$ and the true $p_{[2]} \leqq p_{[2]}^*$.

For example, on the basis of past experience the experimenter may estimate that the probabilities associated with his $k = 4$ processes are all in the neighborhood of 0.60. This constitutes his à priori knowledge. He may then decide that he would like to make a correct selection with probability $P^* = 0.85$ when the best process has a yield of at least 75 per cent and all the others have a yield of at most 60 per cent. Entering column 1 of Table VII we find that $n = 45$ observations per process are required.

It is much more difficult to furnish tables for the alternative specifica-

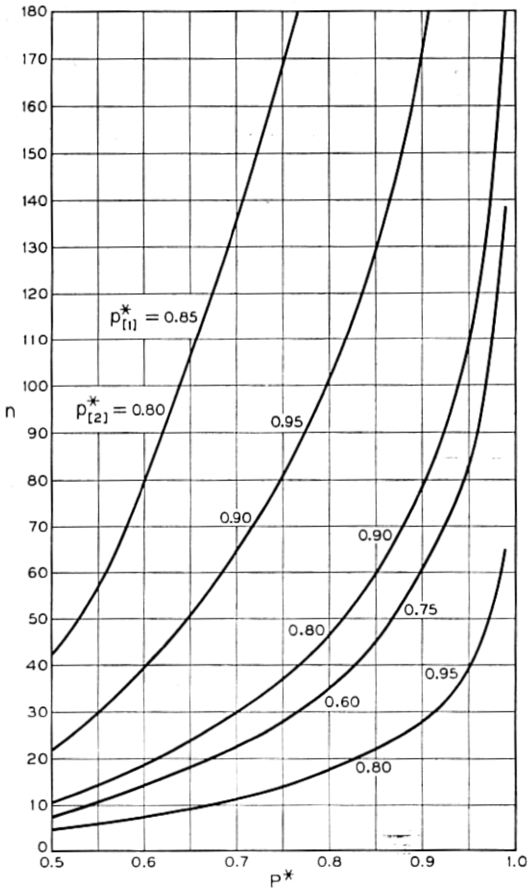

Fig. 16 — Number of units required per process to guarantee a probability of P^* of selecting the best of four binomial processes when the true $p_{[1]} \geqq p^*_{[1]}$ and the true $p_{[2]} \leqq p^*_{[2]}$.

tion since there is an extra parameter to vary and the appropriate tables for the normal approximation are not available.

In the computation of these probabilities the least favorable configuration

$$p_{[1]} = p_{[1]}^* \quad \text{and} \quad p_{[2]}^* = p_{[2]} = p_{[3]} = \cdots = p_{[k]} \tag{14}$$

was used. It follows from Appendix II that if the probability of a correct selection is at least P^* when (14) holds, then it will also be at least P^* when

$$p_{[1]} \geqq p_{[1]}^* \quad \text{and} \quad p_{[2]}^* \geqq p_{[2]} \geqq p_{[3]} \geqq \cdots \geqq p_{[k]} \tag{15}$$

For small values of n, exact calculations were carried out. A typical exact calculation is shown in Appendix IV. The approximations used for large n are given in Appendix III.

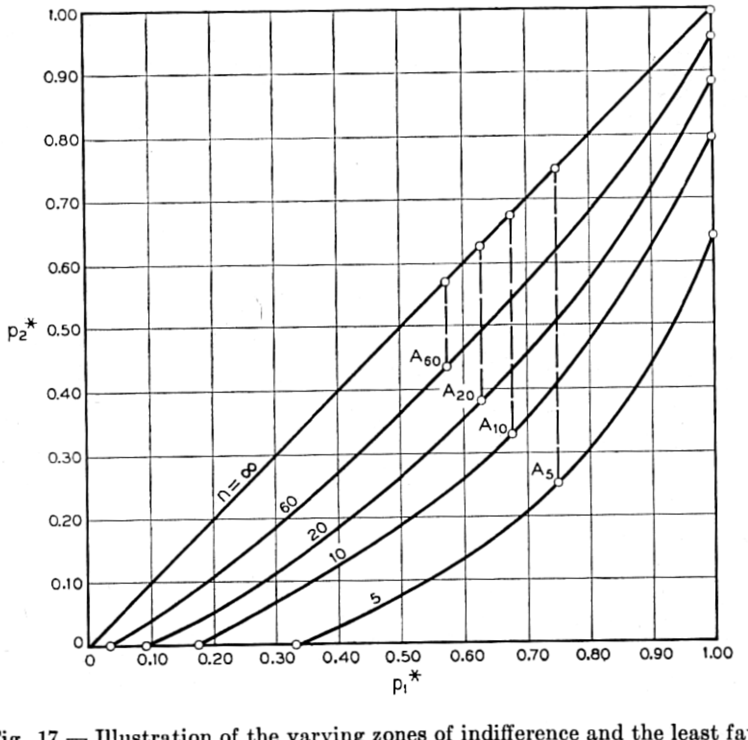

Fig. 17 — Illustration of the varying zones of indifference and the least favorable configuration for $k = 4$ and $P^* = 0.85$. (For $n \geqq 5$ the longest vertical segment occurs at the point A_n where the abcissa and the ordinate are symmetrical about 0.5)

COMPARISON OF THE TWO SPECIFICATIONS

It should be pointed out that for a given k the same value of n would satisfy the specification for different specified triplets

$$P^*, p^*_{[1]}, p^*_{[2]}$$

For example with $k = 4$, $P^* = 0.85$ and n fixed we could vary $p^*_{[1]}$ in the alternative specification and compute for each $p^*_{[1]}$ the corresponding largest value of $p^*_{[2]}$ such that the specification $(P^*, p^*_{[1]}, p^*_{[2]})$ is satisfied. This is shown in Fig. 17 for $n = 5, 10, 20$ and 60. The vertical distance in Fig. 17 between the appropriate curve and the 45° line $(n = \infty)$ is the length of the indifference zone $(p^*_{[1]}, p^*_{[2]})$. The indifference zone widens in the center and narrows at both ends. In fact we find just as in the original specification that for n greater than (say) 4 the indifference zone is widest when $p^*_{[1]}$ and $p^*_{[2]}$ are symmetrical about 0.5. It is clear that the two specifications would coincide if we took d^* in the original specification and set $p^*_{[1]} = \frac{1}{2}(1 + d^*)$, $p^*_{[2]} = \frac{1}{2}(1 - d^*)$

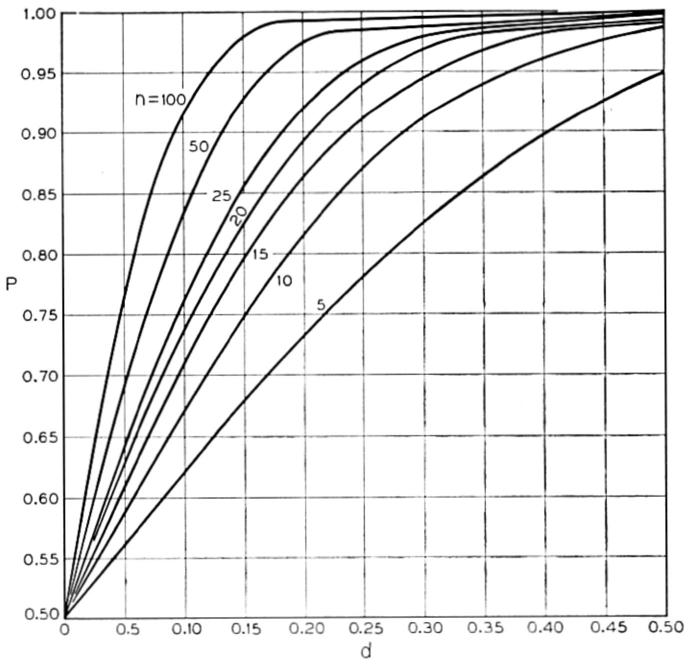

Fig. 18 — Probability of a correct selection as a function of the true difference $d = p_{[1]} - p_{[2]}$ under the least favorable configuration for $k = 2$ and selected values of n.

in the alternative specification. We shall be interested in comparing the alternative specification (P^*, $p^*_{[1]}$, $p^*_{[2]}$) with the original specification with the same P^* and with d^* set equal to $p^*_{[1]} - p^*_{[2]}$. It is clear that the value of n required for the original specification will always be larger.

The original specification is simpler and is preferable to the alternative specification when little or nothing is known about the processes on test, but the price that has to be paid for ignorance is an increase in the

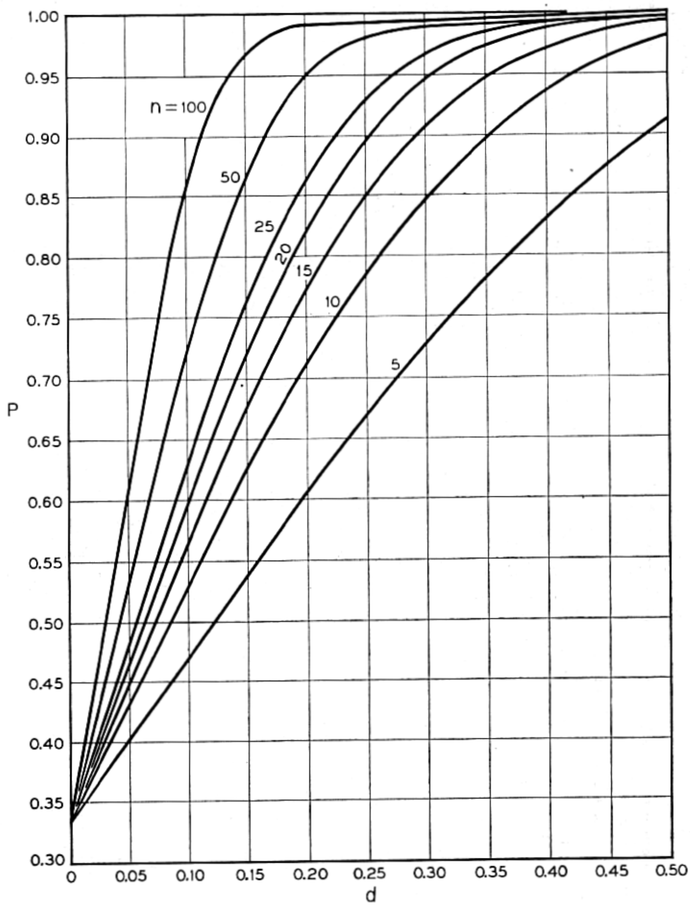

Fig. 19 — Probability of a correct selection as a function of the true difference $d = p_{[1]} - p_{[2]}$ under the least favorable configuration for $k = 3$ and selected values of n.

required number of observations. In the example of the preceding section the value of n required for the alternative specification is 45 as compared to 51 observations per process required to satisfy the original specification with the same P^* and with $d^* = p^*_{[1]} - p^*_{[2]}$. Here the saving is only moderate. The saving will be much larger if $p^*_{[1]}$ and

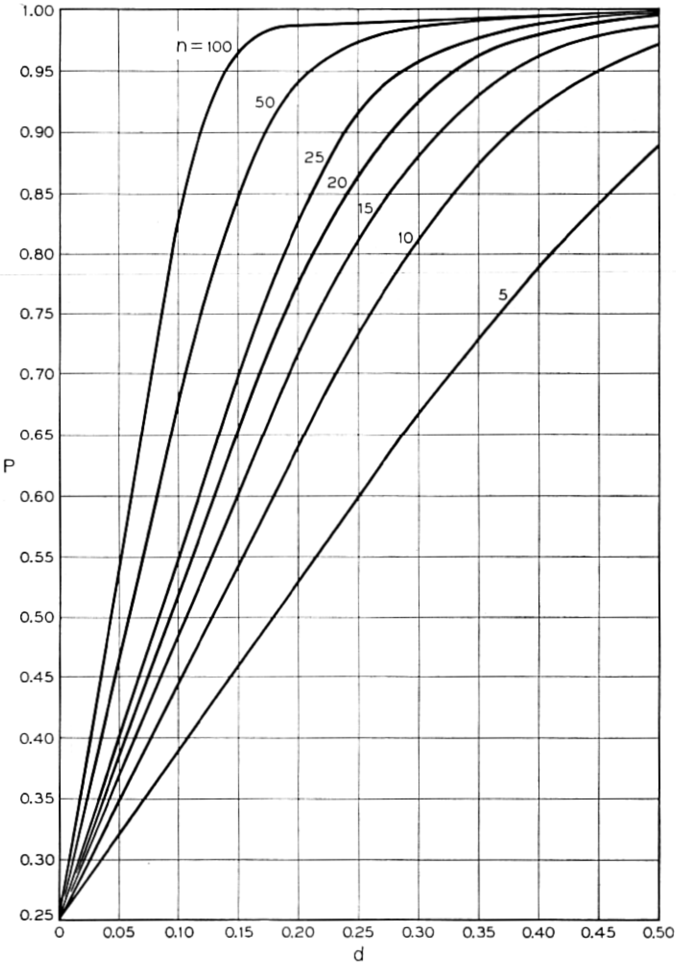

Fig. 20 — Probability of a correct selection as a function of the true difference $d = p_{[1]} - p_{[2]}$ under the least favorable configuration for $k = 4$ and selected values of n.

$p^*_{[2]}$ are further from 0.5 and d^* is small. For example, for $k = 4$, $P^* = 0.95$, $p^*_{[1]} = 0.95$ and $p^*_{[2]} = 0.90$ the value of n required for the alternative specification is 239 as compared to 850 observations per process required to satisfy the original specification with the same P^* and with $d^* = 0.05$. The alternative specification is justified on the basis of à priori or previous information about the approximate values of the p's.

REVERSING THE TABLES

The experimenter may wish to use the tables of this paper in reverse. For example, if n is fixed and d^* is specified by the experimenter, then by using the appropriate table he can find the probability of a correct selection that is guaranteed for $d \geqq d^*$; i.e., a greatest lower bound to the probability of a correct selection for $d \geqq d^*$. This process of reversing the given values and the values to be computed can most easily be carried out on graphs. For example, the above problem of finding the guaranteed probability of a correct selection given d^* and n is most easily carried out on Figs. 18, 19, and 20.

Appendix I

MODIFICATION OF THE ORIGINAL SPECIFICATION

The same value of n will, of course, satisfy the specification for different pairs of specified values (d^*, P^*). From a purely mathematical point of view it is not necessary that d^* should be the *smallest* difference for which the experimenter desires to make a correct selection. For example, if $k = 3$ the experimenter could specify any one of the four pairs (0.10, 0.60), (0.25, 0.90), (0.30, 0.95) or (0.40, 0.99) and obtain the same result, namely $n = 20$. The experimenter may prefer to specify the *curve* or set of points corresponding to a fixed n. Several such curves are given in Figs. 18, 19, and 20 for $k = 2$, 3, and 4, respectively. The experimenter would decide in advance on some property of the curve that he considers desirable and from the appropriate figure he could find the curve with the smallest n-value that satisfies the desired property.

The main point of the above paragraph is to point out that the original specification in the body of the paper is one particular way, but not the only way, of stating a specification that will determine a value of n. The only criterion for a good way to state the specification is that the experimenter should be able to bring his best judgment (or best guesses) to bear on the quantities that have to be specified in advance.

Appendix II

MONOTONICITY PROPERTIES

We shall prove that for any fixed d $(0 \leqq d \leqq 1)$ the probability P_{CS} of a correct selection is smaller for the configuration:

$$p_{[1]} - d = p_{[2]} = p_{[3]} = \cdots = p_{[k]} \tag{A1}$$

than for any configuration given by

$$p_{[1]} - d \geqq p_{[2]} \geqq p_{[3]} \geqq \cdots \geqq p_{[k]} \tag{A2}$$

where $p_{[1]}$ is considered fixed and the $p_{[i]}$ $(i \geqq 2)$ are variables. In other words, for fixed $p_{[1]}$ the probability P_{CS} is a strictly increasing function of each of the differences

$$p_{[1]} - p_{[i]} \qquad (i \geqq 2)$$

We shall need the following lemma.

Lemma 1: For any pair of integers x, n $(0 \leqq x \leqq n)$ *and any* θ $(0 \leqq \theta \leqq 1)$, *not depending on* p, *the function*

$$H(x; p, \theta) = \sum_{j=0}^{x-1} C_j^n p^j (1 - p)^{n-j} + \theta C_x^n p^x (1 - p)^{n-x} \tag{A3}$$

is a decreasing function of p *over the unit interval* $(0 \leqq p \leqq 1)$. *Moreover, it is strictly decreasing unless* $(x = 0$ *and* $\theta = 0)$ *or* $(x = n$ *and* $\theta = 1)$.

Proof: Differentiating (A3) with respect to p gives after telescoping terms

$$(\theta - 1)xC_x^n p^{x-1}(1 - p)^{n-x} - \theta(n - x)C_x^n p^x (1 - p)^{n-x-1} \tag{A4}$$

which is negative for $0 < p < 1$ unless $(x = 0$ and $\theta = 0)$ or $(x = n$ and $\theta = 1)$. Since H is continuous in p at $p = 0$ and $p = 1$ the lemma follows.

Let $X_{(i)}$ denote the chance number of successes that arises from the binomial process associated with

$$p_{[i]} \qquad (i = 1, 2, \cdots, n)$$

the value of the integer n is assumed to be fixed throughout this discussion and it will usually not be listed as an argument. The probability P_{CS} of a correct selection for any configuration with $p_{[1]} > p_{[2]}$ is given by the expression on the top of the next page.

$$P_{CS} = P\{X_{(i)} < X_{(1)} \quad \text{for} \quad i \geqq 2\} + \tfrac{1}{2} \sum_{\alpha=2}^{k} P\{X_{(\alpha)} = X_{(1)} \quad \text{and}$$
$$X_{(i)} < X_{(1)} \quad \text{for } i \geqq 2, i \neq \alpha\} + \cdots \tag{A5}$$
$$+ \frac{1}{k} P\{X_{(1)} = X_{(2)} = \cdots = X_{(k)}\}$$

It will be necessary to write the P_{CS} for any configuration with $p_{[1]} > p_{[2]}$ in another form which is more useful for the purpose at hand. Corresponding to any binomial chance variable X (which takes on integer values from 0 to n) we define a "Continuous Binomial" chance variable Y by letting Y be *uniformly* distributed in the interval $(j - \frac{1}{2}, j + \frac{1}{2})$ with the same total probability in this interval as the ordinary binomial assigns to the integer j, namely

$$C_j^n p^j (1 - p)^{n-j} \qquad (j = 0, 1, \cdots, n)$$

We will now show that the probability P_{CS} of a correct selection is unaltered if we replace each of the k discrete binomials by its corresponding continuous binomial. Let $Y_{(i)}$ denote the continuous binomial (CB) chance variable associated with $p_{[i]}$ and let $y_{(i)}$ denote any value it can take on. Let $X_{(i)}$ denote the nearest integer to $Y_{(i)}$ and let $x_{(i)}$ denote the nearest integer to $y_{(i)}$ $(i = 1, 2, \cdots, k)$. Then $X_{(i)}$ is a discrete binomial (DB) with the same parameters $(p_{[i]}, n)$. Let

$$g(x, p) = C_x^n p^x (1 - p)^{n-x} \qquad (x = 0, 1, \cdots, n)$$

Then the *density* $g(y, p)$ of the continuous binomial (disregarding the half-integers) is given by $g(y, p) = g(x, p)$ where x is the nearest integer to y.

For two continuous binomials (i.e., $k = 2$) the probability P_{CS} of a correct selection for any configuration with $p_{[1]} > p_{[2]}$ is given by

$$P_{CS}(CB) = \int_{-1/2}^{n+1/2} P\{Y_{(2)} < y_{(1)}\} g(y_{(1)}; p_{[1]})\, dy_{(1)} \tag{A6}$$

$$= \sum_{x(1)=0}^{n} \int_{x(1)-1/2}^{x(1)+1/2} P\{Y_{(2)} < y_{(1)}\} g(y_{(1)}; p_{[1]})\, dy_{(1)} \tag{A7}$$

Within any interval $(x_{(1)} - \frac{1}{2}, x_{(1)} + \frac{1}{2})$ we have

$$P\{Y_{(2)} < y_{(1)}\} = P\{X_{(2)} < x_{(1)}\} + P\{X_{(2)} = x_{(1)}\} P\{Y_{(2)} < y_{(1)} \mid X_{(2)} = x_{(1)}\} \tag{A8}$$

$$= P\{X_{(2)} < x_{(1)}\} + \tfrac{1}{2} P\{X_{(2)} = x_{(1)}\} \tag{A9}$$

which depends only on $x_{(1)}$. Hence from (A7)

$P_{CS}(CB)$

$$= \sum_{x_{(1)}=0}^{n} [P\{X_{(2)} < x_{(1)}\} + \tfrac{1}{2}P\{X_{(2)} = x_{(1)}\}]P\{X_{(1)} = x_{(1)}\} \tag{A10}$$

$$= P\{X_{(2)} < X_{(1)}\} + \tfrac{1}{2}P\{X_{(2)} = X_{(1)}\} \tag{A11}$$

$$= P_{CS}(DB) \tag{A12}$$

The above is easily generalized to hold for any $k > 2$. The details of this generalization are omitted. For general k this equality holds not only for the important special case $p_{[1]} > p_{[2]}$ but also for the more general case (2) for any $t < k$. Since the latter result is not needed here, the proof is omitted.

If we let $G(y;p)$ denote the c.d.f. of the continuous binomial then lemma 1 can be restated in the following form.

Lemma 2: For any integer n and any y, the function $G(y;p)$ is a non-increasing function of p. In particular, for $-\frac{1}{2} < y < n + \frac{1}{2}$ it is a strictly decreasing function of p.

Proof: For any y, set $x = x(y)$ and $\theta = \theta(y)$ equal to the integer part and the fractional part of $(y + \frac{1}{2})$, respectively. Then for any y we have the identity in p

$$G(y;p) \equiv H(x;p, \theta) \qquad (0 \leqq p \leqq 1) \tag{A13}$$

For any y_0 such that $-\frac{1}{2} < y_0 < n + \frac{1}{2}$ we have $0 \leqq x(y_0) \leqq n$ and $0 \leqq \theta(y_0) \leqq 1$. The inverse function $y(x,\theta) = x + \theta - \frac{1}{2}$ is a *single-valued* function of the pair (x,θ); the two particular pairs (0,0) and $(n,1)$ correspond to the *unique* values $y = -\frac{1}{2}$ and $y = n + \frac{1}{2}$, respectively. Hence the pair $[x(y_0), \theta(y_0)]$ must be different from these two particular pairs above since it corresponds *only* to y_0 which is in the *interior* of the interval $(-\frac{1}{2}, n + \frac{1}{2})$. Lemma 2 then follows from lemma 1 and the fact that $G(y;p)$ is identically zero in p for $y \leqq -\frac{1}{2}$ and identically one in p for $y \geqq n + \frac{1}{2}$.

The probability P_{CS} of a correct selection for k discrete or k continuous binomials for any configuration with $p_{[1]} > p_{[2]}$ can now be written as

$$P_{CS} = \int_{-1/2}^{n+1/2} \left[\prod_{i=2}^{k} P\{Y_{(i)} < y_{(1)}\}\right] g(y_{(1)}\,;p_{[1]})\,dy_{(1)} \tag{A14}$$

$$= \int_{-1/2}^{n+1/2} \left[\prod_{i=2}^{k} G(y;p_{[i]})\right] g(y;p_{[1]})\,dy. \tag{A15}$$

Clearly if any one or more of the $p_{[i]}$ $(i \geqq 2)$ decreases, holding $p_{[1]}$ fixed, then it follows from lemma 2 that the right member of (A15) is *strictly* increasing, i.e., for fixed $p_{[1]}$ the P_{CS} is a strictly increasing function of each of the differences $p_{[1]} - p_{[i]}$ $(i \geqq 2)$ as was to be shown.

It follows from the above result that in searching for a least favorable configuration among all those in which the experimenter wants his specification satisfied we may restrict our attention to those of the form (A1). Moreover, we may set d in (A1) equal to d^* since, for $d > d^*$ and fixed $p_{[1]}$, the difference $d - d^*$ may be added to each $p_{[i]}$ $(i \geqq 2)$ and the probability of a correct selection is increased. Then (A15) reduces to

$$P_{CS} = \int_{-1/2}^{n+1/2} G^{k-1}(y;p_{[1]} - d^*)g(y;p_{[1]})\,dy \tag{A16}$$

It was shown in the section on the least favorable configuration that there is a value $p_{[1]}^L$ of $p_{[1]}$ which when substituted in (A16) gives the minimum value P_{CS}^L of P_{CS}.

We can now prove the following result in which $p_{[1]}$ is not fixed. For any specified pair $p_{[2]}^* \leqq p_{[1]}^*$ the probability P_{CS} of a correct selection is smaller for the configuration

$$p_{[1]} = p_{[1]}^*; \qquad p_{[2]}^* = p_{[2]} = p_{[3]} = \cdots = p_{[k]} \tag{A17}$$

than for any configuration given by

$$p_{[1]} \geqq p_{[1]}^*; \qquad p_{[2]}^* \geqq p_{[2]} \geqq p_{[3]} \geqq \cdots \geqq p_{[k]} \tag{A18}$$

This is shown by considering two separate steps.

The first step is to increase $p_{[1]}$ holding all the other p's fixed at $p_{[2]}^*$. For any arbitrary set of values of $p_{[i]}$ with $p_{[1]} > p_{[2]}$ the probability of a correct selection can be written as

$$P_{CS} = \sum_{j=2}^{k} \int_{-1/2}^{n+1/2} \left[\prod_{i=2,\, i\neq j}^{k} G(y,p_{[i]}) \right] [1 - G(y,p_{[1]})]g(y,p_{[j]})\,dy \tag{A19}$$

by adding the probabilities that

$$Y_{(1)} > Y_{(j)} > \min\{Y_{(2)}, \cdots, Y_{(j-1)}, Y_{(j+1)}, \cdots, Y_{(k)}\}$$

for

$$j = 2, 3, \cdots, k$$

For

$$p_{[1]} > p_{[2]} = p_{[3]} = \cdots = p_{[k]} = p_{[2]}^*$$

this reduces to

$$P_{CS} = (k - 1) \int_{-1/2}^{n+1/2} [1 - G(y;p_{[1]})]G^{k-2}(y;p_{[2]}^*)g(y,p_{[2]}^*)\,dy \tag{A20}$$

This result can also be obtained by starting with (A16) with

$$p_{[2]} = \cdots = p_{[k]} = p_{[1]} - d^* = p^*_{[2]}$$

and integrating by parts. It is clear from (A20) that for fixed $p_{[i]}$ $(i \geqq 2)$ the P_{CS} is an increasing function of $p_{[1]}$ and is indeed strictly increasing for $p_{[1]}$ in the unit interval.

The second step is to hold $p_{[1]}$ fixed and to decrease the values of $p_{[i]}$ $(i \geqq 2)$. This increases the probability of a correct selection by our previous result above. This proves the monotonicity property for the alternative specification.

Appendix III

LARGE SAMPLE THEORY — ORIGINAL SPECIFICATION

For $p_{[1]} > p_{[2]}$ the probability of a correct selection satisfies the inequalities

$$P\{X_{(1)} > X_{(i)}\ (i = 2, 3, \cdots, k)\} < P_{CS}$$
$$< P\{X_{(1)} \geqq X_{(i)}\ (i = 2, 3, \cdots, k)\} \quad \text{(A21)}$$

unless $p_{[1]} = 1$ *and* $p_{[2]} = 0$ in which case equality signs hold since the three quantities above are all unity. Letting $q_{[1]} = 1 - p_{[1]}$, we can write the left member of (A21) as

$$P\left\{Z_i > \frac{-d^*\sqrt{n}}{\sqrt{p_{[1]}q_{[1]} + (p_{[1]} - d^*)(q_{[1]} + d^*)}}\ (i = 1, 2, \cdots, k-1)\right\} \quad \text{(A22)}$$

where

$$Z_i = \frac{X_{(1)} - X_{(i+1)} - nd^*}{\sqrt{n[p_{[1]}q_{[1]} + (p_{[1]} - d^*)(q_{[1]} + d^*)]}}\ (i = 1, 2, \cdots, k-1) \quad \text{(A23)}$$

For the configuration (A1) with $d = d^*$ the chance variables Z_i tend to normal chance variables $N(0,1)$ with zero mean and unit variance as $n \to \infty$. We have purposely omitted any continuity correction in (A22) in order to get a better approximation for the smaller values of n.

To derive the least favorable configuration for large n we can restrict our attention to those configurations given by (A1) with $d = d^*$. The quantity $p^L_{[1]}$, which minimizes (A22), is obtained by maximizing the expression in (A22)

$$Q(p) = p(1 - p) + (p - d^*)(1 - p + d^*) \quad \text{(A24)}$$
$$= -2p^2 + 2(1 + d^*)p - d^*(1 + d^*) \quad \text{(A25)}$$

The derivative of $Q(p)$ vanishes at

$$p_{[1]}^{L} = \tfrac{1}{2}(1 + d^*); \qquad q_{[1]}^{L} = \tfrac{1}{2}(1 - d^*) \tag{A26}$$

which gives the symmetric configuration. Clearly this value of p gives to $Q(p)$ its maximum value, $\frac{1}{2}(1 - d^{*2})$. This proves that the symmetrical configuration is least favorable in the limit as $n \to \infty$.

Under the configuration (A1) with $d = d^*$ and $n \to \infty$ the distribution of the chance variables Z_i ($i = 1, 2, \cdots, k - 1$) approaches a joint multivariate normal distribution with zero means, unit variances and correlations given by

$$\rho(Z_i Z_j) = \frac{p_{[1]} q_{[1]}}{p_{[1]} q_{[1]} + (p_{[1]} - d^*)(q_{[1]} + d^*)} \qquad (i \neq j) \tag{A27}$$

which do not depend on n. For the symmetric configuration this reduces to the simple form

$$\rho(Z_i Z_j) = \tfrac{1}{2} \qquad (i \neq j). \tag{A28}$$

This is precisely the case which arises in [1] and consequently the tables in [1] can be used for our problem when (the answer) n is large. The constants $C = C(P^*, k)$ tabulated in [1] solve the equation

$$P\left\{Z_i > -\frac{C}{\sqrt{2}}\,(i = 1, 2, \cdots, k - 1)\right\} = P^* \tag{A29}$$

for standard normal chance variables Z_i satisfying (A28). If we equate $C/\sqrt{2}$ and the corresponding member of (A22), then we obtain for the symmetric configuration

$$\frac{C}{\sqrt{2}} \cong \frac{d^* \sqrt{n}}{\sqrt{\frac{1}{2}(1 - d^{*2})}} \tag{A30}$$

or solving for n and letting $B = \frac{1}{4}C^2$ this yields the large sample normal approximation

$$n \cong \frac{B}{d^{*2}}(1 - d^{*2}) \tag{A31}$$

Since d^* is usually small when n is large and since the solution in (A31) is usually somewhat smaller than the true value, then it is of interest to examine the simpler approximation

$$n \cong \frac{B}{d^{*2}} \tag{A32}$$

which is greater than the result in (A31). This is called the straight line approximation since it plots as a straight line on log-log paper as shown in Figs. 2 through 5. As $d^* \to 0$ both the normal approximation and the true value are asymptotically equivalent to the straight line approximation.

The normal approximation to the probability of a correct selection can also be written in another form similar to (A16) which is actually more useful for numerical calculations. The left member of (A21) can be written as

$$\sum_{w_1} \left[\prod_{i=2}^{k-1} P \left\{ W_i < \frac{W_1 \sqrt{p_{[1]} q_{[1]}} + (p_{[1]} - p_{[i]}) \sqrt{n}}{\sqrt{p_{[i]} q_{[i]}}} \right\} \right] P \{ W_1 = w_1 \} \tag{A33}$$

where

$$W_i = \frac{X_{(i)} - np_{[i]}}{\sqrt{np_{[i]} q_{[i]}}} \qquad (i = 1, 2, \cdots, k) \tag{A34}$$

and w_1 is the same function of $x_{(1)}$ as W_1 is of $X_{(1)}$. The outside summation in (A33) is over the values taken on by w_1 as $x_{(1)}$ runs from 0 to n. As $n \to \infty$ the expression in (A33) approaches

$$P_{CS} \cong \int_{-\infty}^{\infty} \left[\prod_{i=2}^{k-1} F \left(\frac{w \sqrt{p_{[1]} q_{[1]}} + (p_{[1]} - p_{[i]}) \sqrt{n}}{\sqrt{p_{[i]} q_{[i]}}} \right) \right] f(w)\, dw \tag{A35}$$

where $f(t)$ is the standard normal density and $F(t)$ is the standard normal c.d.f. For the symmetric configuration, which is least favorable for large n, (A35) reduces to

$$P_{CS}^L \cong \int_{-\infty}^{\infty} F^{k-1} \left(w + \frac{2d^* \sqrt{n}}{\sqrt{1 - d^{*2}}} \right) f(w)\, dw \tag{A36}$$

A straightforward integration by parts gives the alternative form

$$P_{CS}^L \cong (k-1) \int_{-\infty}^{\infty} \left[1 - F \left(w - \frac{2d^* \sqrt{n}}{\sqrt{1 - d^{*2}}} \right) \right] F^{k-2}(w) f(w)\, dw \tag{A37}$$

which corresponds to (A20).

A simple method for computing such integrals based on Hermite polynomials is described by Salzer, Zucker, and Capuano.[3]

LARGE SAMPLE THEORY — ALTERNATIVE SPECIFICATION

The expression corresponding to (A22) for the alternative specification is

$$P \left\{ Z_i > \frac{-(p_{[1]}^* - p_{[2]}^*) \sqrt{n}}{\sqrt{p_{[1]}^* q_{[1]}^* + p_{[2]}^* q_{[2]}^*}} \qquad (i = 1, 2, \cdots, k-1) \right\} \tag{A38}$$

which is already written for the least favorable configuration. The tables[2] are not immediately applicable since the correlations

$$\rho(Z_i Z_j) = \frac{p^*_{[1]} q^*_{[1]}}{p^*_{[1]} q^*_{[1]} + p^*_{[2]} q^*_{[2]}} \qquad (i \neq j) \qquad \text{(A39)}$$

are not, in general, equal to $\frac{1}{2}$. In the cases treated in Tables VI and VII, $p^*_{[2]} \geqq 0.5$ and hence $p^*_{[1]} q^*_{[1]} < p^*_{[2]} q^*_{[2]}$ so that the correlations (A39) are all less than $\frac{1}{2}$. It was found that linear interpolation on the required value of n between the results for $\rho = 0$ and $\rho = \frac{1}{2}$ gives moderately good results when n is large. The result for $\rho = \frac{1}{2}$ is given by

$$n \cong \lambda \frac{(p^*_{[1]} q^*_{[1]} + p^*_{[2]} q^*_{[2]})}{(p^*_{[1]} - p^*_{[2]})^2} \qquad \text{(A40)}$$

with $\lambda = 2B$ where B is given in Table V. The result for $\rho = 0$ is given by (A40) with $\lambda = \lambda_0^2$ where λ_0 is the solution of the equation

$$P\{Z > -\lambda_0\} = P^{*1/(k-1)} \qquad \text{(A41)}$$

which can easily be found from univariate normal probability tables. An explicit expression for the result of this linear interpolation is

$$P^L_{CS} \cong \frac{p^*_{[1]} q^*_{[1]} (4B - \lambda_0^2) + p^*_{[2]} q^*_{[2]} \lambda_0^2}{(p^*_{[1]} - p^*_{[2]})^2} \quad (\text{for } p^*_{[2]} \geqq 0.5) \qquad \text{(A42)}$$

The expressions for the probability of a correct selection for the alternative specification corresponding to (A36) and (A37) are

$$P^L_{CS} \cong \int_{-\infty}^{\infty} F^{k-1}(aw + b) f(w)\, dw \qquad \text{(A43)}$$

$$= a(k - 1) \int_{-\infty}^{\infty} \left[1 - F\left(\frac{w - b}{a}\right)\right] F^{k-2}(w) f(w)\, dw \qquad \text{(A44)}$$

where

$$a = \sqrt{\frac{p^*_{[1]} q^*_{[1]}}{p^*_{[2]} q^*_{[2]}}} > 0 \quad \text{and} \quad b = \frac{(p^*_{[1]} - p^*_{[2]})\sqrt{n}}{\sqrt{p^*_{[2]} q^*_{[2]}}} \geqq 0 \qquad \text{(A45)}$$

These expressions can also be evaluated by the method described by Salzer, Zucker and Capuano.[3]

Appendix IV

TYPICAL EXACT CALCULATION

A. Original Specification

The exact expression (A5) for the probability of a correct selection for any configuration simplifies if the configuration is least favorable. For any pair of integers (j,n) we define

$$b_{1j} = P\{X_{(1)} = j\} = C_j^n (p_{[1]}^L)^j (q_{[1]}^L)^{n-j} \qquad (0 \leqq j \leqq n) \quad \text{(A46)}$$

$$b_{2j} = P\{X_{(2)} = j\} = C_j^n (p_{[1]}^L - d^*)^j (q_{[1]}^L + d^*)^{n-j} \qquad (0 \leqq j \leqq n) \quad \text{(A47)}$$

$$B_{2j} = P\{X_{(2)} \leqq j\} \quad \text{(A48)}$$

Then the exact probability P_{CS}^L of a correct selection for the least favorable configuration can be written as

$$P_{CS}^L = \sum_{j=0}^{n} b_{1j} \sum_{i=0}^{k-1} \frac{C_i^{k-1}}{1+i} b_{2j}^i B_{2,j-1}^{k-1-i} \quad \text{(A49)}$$

where $B_{2,-1}$ is defined to be zero. Here, for each value of $X_{(1)}$, the letter i denotes the number of processes that tie with $X_{(1)}$ for first place and for any given value of i the conditional probability of a correct selection is $1/(1 + i)$. Taking $k = 4$ as a typical case, we can write (A49) more explicitly as

$$P_{CS}^L = \sum_{j=1}^{n} b_{1j} B_{2,j-1}^3 + \frac{3}{2} \sum_{j=1}^{n} b_{1j} b_{2j} B_{2,j-1}^2 + \sum_{j=1}^{n} b_{1j} b_{2j}^2 B_{2,j-1} + \frac{1}{4} \sum_{j=0}^{n} b_{1j} b_{2j}^3 \quad \text{(A50)}$$

If $n \geqq 10$ then we may use the symmetric configuration, i.e., we may set $p_{[1]}^L = \frac{1}{2}(1 + d^*)$, in computing from (A49) or (A50).

B. Alternative Specification

The probability P_{LS}^C of a correct selection for the alternative specification is the same as in (A49) and (A50) except that we now define

$$b_{ij} = P\{X_{(i)} = j\} = C_j^n (p_{[i]}^*)^j (q_{[i]}^*)^{n-j} (i = 1, 2) \quad \text{(A51)}$$

$$B_{2j} = P\{X_{(2)} \leqq j\} \quad \text{(A52)}$$

A typical exact calculation for $k = 4$, using (A50), (A51) and (A52) with

$$p_{[1]} = p_{[1]}^* = 0.75$$

and

$$p_{[2]} = p_{[3]} = p_{[4]} = p^*_{[2]}$$

is given in Table AI. Exact values for the individual and cumulative binomial probabilities were obtained from References 4, 5 and 6.

TABLE AI — CALCULATION OF THE P^L_{CS}

$p_{[1]} = p^*_{[1]} = 0.75$; $p_{[2]} = p_{[3]} = p_{[4]} = p^*_{[2]} = 0.60$; $k = 4$

j	$b_{1,j}$	$b_{2,j}$	$B_{2,j-1}$	$b^2_{2,j}$	$b^3_{2,j}$	$B^2_{2,j-1}$	$B^3_{2,j-1}$
0	0.00000	0.00010	—	0.00000	0.00000	—	—
1	0.00003	0.00157	0.00010	0.00000	0.00000	0.00000	0.00000
2	0.00039	0.01062	0.00168	0.00011	0.00000	0.00000	0.00000
3	0.00309	0.04247	0.01229	0.00180	0.00008	0.00015	0.00000
4	0.01622	0.11148	0.05476	0.01243	0.00139	0.00300	0.00016
5	0.05840	0.20066	0.16624	0.04026	0.00808	0.02764	0.00459
6	0.14600	0.25082	0.36690	0.06291	0.01578	0.13462	0.04939
7	0.25028	0.21499	0.61772	0.04622	0.00994	0.38158	0.23571
8	0.28157	0.12093	0.83271	0.01462	0.00177	0.69341	0.57741
9	0.18771	0.04031	0.95364	0.00162	0.00007	0.90943	0.86727
10	0.05631	0.00605	0.99395	0.00004	0.00000	0.98794	0.98196
Check totals....	1.00000	1.00000					

$$\sum_{j=1}^{10} b_{1,j}\, B^3_{2,j-1} = 0.44715$$

$$\frac{3}{2} \sum_{j=1}^{10} b_{1,j}\, b_{2,j}\, B^2_{2,j-1} = 0.08493$$

$$\sum_{j=1}^{10} b_{1,j}\, b^2_{2,j}\, B_{2,j-1} = 0.01464$$

$$\frac{1}{4} \sum_{j=0}^{10} b_{1,j}\, b^3_{2,j} = 0.00145$$

$$\text{Total} = \overline{0.54817} = P^L_{CS}$$

APPENDIX V

In this appendix it will be shown that for large values of k the value of n required to meet any fixed specification (d^*, P^*) is approximately equal to some constant multiple of ($\ln k$).

Let $n = n(k)$ denote the unique positive decimal solution of the equation

$$\int_{-\infty}^{\infty} F^{k-1}(w + b\sqrt{n}) f(w)\, dw = P^* \tag{A53}$$

where $f(w)$ and $F(w)$ are defined above, P^* and b are known constants with $1/k < P^* < 1$ and $b > 0$ and the argument k is a positive integer. Let ε be a (small) *fixed* number such that $0 < \varepsilon < \text{Min}(P^*, 1 - P^*)$. Then $\varepsilon < P^* - 1/k$ for sufficiently large k. Let $A = A(\varepsilon)$ be defined by

$$\int_{-A}^{A} f(w)\, dw = 1 - \varepsilon \tag{A54}$$

so that

$$0 < \int_{|w|>A} F^{k-1}(w + b\sqrt{n})f(w)\, dw \leqq \varepsilon \tag{A55}$$

for any integer $k \geqq 1$, any $n > 0$ and any $b > 0$. Let n' and n'' be the unique positive decimal solutions, respectively, of the equations

$$\int_{-A}^{A} F^{k-1}(w + b\sqrt{n'})f(w)\, dw = P^* - \varepsilon \tag{A56}$$

$$\int_{-A}^{A} F^{k-1}(w + b\sqrt{n''})f(w)\, dw = P^* \tag{A57}$$

where P^*, b and k are the same as in (A53). It follows from (A55), (A56) and (A57) that for any integer $k \geqq 1$

$$n' \leqq n \leqq n'' \tag{A58}$$

From (A54) and (A57) we have

$$\int_{-A}^{A} F^{k-1}(w + b\sqrt{n''})f_A(w)\, dw = \frac{P^*}{1-\varepsilon} \tag{A59}$$

where $f_A(w)$ is the density of the normal distribution, truncated at A and $-A$. The right hand member of (A59) is positive and less than unity since $\varepsilon < 1 - P^*$. Hence there exists a w_A with $|w_A| \leqq A$ such that

$$\int_{-A}^{A} F^{k-1}(w + b\sqrt{n''})f_A(w)\, dw = F^{k-1}(w_A + b\sqrt{n''}) \tag{A60}$$

Since w_A is bounded and n'' is large for large k we can use the well-known approximation

$$\begin{aligned} F^{k-1}(w_A + b\sqrt{n''}) &\cong \left[1 - \frac{\exp[-(w_A + b\sqrt{n''})^2/2]}{\sqrt{2\pi}(w_A + b\sqrt{n''})}\right]^{k-1} \\ &\cong \exp\left\{-\frac{(k-1)\exp[-(w_A + b\sqrt{n''})^2/2]}{\sqrt{2\pi}(w_A + b\sqrt{n''})}\right\} \end{aligned} \tag{A61}$$

where only the leading term is considered. Hence from (A59), (A60) and (A61)

$$-\ln\ln\left(\frac{1-\varepsilon}{P^*}\right)^{\sqrt{2\pi}} + \ln(k-1) \cong \tfrac{1}{2}(w_A + b\sqrt{n''})^2 + \ln(w_A + b\sqrt{n''}) \tag{A62}$$

Since w_A is bounded and $ln\sqrt{n''} = o(n'')$ it follows that for large k

$$n'' \cong (2/b^2)\ln(k-1) \cong C\ln k \tag{A63}$$

where C is a proportionality factor. Starting with (A54) and (A56) the same argument gives the same result as (A63) for n'. Hence, by (A58), the same result must hold for n.

ACKNOWLEDGMENT

The authors wish to thank R. B. Murphy, J. W. Tukey, E. L. Kaplan, S. S. Gupta, E. Bleicher, S. Monro, all of Bell Telephone Laboratories, and Prof. R. E. Bechhofer of Cornell University for helpful suggestions and constructive criticism in connection with this paper.

REFERENCES

1. Bechhofer, R. E., and Sobel, M., On a Class of Sequential Multiple Decision Procedures for Ranking Parameters of Koopman-Darmois Populations with Special Reference to Means of Normal Populations, in preparation.
2. Bechhofer, R. E., A Single-Sample Multiple Decision Procedure for Ranking Means of Normal Populations with Known Variances, Ann. Math. Stat., **25**, pp. 16–39, 1954.
3. H. E. Salzer, R. Zucker, and R. Capuano, Table of the Zeros and Weight Factors of the First Twenty Hermite Polynomials, Journal of Research of the N. B. S., **48**, pp. 111–116, 1952.
4. National Bureau of Standards, Tables of the Binomial Probability Distribution, App. Math. Series 6, 1950.
5. Ordnance Corps, Tables of the Cumulative Binomial Probabilities, ORDP 20-1, 1952.
6. Harvard Computation Laboratory, Tables of the Cumulative Binomial Probability Distribution, Harvard University Press, Cambridge, 1955.