DESIGN FOR THE CONTROL OF SELECTION BIAS[1]

By David Blackwell and J. L. Hodges, Jr.

University of California, Berkeley

0. Summary. Suppose an experimenter E wishes to compare the effectiveness of two treatments, A and B, on a somewhat vaguely defined population. As individuals arrive, E decides whether they are in the population, and if he decides that they are, he administers A or B and notes the result, until nA's and nB's have been administered. Plainly, if E is aware, before deciding whether an individual is in the population, which treatment is to be administered next, he may, not necessarily deliberately, introduce a bias into the experiment. This bias we call *selection bias*. We propose to investigate the extent to which a statistician S, by determining the order in which treatments are administered, and not revealing to E which treatment comes next until after the individual who is to receive it has been selected, can control this selection bias.

Thus a design d is a distribution over the set T of the $\binom{2n}{n}$ sequences of length $2n$ containing nA's and nB's. We shall measure the bias of a design by the maximum expected number of correct guesses which an experimenter can achieve, knowing d, attempting to guess the successive elements of a sequence $t \varepsilon T$ selected by d, and being told after each guess whether or not it is correct. The distribution of the number G of correct guesses depends both on d and on the prediction method p used by the experimenter. We shall consider particularly two designs, the *truncated binomial*, in which the successive treatments are selected independently with probability $\frac{1}{2}$ each until n treatments of one kind have occurred, and the *sampling* design, in which all $\binom{2n}{n}$ sequences are equally likely. We shall consider particularly two prediction methods, the *convergent* prediction, which predicts that treatment which has hitherto occurred less often, and the *divergent* prediction, which predicts that treatment which has hitherto occurred more often, except that after n treatments of one kind have been administered, the divergent prediction agrees with the convergent predictions that the other treatment will follow; when both treatments have occurred equally often, either method predicts A or B by tossing a fair coin, independently for each case of equality.

We find that among all designs, the truncated binomial minimizes the maximum expected number of correct guesses. For this design, the expected number of correct guesses is independent of the prediction method, and is

$$n + n\binom{2n}{n}\Big/ 2^{2n} \sim n + (n/\pi)^{1/2}$$

Received March 26, 1956; revised November 2, 1956.

[1] This investigation was supported (in part) by research grant from The National Institutes of Health, Public Health Service.

With the truncated binomial design, the variance in the number of correct guesses is largest for the divergence strategy and is

$$3n/2 - D - D^2/4 \sim (3\pi - 2)n/2\pi - 2(n/\pi)^{1/2},$$

where $D = n \binom{2n}{n} \Big/ 2^{2n-1}$, and is smallest for the convergence strategy, and is $n/2 - D^2/4 \sim (\pi - 1)n/2\pi$. For the sampling design, convergent prediction maximizes the expected number of correct guesses; this maximum is

$$n + 2^{2n-1} \Big/ \binom{2n}{n} - \frac{1}{2} \sim n + (\pi n/4)^{1/2}.$$

Finally we note that, if treatments are selected independently at random, bias of the kind we discuss disappears, but the treatment numbers can no longer be preassigned. Three such designs are considered: the *fixed total* design, in which the total number of treatments is a fixed number s, the *fixed factor* design, in which we continue until $1/X + 1/Y \leqq 2/n$, where X is the number of A treatments and Y is the number of B treatments administered, and the *fixed minimum* design, in which we continue until $\min(X, Y) = n$. For the fixed total design, we find that, for $s = 2n + 4$, $\Pr(1/X + 1/Y \leqq 2/n) \sim 0.955$ for large n; at the expense of 4 extra observations, we have a bias-free design whose variance factor will with probability 0.955 be smaller than that in which treatment numbers are preassigned. For the fixed factor design, the additional number of observations required to achieve the given precision has for large n the distribution of the square of a normal deviate. For the fixed minimum design, in which we guarantee precision for the estimated effect of each treatment, the expected number of additional observations is roughly $1.13\,(n)^{1/2}$.

1. Introduction. It is widely recognized that experiments intended to compare two or more treatments may yield biased results if the experimental subjects are selected with knowledge of the treatments they are to receive. Consider as illustration of experiment in cloud seeding. From a sequence of storms the meteorologist selects $2n$ storms deemed suitable for seeding. Of these, n are seeded and we compare the rainfalls they produce with those produced by the other n storms. If the meteorologist knows (or can guess), while considering the suitability of a storm, whether or not the storm will be seeded if he selects it, there exists the possibility that his selection will be biased.

We shall call this *selection bias*. It presents a serious problem when the trials constituting the experiment occur sequentially in time. If it were possible to collect at one time a block of as many subjects as there are treatments, a simple random assignment of treatments to subjects would dispose of the bias. But in many experiments potential subjects occur singly and must be dealt with when they arise. For example, in clinical trials it is often essential to treat the patient as soon as the illness is diagnosed—the physician cannot wait until he has a similar patient merely to permit randomization of the bias.

In some cases it is possible to eliminate selection bias by conducting the experiment in such a way that the person who selects the subjects is not otherwise involved in the experiment or is not able to discover which treatments have been applied. Again, it may be possible to define subject suitability with precision and to accept without conscious selection all subjects meeting the criteria. But oftentimes the exercise of judgment is essential if the treatments are to have a convincing test, and the best or only judges available are those most deeply involved in administering the treatments. Therefore we thought it interesting to see to what extent selection bias can be controlled through design of the experiment—i.e., through the statistician's strategy in choosing the sequence of treatments to be given to the subjects selected by the experimenter.

Admittedly, selection bias will usually operate subconsciously, but to sharpen the problem we imagine an experimenter E who is consciously seeking to produce biased experimental results, while the statistician S is attempting to prevent this. To fix the problem, suppose we wish to compare two treatments, say A and B. It is customary to decide in advance of the experiment how many subjects will receive each treatment, and it is also customary to assign equal numbers of subjects to the two treatments. While we shall return to this question below, at first let us suppose it given that each treatment will be administered to n subjects.

If E wishes to make it appear that A produces a greater response X than does B, and if he knows (or guesses) that S will assign treatment A to the next subject, then E will try to select a subject whose expected response $E(X)$ is high. Conversely, if E anticipates a B treatment, he will select a subject with low $E(X)$. The results of E's guesses and S's assignments can be displayed in a two-by-two table:

		Number of times when S assigns	
		A	B
and E guesses	A	α	$n - \beta$
	B	$n - \alpha$	β

Suppose that the treatment effects do not differ, but that when E anticipates an $A(B)$ treatment he selects a subject with expected response $\mu + \Delta(\mu - \Delta)$. Then the expected difference of treatment sums is

$$2\Delta(\alpha + \beta - n) = 2\Delta(G - n). \tag{1.1}$$

The quantity G is thus the total number of correct guesses. If he guesses at random, E would on the average be right half the time, giving $E(G - n) = 0$. His ability to bias the experiment depends on getting $E(G)$ above n.

In accordance with the foregoing analysis we now formulate our design problem as a two-person game. The game is played in $2n$ moves. On each move, each of the players E and S privately selects one of the letters A and B, with the restriction that exactly n of S's choices must be A. They then compare selections; if they agree S pays E one unit. The total payoff is G, and we wish to know S's minimax strategy for minimizing $E(G)$, i.e. the optimum design for controlling selection bias. The value of the game will indicate to what extent selection bias can be controlled through design.

2. The biases of three designs. Before giving the solution of our game, we shall illustrate the ideas by deriving the optimum strategies for E and the corresponding biases for three designs used in experimental work.

(i) A very common practice is the alternation of treatments, producing the treatment sequence $ABAB \cdots AB$ or $BABA \cdots BA$. While this design is exceedingly simple and does an optimum job of spreading the treatments over time, it is about the worst possible design from the standpoint of selection bias. Since E can correctly guess every treatment, $E(G) = 2n$. Even if, as is sometimes done, S selects one of the two patterns at random, E can guess all but the first trial and has half a chance for that, so that $E(G) = 2n - \frac{1}{2}$. (Exactly the same conclusions apply to the "Student" sandwich design $ABBAABBA \cdots ABBA$).

(ii) As just remarked, S can insure that E's expectation of correct guessing on the initial trial is only $\frac{1}{2}$ by simply choosing a treatment at random. Further, S cannot do any better than this, since E can guarantee himself half a chance, whatever S may do, by guessing at random. A similar analysis applies to the second trial, and to all trials until one treatment has been given to n subjects. At that point the requirement that each treatment be given to just n subjects takes over, and the remainder of the subjects must be given the unexhausted treatment. We shall refer to this as the *truncated binomial* design.

Suppose S has announced that he will adopt the truncated binomial design. What should E do, and how large can he make $E(G)$? In the tail of the experiment, consisting of the terminal sequence of trials having like treatment, E knows which treatment will be assigned, so he is sure to guess all of these correctly—let R denote the number of trials in the tail. We take advantage of the fact that $E(G)$ is independent of E's strategy except in the tail, and give to E a strategy which simplifies the calculation. Suppose E guesses A every time, except of course in a tail of B's. Then G must be at least n (since n A's are used), and may in addition contain a B tail. By symmetry, $E(G) = n + E(R)/2$. We must now discuss the distribution of R.

To calculate the probability that $R = r$, notice that this may occur in two ways: the nth A treatment, or the nth B treatment, is assigned on the $(2n - r)$th trial. These events have equal probability $\binom{2n-r-1}{n-1} \Big/ 2^{2n-r}$ according to the negative binomial distribution. Therefore R has a truncated negative binomial distribution,

$$\text{Pr}\,(R = r) = \binom{2n-r-1}{n-1} \Big/ 2^{2n-r-1}, \qquad r = 1, 2, \cdots, n. \tag{2.1}$$

By calculating $E(2n - R)$, it is easy to establish

$$E(R) = n\binom{2n}{n}\Big/2^{2n-1} \approx \frac{2(n)^{1/2}}{\pi^{1/2}} - \frac{1}{4(\pi n)^{1/2}} + \cdots. \tag{2.2}$$

In a similar way, calculation of $E[(2n - R)(2n - R + 1)]$ yields $E(R^2) = 2n - E(R)$. Furthermore, $R(2n)^{1/2}$ has asymptotically the distribution of the absolute value of a normal deviate. For example, there is about one chance in ten that R will exceed $2.32\ n^{1/2}$. Combining with the result of the previous paragraph, we see that the value of the truncated binomial design is $E(G) = n + n\binom{2n}{n}\Big/2^{2n}$. The excess $E(G) - n \approx n^{1/2}/\pi^{1/2} - 1/8(\pi n)^{1/2} + \cdots$ is shown in Table II for a number of values of n.

(iii) In the *random allocation* design, S selects n of the first $2n$ positive integers at random without replacement, and then assigns treatment A to those subjects whose ordinal numbers have been selected. Another way of expressing this strategy is to say that on each move S selects a treatment with probability proportional to the number of subjects still to receive that treatment.

It is intuitively clear that against this strategy, E should always use the *convergence* strategy, i.e., he should guess that treatment which has previously been less used; when there is a tie in past use, S will choose A or B with equal probability so E's choice is arbitrary. In calculating $E(G)$ for the sampling design it is very convenient to picture the results of S's choices as a walk on the lattice points of the plane. We start at the point (0, 0), and move one unit to the right (or up) when S picks treatment A (or B). The experiment terminates when we reach the point (n, n). In terms of this walk, E will always guess that the walk will move toward the diagonal—if the walk is on the diagonal his guess is arbitrary. Since the walk starts and stops on the diagonal, it must move towards it exactly n times and away exactly n times. Therefore E's strategy assures him of n correct guesses. In addition, there will be a number of steps originating on the diagonal, say T of them.

If we denote generically by $B(k)$ the number of successes in k binomial trials of success probability one-half, we see that $G = n + B(t)$ when $T = t$. Thus $E(G \mid T = t) = n + t/2$, and $E(G) = n + E(T)/2$.

The distribution of T has been studied by Feller, and it is apparent from formula (6.15) of Chapter 12 of [1] that

$$\Pr(T = t) = 2^t\left[\binom{2n - t - 2}{n - t} - \binom{2n - t - 2}{n - t - 2}\right]\Big/\binom{2n}{n} \tag{2.3}$$

from which it follows that, for large n, $T/n^{1/2}$ has asymptotically the distribution of an absolute normal deviate. If we consider the probability of the walk passing through the point (j, j), we see that

$$E(T) = \sum_{j=0}^{n-1}\binom{2j}{j}\binom{2n - 2j}{u - j}\Big/\binom{2n}{n}. \tag{2.4}$$

The identity (see [2], p. 252)

(2.5)
$$\sum_{j=0}^{n} \binom{2j}{j}\binom{2n-2j}{n-j} = 2^{2n}$$

gives

(2.6)
$$E(T) = 2^{2n} \Big/ \binom{2n}{n} - 1 \approx n^{1/2} - 1 + \frac{\pi^{1/2}}{8(n)^{1/2}} + \cdots.$$

The asymptotic approximation has error of 0.03 per cent at $n = 5$.

Table 1 gives some values of $E(G) - n$ computed with the aid of these formulas. Notice that the truncated binomial design has in each case a smaller value of $E(G) - n$ than does the random allocation design. This is not an accident, as we now proceed to show.

3. Solution of the game.

THEOREM 1. *The truncated binomial design is the solution of our game.*

In proving this theorem, it is helpful to generalize the problem to permit different preassigned numbers of subjects for the treatments. Let $D(m, k)$ denote the design problem when we are required to use A just m times and B just k times. By analogy, we say S uses the truncated binomial design if he chooses treatments independently and at random until one of the treatments is exhausted. As in the special case $D(n, n)$ it is easy to see that $E(G)$ does not depend on E's strategy (provided always that he guesses the obvious in the tail) when S uses the truncated binomial strategy. If we denote this invariant value of $E(G)$ by $\phi(m, k)$ for the problem $D(m, k)$, we easily find that

$$\phi(m, 0) = m; \qquad \phi(0, k) = k;$$
$$\phi(m, k) = [1 + \phi(m, k-1) + \phi(m-1, k)]/2 \quad \text{for } m, k > 0.$$

For future reference we note that

(3.1)
$$|\phi(m-1, k) - \phi(m, k-1)| < 1.$$

TABLE 1
$E(G) - n$

n	Truncated Binomial Design	Sampling Design
0	1.23	1.53
10	1.76	2.24
15	2.17	2.96
20	2.51	3.49
25	2.81	3.95
30	3.08	4.37
40	3.56	5.12
50	3.98	5.78
100	5.63	8.37
∞	$0.564(n)^{1/2}$	$0.886(n)^{1/2}$

This is obvious for $m + k = 2$, and since ϕ on the line $m + k = s + 1$ is just one-half more than the average of consecutive values on the line $m + k = s$, (3.1) holds in general.

Our design problems D are inductively related. Suppose we have checked our theorem for the design problems $D(m - 1, k)$ and $D(m, k - 1)$, showing that the truncated binomial strategy solves these, yielding values $\phi(m - 1, k)$ and $\phi(m, k - 1)$ respectively. We now consider the game $D(m, k)$. After the first move we shall be faced with one of the former games. Therefore the payoff matrix can be expressed in terms of the choices of E and S on the first move only. In fact, the expected payoffs are given by

		S	
		A	B
E	A	$1 + \phi(m - 1, k)$	$\phi(m, k - 1)$
	B	$\phi(m - 1, k)$	$1 + \phi(m, k - 1)$

Now we hope to show that S should choose the columns with equal probabilities. Therefore, let us try to find a strategy for E which will make these columns equally attractive to S. This leads at once to having E choose the first row with probability

$$[1 + \phi(m, k - 1) - \phi(m - 1, k)]/2. \tag{3.2}$$

(That (3.2) is indeed a probability follows from (3.1)). The game is now solved, since (a) when E uses (3.2), the options are equally attractive to S who is then content to choose them at random, while (b) S's random choice makes E indifferent and hence content with (3.2).

Incidentally, our game has an interesting feature. When either player uses his minimax strategy, the expected outcome of the game is independent of the strategy adopted by the other player. Notice also that we have shown the truncated binomial design to be the solution of the general design problem $D(m, k)$, with preassigned but possibly different treatment numbers.

We remark that although this design minimizes $E(G)$, the minimized value is disturbingly large. If we divide the difference of treatment sums (1.1) by $n^{1/2}$ as is customary in standardizing it, the expected value is about $2\Delta/\pi^{1/2}$, which does not tend to 0 as $n \to \infty$. In many experimental situations Δ could be large enough to produce a serious distortion.

4. The variance of G. When S uses the truncated binomial design, the value of $E(G)$ is independent of E's strategy, but it should not be thought that E is unable to influence other aspects of the distribution of G. For example, if E guesses the treatment A, as long as that treatment is possible, he is assured of at least n correct guesses, while if he guesses at random, G can be as low as 1. We shall in particular examine the influence of E's strategy on the variance $V(G)$. This would be an essential quantity in computing the expectation of a payoff function which can be represented by a quadratic function of G, or in approximating the probability that the estimated treatment effect exceeds a specified critical value. (If E believes that the treatments are not different, but wishes as large as possible a probability of having the difference appear highly significant, he would want $V(G)$ large.)

Our methods permit us to find the strategy for E which will maximize (minimize) $V(G)$. We have introduced above the convergence strategy, according to which E always guesses the hitherto least frequently used treatment. Opposite to this is the *divergence* strategy: as long as both treatments are available, E guesses the one which has been most used; when there is a tie, he guesses at random; while as always in the tail he guesses the treatment certain to be used.

THEOREM 2. *Against the truncated binomial design, $V(G)$ is maximized (minimized when E uses the divergence (convergence) strategy.*

Since $E(G)$ is constant it will suffice to prove the corresponding assertion for $E(G^2) = V(G) + E^2(G)$. Consider the problem $D(m, k)$ where to avoid obvious cases we assume $mk > 0$. Let E employ the pure strategy of guessing A on the first trial. If S assigns A, E wins 1 on that trial and is faced with the game $D(m - 1, k)$, in which E wins, say, H. If S assigns B, E wins nothing on the first trial and then must play $D(m, k - 1)$, winning K. As the assignments are equally likely,

$$E(G^2) = [E(1 + H)^2 + E(K^2)]/2 = \tfrac{1}{2} + \phi(m - 1, k) + E(H^2 + K^2).$$

Similarly, if E adopts pure strategy B on the first trial,

$$E(G^2) = \tfrac{1}{2} + \phi(m, k - 1) + E(H^2 + K^2).$$

Now the distributions of H and K depend on the strategies adopted in playing $D(m - 1, k)$ and $D(m, k - 1)$ respectively, but not on the guess which E makes in the first trial. Therefore, $E(G^2)$ will be maximized when E guesses A on the first trial if $\phi(m - 1, k) > \phi(m, k - 1)$, and when E guesses B if the inequality is reversed. As $\phi(m, s - m)$, viewed as a function of m, is an increasing function of $| m - s/2 |$, we see that the divergence strategy will maximize $E(G^2)$. A similar argument shows that $E(G^2)$ is minimized by the convergence strategy.

As argued in Section 2, when E adopts the convergence strategy, and the walk has t ties, $G = n + B(t)$. Thus G has as its distribution a mixture of binomials, the mixing coefficients being given by the distribution of T. We shall derive this by considering first the joint distribution of T and R. Denote

$$\Pr(T = t \quad \text{and} \quad R = r)$$

by $\pi(t, r)$, and observe that these variables have the range $3 \leqq t + r \leqq n + 1$, $1 \leqq t, r$.

It is remarkable that $\pi(t, r)$ depends only on $t + r$. This can be seen by establishing a two-to-one mapping of the walks with values $(t + 1, r)$ onto walks with values $(t, r + 1)$. Consider any walk W with $T = t + 1$, $R = r$. Let W' be the walk identical with W except that the part after the last tie has been reflected about the diagonal; W' also has $T = t + 1$, $R = r$. Each of these walks has probability $1/2^{2n-r}$ under the truncated binomial design. Now locate on W(or W') the point immediately preceding the last tie, and denote its coordinates by (x, y). We shall assume $y = x + 1$ with the case $x = y + 1$ being argued similarly. Suppose W is the walk having its last part above the diagonal. We now create from W a new walk W^*, by (a) eliminating the step from (x, y) to $(x + 1, y)$, (b) shifting one step to the left that part of W from $(x + 1, y)$ to the boundary, and (c) closing the gap thus created by adding a step to the tail. Note that the correspondence between the pair (W, W') and W^* is one-to-one, that W^* has probability $1/2^{2n-r-1}$, and that it has $T = t$, $R = r + 1$.

As a corollary we observe that T and R are identically distributed. Since we have already obtained the distribution of R (2.1), we can now give that of G:

$$\Pr(G = g) = \sum_{t=1}^{n} \binom{t}{g - n} \binom{2n - t - 1}{n - 1} \Big/ 2^{2n-1}.$$

$V(G)$ can also easily be calculated. We have

$$E(G^2 \mid T = t) = n^2 + nt + (t + t^2)/4$$

so that $E(G^2) = n^2 + (n + \frac{1}{4})ET + \frac{1}{4}ET^2 = n^2 + n/2 + nE(R)$. Since $E(G) = n + E(R)/2$, we get

$$V(G) = n/2 - E^2(R)/4 \approx \frac{\pi - 1}{2\pi} n + \frac{1}{4\pi} + \cdots . \tag{4.1}$$

This is the smallest value which $V(G)$ can have.

Since $\pi(t, r)$ depends only on the sum of its arguments,

$$\pi(t, r) = \pi(t + r - 1, 1).$$

A walk which has $T = t + r - 1$ and $R = 1$ must have just $t + r - 2$ ties before reaching the point $(n - 1, n - 1)$. Each such walk has probability $\frac{1}{2}^{2n-1}$ and the number of them can be read at once from (2.3). It follows that, for $n > 1$,

$$\pi(t, r) = \left[\binom{2n - t - r - 2}{n - t - r + 1} - \binom{2n - t - r - 2}{n - t - r - 1}\right] \Big/ 2^{2n-t-r} \tag{4.2}$$

We shall need $E(RT)$. If we let U_k indicate a tie at (k, k), so that $T = U_0 + U_1 + \cdots + U_{n-1}$, we see that

$$E(RT) = \sum_{k=0}^{n-1} P(U_k = 1)\, E(R \mid U_k = 1) = \sum_{k=0}^{n-1} (n - k) \binom{2k}{k} \binom{2n - 2k}{n - k} \Big/ 2^{2n-1}.$$

Again making use of (2.4), we find that $E(RT) = n$.

In computing $V(G)$ for the divergence strategy, note that when $T = t$ and $R = r$, $G = n - t + r + B(t)$. Therefore,

$$E(G^2 \mid T = t, R = r) = (n + r - t)(n + r) + (t + t^2)/4.$$

Using the relations $E(T) = E(R)$, $E(T^2) = E(R^2) = 2n - E(R)$, and $E(RT) = n$, we find after simplifying

$$(4.3)\quad V(G) = 3n/2 - E(R) - E^2(R)/4 \approx \frac{3\pi - 2}{2\pi} n - \frac{2(n)^{1/2}}{\pi^{1/2}} + \frac{1}{4\pi} + \cdots .$$

This is the greatest value which $V(G)$ can attain.

Note that the range of $V(G)$ is quite large. The ratio of maximum to minimum values tends with large n to $(3\pi - 2)/(\pi - 1) = 3.467$.

5. Completely binomial designs. Since it is not possible to find a design with adequate bias control when the treatment numbers are preassigned, we shall now examine some designs free of this restriction. In the present section we shall assume that each subject has probability $\frac{1}{2}$ of receiving each treatment, and that the assignments are independent. Such *completely binomial* designs are bias-free, in the sense that every guessing strategy will produce a G whose expectation, given the number s of trials, is exactly $s/2$. We can still exercise a measure of control over the experiment through the decision to terminate it. In our geometrical picture, the design of the experiment now consists in specifying a set of points in the plane at which experimentation will stop. Each such sequential stopping rule will provide a distribution of the numbers X and Y of subjects receiving treatments A and B, respectively, leading to distributions of the total number of trials $X + Y = S$ and of the variance factor $1/X + 1/Y = V$.

(i) *Fixed total design.* In some experiments it may be necessary or desirable to know in advance the total number s of trials to be performed. This leads to the stopping rule $x + y = s$, for which the variance factor V is variable and indeed unbounded: if x or y is 0, $V = \infty$. However, since $X = B(s)$, if s is large it is unlikely that X will be far from $s/2$ and V will probably not much exceed its minimum value $4/s$. In fact, if we expand V in powers of $(x - s/2)$, we find

$$v = \frac{4}{s} + \frac{16}{s^3}\left(X - \frac{s}{2}\right)^2 + \cdots .$$

Here $2(X - s/2)/s^{1/2}$ is approximately distributed as a normal deviate. If K_α denotes the upper $\alpha/2$ point on the normal distribution, and if we want to have $V \leqq 2/n$ with probability $1 - \alpha$, we should choose s so that

$$\frac{2}{n} = \frac{4}{s} + \frac{4}{s^2} K_\alpha^2 ,$$

or $s = 2n + K_\alpha^2$. For example, if we set $s = 2n + 4$, we shall be for large n about 95.5 per cent sure of obtaining a bias-free experiment with variance factor smaller than that obtainable with $X = Y = n$ preassigned.

(ii) *Fixed factor design.* Instead of fixing S and permitting V to vary, we might often prefer to fix V and permit S to vary. For example, we could continue taking observations until X and Y satisfy

$$\frac{1}{X} + \frac{1}{Y} \leqq \frac{2}{n}. \tag{5.1}$$

Write $X + Y = S = 2n + U$, so that $U \geqq 0$ may be viewed as the number of additional observations required to obtain freedom from bias.

For a given value $U = u$, let x_u denote the smallest value of X for which (5.1) holds; i.e.,

$$\frac{1}{x_u} + \frac{1}{y_u} \leqq \frac{2}{n} < \frac{1}{x_u - 1} + \frac{1}{y_u + 1}, \qquad x_u + y_u = 2n + u. \tag{5.2}$$

A path will yield $U \leqq u$ if and only if, at the $(2n + u)$th step it is at, say, (X^*, Y^*) with $x_u \leqq X^* \leqq y_u$. As $X^* = B(2n + u)$, distribution of U is now easily calculated.

As $n \to \infty$, U has a simple limit law. From (5.2) it appears that for large n, $y_u \sim (n + u/2) + u^{1/2}(2n + u)^{1/2}/2$. Since the binomial X^* has $EX^* = n + u/2$ and $\sigma_{X^*} = (2n + u)^{1/2}/2$, we have

$$\Pr(U \leqq u) \approx \Phi(u^{1/2}) - \Phi(-u^{1/2}); \qquad u = 0, 1, 2, \cdots.$$

Table 2 compares the distributions of U for $n = 5$, 10, 20, and ∞. We see that on the average it costs about one and one-half observations, and is practically certain not to cost as many as ten observations, to eliminate selection bias entirely. (Even this comparison is unfair to the bias-free design, as the inequality (5.1) is usually strict and we are obtaining a somewhat more accurate estimate. If we were to take the final step with a probability adjusted to make $E(1/X + 1/Y) = 2/n$, we should find $EU \to 1$ as $n \to \infty$.)

TABLE 2

Fixed factor design

Distribution of U = excess observations required

	$P(U \leqq u)$			
u	$n = 5$	10	20	∞
0	.246	.176	.125	0
1	.773	.617	.651	.6827
2	.854	.866	.836	.8427
3	.908	.907	.934	.9167
4	.943	.936	.951	.9545
5	.965	.985	.984	.9747
6	.979	.991	.989	.9857
7	.987	.994	.992	.9918
8	.999	.996	.994	.9953
9	.999	.998	.999	.9973
10	1.000	.999	.999	.9984
$E(U)$	1.338	1.464	1.535	1.6625

In practice, the fixed factor design would be used in a truncated form. For example, we could continue the binomial choice until (5.1) holds, or until $X + Y \geqq 2n + u$. If the latter eventuates first, the deficient treatment could be applied until (5.1) holds. By setting $u = 10$, we would have practical certainty of a bias-free design, without the theoretical possibility of an infinite sequence of trials.

(iii) *Fixed minimum design.* In some cases we might wish to guarantee the precision of estimation for each treatment effect separately, rather than for their difference. We should then need

$$\min (X, Y) = n.$$

By symmetry, we are equally likely to stop at (n, x) and at (x, n), so it will suffice to consider the probabilities of stopping at points (x, n) for $x = n, n + 1, \cdots$. These probabilities are easily seen to be proportional to those of the single negative binomial design, which is stopped by $y = n$. Thus our X has a truncated negative binomial distribution, with range just the complement of that of $n - R$ considered in Section 2. As each of the ranges is equally likely, we must have $\frac{1}{2}ES + \frac{1}{2}E(2n - R)$ equal to the expected number of steps in the single negative binomial, which is $2n$. Therefore $ES = 2n + ER$, where $E(R)$ is given by (2.2). Roughly, we must expend on the average $1.13\ n^{1/2}$ additional observations in this case.

6. Extensions. A good deal of the preceding argument generalizes rather easily to experiments involving more than two treatments. In particular, the minimax design for preassigned treatment numbers consists in choosing at each step among the remaining treatments with equal probabilities. A simple bias-free design, which generalizes 5(ii), consists in choosing a treatment at each step, with equal probabilities, and terminating the experiment when the sum of reciprocals of treatment numbers falls below a preassigned level.

REFERENCE

[1] Feller, William, *An Introduction to Probability Theory and Its Applications*, Vol. 1, John Wiley & Sons, 1950.