Sequential Analysis, 25: 167–178, 2006

Contribution of Milton Sobel in Selection Problem Following Ethical Allocation

Atanu Biswas
Applied Statistics Unit, Indian Statistical Institute, Kolkata, India

Abstract: In his research career of nearly six decades, Milton Sobel has made substantial contributions in several areas of statistics and mathematics. In this present discussion, I focus on the contribution of Sobel to a specific aspect of the selection problem, namely, selection following a sequential adaptive allocation. Follow-up works by Sobel himself and other researchers are also discussed in detail. The scope for further research in this area is indicated. This is only a small part of the vast contribution of Milton Sobel.

Keywords: Adaptive design; Inverse sampling; Inverse stopping rule; Play-the-winner rule; Randomization; Three hypotheses problem; Vector-at-a-time.

Subject Classification: 62F07; 62F35.

1. INTRODUCTION

The research career of Milton Sobel spans nearly six decades, starting from 1946 when he began to work for his Ph.D. at Columbia University under the supervision of Abraham Wald. In his research career, he worked in different areas of statistics and mathematics, including decision theory and sequential analysis, selection and ranking methodology, group sequential tests of hypotheses, reliability analyses, combinatorial problems, Dirichlet processes, statistical tables, and computations. His contributions and related stories are briefly described in the interview with Milton Sobel by Nitis Mukhopadhyay (2000). In this paper, I would like to highlight only one of the major contributions of Sobel to the solution of the sequential selection problem. This is an ethical-allocation-based selection problem discussed by Sobel and Weiss (1970), which was followed by a number of papers by them using a play-the-winner allocation. Here the goal was to allocate a larger number of subjects to the eventual better treatment.

Received July 22, 2004, Revised October 8, 2004, November 24, 2004, January 31, 2005, Accepted February 15, 2005

Recommended by D. S. Coad

Address correspondence to Atanu Biswas, Applied Statistics Unit, Indian Statistical Institute, 203 B.T. Road, Kolkata 700108, India; E-mail: atanu@isical.ac.in

The rest of this paper is organized as follows. Section 2 provides some background and details of the Sobel–Weiss procedure (1970). In Section 3, extensions of the Sobel–Weiss procedure in different directions by authors like Hoel (1972), Fushimi (1973), and Bandyopadhyay and Biswas (1999a,b) are also discussed. Section 4 provides a comparative study of the Sobel–Weiss procedure, the Hoel procedure, and the Fushimi procedure by using some randomized adaptive allocation design. Some numerical results are also reported in this context. Section 5 ends with the indication of some important possible extensions of the Sobel–Weiss procedure.

2. AN ETHICAL SELECTION PROBLEM: THE SOBEL–WEISS PROCEDURE

2.1. Sobel's Work on the Selection Problem: A Brief Appraisal

Sequential ranking and selection is a major application of sequential analysis pioneered by Milton Sobel, among others. Sobel worked with C. Dunnett, S. S. Gupta, J. Kiefer, H. Robbins, N. Starr, P. Chen, and S. Panchapakesan in this interesting area. Sobel's book with Gibbons and Olkin (1977) was the first textbook on ranking and selection. However, a research monograph in this area was written by Bechhofer et al. (1968) (BKS) nine years before that. This book gives an account of the initial work of Sobel in this research area.

The book by Bechhofer et al. (1995) (BSG) discusses the legacy of the BKS procedure in greater detail. Chapter 7.3 of BSG describes a closed adaptive sequential procedure, due to Bechhofer and Kulkarni (1982), which is adaptive in the sense that earlier outcomes are used to decide which treatment is to be sampled next. However, it has the same probability of correct selection as the corresponding single-stage Sobel–Huyett procedure (see Sobel and Huyett, 1957), which is a fixed-sample-size procedure where the treatment with the highest number of successes among an equal number of allocations to each arm emerges as the winner. Chapter 7.4 of BSG describes some open sequential procedures for the odds-ratio indifference zone (see BSG for details). The open sequential procedure without early elimination of arms, suggested by BKS, is discussed and compared with the open sequential procedure with elimination, suggested by Paulson in 1993 in a personal communication. The design of Paulson is based on his earlier work (Paulson, 1964, 1969).

2.2. The Play-the-Winner Allocation Rule

In 1970, Sobel published a paper with G. E. Weiss that deals with the problem of selecting the better of two treatments in a clinical-trial setup. The problem of comparing and selecting the better of two independent Bernoulli populations has been formulated in different ways both in selection problems and clinical trials. The treatment responses are assumed to be instantaneous, i.e., the response of a patient is available before the entry of the next patient. Bechhofer et al. (1968) provide a detailed bibliography for this area of research. A number of early papers were written using "vector-at-a-time" (VT) sampling where the number of patients receiving each treatment is the same. On the other hand, the adaptive designs have the simultaneous goal of (1) selecting a better treatment with probability at least P^* and (2) treating a larger number of patients by the better treatment. If the horizon contains N patients, a procedure is used that uses the first n patients to select the better treatment, and the remaining $(N - n)$ patients are treated by the selected treatment. This has been studied by Armitage (1960), Anscombe (1963), Colton (1963), and Zelen (1969). An adaptive design is a sequential sampling design where we adapt the (allocation) design by using the accumulated data. Robbins (1952) is the forerunner in adaptive design, followed by Anscombe (1963) and Colton (1963).

One easily applied and appealing approach, much used in subsequent studies, is the play-the-winner (PW) rule, suggested by Robbins (1956). Here the first patient receives one of the two competing treatments on the basis of a toss of a fair coin. For subsequent patients, a success using a treatment results in the next entering patient receiving the same treatment, whereas the effect of a failure is that the next entering patient receives the other treatment. Zelen (1969) provided an excellent discussion of the play-the-winner rule. Iglewicz (1983) reports one unpublished application of the PW rule in a lung-cancer trial by Marvin Zelen. The most famous and controversial application of the PW sampling design is the extracorporeal membrane oxygenation (ECMO) trial (see Ware, 1989). Rout et al. (1993) described another application of the PW rule where the first 40 patients were prerandomized in pairs to two treatments and PW was carried out thereafter and evaluated after 50 patients. The study was terminated after 140 patients were studied, of which 100 were treated by the PW rule with an allocation of 58:42 between treatments, with a larger allocation in favor of the better treatment.

This has, in principle, similarity to the bandit problem. There are adaptive mechanisms like the "switch-from-a-loser" and "stay-with-a-winner" in the bandit-problem literature; see the book by Berry and Fristedt (1985), for example. Bandit allocations optimize the benefits of the patients under study, whereas a clinical trial is motivated toward the benefits of the future patients as well as the current patients. See Berry (1972, 1978), Hayre and Turnbull (1981), and Berry and Fristedt (1985) for a detailed discussion of the bandit problem. The seminal paper on the multiarmed bandit problem with infinite horizon and discounting was written by Gittins (1979).

2.3. The Sobel–Weiss Procedure

Sobel, along with G. H. Weiss, worked in the area of PW-type adaptive sampling followed by selection. The paper, published in 1970, was a major breakthrough in this area. Let the two treatments be A and B with success probabilities p_A and p_B, and let S_A and S_B be the number of successes by the two treatments. Writing

$$\Delta = p_A - p_B,$$

Sobel and Weiss (1970) proposed a sampling design that stops sampling and declares treatment i to be better than j, for $i, j = A, B$, where $i \neq j$, if $S_i - S_j = r$, where r is chosen so that

$$P(CS) \geq P^* \quad \text{for } \Delta \geq \Delta^*,$$

with preassigned P^* and Δ^*, and CS being "correct selection." Sobel and Weiss (1970) studied the properties of the procedure, including the probability model and average sample number, and compared the results with the corresponding VT procedure. They observed that the PW procedure arrives at a decision at the expense of fewer patients using the poorer treatment. This work is one of the first of this kind to link adaptive procedures with the selection problem. In a sequel paper, Sobel and Weiss (1971b) derived the exact and approximate expressions of the expected total number of patients and also the expected number of patients to either treatment for both PW and VT procedures. The expressions for $P(CS)$ are also obtained. If the incomplete beta function $I_q(j, r)$ is defined by

$$I_q(j, r) = \frac{\Gamma(j+r)}{\Gamma(j)\Gamma(r)} \int_0^q t^{j-1}(1-t)^{r-1} dt,$$

and X be a negative binomial random variable with

$$E_r[f(X)] = p^r \sum_{j=0}^{\infty} \binom{j+r-1}{j} q^j f(j),$$

then

$$P(CS \mid PW) = P(CS \mid VT) = \frac{1}{2} E_r[I_{q_B}(X, r) + I_{q_A}(X+1, r)],$$

where $q_A = 1 - p_A$, $q_B = 1 - p_B$. Denoting by N and N_B respectively the total number of allocations and the number of allocations to the inferior treatment (assuming $p_A \geq p_B$),

$$\begin{aligned}
E[N_B \mid PW] &= \frac{rq_A}{2q_B p_A} E_{r+1}[I_{q_B}(X+2, r) + I_{q_B}(X+3, r)] + \frac{1}{2q_B} E_r[I_{q_B}(X+2, r)] \\
&\quad + \frac{r}{2p_B} E_r[I_{p_B}(r+1, X) + I_{p_B}(r+1, X+1)], \\
E[N_B \mid VT] &= \frac{r}{p_A}\left\{1 - \frac{1}{2} E_{r+1}[I_{p_B}(r, X) + I_{p_B}(r, X+1)]\right\} \\
&\quad + \frac{r}{2p_B} E_r[I_{p_B}(r+1, X) + I_{p_B}(r+1, X+1)], \\
E[N \mid PW] &= E[N_B \mid PW] + \frac{r}{p_A}\left\{1 - \frac{1}{2} E_{r+1}[I_{p_B}(r, X) + I_{p_B}(r, X+1)]\right\} \\
&\quad + \frac{rq_B}{2q_A p_B} E_r[I_{p_B}(r+1, X-1) + I_{p_B}(r+1, X-2)] \\
&\quad + \frac{1}{2q_A} E_r[I_{p_B}(r, X-1)],\ E[N \mid VT] = 2E[N_B \mid VT].
\end{aligned}$$

An approximation is obtained when $r \to \infty$. If $\Phi(\cdot)$ is the standard normal cumulative distribution function, then, writing $D = q_A p_B^2 + q_B p_A^2$, $y = \Delta\sqrt{r/D}$, we have

$$P(CS) \sim \Phi(y) + O(r^{-1/2}),$$

$$E[N_B \mid PW] \sim \frac{r}{q_B}\left\{\frac{q_A}{p_A}\Phi(y) + \frac{q_B}{p_B}[1 - \Phi(y)]\right\}, \quad E[N \mid PW] \sim \frac{r(q_A + q_B)}{p_A q_B} + \frac{1}{2q_B},$$

$$E[N_B \mid VT] \sim \frac{r}{p_A}\Phi(y) + \frac{r}{p_B}[1 - \Phi(y)].$$

See Sobel and Weiss (1971b) for details, and Sobel and Weiss (1971a, 1972) for related work.

3. INVERSE STOPPING RULE IN THE SOBEL–WEISS-TYPE PROCEDURE

3.1. The Hoel Procedure

The Sobel–Weiss procedure drew the immediate attention of some excellent researchers. As a result, some modifications of the procedure were suggested. David Hoel contributed to research in such selection problems. He worked jointly with Sobel and wrote a paper on comparison of sequential procedures for selecting the best binomial population (see Hoel and Sobel, 1971), and another on two-stage procedures (see Hoel et al., 1972). The Sobel–Weiss procedure has an apparent drawback of having very large expected sample sizes for small values of p_A and p_B, which is clear from the formulation of their stopping rule, and also from the expression of $E(N \mid PW)$ given in Section 2. In his solo paper in 1972, Hoel modified the Sobel–Weiss procedure and also incorporated the number of failures in the inverse stopping rule. If F_A and F_B are the number of failures by treatments A and B, respectively, in the procedure, Hoel (1972) suggested continuing to sampling until either $S_A + F_B$ or $S_B + F_A$ is equal to r, where r is preassigned, thus controlling the total sample size. If $S_A + F_B = r$, treatment A is selected as the better treatment. This is a truncated procedure with more or less stable expected sample sizes for different possible values of p_A and p_B. Also, Hoel observed that with this modification, for the PW allocation design, the very large expected sample sizes of the Sobel–Weiss procedure for small values of p_A and p_B are avoided. However, the Sobel–Weiss procedure works better when p_A and p_B are large.

3.2. The Fushimi Procedure

Fushimi (1973) introduced an improved version of the Sobel–Weiss procedure. He suggested stopping sampling when either the condition $|S_A - S_B| = r$ or the condition $F_A + F_B = s$ is satisfied for preassigned integers r and s. After the sampling is stopped, the population with a larger number of successes is selected as the better. If there is a tie, a fair coin is used to select one population. Fushimi obtained the expressions for the probability of correct selection and expected sample sizes. This procedure is then compared with the Sobel–Weiss procedure, the VT design, and Hoel's modification. Although none of them is uniformly better than the others, Fushimi's approach emerges as a fairly good procedure when we have no prior information on p_A and p_B. In fact, the Fushimi procedure, using the PW rule, is much better than others for some parameter values, and never drastically bad when it is not the best.

3.3. Other Modifications

Some modifications of these approaches have since been proposed. Some of them are indicated below.

Nebenzahl and Sobel (1972) considered both the PW and the VT sampling procedures with a termination rule based on $N = n$. Kiefer and Weiss (1971) considered the VT sampling procedure with a termination rule based on $|S_A - S_B| = r$ or $N = n$, whichever comes first. Kiefer and Weiss (1974) extended that approach for the PW sampling scheme. Berry and Sobel (1973) considered the PW sampling procedure with a termination rule based on $\max(S_A, S_B) = c$ or $F_A = F_B = r$, whichever comes first. A further contribution in this context is due to Pradhan and Sathe (1973), who considered the Nebenzahl–Sobel PW sampling procedure with curtailment. In a subsequent work, Pradhan and Sathe (1974) proved the equivalence between the Nebenzahl–Sobel procedure and the Hoel procedure. Taheri and Young (1974) studied the Sobel–Weiss procedure using $\Delta = p_A/p_B$.

Sobel and Weiss (1971b) proposed an alternative procedure (say SW^*) that is more efficient than their standard rule (SW) when p_A and p_B are small. They used the PW sampling scheme with a termination rule based on $F_A = F_B = r$, where $r = r(P^*, \Delta^*)$. They showed that for small Δ^*, the SW procedure is preferable when $p_A > 0.5$, and SW^* is preferable when $p_A < 0.5$. Liu (1993) proposed a procedure (say L) that is a modified version of the SW^* procedure. The procedure L uses the PW sampling rule and stops sampling when either $F_A = F_B = r$, or $F_i = r$, $F_j = r - 1$, $S_j > S_i$, if we start sampling with population i, where $i, j = A, B$, and $i \neq j$. Liu (1993) showed that $P(CS \mid L) = P(CS \mid SW^*)$, and $E(N_i \mid L) < E(N_i \mid SW^*)$ for $i = A, B$. The same rule was studied by Liu (1991) for the VT procedure.

4. RANDOMIZED ADAPTIVE ALLOCATION

Selection using the PW allocation is a pioneering concept of Sobel and Weiss (1969). However, it is well accepted that randomization is a *must do* concept in clinical trials for ethical reasons to avoid allocation bias. Moreover, the treatment of a patient completely determined by the response of only one previously allocated patient's response is unacceptable. Book-length treatments of randomization in clinical trials include Matthews (2000) and Rosenberger and Lachin (2002). Several randomized adaptive designs are available in the literature as the modifications of the PW rule. These include the randomized play-the-winner (RPW) rule (see Wei and Durham, 1978), success-driven design (see Durham et al., 1998), birth and death urn design (see Ivanova et al., 2000), and drop-the-loser rule (see Ivanova, 2003), among others. Following the approach of Sobel and Weiss (1970) and the modifications by several authors thereafter, one can easily use any such design in place of the PW rule for the Sobel–Weiss-type selection problem. For example, Bandyopadhyay and Biswas (1999b) considered the selection of the best subset among k treatments using the RPW rule. They, of course, adopted a different kind of stopping and elimination rule, as it was a multitreatment procedure.

The RPW rule can be illustrated by an urn model as follows. We start with an urn having two kinds of balls, A and B, representing the two treatments, α balls of each kind. Any entering patient is treated by drawing a ball from the urn, with replacement. If the response (of Bernoulli type) is a success, we add an additional β balls of the same kind in the urn. On the other hand, if the response is a failure, we add an additional β balls of the other kind in the urn. The idea is to skew the urn and hence the allocation in favor of the treatment doing better. Some real applications of the RPW rule are the Michigan ECMO trial (Bartlett et al., 1985); the fluoxetine trial, sponsored by Eli Lilly, to treat outpatients of depressive disorder (Tamura et al., 1994); and the rheumatoid-arthritis trial (Biswas and Dewanji, 2004a,b).

As an illustration, we consider the selection rule of Sobel and Weiss (1970) using the RPW allocation rule in place of the *PW* rule. The results are given in Table 1. It is clear that the Sobel–Weiss procedure is flexible enough to incorporate randomization. The randomized designs are the present and future mode for clinical trials. With that in mind, in the present discussion, we present detailed simulation results of the average sample numbers (ASNs), their standard deviations (SDs), and the expected allocation to treatment A (ASN-A) for the Sobel–Weiss, Hoel, and Fushimi procedures using an RPW allocation rule in place of the *PW* rule. The computations are carried out using 30,000 simulations. This provides a comparison of several available procedures that might help the possible application. It is observed that the Hoel procedure performs well in terms of ASN and ASN-A when p_A and p_B are large. For smaller values of p_A and p_B, the Sobel–Weiss procedure has smaller ASN and ASN-A values, as far as the RPW allocation rule is concerned. Again, the SDs of the total number of samples required are less for the Hoel procedure unless p_A and p_B are too small. Similarly, we can implement the selection procedures for other available randomized adaptive allocation designs. Note that the combination of (r, s), which is reported, is not the optimal one. It can be changed to get possibly better results. The theoretical developments and properties

Table 1. Comparison of average sample numbers (ASNs), with SD in parentheses, expected sample size for treatment A (ASN-A) for the Sobel–Weiss (*SW*), Hoel (*H*), and Fushimi (*F*) procedures for different values of (p_A, p_B) when $P^* = 0.8$ and an RPW rule with $\alpha = \beta = 1$ is employed for allocation. The actual values of r (which has different interpretations for the three procedures), s, and $P(CS)$ are also given

(p_A, p_B)	Procedures	ASN (SD)	ASN-A	r	s	$P(CS)$
(0.4, 0.8)	*SW*	4.48 (3.25)	2.11	2		0.83
	H	3.76 (0.78)	1.68	3		0.83
	F	4.37 (2.42)	1.90	2	5	0.80
(0.6, 0.8)	*SW*	17.98 (15.65)	8.09	5		0.80
	H	14.84 (2.56)	6.44	10		0.80
	F	24.33 (11.64)	10.50	8	12	0.81
(0.2, 0.6)	*SW*	6.68 (5.03)	2.94	2		0.91
	H	3.96 (0.78)	1.78	3		0.85
	F	4.93 (1.98)	2.17	2	5	0.82
(0.4, 0.6)	*SW*	19.32 (16.45)	8.84	4		0.84
	H	14.10 (2.05)	6.29	9		0.80
	F	20.55 (7.25)	9.12	6	14	0.81
(0.1, 0.4)	*SW*	10.44 (7.91)	4.63	2		0.95
	H	5.83 (0.96)	2.64	4		0.80
	F	8.06 (1.73)	3.48	3	7	0.81
(0.3, 0.4)	*SW*	47.56 (41.02)	22.77	5		0.81
	H	63.09 (4.21)	29.43	35		0.81
	F	76.26 (7.22)	35.32	18	50	0.81
(0.1, 0.2)	*SW*	22.22 (18.63)	10.67	2		0.80
	H	77.56 (3.95)	36.73	42		0.81
	F	41.24 (2.72)	19.48	10	35	0.80

Table 2. Simulations for the fluoxetine example. Comparison of average sample numbers (ASNs), with SD in parentheses, expected sample size for treatment A (ASN-A) for the Sobel–Weiss (*SW*), Hoel (*H*), and Fushimi (*F*) procedures for different values of r and s (for Fushimi-procedure only) when $P^* = 0.8$ and an RPW rule with $\alpha = \beta = 1$ is employed for allocation. The actual values of $P(CS)$ are also given

Procedures	r	s	ASN	(SD)	ASN-A	$P(CS)$
SW	4		18.60	(15.14)	8.48	0.84
	5		25.60	(21.66)	11.61	0.89
	6		31.82	(25.80)	14.24	0.92
H	9		14.00	(2.06)	6.21	0.81
	12		19.06	(2.58)	8.37	0.83
	20		32.61	(3.86)	13.95	0.90
F	6	14	20.57	(7.44)	9.12	0.81
	6	30	28.34	(16.77)	12.61	0.89
	8	25	33.67	(14.14)	14.73	0.87

of the selection rules with several randomized allocation designs (including the RPW rule) are an exciting research area that still needs to be explored.

As an example of the construction of selection rules, we use part of the data from Tamura et al. (1994) on the treatment of patients of depressive disorder. In order to correspond to our formalization of the present methods, we denote treatment A control and B fluoxetine. The response is the change in $HAMD_{17}$, a measure of depression. Since $HAMD_{17}$ is measured on a 53-point scale, a more than 50% reduction in $HAMD_{17}$ is treated as a success, and failure otherwise. An RPW rule was used for the allocation. There are 83 complete observations since a few observations in the data set do not have the final response. From the data we observe that $\hat{p}_A = 0.405$ and $\hat{p}_B = 0.609$. Using these as the true values we carry out 30,000 simulations for the Sobel–Weiss, Hoel, and Fushimi procedures using the same RPW allocation rule. The results are given in Table 2. From Table 2 we observe that if an appropriate procedure were adopted and the sequential sampling were carried out in the fluoxetine trial, the trial could terminate with a much smaller sample size but with a reasonably high $P(CS)$. For the present problem with the RPW allocation rule, the Hoel procedure performs very well in terms of ASN and SD. The Sobel–Weiss procedure is reasonable in terms of ASN, but the SD for that procedure is too high.

5. FURTHER RESEARCH

In this section, we discuss some related problems for future research that stem from the legacy of Sobel and Weiss (1970).

5.1. Extending the Allocation and Selection Problem to the Case of More than Two Treatments

Quite often, in phase III clinical trials, the number of treatments is more than two. So one needs some sort of modification of the Sobel–Weiss procedure to incorporate this. At the same time as their initial work in this direction, Sobel and Weiss (1969) proposed using the *PW* rule and inverse sampling for selecting $k \geq 3$ binomial populaions. Later Bandyopadhyay and Biswas (1999b) suggested an extension of the Sobel–Weiss procedure where they used the RPW rule in place of the *PW* rule. They also considered more than two treatments in their setup. In such a situation the problem may be to find the best treatment (see Bandyopadhyay and Biswas, 1999a) or the best subset of treatments (see Bandyopadhyay and Biswas, 1999b).

5.2. Three Hypotheses Problem in the Sobel–Weiss Procedure

As a research student, Sobel published one of his best papers in 1949, jointly with his advisor, Abraham Wald. The paper deals with a sequential test for choosing one of three hypotheses concerning the unknown mean of a normal distribution. The unknown mean may belong to any one of three disjoint intervals. The idea is to find the correct interval. Based on the pioneering contribution of Sobel and Wald (1949), Robbins (1970) provided a sequential approach to choose one among a countably infinite number of hypotheses concerning the unknown mean of a normal distribution and developed the concept of "sequential distinguishability." Subsequent works in this direction are due to Khan (1973) and McCabe (1973). An interesting extension would be to pose the ethical selection problem of Sobel–Weiss in such a multidecision framework. This is an area that needs to be explored.

5.3. Comparing Two Non-Bernoulli Populations

What happens if the treatment responses are not Bernoulli, say, categorical or continuous? In principle, the selection procedure of Sobel–Weiss can still be carried out by using suitable adaptive allocation designs for those types of responses (see Bandyopadhyay and Biswas, 2001 for an adaptive design with continuous responses). Theoretical developments in this direction along with a detailed numerical study can have tremendous practical applicability.

5.4. More than Two Non-Bernoulli Populations

Adaptive allocation and selection for more than two non-Bernoulli populations is also an interesting and important research area. There are some papers available, motivated by BKS. The works of Turnbull et al. (1978) and Jennison et al. (1982) are noteworthy in this context. Bechhofer and Sobel (1958) discussed the selection problem based on the vector of observations of the competing treatments. It is also interesting to note that Paulson (1964) considered an adaptive procedure for selecting the population with the largest mean from k normal populations. In principle, one can start with $k(>2)$ treatments having responses other than Bernoulli, for example, continuous. An adaptive allocation design can be proposed as in Biswas and Coad (2005) or Atkinson and Biswas (2005). One can carry out the sequential adaptive allocation and monitor the accumulated data to (possibly) drop some of the poorly performing treatments in between by using some preset rule. There is recent interest in this area that draws from the legacy of Sobel's work.

ACKNOWLEDGMENTS

The author wishes to acknowledge the encouragement of the editor for writing this review paper. The author also thanks two anonymous referees for their valuable suggestions and references, which enriched the paper over some earlier versions.

REFERENCES

Anscombe, F. J. (1963). Sequential Medical Trials, *Journal of American Statistical Association* 5: 365–383.

Armitage, P. (1960). *Sequential Medical Trials*, Oxford: Blackwell.

Atkinson, A. C. and Biswas, A. (2005). Bayesian Adaptive Biased-Coin Designs for Clinical Trials with Normal Responses, *Biometrics* 61: 118–125.

Bandyopadhyay, U. and Biswas, A. (1999a). Selection Procedures in Multi-treatment Clinical Trials, *Metron* 60: 143–157.

Bandyopadhyay, U. and Biswas, A. (1999b). Sequential-Type Procedures to Select the Best Treatments, *Statistics and Decisions*, supplement issue no. 4: 139–164.

Bandyopadhyay, U. and Biswas, A. (2001). Adaptive Designs for Normal Responses with Prognostic Factors, *Biometrika* 88: 409–419.

Bartlett, R. H., Roloff, D. W., Cornell, R. G., Andrews, A. F., Dillon, P. W., and Zwischenberger, J. B. (1985). Extracorporeal Circulation in Neonatal Respiratory Failure: A Prospective Randomized Trial, *Pediatrics* 76: 479–487.

Bechhofer, R. E., Kiefer, J., and Sobel, M. (1968). *Sequential Identification and Ranking Problems*, Chicago: University of Chicago Press.

Bechhofer, R. E. and Kulkarni, R. V. (1982). Closed Adaptive Sequential Procedures for Selecting the Best of $k \geq 2$ Bernoulli Populations, in *Statistical Decision Theory and Related Topics, III*, S. S. Gupta and J. O. Berger, eds., vol. 1, pp. 61–108, New York: Academic Press.

Bechhofer, R. E., Santner, T. J., and Goldsman, D. M. (1995). *Design and Analysis of Experiments for Statistical Selection, Screening, and Multiple Comparisons*, New York: Wiley.

Bechhofer, R. E. and Sobel, M. (1958). Nonparametric Multiple-Decision Procedures for Selecting that one of k Populations Which has the Highest Probability of Yielding the Largest Observation (Abstract), *Annals of Mathematical Statistics* 29: 325.

Berry, D. A. (1972). A Bernoulli Two-Armed Bandit, *Annals of Mathematical Statistics* 43: 871–897.

Berry, D. A. (1978). Modified Two-Armed Bandit Strategies for Certain Clinical Trials, *Journal of American Statistical Association* 73: 339–345.

Berry, D. A. and Fristedt, B. (1985). *Bandit Problems: Sequential Allocation of Experiments*, London: Chapman and Hall.

Berry, D. A. and Sobel, M. (1973). An Improved Procedure for Selecting the Better of Two Binomial Populations, *Journal of American Statistical Association* 68: 979–984.

Biswas, A. and Coad, D. S. (2005). A General Multi-Treatment Adaptive Design for Multivariate Responses, *Sequential Analysis* 24: 139–158.

Biswas, A. and Dewanji, A. (2004a). Inference for a RPW-Type Clinical Trial with Repeated Monitoring for the Treatment of Rheumatoid Arthritis, *Biometrical Journal* 46: 769–779.

Biswas, A. and Dewanji, A. (2004b). A Randomized Longitudinal Play-the-Winner Design for Repeated Binary Data, *Australian and New Zealand Journal of Statistics* 46: 675–684.

Colton, T. (1963). A Model for Selecting One of Two Medical Treatments, *Journal of American Statistical Association* 58: 388–400.

Durham, S., Fluornoy, N., and Li, W. (1998). A Sequential Design for Maximizing the Probability of a Favorable Response, *Canadian Journal of Statistics* 26: 479–495.

Fushimi, M. (1973). An Improved Version of a Sobel–Weiss Play-the-Winner Procedure for Selecting the Better of Two Binomial Populations, *Biometrika* 60: 517–523.

Gibbons, J. D., Olkin, I., and Sobel, M. (1977). *Selecting and Ordering Populations: A New Statistical Methodology*, New York: Wiley.

Gittins, J. C. (1979). Bandit Processes and Dynamic Allocation Indices, *Journal of Royal Statistical Society, Series B* 41: 148–164.

Hayre, L. S. and Turnbull, B. W. (1981). A Class of Simple Approximate Sequential Tests for Adaptive Comparison of Two Treatments, *Communications in Statistics, Series A* 10: 2339–2360.

Hoel, D. G. (1972). An Inverse Stopping Rule for Play-the-Winner Sampling, *Journal of American Statistical Association* 67: 148–151.

Hoel, D. G. and Sobel, M. (1971). Comparison of Sequential Procedures for Selecting the Best Binomial Population, *Proceedings of Sixth Berkeley Symposium on Mathematical Statistics and Probability* 4: 53–69.

Hoel, D. G., Sobel, M., and Weiss, G. H. (1972). A Two-Stage Procedure for Choosing the Better of Two Binomial Populations, *Biometrika* 59: 317–322.

Iglewicz, B. (1983). Alternative Designs: Sequential, Multi-Stage, Decision Theory and Adaptive Designs, in *Cancer Clinical Trials: Methods and Practice*, M. E. Buyse, J. Staquet, and R. J. Sylvester, eds., pp. 312–334, Oxford: Oxford University Press.

Ivanova, A. (2003). A Play-the-Winner-Type Urn Design with Reduced Variability, *Metrika* 58: 1–13.

Ivanova, A., Rosenberger, W. F., Durham, S. D., and Flournoy, N. (2000). A Birth and Death Urn for Randomized Clinical Trials, *Sankhyā, Series B* 62: 104–118.

Jennison, C., Johnstone, I. M., and Turnbull, B. W. (1982). Asymptotically Optimal Procedures for Sequential Adaptive Selection of the Best of Several Normal Means, in *Statistical Decision Theory and Related Topics, III*, S. S. Gupta and J. O. Berger, eds., vol. 2, pp. 55–86, New York: Academic Press.

Khan, R. A. (1973). On Sequential Distinguishability, *Annals of Statistics* 1: 838–850.

Kiefer, J. E. and Weiss, G. H. (1971). A Truncated Test for Choosing the Better of Two Binomial Populations, *Journal of American Statistical Association* 66: 867–871.

Kiefer, J. E. and Weiss, G. H. (1974). Truncated Version of a Play-the-Winner Rule for Choosing the Better of Two Binomial Populations, *Journal of American Statistical Association* 69: 807–809.

Liu, H.-C. (1991). An Improved Procedure for Selecting the Better of Two Bernoulli Populations with Small Probabilities of Success, *Chinese Journal of Mathematics* 19: 321–329.

Liu, H.-C. (1993). A Modified Version of a Sobel–Weiss Play-the-Winner Procedure for Selecting the Better of Two Bernoulli Populations, *Chinese Journal of Mathematics* 21: 151–157.

Matthews, J. N. S. (2000). *An Introduction to Randomized Controlled Clinical Trials*, London: Edward Arnold.

McCabe, G. P., Jr. (1973). Estimation of the Number of Terms in a Sum, *Journal of American Statistical Association* 68: 452–456.

Mukhopadhyay, N. (2000). A Conversation with Milton Sobel, *Statistical Science* 15: 168–190.

Nebenzahl, E. and Sobel, M. (1972). Play-the-Winner Sampling for a Fixed Sample Size Binomial Selection Problem, *Biometrika* 59: 1–8.

Paulson, E. (1964). A Sequential Procedure for Selecting the Population with the Largest Mean from k Normal Populations, *Annals of Mathematical Statistics* 35: 174–180.

Paulson, E. (1969). A New Sequential Procedure for Selecting the Best One of k Binomial Populations (Abstract), *Annals of Mathematical Statistics* 40: 1865–1866.

Pradhan, M. and Sathe, Y. S. (1973). Play-the-Winner Sampling for a Fixed Sample Size with Curtailment, *Biometrika* 60: 424–427.

Pradhan, M. and Sathe, Y. S. (1974). Equivalence Between Fixed Sample Rule and Hoel's Inverse Sampling Rule for Play-the-Winner, *Journal of American Statistical Association* 69: 475–476.

Robbins, H. (1952). Some Aspects of the Sequential Design of Experiments, *Bulletin of American Mathematical Society* 58: 527–535.

Robbins, H. (1956). A Sequential Decision Problem with a Finite Memory, *Proceedings of National Academy of Sciences, U.S.A.* 42: 920–923.

Robbins, H. (1970). Sequential Estimation of an Integer Mean, in *Scientists at Work, Festschrift Honoring Herman Wold*, T. Dalenius, G. Karlsson, and S. Malmquist, eds., pp. 205–210, Uppsala: Almqvist and Wiksells.

Rosenberger, W. F. and Lachin, J. L. (2002). *Randomization in Clinical Trials: Theory and Practice*, New York: Wiley.

Rout, C. C., Rocke, D. A., Levis, J., Gouws, E., and Reddy, D. (1993). A Reevaluation of the Role of Crystalloid Preload in the Prevention of Hypotension Associated with Spinal Anesthesia for Elective Cesarean Section, *Anesthesiology* 79: 262–269.

Sobel, M. and Huyett, M. J. (1957). Selecting the Best One of Several Binomial Populations, *Bell System Technical Journal* 36: 537–576.

Sobel, M. and Wald, A. (1949). A Sequential Decision Procedure for Choosing One of Three Hypotheses Concerning the Unknown Mean of a Normal Distribution, *Annals of Mathematical Statistics* 20: 502–522.

Sobel, M. and Weiss, G. H. (1969). Play-the-Winner Rule and Inverse Sampling for Selecting the Best of $k \geq 3$ Binomial Populations, University of Minnesota, Department of Statistics, Technical Report No. 126, pp. 1–25.

Sobel, M. and Weiss, G. H. (1970). Play-the-Winner Sampling for Selecting the Better of Two Binomial Populations, *Biometrika* 57: 357–365.

Sobel, M. and Weiss, G. H. (1971a). A Comparison of Play-the-Winner and Vector-at-a-Time Sampling for Selecting the Better of Two Binomial Populations with Restricted Parameter Values, *Trabajos de Estadistica Y de Investigation Operativa* 22: 195–206.

Sobel, M. and Weiss, G. H. (1971b). Play-the-Winner Rule and Inverse Sampling in Selecting the Better of Two Binomial Populations, *Journal of American Statistical Association* 66: 545–551.

Sobel, M. and Weiss, G. H. (1972). Recent Results on Using Play the Winner Sampling Rule with Binomial Selection Problems, *Proceedings of Sixth Berkeley Symposium on Mathematical Statistics and Probability* 1: 717–736.

Taheri, H. and Young, D. H. (1974). A Comparison of Sequential Procedure for Selecting the Better of Two Binomial Populations, *Biometrika* 61: 585–592.

Tamura, R. N., Faries, D. E., Andersen, J. S., and Heiligenstein, J. H. (1994). A Case Study of an Adaptive Clinical Trial in the Treatment of Out-Patients with Depressive Disorder, *Journal of American Statistical Association* 89: 768–776.

Turnbull, B. W., Kaspi, H., and Smith, R. L. (1978). Sequential Adaptive Sampling Rules for Selecting the Best of Several Normal Populations, *Journal of Statistical Computation and Simulation* 7: 133–150.

Ware, J. H. (1989). Investigating Therapies of Potentially Great Benefit: ECMO (C/R: p306–340), *Statistical Science* 4: 298–306.

Wei, L. J. and Durham, S. (1978). A Randomized Play-the-Winner Rule in Medical Trials, *Journal of American Statistical Association* 73: 838–843.

Zelen, M. (1969). Play the Winner Rule and the Controlled Clinical Trial, *Journal of American Statistical Association* 64: 131–146.