Probability in the Engineering and Informational Sciences, **20**, 2006, 383–411. Printed in the U.S.A.

THE SCIENTIFIC CONTRIBUTIONS OF MILTON SOBEL

INGRAM OLKIN
Department of Statistics
Stanford University
Stanford, CA
E-mail: iolkin@stat.stanford.edu

THOMAS SANTNER
Department of Statistics
The Ohio State University
E-mail: tjs@stat.ohio-state.edu

Y. L. TONG
School of Mathematics
Georgia Institute of Technology
E-mail: yung.tong@earthlink.com

This article describes the scientific contributions of Milton Sobel. It motivates his research by considering his family background, his war experiences, and his mentors and fellow students at Columbia University. His research in sequential analysis, selection, ranking, group testing, and probabilistic combinatorics are highlighted.

1. INTRODUCTION

Milton Sobel was born on August 30, 1919, in Manhattan, New York and died at his home in Santa Barbara, California on December 31, 2002. Milton's parents were Hungarian immigrants who arrived in New York City in 1915 and, over time, lived in four of the five boroughs. Milton was one of four children; he had a younger and an older brother, and a sister who died at an early age. The Great Depression was ongoing during his first 20 years, and he had to help finance his education by working in the family grocery store. The goal of attending medical school was unfulfilled, and he turned to mathematics, which was a fortunate decision for mathematics.

© 2006 Cambridge University Press 0269-9648/06 $16.00

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press



Milton Sobel

Milton's academic background started at the College of the City of New York (CCNY), and it was there that his future in mathematics became certain. Milton was influenced by Emil Post, who was a well-known logician, by Bennington P. Gill, who taught number theory, and by John Firestone, who taught statistics. At CCNY he joined the Math Club, where he came into contact with a stellar group: Kenneth Arrow, Herman Chernoff, Herbert Solomon, and Harvey Cohn, with whom he interacted for many years. He was awarded a Bachelor of Science degree in 1940.

After a short period at the Census Bureau working with W. Edwards Deming on sampling projects, Milton was drafted in 1940 and served first at the Army War College and later under General George E. Patton, Jr. in France. Milton was one of the first U.S. soldiers to enter the Nazi concentration camp at Mauthausen, Austria and to take part in its liberation.

During all of his adult life, Milton was never far from mathematics. Even while in the Army, he took correspondence courses from the University of Chicago. After his discharge in 1946 he attended graduate school at Columbia University, receiving a master's degree in mathematics in 1946 and a doctorate in mathematical statistics in 1951. The period of his graduate education coincided with the development of decision theory, due in part to Abraham Wald. Milton's dissertation was on a complete class sequential problem, but earlier he wrote a joint paper with Wald on a three-decision problem. It might very well be that this early work was the germ of his later work on ranking and selection problems.

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press

Milton wrote over 125 papers in a wide range of areas: decision theory, ranking and selection models, life testing, sequential analysis, group testing, combinatorics, and Dirichlet distributions. Starting in 1954, together with Robert Bechhofer, Jack Kiefer, Shanti Gupta, and Charles Dunnett, he began to develop what are called "ranking and selection models." The point of view is that of experimental design in that one seeks to determine a sample size that will choose the population with the largest parameter value with a given precision, at least when the parameters of the populations under study are separated by a given amount. Although we often think of choosing a best "treatment" in some context, the general culture attempts to provide alternatives that are satisfactory, and not to single out individual treatments. These procedures are more common in engineering and, in particular, in the context of reliability.

It might have been the ranking and selection methodology that led Milton to start conducting research in the area of group testing from 1966 to 1968. The basic idea is illustrated by the example in which each of N individuals is to be tested by a blood test with a positive or negative outcome. If none of the N individuals is positive, then we could have pooled the blood and conducted a single test. At the other extreme, if each individual is positive, then we would need N tests. The group testing problem is to design an "optimal" test for minimizing the expected number of tests.

An early paper in 1965 recognized a connection between the tails of multinomial distributions and Dirichlet distributions. This connection surfaced continually throughout Milton's life, with papers in 1973, 1977, and 1980, and, ultimately, in several books. The key point is that the multinomial and negative multinomial distributions provide models for a very diverse set of sampling and stopping procedures. Consider, for example, a multinomial model with K cells in which we sample until every cell has at least one observation, and we stop as soon as the difference between the maximum and minimum cell counts is greater than a preassigned number.

Milton displayed the "Paul Erdos style" in that he was an inveterate collaborator. One count lists 55 collaborators. He enjoyed collaboration and would travel to distant places to work with colleagues. Milton always gave his students due credit, and he was a devoted mentor. Some students stayed at his house, and almost all students partook in a least one meal at his house.

As we will see, Milton's research displays an underlying theme of how to conserve observations in a statistical design. This theme appears in his research on sequential analysis, ranking and selection, group testing, combinatorics, and applications of the Dirichlet distribution.

2. SEQUENTIAL ANALYSIS

As a Ph.D. student of Abraham Wald, Milton naturally possessed a strong research interest in the area of sequential analysis, especially during the early stage of his academic career. If we loosely interpret "sequential analysis" as theory and meth-

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press

ods for statistical sampling procedures and terminal decision rules under which the sample size is random, then Milton's contributions to sequential analysis include the following topics: (1) classical sequential analysis in the spirit of Wald's sequential probability ratio test (SPRT), (2) decision-theoretic approach to sequential analysis, (3) generalized SPRT (GSPRT): theory and applications, (4) sequential life testing, and (5) others, including two-stage sampling and the Chow–Robbins-type stopping rules, inverse-sampling and play-the-winner rules, and their applications.

In this section we provide a general overview of his contributions concerning these topics. The selection of the material is not intended to be exhaustive; we will try to state the problems he (and his coauthors) studied and the main results they obtained and to discuss the fundamental ideas for solving the problems. However, for the detailed mathematical proofs, the reader is referred to the original publications. For notational consistency within this section, the notation and symbols used here are not always in agreement with those in the original publications.

2.1. Classical Sequential Analysis: Sequential Multihypothesis Testing*

Sobel and Wald (1949) proposed a new sequential method for testing three hypotheses concerning a normal mean. This work was clearly related to Milton's thesis at Columbia University, and to our knowledge, it is the first published research paper in the history of sequential analysis that extends the two-decision problem to multihypothesis testing.

For a sequence $\{X_i\}_{i=1}^{\infty}$ of independent and identically distributed normal random variables with mean θ and known variance σ^2, consider the following hypothesis-testing problem in a sequential setting:

$$H_1 : \theta < a_1 \text{ versus } H_2 : \theta \in [a_1, a_2] \text{ versus } H_3 : \theta > a_2, \tag{1}$$

where a_1 and a_2 are two given constants.

2.1.1. The Fundamental Approach: Two Separate SPRTs. The three hypotheses involved in (1) are composite hypotheses, but the theory and methods in sequential analysis that were available at that time (see, e.g., Wald [8]) dealt mostly with simple hypotheses and were for testing two hypotheses only. In view of this fact, Sobel and Wald (1949) solved this problem by first applying a combination of two separate Wald SPRTs, and this was a key idea in the paper. More specifically, one chooses four constants $\theta_1 < \theta_2 \leq \theta_3 < \theta_4$ such that

(i) $\theta_1 < a_1 < \theta_2 \leq \theta_3 < a_2 < \theta_4$,
(ii) $\theta_1 + \theta_2 = 2a_1$, $\theta_3 + \theta_4 = 2a_2$,
(iii) $\theta_2 - \theta_1 = \theta_4 - \theta_3 = \Delta$ (say).

One then carries out two SPRTs, R_1 and R_2, for testing the hypotheses:

$$H_1' : \theta = \theta_1 \text{ versus } H_2' : \theta = \theta_2 \tag{2}$$

*Publications by Sobel (as sole author and as coauthor) are listed in part A of the Bibliography.

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press

and

$$H_1'' : \theta = \theta_3 \text{ versus } H_2'' : \theta = \theta_4, \tag{3}$$

respectively, with type I error α and type II error β, where $\alpha, \beta \in (0,1)$ are preassigned. By applying the fundamental results in Wald [8, p. 119], it follows that test R_j ($j = 1,2$) calls for continuing sampling if and only if

$$(\sigma^2/\Delta)\log(B) + a_j n < \sum_{i=1}^{n} x_i < (\sigma^2/\Delta)\log(A) + a_j n,$$

where, as in Wald (1947),

$$A \approx (1-\beta)/\alpha \quad \text{and} \quad B \approx \beta/(1-\alpha).$$

For $j = 1$, when sampling terminates and the inequality on the left-hand side (right-hand side) is violated, then H_1' (H_2') is accepted. A similar statement applies for $j = 2$.

Note that both tests depend on the same sequence of observations $\{x_i\}_{i=1}^{\infty}$ and that when one of them calls for the termination of sampling, then that test is considered completed, and the remaining observations in the sequence apply only to the other test, which is not yet completed. Their paper includes a neatly drawn figure (Fig. 1) for illustrating the combined sampling procedure and the terminal decision rule.

2.1.2. The Operating Characteristic Function and the Probability of Correct Decision Function. Let R denote the combined sequential test, which consists of R_1 and R_2 as described earlier. In order to apply the test R to solve the original problem as defined in (1), Sobel and Wald (1949) proposed the following. In the test R, we have the following:

(i) If H_1' and H_1'' are accepted, then accept H_1 (in the original hypothesis-testing problem).
(ii) if H_2' and H_1'' are accepted, then accept H_2.
(iii) if H_2' and H_2'' are accepted, then accept H_3.

They show analytically that, under certain regularity conditions on (A, B), it is not possible to accept H_1' and H_2'' at the same time. With this understanding, a solution for the original problem becomes possible via the combined test R. Now let $L_\theta(H, R)$ denote the probability that the hypothesis $H \in \{H_1, H_2, H_3\}$ is accepted by R when θ is the true value of the parameter. Sobel and Wald (1949) provided an extensive analysis of the function $L_\theta(H, R)$ and its properties and gave analytical expressions for it in terms of A, B, θ, and $(\theta_1, \theta_2, \theta_3, \theta_4)$ for certain cases.

Because the intervals $[\theta_1, \theta_2]$ and $[\theta_3, \theta_4]$ can be regarded as indifference zones, they proceed to define the correct decision of a test for the hypotheses in (1) as follows:

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press

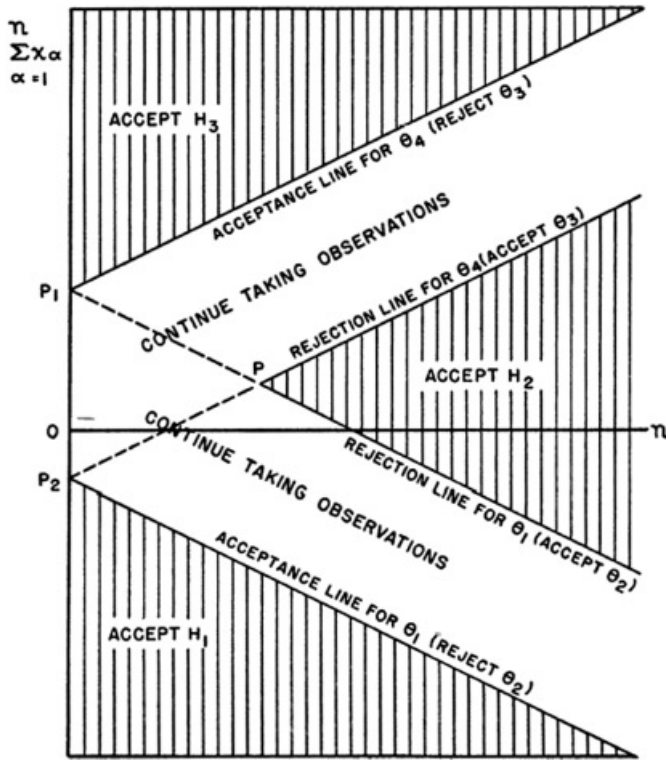


FIGURE 1. Decision and continuation regions for the sequential procedure for the three-hypothesis problem.

(i) When $\theta < \theta_1$, H_1 is accepted.
(ii) When $\theta \in [\theta_1, \theta_2]$, either H_1 or H_2 is accepted.
(iii) When $\theta \in (\theta_2, \theta_3)$, H_2 is accepted.
(iv) When $\theta \in [\theta_3, \theta_4]$, either H_2 or H_3 is accepted.
(v) When $\theta > \theta_4$, H_3 is accepted.

With this definition, the probability of a correct decision (PCD) of the test under the procedure R can be obtained in terms of the operating characteristic (OC) function $L_\theta(H, R)$, as given in their paper.

2.1.3. The Average Sample Number Function and Its Bounds. Let N, N_1, and N_2 be the random sample sizes required by the procedures R, R_1, and R_2, respectively; then, clearly, $N = \max\{N_1, N_2\}$. Thus,

$$E_\theta\{N\} \geq \max\{E_\theta\{N_1\}, E_\theta\{N_2\}\}.$$

By applying Wald's well-known first and second equations in sequential analysis, $E_\theta\{N_j\}$ can be obtained for $j = 1,2$, at least for certain values of θ. Thus the above

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press

inequality yields a lower bound on the average sample number (ASN) function of R. The upper bound given in the paper appears more involved. First, a new stopping variable N^* was introduced such that $N \leq N^*$ holds with probability 1 and for which the ASN function for N^* is more manageable. Based on the numerical results given in the paper, it appears that the lower bound and the upper bound on $E_\theta\{N\}$ are very close, at least for certain specified values of θ.

2.2. Decision Theory in Sequential Analysis

In addition to sequential analysis, decision theory was another subject that was developed and established by Wald. Thus, Milton naturally was led to work in this area. In fact, his 1953 paper (Sobel 1953) contained a very fundamental result on the decision-theoretical aspects of a sequential test described below. In addition to that main theorem, the paper also contained results on some terminating properties of Bayes rules. The work represents one of the earliest papers on studying properties of sequential procedures from a decision-theoretic viewpoint.

Let $\{X_i\}_{i=1}^{\infty}$ be a sequence of independent and identically distributed random variables with distribution function $F(x)$, where

$$dF(x) = \psi(\theta)e^{\theta x}\, d\nu(x), \qquad x \in \chi, \tag{4}$$

for some measure ν, and $\theta \in \Omega$ is an unknown parameter. If ν is chosen to be the Lebesgue measure or the counting measure and if the corresponding density function is of the form

$$f_\theta(x) = \psi_1(\theta)\psi_2(x)e^{\psi_3(\theta)\psi_4(x)}, \qquad x \in \chi, \tag{5}$$

where both ψ_3 and ψ_4 are increasing functions, then the model in (4) still applies after a reparameterization process. Consequently, as indicated in Sobel (1953), the results in the paper apply to the familiar distributions in the exponential family, including the binomial, Poisson, exponential, gamma, Normal, and a number of other families.

Now consider the one-sided hypothesis-testing problem:

$$H_1 : \theta \leq \theta_0 \text{ versus } H_2 : \theta > \theta_0,$$

where θ_0 is a given point in Ω, and consider a class of sequential decision rules A such that $\delta \in A$ if and only if δ possesses the following properties:

(i) After taking n observations $(x_i, \ldots, x_n) \equiv \boldsymbol{x}_n$, δ depends on $\boldsymbol{x}_n$ only through n and $t_n = \sum_{i=1}^{n} x_i$.

(ii) δ specifies a sequence of closed intervals $I_n = [a_{1n}, a_{2n}]$, $n = 1,2,\ldots,$ such that the stopping variable N is defined by

$$N = \inf_{n \geq 1}\{n : t_n \notin I_n\}.$$

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press

(iii) When sampling terminates, accept H_1 if $t_n < a_{1n}$, and accept H_2 if $t_n > a_{2n}$. (There is also a randomization process involved if equality holds.)

A main result in Sobel (1953) is that

Under certain regularity conditions, the class of sequential rules A is essentially complete.

Some of the regularity conditions are that the loss function has certain monotonic properties, the sampling-cost function is bounded below, and the risk function is uniformly bounded for $\theta \in \Omega$.

The proof of this result depends on a key application of a theorem of Wald (see Wald [9, Thm. 3.19]) and is quite involved mathematically. Let $\bar{A}$ denote the closure of A (in the regular sense) and let D_b be the class of all decision rules whose risk functions are uniformly bounded in Ω. The paper shows the following:

(i) A is essentially complete related to $\bar{A}$.
(ii) $\bar{A}$ is essentially complete related to D_b.

2.3. Sequential Multihypothesis Testing

When applying the GSPRT to multihypothesis testing, the observations involved are not required to be independent and identically distributed random variables. More specifically, for $n = 1,2,\ldots,$ let $\boldsymbol{X}_n = (X_1,\ldots,X_n)$ have density function $f_n(\boldsymbol{x}_n)$, where $\boldsymbol{x}_n = (x_1,\ldots,x_n)$. Let

$$\boldsymbol{f} = (f_1(x_1), f_2(\boldsymbol{x}_2),\ldots)$$

be the true sequence of the density functions of $X_1, \boldsymbol{X}_2, \ldots$. For $j = 1,\ldots,K$, let

$$\boldsymbol{f}_j = (f_{1,j}(x_1), f_{2,j}(\boldsymbol{x}_2),\ldots)$$

be K possible sequences of $\boldsymbol{f}$ density functions. Consider the sequential multihypothesis testing problem

$$H_1 : \boldsymbol{f} = \boldsymbol{f}_1 \text{ versus } H_2 : \boldsymbol{f} = \boldsymbol{f}_2 \text{ versus } \cdots \text{ versus } H_K : \boldsymbol{f} = \boldsymbol{f}_K.$$

Let $P^* \in (0,1)$ be preassigned and let $\pi_1,\ldots,\pi_K$ be given positive constants with sum one. [Note that π_j might be interpreted as the a priori probability that H_j is true in a Bayesian setting.] Let R denote the sequential procedure described below and let

$$\alpha_{jj^\star} = P_{H_j}\{H_{j^\star} \text{ is accepted under } R\}$$

denote the $(j, j^\star)$th element of the error matrix. Then, it follows from the general results in Bechhofer, Kiefer, and Sobel (BKS) (1968, Chap. 3) that under certain regularity conditions

(i) $P_{H_j}\{N < \infty\} = 1$ for each $j = 1,\ldots,K$
(ii) $\sum_{j=1}^{K} \pi_j \alpha_{jj} \geq P^*$.

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press

The procedure R, derived as a special case from BKS (1968, Sect. 3.1), consists of the following stopping rule and terminal decision rule:

Stopping Rule: Stop for the first time n when the inequality

$$P^* < \max_{1\le j\le K}\left\{\pi_j f_{n,j}(\boldsymbol{x}_n)\Big/\left[\sum_{j^\star=1}^{K}\pi_{j^\star} f_{n,j^\star}(\boldsymbol{x}_n)\right]\right\} \tag{6}$$

holds; otherwise, continue to observe x_{n+1}.

Terminal Decision Rule: When sampling terminates and the maximum on the right-hand side of (6) occurs with $j = J$, accept H_J and reject all other hypotheses. [For simplicity, we ignore the possibility of ties and the needed randomization process.]

2.4. Sequential Identification of Populations

As shown in BKS (1968), the solution of the sequential multihypothesis testing problem leads to that of a *sequential identification problem*, and we illustrate their approach below by considering this problem for the exponential family.

For $k \ge 2$, let there be k independent populations with density functions belonging to the exponential family and let $\theta_1, \ldots, \theta_k$ be given parameter values in Ω. The problem of interest is to identify the populations associated with $\theta_1, \ldots, \theta_k$. There are $K \equiv k!$ possible associations of populations with $\theta_1, \ldots, \theta_k$. Let $\boldsymbol{R}_0$ be the true (unknown) order and let $\boldsymbol{R}_1, \ldots, \boldsymbol{R}_K$ be the K permutations of the vector $(1, \ldots, k)$. Then this identification problem can be solved by testing the hypotheses

$$H_1 : \boldsymbol{R}_0 = \boldsymbol{R}_1 \text{ versus } H_2 : \boldsymbol{R}_0 = \boldsymbol{R}_2 \text{ versus } \cdots \text{ versus } H_K : \boldsymbol{R}_0 = \boldsymbol{R}_K.$$

Now, for $i = 1,2,\ldots$, let $\boldsymbol{X}^{(i)} = (X_{1i}, \ldots, X_{ki})$ be a sequence of independent and identically distributed random vectors such that for $\boldsymbol{R}_j = (r_j(1), \ldots, r_j(k))$, the joint density function of $\boldsymbol{X}_n = (\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(n)})$ under H_j is of the form

$$\begin{aligned} f_{n,j}(\boldsymbol{x}_n) &= \prod_{i=1}^{n}\prod_{m=1}^{k}\psi_1(\theta_{r_j(m)})\psi_2(x_{m,i})\exp[\psi_3(\theta_{r_j(m)})\psi_4(x_{m,i})] \\ &= \left[\prod_{m=1}^{k}\psi_1(\theta_m)\right]^n\left[\prod_{i=1}^{n}\prod_{m=1}^{k}\psi_2(x_{m,i})\right]\exp\left[\sum_{m=1}^{k}\psi_3(\theta_{r_j(m)})t_{m,n}(\boldsymbol{x}_n)\right], \end{aligned}$$

where $t_{m,n}(\boldsymbol{x}_n) = \sum_{i=1}^{n}\psi_4(x_{m,i})$ for $m = 1, \ldots, k$.

Consequently, when ψ_3 and ψ_4 are both monotonically increasing functions (as in most applications), the stopping rule in (6) reduces to the following: Stop for the first time n for which

$$P^* < \max_{1\le j\le K}\pi_j \exp\left[\sum_{m=1}^{k}\psi_3(\theta_{r_j(m)})t_{m,n}(\boldsymbol{x}_n)\right]\Big/\Delta, \tag{7}$$

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press

where

$$\Delta = \sum_{j^\star=1}^{K} \pi_{j^\star} \exp\left[\sum_{m=1}^{k} \psi_3(\theta_{r_{j^\star}(m)}) t_{m,n}(\boldsymbol{x}_n) \right].$$

Now the question is when (i.e., for which permutation) the right-hand side of (7) reaches a maximum. This question can be answered by applying an algebraic inequality, known as the Hardy–Littlewood–Pólya inequality (see Hardy, Littlewood, and Pólya [5] or Marshall and Olkin [6]): Let $\boldsymbol{a} = (a_1, \ldots, a_k)$ and $\boldsymbol{b} = (b_1, \ldots, b_k)$ be two real vectors and let

$$a_{(1)} \le \cdots \le a_{(k)}, \qquad b_{(1)} \le \cdots \le b_{(k)}$$

be their ordered components. Then for the dot product of $\boldsymbol{a}$ and $\boldsymbol{b}$,

$$\sum_{m=1}^{k} a_{(m)} b_{(k-m+1)} \le \sum_{m=1}^{k} a_m b_{r(m)} \le \sum_{m=1}^{k} a_{(m)} b_{(m)}$$

holds for all permutations $\boldsymbol{R} = (r(1), \ldots, r(k))$. It then follows from this result that in the sequential identification problem, the maximum (over all of the permutations) occurs when the θ's and the t's are similarly ordered. This result, of course, has an obvious but important implication on the stopping rule defined in (7) and the corresponding terminal decision rule.

In a related paper, Sobel (1954) provided a generalization of the Hardy–Littlewood–Pólya inequality. That generalization gives a refinement of that well-known result and yields a partial ordering on the values of the likelihood function for this sequential identification problem. This then, together with the introduction of the indifference-zone formulation, provides a sequential solution for the ranking and selection problem for the exponential family, as shown in BKS (1968).

2.5. Sequential Life Testing

Although the exponential distribution has been known for a long time, properties of order statistics and truncated data had not been explored. The paper by Epstein and Sobel (1954) had a great impact, and the exponential distribution became an important model in life testing after the appearance of this paper.

For the nonsequential case, assume that the life lengths of items in a population follow an exponential distribution with mean θ; that is, with density function

$$f_\theta(x) = \frac{1}{\theta} e^{-x/\theta}, \qquad x > 0, \theta > 0.$$

The estimation of θ is carried out by drawing n items at random from the population and placing them all together on a life test. Let $X_1, \ldots, X_n$ be the life lengths of the items, let

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press

$$X_{(1)} \leq \cdots \leq X_{(n)}$$

be their order statistics, and let

$$Y_i = X_{(i)} - X_{(i-1)}, \qquad i = 1, \ldots, n,$$

denote the spacings (where $X_{(0)} \equiv 0$). Then $Y_1, \ldots, Y_n$ are independent exponential random variables with means $(n - i + 1)^{-1}\theta$ for $i = 1, \ldots, n$, respectively. Epstein and Sobel (1954) showed that by writing the $X_{(i)}$'s in terms of the Y_i's, the random variables

$$\hat{\theta}_r = \frac{1}{r}\left[\sum_{i=1}^{r} X_{(i)} + (n - r)X_{(r)}\right], \qquad 1 \leq r \leq n,$$

are unbiased estimators of θ, and this leads to the possibility of estimating θ based on truncated data only. This idea was a major contribution in life testing.

Epstein and Sobel (1955) proposed a sequential test for θ for testing

$$H_1 : \theta = \theta_1 \text{ versus } H_2 : \theta = \theta_2$$

with a preassigned type I error α and type II error β. The procedures considered include both the cases of replacement and nonreplacement of units being tested. They applied Wald's fundamental theorems to obtain the sequential likelihood ratio tests, gave approximate formulas for the OC function, and obtained results on the expected number of failures ($E_\theta\{V\}$) and the expected waiting time ($E_\theta\{T\}$) before a decision is made. For the replacement case, a fundamental identity is that

$$E_\theta\{T\} = (\theta/n)E_\theta\{V\}.$$

We note that this result is remarkably similar in spirit to Wald's celebrated first equation in sequential analysis.

2.6. Other Topics

In addition to the topics discussed above, Milton Sobel's contributions to the general area of sequential analysis include two-stage sampling (Bechhofer, Dunnett, and Sobel (1954)) and multistage sampling (Robbins, Sobel, and Starr (1968); Sobel and Starr (1975)), inverse-sampling and play-the-winner rules (Sobel and Weiss (1970b, 1972a, 1972b)), and others. Some of those papers involve more on applications to ranking and selection problems and are discussed in Section 3.

3. SELECTION AND SCREENING

For simplicity of exposition, initially consider an experiment involving t treatments. Suppose that each treatment produces Gaussian distributed observations with possibly different means but common variability. Let $\{X_{ij}\}_{j \geq 1}$ denote the observations from the ith treatment; in the notation of Section 2, $X_{ij} \sim N(\mu_i, \sigma^2)$, for $j \geq 1$.

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press

In this simple setup, the so-called one-way layout, only the means $\mu_1, \ldots, \mu_t$ distinguish the treatments.

The earliest statistical interest in this model was concerned with testing the homogeneity hypothesis

$$H_0 : \mu_1 = \cdots = \mu_t . \tag{8}$$

From a practical viewpoint, testing *exact* equality of means is seldom, if ever, of interest to researchers; most experimentalists would never regard the means of their experimental treatments as precisely equal. Instead, they are concerned with some method of quantifying the differences among the treatment means or with the magnitude of the range of the means. In this spirit, selection and screening methods are concerned with two practical experimental objectives. To describe these approaches, let

$$\mu_{(1)} \leq \cdots \leq \mu_{(t)}$$

denote the ordered means and call the treatment having mean $\mu_{(t)}$ the *best treatment*.

In its simplest formulation, *selection* is that branch of statistical science that is concerned with identifying (selecting) the best treatment. More sophisticated selection goals are also of interest to researchers; for example, a researcher might wish to identify the best s of the t treatments (i.e., those with means $\mu_{(t-s+1)}, \ldots, \mu_{(t)}$). Selection procedures often suggest themselves quite naturally. For example, the statistical method that selects the treatment having the largest sample mean, $\max_{1 \leq i \leq t} \overline{X}_i$, as being the treatment having the largest population mean $\mu_{(t)}$ is intuitive. Much of the early work in the selection literature is concerned with issues of design. How many (common) observations must be taken per treatment in order to achieve some design criteria? In a 1954 paper, Bechhofer [1] introduced the "indifference zone" criterion for this problem when variability σ^2 is known; the best treatment should be correctly identified by the above naive procedure with a *given* probability $P^\star$ whenever $\mu_{(t)} - \mu_{(t-1)} \geq \delta^\star$, where $\delta^\star > 0$ is a practically important treatment difference specified by the experimenter.

Screening procedures, also called *subset selection* or *ranking* procedures, are concerned with forming a subset of the t treatments in such a way that the subset contains the best treatment. Screening procedures were popularized by Gupta (see Gupta [3]; Gupta and Panchapakesan [4] and the references therein). As with selection, there are many other goals of research interest such as identifying all treatments whose means μ_i exceed a (known) standard or identifying the treatments whose means exceed the (unknown) mean of a control treatment (Gupta and Sobel 1958a). Screening procedures have some similarities to confidence interval methods. The latter form a random-size set of values that contain a true unknown parameter with a given probability, regardless of what the value of that unknown parameter might be. Practical screening procedures choose a random-size subset of the t treatments in such a way that the subset contains the best treatment with a given probability $P^\star$, regardless of what the underlying vector of treatment means. The

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press

(expected) number of treatments in the selected subset depends on the true configuration of the t treatment means. For any reasonable screening procedure, the expected number of treatments in the selected subset decreases as the distance between the best treatment mean and the remaining means, $\mu_{(t)} - \mu_{(t-1)}$, increases.

Both selection and screening procedures are multiple-decision procedures. The reader will recall that Milton already had experience with multiple-decision problems, sequential analysis, and exponential families in his thesis and in other research with Abraham Wald. In addition, he was the fellow Columbia student of Robert Bechhofer and Charles Dunnett, both of whom were also working on advanced degrees and had a commitment to a more pragmatic, nontesting orientation. Milton's exposure to this alternative statistical philosophy, his natural curiosity, and previous research expertise were a potent mix. These influences led him to publish over 55 solely or jointly authored journal articles and two jointly authored books and to guide numerous Ph.D. dissertations in the fields of selection and ranking. Milton Sobel, his students, and other co-workers contributed to both the selection and screening literatures. Among the themes that occur in these papers are proposals for sequential and related methods for designing experiments that better conserve scarce experimental resources. They developed nonparametric procedures for both selection and screening. They proposed numerous novel formulations of problems that attempt to unify the selection and screening literatures. Of special note are Milton's many contributions to multinomial data procedures. Some examples of these ideas are described in the following subsections.

3.1. Sequential and Related Methods

In some cases, sequential procedures were a necessity because single-stage procedures do not exist for every design objective. For example, there does not exist a single-stage sampling procedure that allows one to achieve a given probability $P^\star$ of correct selection for the Bechhofer design objective described in the previous paragraph when the common variance σ^2 is unknown. Bechhofer et al. (1954) devised a two-stage sampling procedure that solves this problem.

More generally, Milton was interested in the use of sequential procedures in order to conserve scarce experimental resources. The book written by Bechhofer, Kiefer, and Sobel (1968) was (and remains) a major addition to the area of sequential analysis, especially on the topics of sequential identification of populations (see Section 2.4) and sequential ranking and selection problems. This research monograph summarizes 15 years of research on both the theory and applications of a class of GSPRTs for any exponential family of distributions, as specified in Eq. (5), along the line of Wald's SPRT, but it goes far beyond that. Thus, the theory is developed simultaneously for sequential procedures for many important parametric families. Recall that Wald's SPRT can be interpreted as a sequential version of the celebrated Neyman–Pearson test, and its stopping rule depends on the sequential likelihood ratio defined by the two hypotheses under consideration. In BKS (1968), a GSPRT is given that depends on the generalized sequential likelihood ratios. It is

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press

based on a decision-theoretic approach, and its fundamental properties are carefully investigated. In particular, the earlier chapters of the book provide analytical results on the PCD function, the ASN functions, and their bounds. [A lower bound on the ASN, known as Hoeffding's inequality (BKS 1968, p. 33), is fundamentally different from the bounds given in Sobel and Wald (1949).] After this is accomplished, the latter parts of the book deal with applications of the basic theorems, especially to the sequential identification and ranking and selection problems for the exponential family.

In addition to generalizing the sequential likelihood ratio heuristic, Milton used a number of other intuitive devices as the basis for adaptive sequential sampling and selection decision rules, particularly in the case of discrete observations. One important tool that he emphasized was "play-the-winner" (PW) adaptive sampling, used in settings that yield binomial or multinomial data. For example, Sobel and Weiss (1970a) studied the case of $t = 2$ treatments, each of which produces 0/1 Bernoulli responses X_{ij}, where $p_i = P\{X_{ij} = 1\}$ for $i = 1,2$ and $j \geq 1$. This is the model for the so-called "two-armed bandit" problem. It is also appropriate in medical settings where two treatments are compared to determine the treatment having the higher probability of a "cure."

Let $p_{(1)} \leq p_{(2)}$ denote the ordered success probabilities. Sobel and Weiss's (1970a) goal was to select the treatment having success probability $p_{(2)}$. The design objective was to correctly make this decision with at least a given probability $P^\star$ whenever $p_{(2)} - p_{(1)} \geq \delta^\star$, where $\delta^\star > 0$ is given. They considered the intuitive stopping rule that ceases sampling when the absolute value of the difference between the total number of successes of the two treatments is equal to a given number r. Upon stopping, their decision rule chooses the treatment having the greater number of successes. The only aspect of the statistical procedure remaining to be defined is the sampling scheme—in this case, PW sampling. In the setting of the two-armed slot machine, the PW rule assumes that the choice of arm to pull next, say the left or the right arm, is made after seeing the win/lose result for the previous pull of the left or right arm. In a medical setting, one of the two treatments must be assigned to the next in a sequence of identical experimental units after seeing the 0/1 result when using the previous treatment assignment. PW chooses the same arm (treatment) that was used in the previous observation if the previous observation resulted in a success, and PW switches to the other treatment if a failure has been observed in the previous observation. PW sampling myopically attempts to increase the number of successes produced in the course of the experiment; this is clearly an ethical imperative in the case of medical experiments and an economic desirability in the case of two-armed bandits. (See also Nebenzahl and Sobel (1972); Sobel and Weiss (1971a, 1971b, 1972a, 1972b).)

Another adaptive sampling scheme for discrete response problems that Milton explored in numerous papers was inverse sampling (Cacoullos and Sobel 1966; Sobel and Weiss 1970a, 1971b, 1972a, 1972b; Chen and Sobel 1984a, 1984b; Ebneshahrashoob and Sobel 1994; Sobel and Chen 1992). As an example, consider a multinomial distribution with t cells, having cell probabilities $p_1, \ldots, p_t$ and cell

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press

counts $(X_1, \ldots, X_t)$. Suppose that the goal is to identify the cell having the largest success probability. The cell with largest p_i will (stochastically) have the largest cell count and an intuitive decision rule chooses the cell associated with the largest cell count. A fixed-sample rule chooses a total number of multinomial observations to sample and uses the decision rule stated in the previous sentence. A typical inverse sampling rule would be to continue sampling from the multinomial distribution until the difference between two largest cell counts is a given value. For many design objectives, it can be shown that appropriate inverse-sampling rules have smaller expected sample sizes than fixed-sample rules that achieve the same design objective.

Another important tool that Milton used to gain efficiency in sampling is *curtailment*. The curtailment philosophy is that an experiment should be stopped (curtailed) as soon as the final decision is clear based on the data collected so far. As a simple example of curtailment, suppose that $n = 4$ observations are to be collected from each of $t = 3$ treatments. The decision rule is to select the treatment having the largest proportion of successes. Suppose that at a certain point in the sampling, the following results have been obtained.

Treatment		
1	2	3
S	F	S
F	F	S
F	?	S
?	?	?

In this case $\hat{p}_3 \geq 0.75$ (with equality if the fourth observation on Treatment 3 is an F) whereas $\hat{p}_2 \leq 0.50$ (with equality if the third and fourth observations on Treatment 2 are S) and $\hat{p}_1 \leq 0.50$ (with equality if the fourth observation on Treatment 1 is an S). Thus, the decision rule can be implemented to choose Treatment 3 without collecting further data. The book by Gibbons, Olkin, and Sobel (1977) emphasizes the benefits of curtailment in decreasing sampling costs. Milton emphasized the use of curtailment in conjunction with adaptive sequential sampling plans to further reduce the number of observations required in discrete model problems.

3.2. Nonparametric Methods

One weakening of the one-way layout specifies that the observations $\{X_{ij}\}_{j \geq 1}$ from the ith treatment, $1 \leq i \leq t$, are independent but have arbitrary continuous distributions $F_i(\cdot)$. Let ξ_i^{α} denote the α-quantile of $F_i(\cdot)$; for example, the population median of $F_i(\cdot)$ corresponds to $\alpha = \frac{1}{2}$. One analog of the problem of selecting the treatment with the largest mean is that of selecting the treatment having the largest α-quantile; this is clearest in the case of $\alpha = \frac{1}{2}$.

Beginning in the late 1960s, Sobel used this general α-quantile model for $\{X_{ij}\}_{j \geq 1}$ as the basis for a series of extensions of the original normal theory selection and screening methods. The key technical difficulty of these papers is in forming a rea-

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press

sonable design goal. Recall that the normal-theory design requirement is that a correct selection must occur with at least a given probability $P^\star$ whenever the two largest means are $\mu_{(t)} - \mu_{(t-1)} > \delta^\star > 0$ or more apart. The technical issue that occurs in the nonparametric version of this problem is that two arbitrary distribution functions can have α-quantiles whose difference is arbitrarily large and yet be within any $\epsilon > 0$ of one another for all arguments x. (Figure 2 illustrates such a pair of cumulative distribution functions.)

Thus, design objectives that require that the probability of correct selection be at least $P^\star$ cannot be based solely on the requirement that the vector of underlying distributions $\boldsymbol{F} = (F_1(\cdot), \ldots, F_t(\cdot))$ satisfy $\xi^\alpha_{(t)} - \xi^\alpha_{(t-1)} > 0$, where $\xi^\alpha_{(i)}$ denotes the ith ordered α-quantile, but must specify additional restrictions on $\boldsymbol{F}$.

Sobel (1967) introduces a clever method of adding an additional separation requirement to $\boldsymbol{F}$ so that when added to the $\boldsymbol{F}$ restriction $\xi^\alpha_{(t)} - \xi^\alpha_{(t-1)} > 0$, the combination can be used as a design requirement for procedures that correctly identify the s treatments having the largest α-quantiles in a set of t treatments. Rizvi and Sobel (1967) constructed screening procedures for selecting a (random-sized) subset of the t treatments so as to include the treatment having the largest α-quantile. Rizvi, Sobel, and Woodworth (1968) studied the selection of a subset of treatments

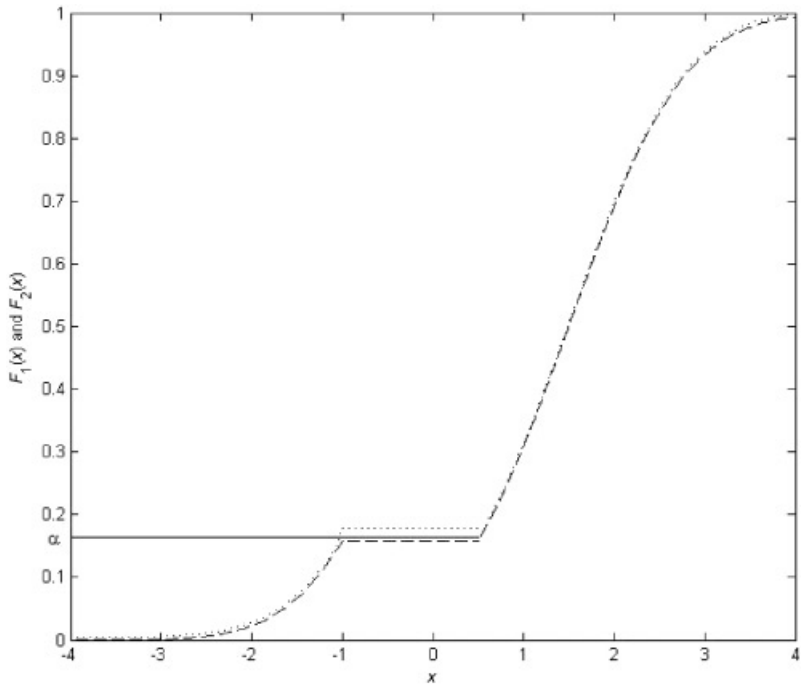


FIGURE 2. Two distributions $F_1(\cdot)$ and $F_2(\cdot)$ with α-quantiles that are $\delta = 1.5$ apart and for which $\max_x |F_1(x) - F_2(x)|$ can be made arbitrarily small.

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press

whose α-quantiles exceed the α-quantile of a control treatment. (For additional nonparametric contributions, see Desu and Sobel (1971); Sobel (1985b).)

3.3. Novel Formulations of Selection and Ranking Problems

Milton and his collaborators proposed numerous generalizations of the original Bechhofer selection/design requirement and Gupta screening rule. Desu and Sobel (1968) proposed statistical rules and design criteria for selecting a subset of fixed size s of t treatments so as to contain the r best treatment means ($r \leq s \leq t$). Sobel and Tong (1971) studied a statistical method for partitioning a set of t treatments with respect to a control treatment with (unknown) mean. Sobel and Wu (1984) introduced a sequential subset selection rule.

The experimental design requirement that was originally emphasized by Bechhofer, and in other settings by later authors, mandated that correct identification of the best treatment occurs with a given probability $P^\star$ as long as the best and nonbest treatments were sufficiently separated (in the motivating example of this section for $\boldsymbol{\mu}$ such that $\mu_{(t)} - \mu_{(t-1)} \geq \delta^\star$ where $\delta^\star > 0$ is specified by the experimenter). Such design requirements do not care about the decision a procedure makes when the true $\boldsymbol{\mu}$ has the best treatment mean that is "close" to one or more other treatment means; these $\boldsymbol{\mu}$ (i.e., those satisfying $\mu_{(t)} - \mu_{(t-1)} < \delta^\star$), form the so-called *indifference zone* of the parameter space. A number of Sobel's papers were motivated by the desire to better understand the performance of existing selection rules when the true parameter lay in the indifference zone and to devise new procedures whose performance was "reasonable" for any true parameter vector.

As an example, Chen and Sobel (1985) studied the naive procedure for selecting the best cell from a t-cell multinomial distribution where the best cell is that cell having the largest cell probability. Upon completing sampling, this procedure selects the cell having the largest cell count. They introduced a new design requirement in the context of this problem. Given a $\delta^\star < 1$, they required that a sufficiently large number of multinomial vector observations be taken so that, with a given probability $P^\star$, the selected cell have cell probability p_i that satisfies

$$p_i > \delta^\star p_{(t)}, \tag{9}$$

where $\delta^\star < 1$ and $p_{(1)} \leq \cdots \leq p_{(t)}$ denote the ordered cell probabilities. Notice that $p_{(t)} > \delta^\star p_{(t)}$ for every $\boldsymbol{p}$; hence, every $\boldsymbol{p}$ must have one or more of these near-best cells (i.e., cells that satisfy (9)). If

$$\delta^\star p_{(t)} \geq p_{(t-1)},$$

then only the cell having maximum cell probability $p_{(t)}$ satisfies (9) and the design requirement dictates that it must be selected with probability $P^\star$.

The design philosophy of Chen and Sobel (1985) is that any cell whose p_i is "close" to $p_{(t)}$ is, from a practical viewpoint, equivalent to the best cell. *This means that the definition of correct selection should depend on the true* $\boldsymbol{p}$. This type of

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press

reasoning suggested statistical methods that combine the idea of choosing a fixed number of treatments, as in classical selection procedures, with that of choosing a random number of treatments, as in classical screening procedures. As an example of this final idea, Chen and Sobel (1987a) devised procedures that select the single best treatment with a given probability $P_1^\star$ whenever $\boldsymbol{\mu}$ contains a best treatment mean that is sufficiently better than all remaining means (i.e., $\boldsymbol{\mu}$ satisfies $\mu_{(t)} - \mu_{(t-1)} \geq \delta^\star$). In addition, they required that their rule select a subset of treatments so as to include the best treatment with a given probability $P_0^\star$ whenever $\boldsymbol{\mu}$ satisfies $\mu_{(t)} - \mu_{(t-1)} < \delta^\star$.

4. PROBABILISTIC COMBINATORICS

Milton was always attracted to combinatorial problems, especially those that had a probabilistic basis. This generally was the case in both ranking and selection methods and in some sequential procedures. This section contains other combinatorial illustrations.

4.1. A Partitioning Model

The paper by Arrow, Pesotchinsky, and Sobel (1981) presented a problem in which from a sample of n individuals we wish to search for the k largest observations. The restriction is that the data consist of responses to dichotomous questions; namely Is $x > c_1$ or $x \leq c_1$, where c_1 is a threshold at the option of the researcher? The defined procedure is to create two subsets: those who answer positively and those who answer negatively. From the subset of those answering positively, we again ask a dichotomous question: Is $x > c_2$ or $x \leq c_2$, and so on? The problem is how to choose the thresholds.

Let π_{ij} denote the probability that j responses out of i are positive for any single question. If the X's are independent, then the optimal procedure that minimizes the expected number of questions is obtained by a Markov decision procedure with transition probabilities π_{ij} not dependent on the space of trajectories of the underlying random process.

The authors illustrated this procedure for the case $n = 9$ and $k = 3$ with measures independently and uniformly distributed in the interval [500, 980]. First note that $\max \pi_{9,3} = \max\binom{9}{3} p^3(1-p)^6$ occurs at $p = \frac{1}{3}$, which is used to divide the interval into segments 500 to 820, and 820^+ to 980. The first question is whether $x > 820$ or $x \leq 820$. Suppose that one person responds positively; at the next stage, we divide the interval 500 to 820, into 500 to 740 and 740^+ to 820 based on $\max \pi_{8,2}$, which occurs at $p = \frac{1}{4}$, and the process then continues.

The important point is that if the X's come from a distribution F, the optimal solution in terms of the binomial coefficients do not depend on F. The surprising result is that for the case $k = 1$ and any n, the minimum expected number of questions is 2.43, and the procedure that maximizes the probability of terminating at the very next step has an expected number of questions equal to 2.44.

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press

4.2. Group Testing†

A group-testing procedure was first proposed by Dorfman [2] during World War II. Inductees in the military were required to pass a test for syphilis by means of a blood test that showed the presence or absence of the "syphilitic antigen." The unique aspect of this test is that if blood were pooled in groups, then a negative result indicated that everyone in the group was negative. On the other hand, a positive outcome indicated that at least one member of the group was positive, in which case each member of the group would be tested.

Dorfman [2] proposed dividing the total sample of size N into groups of a fixed size n using an estimate of the prevalence. If a group tested positively, then every member of the group would be tested individually. Sterrett [7] provided a modification of this method in the second stage by repeating the group testing model if a group tested positively. Sobel (1960, 1968) and Sobel and Groll (1959, 1966) attacked this problem in a very general way; in particular, when the underlying incidence is known and when it is unknown. They also treated variations that considerably complicate the problem. For example, if the blood sample can be separated into several samples, then an individual's blood can be used in different groups. They also are able to insert the group-testing procedure in a design context in screening experiments. In another modification, they considered the case in which the tests themselves exhibit some uncertainty.

They described a simplified and intuitive procedure that gives reasonable results when the probability of a defective (or positive) is small. First, test all N units in a single group, and if the test fails, test $N/2$ chosen at random, and continue halving in this group and similarly for the untested group. The optimal results are in the form of recurrence relations, and some tables help in the calculations.

All in all, Milton wrote more than 20 papers on the group testing problem over a period of 30 years. In addition to the papers listed above, we mention Pasternack, Sobel, and Thomas (1987a), Blumenthal, Kumar, and Sobel (1971c), Elashoff and Sobel (1975b), and Chen, Hsu, and Sobel (1987b).

4.3. Reliable Networks

Networks provided a natural setting that combined statistics and combinatorics that intrigued Milton. This was the subject that he was working on just before his death.

Consider a network (an undirected simple graph $G(n, k)$) consisting of n nodes and k edges, and suppose that the probability, p, of transmission from each node to a connected node is constant. Then the reliability of the graph $G(n, k)$ is

$$R(p) = \sum_{i=o}^{k} N_i \, p^i (1-p)^{k-i},$$

†Publications by Sobel (and coauthors) on Group Testing are listed chronologically in part B of the Bibliography.

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press

where N_i is the number of i-edge connected spanning subgraphs of G. Thus, each graph generates a reliability polynomial in p.

For example, for $n = 6$ and $k = 8$, two potential graphs are provided in Figure 3. The reliabilities of these graphs are

$$R(A) = 36p^5(1-p)^3 + 26p^6(1-p)^2 + 8p^7(1-p) + p^8,$$

$$R(B) = 32p^5(1-p)^3 + 24p^6(1-p)^2 + 8p^7(1-p) + p^8.$$

In this example $R(A) > R(B)$ for all $0 < p < 1$, so that the network G_1 is more reliable than the network G_2.

The object is to find uniformly most reliable graphs (UMRGs) for different choices of n and k. A first question is whether a UMRG exists. A conjecture that they do exist was proved false. Ath and Sobel (2000a) provided counterexamples to conjectures about the minimum diameter and maximum girth of UMRG. They also provided examples of UMRG; two such are given in Figure 4. In a second paper, Ath and Sobel (2000b) present a set of graphs that they conjectured are optimal.

4.4. Dirichlet Distribution

Dirichlet distributions arise in a variety of contexts, and most recently in a Bayesian context as a prior distribution for the parameters of the multinomial distribution. However, the origin of the Dirichlet distribution is considerably older, and it too has a combinatorial basis. Recall that the tail of a binomial distribution can be represented in terms of the incomplete beta integral (type I):

$$\sum_{j=k}^{n} \binom{n}{j} p^j (1-p)^{n-j} = \int_0^p \frac{t^{k-1}(1-t)^{n-k}}{B(k, n-k+1)}\, dt,$$

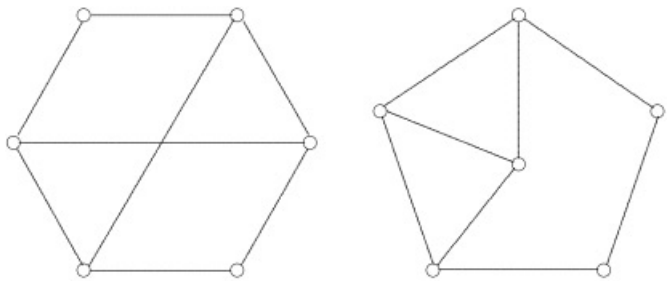


FIGURE 3. Two graphs, each with six nodes and eight edges.

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press

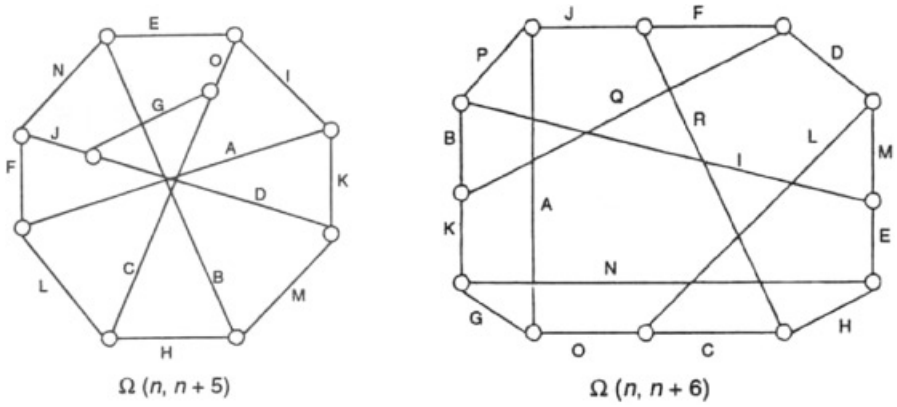


FIGURE 4. Two examples of UMRG graphs with n nodes and $n + 5$ and $n + 6$ edges.

where $B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ is the beta function. Similarly, the tail of the negative binomial distribution can be represented in terms of the incomplete beta integral (type II):

$$\sum_{j=s}^{\infty} \frac{\Gamma(r+j)}{\Gamma(r)j!} p^j (1-p)^r = \int_0^{p/(1-p)} \frac{t^{s-1}}{(1+t)^{r+s} B(r,s)}\, dt$$

$$= \int_0^p \frac{t^{s-1}(1-t)^{r-1}}{B(r,s)}\, dt.$$

Because the multinomial distribution is a multivariate version of the binomial distribution and the Dirichlet distribution is a multivariate version of the beta distribution, we can expect similar representations of the tails of a multinomial distribution as integrals of the Dirichlet distribution. However, now there are more options for sampling procedures.

Consider a multinomial model with $k + 1$ cells $C_1, \ldots, C_{k+1}$ with respective probabilities $p_1, \ldots, p_{k+1}$, where $p_{k+1} > 0$ and $\sum p_i = 1$. Let $s_1, \ldots, s_k$ denote nonnegative integers, and s a positive integer. Sampling is continued until cell C_{k+1} contains s observations,

For this sampling procedure, we denote two events of interest:

$\mathcal{E}_1$: At time of stopping, X_j, the number of items in cell C_j, is $\geq s_j, j = 1, \ldots, k$.

$\mathcal{E}_2$: At time of stopping, X_j, the number of items in cell C_j, is $\leq s_j - 1$, $j = 1, \ldots, k$.

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press

Olkin and Sobel (1965) show that

$$P\{\mathcal{E}_1\} = p_{k+1}^{s} \sum_{\substack{x_i = s_i \\ i=1,\ldots,k}} \frac{\Gamma(T+s)}{\Gamma(s)\prod x_i!} \prod p_i^{x_i}$$

$$= \int_0^{\theta_1} \cdots \int_0^{\theta_k} \frac{\prod z_j^{s_j - 1}}{\left(1 + \sum_1^k z_i\right)^{s+s_o}} \frac{\prod dz_i}{B(s_1,\ldots,s_k,s_o)},$$

where $\sum_1^{k+1} p_i = 1, T = \sum_1^k x_i, s_0 = \sum_1^k s_i$, and $B(s_1,\ldots,s_k,s_0) = \prod_0^k \Gamma(s_i)/\Gamma(\sum s_i)$ is the generalized beta function.

Similarly, for the lower tail,

$$P\{\mathcal{E}_2\} = \sum{}^{*} \frac{\prod_1^k p_i^{x_i}}{\prod_1^n x_i!}$$

$$= \int_{p_1}^{A_1} \int_{p_2}^{A_2} \cdots \int_{p_k}^{A_k} \frac{\prod_1^k z_i^{s_i - 1}(1 - \Sigma z_i)^{n - s_0}}{B(s_1,\ldots,s_k,n - s_0 + 1)} \prod dz_i,$$

$$A_1 = 1 - \sum_2^k p_i, A_2 = 1 - z_1 - \sum_3^k p_i, \ldots, A_k = 1 - \sum_1^{k-1} z_i,$$

where $\sum^{*}$ denotes the summation over all k-tuples with $0 \le x_j \le s_j - 1, j = 1,\ldots,k$, and $n = \sum_1^k x_i$.

From this beginning, Sobel, Uppuluri, and Frankowski (1977, 1985) published two books in which they provided detailed discussions of recursive relations, tables, and applications. Even after 20 years, integration in multidimensions is not an easy task, and their recursions have proved most useful.

Acknowledgments

The authors thank the Institute of Mathematical Statistics for permission to reproduce Figure 1 from *The Annals of Mathematical Statistics*, volume 20, 1949, pp. 502–522. We are greatly indebted to Richard Cottle for his comments and for so graciously helping us with the preparation of this manuscript.

References[‡]

1. Bechhofer, R.E. (1954). A single-sample multiple decision procedure for ranking means of normal populations with known variances. *The Annals of Mathematical Statistics* 25: 16–39.
2. Dorfman, R. (1943). The detection of defective members of large populations. *The Annals of Mathematical Statistics* 14: 436–440.

[‡]References in which Sobel was an author are contained in the Bibliography.

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press

3. Gupta, S.S. (1956). On a decision rule for a problem in ranking means. Ph.D. dissertation (Mimeo. Ser. No. 150), Institute of Statistics, University of North Carolina, Chapel Hill.
4. Gupta, S.S. & Panchapakesan, S. (1979). *Multiple decision procedures: Theory and methodology of selecting and ranking populations*. New York: Wiley.
5. Hardy, G.H., Littlewood, J.E., & Pólya, G. (1934). Some simple inequalities satisfied by convex functions, *Messenger Mathematics* 58: 145–152.
6. Marshall, A.W. & Olkin, I. (1979). *Inequalities—Theory of majorization and its applications*. New York: Academic Press.
7. Sterrett, A. (1957). On the detection of defective members of large populations. *The Annals of Mathematical Statistics* 28: 1033–1036.
8. Wald, A. (1947). *Sequential analysis*. New York: Wiley.
9. Wald, A. (1950). *Statistical decision functions*. New York: Wiley.

BIBLIOGRAPHY OF MILTON SOBEL

A. Sequential Analysis, Ranking and Selection, Combinatorics, and Other Topics

Anderson, K., Sobel, M., & Uppuluri, V.R.R. (1982). Quota fulfillment times. *The Canadian Journal of Statistics* 10: 73–88.

Arrow, K.J., Pesotchinsky, L. & Sobel, M. (1981). On partitioning a sample with binary-type questions in lieu of collecting observations. *Journal of the American Statistical Association* 76: 402–409.

Ath, Y., Ebneshahrashoob, M., Gao, T., & Sobel, M. (2002). Accessibility and reliability for connect bipartite graphs. *Methodology and Computing in Applied Probability* 3: 153–161.

Ath, Y. & Sobel, M. (2000a). Counterexamples to conjectures for uniformly optimally reliable graphs. *Probability in the Engineering and Informational Sciences* 14: 173–177.

Ath, Y. & Sobel, M. (2000b). Some conjectured uniformly optimal reliable networks. *Probability in the Engineering and Informational Sciences* 14: 375–383.

Barndorff-Nielsen, O. & Sobel, M. (1966). On the distribution of the number of admissible points in a vector random sample. *Theory of Probability and its Applications* 11: 249–269.

Bechhofer, R.E., Dunnett, C.W., & Sobel, M. (1954). A two-sample multiple decision procedure for ranking means of normal populations with a common unknown variance. *Biometrika* 41: 170–176.

Bechhofer, R.E., Gupta, S.S., & Sobel, M. (1986). *Modern statistical selection, part 1*. Columbus, OH: American Sciences Press.

Bechhofer, R.E., Gupta, S.S., & Sobel, M. (1988). Review of "Modern statistical selection, part 1" by H.R.G. Ahrens. *Biometrical Journal* 30: 46.

Bechhofer, R.E., Kiefer, J., & Sobel, M. (1968). *Sequential identification and ranking procedures.* Chicago: University of Chicago Press.

Bechhofer, R.E. & Sobel, M. (1954). A single-sample multiple decision procedure for ranking variances of normal populations. *The Annals of Mathematical Statistics* 25: 273–289.

Bernard, S.R., Sobel, M., & Uppuluri, V.R.R. (1981). A two-urn model of Pólya type. *Bulletin of Mathematical Biology* 43: 33–46.

Berry, D.A. & Sobel, M. (1973). An improved procedure for selecting the better of two Bernoulli populations. *Journal of the American Statistical Association* 68: 979–984.

Blumenthal, S., Christie, T., & Sobel, M. (1971). Best sequential tests for an identification problem with partially overlapping distributions. *Sankhyā, Series A* 33: 185–192.

Cacoullos, T. & Sobel, M. (1966). An inverse-sampling procedure for selecting the most probable event in a multinomial distribution. In P.R. Krishnaiah (ed.), *Multivariate analysis I*. New York: Academic Press, pp. 423–455.

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press

Chen, P., Panchapakesan, S., & Sobel, M. (1994). Selecting among the multinomial losers. *Sequential Analysis* 13: 177–200.

Chen, P. & Sobel, M. (1984a). Selecting among multinomial cells using inverse sampling: A generalized goal. *Statistics & Decisions* (Suppl. 1): 285–295.

Chen, P. & Sobel, M. (1984b). Selecting the *t* best cells of a multinomial using inverse sampling. In Y.L. Tong (ed.), *Inequalities in statistics and probability*. Hayward, CA: Institute of Mathematical Statistics, pp. 206–210.

Chen, P. & Sobel, M. (1985). On a new criterion for selecting the most probable cell in a multinomial distribution. In E.J. Dudewicz (ed.), *The frontiers of modern statistical inference procedures*. Columbus, OH: American Sciences Press, pp. 215–236.

Chen, P. & Sobel, M. (1987a). An integrated formulation for selecting the *t* best of *k* normal populations. *Communications in Statistics, Part A: Theory and Methods* 16: 121–146.

Chen, P. & Sobel, M. (1987b). A new formulation for the multinomial selection problem. *Communications in Statistics, Part A: Theory and Methods* 16: 147–180.

Chen, P., Sobel, M.J., & Sobel, M. (1988). Selection problem for a modified multinomial (voting) model. In S.S. Gupta & J.O. Berger (eds.), *Statistical decision theory and related topics IV*, Vol. 2. New York: Springer-Verlag, pp. 23–32.

Das Gupta, S., Eaton, M., Olkin, I., Perlman, M., Savage, L.J., & Sobel, M. (1972). Inequalities on the probability content of convex regions for elliptically-contoured distributions. In L.M. LeCam, J. Neyman, & E.L. Scott (eds.), *Proceedings of the sixth Berkeley symposium on mathematical statistics and probability*, Vol. II. Berkeley: University of California Press, pp. 241–265.

Desu, M.M. & Sobel, M. (1968). A fixed subset-size approach to the selection problem (correction 63: 685). *Biometrika* 55: 401–410.

Desu, M.M. & Sobel, M. (1971). Nonparametric procedures for selecting fixed-size subsets. In S.S. Gupta & J. Yackel (eds.), *Statistical decision theory and related topics*, Vol. 1. New York: Academic Press, pp. 255–274.

Dunnett, C.W. & Sobel, M. (1954a). A bivariate generalization of Student's t-distribution with tables for certain special cases. *Biometrika* 41: 153–169.

Dunnett, C.W. & Sobel, M. (1954b). A two-sample multiple decision procedure for ranking means of normal populations with a common unknown variance. *Biometrika* 41: 170–176.

Dunnett, C.W. & Sobel, M. (1955). Approximations to the probability integral and certain percentage points of a multivariate analogue of Student's t-distributions. *Biometrika* 42: 258–267.

Ebneshahrashoob, M., Gao, T., & Sobel, M. (2004). Double window acceptance sampling. *Naval Research Logistics* 51: 297–306.

Ebneshahrashoob, M. & Sobel, M. (1990). Sooner and later waiting time problems for Bernoulli trials: Frequency and run quotas. *Statistics & Probability Letters* 9: 5–11.

Ebneshahrashoob, M. & Sobel, M. (1994). Probability, waiting-time results for pattern and frequency quotas in the same inverse sampling problem via the Dirichlet. In G. Anastassiou & S.T. Rachev (eds.), *Proceedings of the conference on approximation, probability, and related fields*. London: Plenum, pp. 179–186.

Ebneshahrashoob, M. & Sobel, M. (1995). Dirichlet analysis for inverse multinomial with both quota and quota-free cells. *Computational Statistics and Data Analysis* 19: 293–307.

Ebneshahrashoob, M. & Sobel, M. (1997). Dirichlet applications to network problems. In *Computing Science and Statistics. Proceedings of the Second World Congress of IASC*, Vol. 29, pp. 492–497.

Epstein, B. & Sobel, M. (1953). Life testing. *Journal of the American Statistical Association* 48: 486–502.

Epstein, B. & Sobel, M. (1954). Some theorems relevant to life testing from an exponential distribution. *The Annals of Mathematical Statistics* 25: 373–381.

Epstein, B. & Sobel, M. (1955). Sequential life tests in the exponential case. *The Annals of Mathematical Statistics* 26: 82–93.

Frankowski, K., Sobel, M., & Uppuluri, V.R.R. (1975). Dirichlet type 1 distribution: Properties and applications. *Bulletin of the International Statistical Institute* 46: 267–271.

Gibbons, J.D., Olkin, I., & Sobel, M. (1999). *Selecting and ordering populations: A new statistical meth-*

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press

odology. Philadelphia: Society for Industrial and Applied Mathematics (originally published 1977, New York: Wiley).

Gibbons, J.D., Olkin, I., & Sobel, M. (1978). Baseball competitions: Are enough games played? *The American Statistician* 32: 89–95.

Gibbons, J.D., Olkin, I., & Sobel, M. (1979a). An introduction to ranking and selection. *The American Statistician* 33: 185–195.

Gibbons, J.D., Olkin, I., & Sobel, M. (1979b). A subset selection technique for scoring items on a multiple choice test. *Psychometrika* 44: 259–270.

Gupta, S.S., Huyett, M.J., & Sobel, M. (1957). Selection and ranking problems with binomial populations. *Transactions of the American Society for Quality Control* 11: 635–644.

Gupta, S.S. & Sobel, M. (1957). On a statistic which arises in selection and ranking problems. *The Annals of Mathematical Statistics* 28: 957–967.

Gupta, S.S. & Sobel, M. (1958a). On selecting a subset which contains all populations better than a standard. *The Annals of Mathematical Statistics* 29: 235–244.

Gupta, S.S. & Sobel, M. (1958b). On the distribution of a statistic based on ordered uniform chance variables. *The Annals of Mathematical Statistics* 29: 274–281.

Gupta, S.S. & Sobel, M. (1960). Selecting a subset containing the best of several binomial populations. In I. Olkin, S.G. Ghurye, W. Hoeffding, W.G. Madow, & H.B. Mann (eds.), *Contributions to probability and statistics, essays in honor of Harold Hotelling*. Stanford, CA: Stanford University Press, pp. 224–248.

Gupta, S.S. & Sobel, M. (1962a). On selecting a subset containing the population with the smallest variance. *Biometrika* 49: 495–507.

Gupta, S.S. & Sobel, M. (1962b). On the smallest of several correlated F-statistics. *Biometrika* 49: 509–523.

Hadian, A. & Sobel, M. (1970a). Selecting the t-th largest using binary errorless comparisons. In P. Erdös, A. Rényi, & V.T. Sós (eds.), *Combinatorial theory and its applications II*. Amsterdam: North-Holland, pp. 585–599.

Hadian, A. & Sobel, M. (1970b). Ordering the t largest of n items with binary comparisons. In R.C. Bose & D.A. Dowling (eds.), *Proceedings of the second Chapel Hill conference on combinatorial mathematics and its applications*. Chapel Hill: University of North Carolina, pp. 238–251.

Hoel, D. & Sobel, M. (1971). Comparisons of sequential procedures for selecting the best binomial populations, with tables and comparisons. In L.M. LeCam, J. Neyman, & E.L. Scott (eds.), *Proceedings of the sixth Berkeley symposium on mathematical statistics and probability*, Vol. IV. Berkeley: University of California Press, pp. 53–69.

Hoel, D.G., Sobel, M., & Weiss, G.H. (1972). A two-stage procedure for choosing the better of two binomial populations. *Biometrika* 59: 317–322.

Hoel, D.G., Sobel, M., & Weiss, G.H. (1975a). A survey of adaptive sampling for clinical trials. In R.M. Elashoff (ed.), *Perspectives in biometrics*, Vol. 1. New York: Academic Press, pp. 29–62.

Hoel, D.G., Sobel, M., & Weiss, G.H. (1975b). Comparison of sampling methods for choosing the best binomial population with delayed observations. *Journal of Statistical Computation and Simulation* 3: 299–313.

Katz, L. & Sobel, M. (1972). Coverage of generalized chess boards by randomly placed rooks. In L.M. LeCam, J. Neyman, & E.L. Scott (eds.), *Proceedings of the sixth Berkeley symposium on mathematical statistics and probability*, Vol. III. Berkeley: University of California Press, pp. 555-564.

Mulekar, M.S. & Sobel, M. (1999). Fixed-sample-size selection problem for Poisson populations. *Statistics and Decision* (Suppl. 4) 17: 69–85.

Nebenzahl, E. & Sobel, M. (1972). Play-the-winner sampling for a fixed sample size binomial selection problem. *Biometrika* 59: 1–8.

Olkin, I. & Sobel, M. (1965). Integral expressions for tail probabilities of the multinomial and negative multinomial distributions. *Biometrika* 52: 167–179.

Olkin, I. & Sobel, M. (1979). Admissible and minimax estimation for the multinomial distribution and for *k* independent binomial distributions. *The Annals of Statistics* 7: 284–290.

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press

Olkin, I. & Sobel, M. (1987). A model for interlaboratory differences. In A.K. Gupta (ed.), *Advances in multivariate statistical analysis: Pillai memorial volume*. Dordrecht: D. Reidel, pp. 303–314.

Olkin, I., Sobel, M., & Tong, Y.L. (1982). Bounds for a k-fold integral for location and scale parameter models with applications to statistical ranking and selection problems. In S.S. Gupta & J.O. Berger (eds.), *Statistical decision theory and related topics III*, Vol. 2. New York: Academic Press, pp. 193–212.

Rao, V.R., Savage, I.R., & Sobel, M. (1960). Contributions to the theory of rank order statistics: The two-sample censored case. *The Annals of Mathematical Statistics* 31: 415–426.

Rao, J.S. & Sobel, M. (1980). Incomplete Dirichlet integrals with applications to ordered uniform spacings. *Journal of Multivariate Analysis* 10: 603–610.

Rizvi, M.H. & Sobel, M. (1967). Nonparametric procedures for selecting a subset containing the population with the largest α-quantile. *The Annals of Mathematical Statistics* 38: 1788–1803.

Rizvi, M.H., Sobel, M., & Woodworth, G.G. (1968). Nonparametric ranking procedures for comparison with a control. *The Annals of Mathematical Statistics* 39: 2075–2093.

Robbins, H., Sobel, M., & Starr, N. (1968). A sequential procedure for selecting the largest of k means. *The Annals of Mathematical Statistics* 39: 88–92.

Roth, A.J. & Sobel, M. (1980). A probabilistically attained set of polynomials that generate Stirling numbers of the second kind. *Discrete Mathematics* 32: 77–84.

Savage, I.R., Sobel, M., & Woodworth, G. (1966). Fine structure of the ordering of probabilities of rank orders in the two sample case. *The Annals of Mathematical Statistics* 37: 98–112.

Shapiro, S.B. & Sobel, M. (1981). Two multinomial random sociometric voting models. *Journal of Educational Statistics* 6: 287–310.

Sobel, M. (1953). An essentially complete class of decision functions for certain standard sequential problems. *The Annals of Mathematical Statistics* 24: 319–337.

Sobel, M. (1954). On a generalization of an inequality of Hardy, Littlewood and Pólya. *Proceedings of the American Mathematical Society* 5: 596–602.

Sobel, M. (1956a). Sequential procedures for selecting the best exponential population. In J. Neyman (ed.), *Proceedings of the third Berkeley symposium on mathematical statistics and probability*, Vol. 5. Berkeley: University of California Press, pp. 99–110.

Sobel, M. (1956b). Statistical techniques for reducing the experiment time in reliability studies. *Bell System Technical Journal* 35: 179–202.

Sobel, M. (1967). Nonparametric procedures for selecting the t populations with the largest α-quantile. *The Annals of Mathematical Statistics* 38: 1804–1816.

Sobel, M. (1969). Selecting a subset containing at least one of the t best populations. In P.R. Krishnaiah (ed.), *Multivariate analysis II*. New York: Academic Press, pp. 515–540.

Sobel, M. (1972). Procédures optimales pour le classement des t meilleurs articles parmi n au moyen de comparaisons binaires. *Revue Française d'Automatique, Informatique et Recherche Opérationnelle* 6: 23–68 (transl. by A. Lahoute, G. Louit, & J.-P. Roubelat).

Sobel, M. (1977). The Dirichlet type 1 integral and its applications to multinomial-related problems. In S.S. Gupta & D.S. Moore (eds.), *Statistical decision theory and related topics*, Vol. 2. New York: Academic Press, pp. 357–374.

Sobel, M. (1985a). Likelihood procedures for selecting a subset. In E.J. Dudewicz (ed.), *The frontiers of modern statistical inference procedures*. Columbus, OH: American Sciences Press, pp. 251–270.

Sobel, M. (1985b). Nonparametric selection of the smallest-dispersion population. *American Journal of Mathematical and Management Sciences* 5: 313–329.

Sobel, M. (1987). Further paradoxes in admissibility: Part II. *Statistics and Decisions* 5: 297–306.

Sobel, M. (1989). Multinomial problems in geometric probability with Dirichlet analysis. In L.J. Gleser, M.D. Perlman, S.J. Press, & A.R. Sampson (eds.), *Contributions to probability and statistics. Essays in honor of Ingram Olkin*. New York: Springer-Verlag, pp. 107–145.

Sobel, M. (1992). New formulation for normal selection. In E.J. Dudewicz (ed.), *The frontiers of modern statistical inference procedures, II. Proceedings of the IPASRAS-II conference*. Columbus, OH: American Sciences Press, pp. 349–374.

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press

Sobel, M. (1996). Review of "Randomization tests (third edition)" by E.S.A. Edgington. *Short Book Reviews* 16: 26.
Sobel, M. (1997). I and J polynomials in a potpourri of probability problems. In N. Balakrishnan (ed.), *Proceedings of the conference on advances in combinatorial methods and applications to probability and statistics*. Boston: Birkhäuser, pp. 165–187.
Sobel, M., Bernard, S.R., & Uppuluri, V.R.R. (1985). On the distribution of the time required to remove white balls from an urn. *Bulletin of Mathematical Biology* 47: 791–797.
Sobel, M. & Chen, P. (1992). An optimal stopping rule for multinomial inverse sampling problems. In F.M. Hoppe (ed.), *Multiple comparisons, selection, and applications in biometry. A festschrift in Honor of Charles W. Dunnett*. New York: Marcel Dekker, pp. 531–548.
Sobel, M.J., Chen, P., & Sobel, M. (1991). A two-stage procedure for selecting the largest mean based on conditional probability analysis. *American Journal of Mathematical and Management Sciences* 11: 325–350.
Sobel, M. & Childs, A. (2002). Dirichlet-related probability problems and new tools for solving them. *Communications in Statistics, Part A: Theory and Methods* 31: 2027–2050.
Sobel, M. & Ebneshahrashoob, M. (1989). Optimal sequential stopping rules for certain binomial sampling problems (how to stop having babies). In J.P. Klein & J.C. Lee (ed.), *Recent developments in statistics and their applications*. Seoul: Freedom Academy, pp. 333-362.
Sobel, M. & Ebneshahrashoob, M. (1992). Quota sampling for multinomial via Dirichlet. *Journal of Statistical Planning and Inference* 33: 157–164.
Sobel, M. & Ebneshahrashoob, M. (1993). Dirichlet integrals for inverse sampling and quota sampling. In *Proceedings of the section on physical and engineering sciences*. Alexandria, VA: American Statistical Association, pp. 154–159.
Sobel, M., Ebneshahrashoob, M., & Lin, J.-Y. (1989). Application of Dirichlet integrals for curtailment problems in sampling inspection. *Naval Research Logistics* 36: 215–238.
Sobel, M. & Frankowski, K. (1994a). The 500th anniversary of the sharing problem (the oldest problem in the theory of probability). *The American Mathematical Monthly* 101: 833–847.
Sobel, M. & Frankowski, K. (1994b). A duality theorem for solving multiple-player multivariate hypergeometric problems. *Statistics & Probability Letters* 20: 155–162.
Sobel, M. & Frankowski, K. (1994c). Hypergeometric analogues of multinomial type-I Dirichlet problems. *Congressus Numerantium* 101: 65–82.
Sobel, M. and Frankowski, K. (1997). Venn diagrams, coupon collections, bingo games and Dirichlet distributions. In S. Panchapakesan (ed.), *Advances in statistical decision theory and applications*. Basel: Birkhäuser, pp. 343–367.
Sobel, M. & Frankowski, K. (2002). Extensions of gambler's ruin with even odds: Multinomial and other models. *Journal of Statistical Planning and Inference* 101: 291–310.
Sobel, M. & Huyett, M.J. (1957). Selecting the best one of several binomial populations. *Bell System Technical Journal* 36: 537–576.
Sobel, M. & Huyett, M.J. (1958). Nonparametric definition of representativeness of a sample—with tables. *Bell System Technical Journal* 37: 135–161.
Sobel, M. & Kelemen, E. (1984). On an application of statistical ranking and selection tables to the identification of the learning disabled. *Journal of Educational Statistics* 9: 215–226.
Sobel, M. & Starr, N. (1975). Selecting the coin with the greatest bias. *Sankhyā, Series A* 37: 197–210.
Sobel, M. & Tong, Y.L. (1971). Optimal allocation of observations for partitioning a set of normal populations in comparison with a control. *Biometrika* 58: 177–181.
Sobel, M. & Tong, Y.L. (1976). Estimation of a normal percentile by grouping. *Journal of the American Statistical Association* 71: 189–192.
Sobel, M. & Uppuluri, V.R.R. (1972). On Bonferroni-type inequalities of the same degree for the probability of unions and intersections. *The Annals of Mathematical Statistics* 43: 1549–1558.
Sobel, M. & Uppuluri, V.R.R. (1973). Sparse and crowded cells and Dirichlet distributions. *Bulletin of the International Statistical Institute* 45: 440–449.

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press

Sobel, M. & Uppuluri, V.R.R. (1974). Sparse and crowded cells and Dirichlet distributions. *The Annals of Statistics* 2: 977–987.
Sobel, M., Uppuluri, V.R.R., & Frankowski, K. (1977). Dirichlet distribution–type I. *Selected tables in mathematical statistics*, Vol. IV. Providence, RI: American Mathematical Society.
Sobel, M., Uppuluri, V.R.R., & Frankowski, K. (1978). Review of "Selected tables in mathematical statistics, Vol. IV" by N.L.R. Johnson. *Journal of the American Statistical Association* 73: 229.
Sobel, M., Uppuluri, V.R.R., & Frankowski, K. (1985). Dirichlet integrals of type II and their applications. *Selected tables in mathematical statistics*, Vol. 9. Providence, RI: American Mathematical Society.
Sobel, M., Uppuluri, V.R.R., & Frankowski, K.A. (1986). Review of "Selected tables in mathematical statistics (vol. 9)" by B.C.R. Arnold. *Journal of the American Statistical Association* 81: 860.
Sobel, M. & Wald, A. (1949). A sequential decision procedure for choosing one of three hypotheses concerning the unknown mean of a normal distribution. *The Annals of Mathematical Statistics* 20: 502–522.
Sobel, M. & Weiss, G.H. (1969). Comparison of sampling procedures for selecting the better of 2 binomial populations. *Bulletin of the International Statistical Institute* 43: 308–310.
Sobel, M. & Weiss, G.H. (1970a). Inverse sampling and other selection procedures for tournaments with 2 or 3 players. In M.L. Puri (ed.), *Nonparametric techniques in statistical inference*. Cambridge: Cambridge University Press, pp. 515–543.
Sobel, M. & Weiss, G.H. (1970b). Play-the-winner sampling for selecting the better of two binomial populations. *Biometrika* 57: 357–365.
Sobel, M. & Weiss, G.H. (1971a). A comparison of play-the-winner and vector-at-a-time sampling for sectecting the better of 2 binomial populations with restricted parameter values. *Trabajos de Estadística* 22: 195–206.
Sobel, M. & Weiss, G.H. (1971b). Play-the-winner rule and inverse sampling in selecting the better of two binomial populations. *Journal of the American Statistical Association* 66: 545–551.
Sobel, M. & Weiss, G.H. (1972a). Play-the-winner rule and inverse sampling for selecting the best of $k \geq 3$ binomial populations. *The Annals of Mathematical Statistics* 43: 1808–1826.
Sobel, M. & Weiss, G.H. (1972b). Recent results on using the play-the-winner sampling rule with binomial selection problems. In L.M. LeCam, J. Neyman, & E.L. Scott (eds.), *Proceedings of the sixth Berkeley symposium on mathematical statistics and probability,* Vol. I. Berkeley: University of California Press, pp. 717–736.
Sobel, M. & Wells, M. (1990). Dirichlet integrals and moments of gamma distribution order statistics. *Statistics & Probability Letters* 9: 431–437.
Sobel, M. & Wu, H.-F. (1984). Sequential likelihood procedures for selecting a subset and fixed sample size analogues. In T.J. Santner & A.C. Tamhane (eds.), *Design of experiments: Ranking and selection—Essays in honor of Robert E. Bechhofer*. New York: Marcel Dekker, pp. 199-211.
Sobel, M. & Yen, S.P. (1972). An asymptotic comparison of subset selection procedures. *Institute of Statistical Mathematics* 24: 465–468.
Sobel, M.J., Arnold, J., & Sobel, M. (1986). Statistics of natural populations, III. Sequential sampling plans for the estimation of gene frequencies. *Biometrics* 42: 45–65.

B. Group Testing

1959. Group testing to eliminate efficiently all defectives, in a binomial sample. (with P.A. Groll). *Bell System Technical Journal* 38: 1179–1252.
1960. Group testing to classify all defectives in a binomial sample. In R. Machol (ed.), *Information and decision processes*. New York: McGraw-Hill, pp. 127–161.
1966. Binomial group-testing with an unknown proportion of defectives. (with P.A. Groll). *Technometrics* 8: 631–656.
1967–68. Optimal group testing. In A. Rényi (ed.), *Proceedings of the colloquium on information theory*, Vol. II, Budapest: Jănos Bolyai Mathematical Society, pp. 441–488.
1968. Binomial and hypergeometric group-testing. *Studia Scientiarum Mathematicarium Hungarica* 3: 19–42.

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press

1970. Group-testing with at most c tests for finite c and c approaching ∞. (with S. Kumar). Technical report No. 146, Department of Statistics, University of Minnesota.
1971a. Finding a single defective in binomial group-testing. (with S. Kumar). *Journal of the American Statistical Association* 66: 824–828.
1971b. A characterization of binary codes that correspond to a class of group testing procedures. Technical Report No. 148, Department of Statistics, University of Minnesota.
1971c. Symmetric binomial group-testing with three outcomes. (with S. Kumar and S. Blumenthal). In *Statistical decision theory and related topics*, Vol. 1 (ed. by S.S. Gupta and J. Yackel). Academic Press, New York.
1972. Dorfman and R_1-type procedures for a generalized group-testing problem. (with J.K. Lee). *Mathematical Biosciences* 15: 317–340.
1973. Group-testing with at most D defectives: Binomial and hypergeometric models. Technical Report No. 204, School of Statistics, University of Minnesota.
1974. Finite and infinite models for generalized group-testing with unequal probabilities of success for each item. (with E. Nebenzahl). In J.N. Srivastava (ed.), *Discriminant analysis and applications*. New York: Academic Press, pp. 239–289.
1975a. Optimal group-testing given a sample which contains at most D defectives. (with G.L. Lin). Technical report, School of Statistics, University of Minnesota.
1975b. Group-testing with a new goal: Estimation. (with R.M. Elashoff). *Biometrika* 62: 181–194.
1975c. An asymptotically optimal Bayes solution for group testing. (with S. Kumar). In J.N. Srivastava (ed.), *A survey of statistical design and linear models*. Amsterdam: North-Holland, pp. 367–382.
1977a. Group-testing with preassigned confidence levels for a correct classification of all units. Technical Report No. 223, School of Statistics, University of Minnesota.
1977b. Enumerating a class of nested group-testing procedures. (with J.W. Moon). *Journal of Combinatorial Theory, Series B* 23: 184–188.
1987a. Group-testing, halving procedures and binary search. (with B.S. Pasternack and J. Thomas). *Communications in Statistics, Part A: Theory and Methods* 16: 2851–2871.
1987b. Entropy-based optimal group-testing procedures. (with P. Chen and L. Hsu). *Probability in the Engineering and Informational Sciences* 1: 497–509.

https://doi.org/10.1017/S0269964806060244 Published online by Cambridge University Press