Program Abstracts/Algorithms

Simulation of complex survival distributions

MICHAEL BORENSTEIN

Hillside Hospital, Glen Oaks, New York, and
Biostatistical Programming Associates,
Teaneck, New Jersey

This paper describes a method for generating sample survival distributions from a hypothetical population, as would be required for running Monte Carlo simulations. The method is based on the concept of a quincunx. Cases are entered into a life table and allowed to drop out or die during each interval with probabilities that mirror the hypothetical population. By repeating this process many times and tracking the results, the researcher is able to study the sampling distribution of effect size indices and test statistics, and can generate empirical estimates of power and precision for planned studies. Unlike other methods that are commonly used for this purpose, the model proposed here allows the researcher to define a population in which the hazard rates and/or attrition rates vary substantially from one time point to the next, as may be the case in clinical trials or studies of processing times. The method requires less than 100 lines of code and runs some 10,000 simulations per hour on a microcomputer.

Monte Carlo simulations are commonly used in survival analysis to study the sampling distribution of effect size indices and test statistics and to generate empirical estimates of power and precision for planned studies (see, e.g., Berry, Kitchin, & Mock, 1991; Breslow & Liang, 1982; Flanders, 1985; Hsieh, 1987; Sato, 1990). These simulations typically use mathematical formulas to generate survival and/or censoring times for individual observations, which are then grouped into time intervals. This approach functions well when the survival distributions are exponential, Weibull, or Gamma in form, and when the attrition (dropout) rate is constant (Beltangady & Frankowski, 1989; Ciampi, Chapman, Hogg, & Thiffault, 1989; Johnson & Mathe-son, 1992; Rubinstein, Gail, & Santner, 1981). However, as the shape of the distribution and/or the nature of the attrition become complex, it becomes increasingly difficult to run simulations by these methods. Such would be the case, for example, if we assumed that the survival

curves for two groups diverged slightly for a period of time, then diverged sharply for a period of time, and then began to converge—as might be hypothesized for certain types of clinical trials. Similarly, this approach could not be used to study the impact of a complex attrition process on the behavior of the various statistics. Ironically, it is precisely in these nonstandard cases that the behavior of the statistics is either likely to require investigation, or it is necessary to empirically determine statistical power or precision. Therefore, it would be helpful to have access to a simulation method that could be applied in these cases.

One such approach is described by Kane and Borenstein (1985), who used a simulation study to determine the power for finding a difference in relapse rates between groups of psychiatric patients. That simulation employed a brute force approach by actually building a data set of 20,000 cases that mirrored the population and then drawing random samples from this population. Although this method served its intended purpose, the process of drawing samples from the larger data set required prohibitive amounts of computer time. An analogous approach was employed by Halpern and Brown (1987).

The method is modeled on the concept of an electronic quincunx. The traditional quincunx is a sort of vertical pinball machine, in which Ping-Pong balls are dropped into the top and hit a pin. Then they bounce to either side, hit another pin, bounce to either side, and so on until they reach the bottom of the mechanism and fall into one of several bins. By studying the distribution of balls in the bins and how this distribution differs from one replication to the next, we can learn about the nature of the sampling distribution. Galton (cited in Stigler, 1989) used one of these machines as far back as 1873 to develop his theories of heredity, the central limit theorem, and correlation. Stigler (1986, 1989) offers an illustrated and illustrative narrative of Galton's use of this mechanism, which, as Stigler notes, was in effect an *ingenious analogue* computer. IBM built a quincunx at the 1963–1964 New York World's Fair (this particular model was three times as tall as the author, if memory serves), which demonstrated the creation of a binomial distribution six times an hour for a period of 2 years. Madigan (1990) discusses the use of an electronic version for teaching statistical concepts. The same basic idea can be applied to the problem at hand.

Life Table Formulas

Before proceeding to this application, we review the basic formulas for computing hazard rates and cumula-

This work was supported in part by Grants NIMH/SBIR 1-R43/44-MH-43083 and NIMH MH-41960. The author would also like to acknowledge the gracious help and support provided by the late N. John Castellan, Jr. Correspondence and reprint requests should be addressed to M. Borenstein, Director of Biostatistics, Hillside Hospital, P.O. Box 38, Glen Oaks, NY 11004.

tive survival in life table analysis (see, e.g., Kahn & Sempos, 1989; Lawless, 1982; Miller, 1981; Peto, 1984; Selvin, 1991; Simon, 1986; Simon & Lee, 1982). In the equations that follow, lowercase letters are used to denote values computed within a single time interval, k , and uppercase letters are used to denote related values derived from all periods, from the beginning of the study through interval k . Roman letters (atrisk, ATRISK) denote *numbers* of cases, and Greek letters (λ , Λ) denote *rates*. Subscripts are used to denote the group (where required) and the time point.

Survival analyses typically are carried out by the Kaplan-Meier method (Kaplan & Meier, 1958) or by the Cutler-Ederer method (Cutler & Ederer, 1958). The two methods utilize essentially the same formulas, but differ in that the former computes a new segment in the survival curve whenever there is a death (and is therefore suitable for relatively small sample sizes), and the latter calculates the segments of the survival curve in fixed increments of time (e.g., monthly—and is therefore suitable for relatively large samples). Computationally, the two methods differ in their treatment of censored cases (i.e., patients whose time of death is unknown because they dropped out of the study or are still alive at the time of the analysis). When a censored case and a death occur at the same time, the Kaplan-Meier model assumes that the death preceded the loss, but the Cutler-Ederer model assumes that the losses that occur during an interval are evenly spaced across that interval. For any given interval k , the number at risk, atrisk_k , is computed as

$$\text{atrisk}_k = \begin{cases} \text{enter}_k - \text{censor}_k / 2 & \text{for Cutler-Ederer} \\ \text{enter}_k & \text{for Kaplan-Meier,} \end{cases}$$

where censor_k is the number of cases lost to follow-up (and last seen alive) during this interval. The number remaining alive at the conclusion of the intervals is defined for computational purposes as

$$\text{alive}_k = \text{atrisk}_k - \text{died}_k,$$

and is distinct from the number of cases entering the subsequent interval, which is defined as

$$\text{enter}_{k+1} = \text{enter}_k - \text{censor}_k - \text{died}_k.$$

The hazard rate for the interval, λ_k , is computed as

$$\lambda_k = \text{died}_k / \text{atrisk}_k.$$

If the survival rate within any given interval k , ϕ_k , is defined as

$$\phi_k = 1 - \lambda_k,$$

the cumulative survival through interval k , Φ_k , would be computed as

$$\Phi_k = \prod_1^k \phi_k.$$

Model for Running Simulations

As noted earlier, the proposed method for running simulations is based on the concept of an electronic quin-

cunx. The method is shown schematically in Figure 1, which is composed of three sections—A, B, and C.

Section A in Figure 1 is used to hold the population parameters. The k rows in this matrix correspond to the time intervals, and the columns hold the attrition rate (ω) and hazard rate (λ) for each interval. Each rate may be constant across time, may vary systematically across time, or be set independently for each interval. In this example, ω for the initial interval (ω_1) is displayed as 0.10, ω_2 is 0.08, and so on. Similarly, λ for the initial interval (λ_1) is 0.04, λ_2 is 0.02, and so on, through the 10th interval.

Section B is the matrix holding the *observed* data, as it were, for a single sample. The structure of this matrix is similar to that of the population matrix, but in this case the cells hold the observed numbers of cases that have either dropped out or died during each interval. Cases enter section B at the upper left corner and then either fall into or bypass each cell on the first row, the second row, and so on. Specifically, a case is entered into the sample matrix at row 1 and is assigned a tag, $X1$, from a random uniform distribution in the range of 0–1 and (independently) a tag $X2$ from the same distribution. The $X1$ tag is used to identify dropouts: If the value of $X1$ is less than or equal to ω_1 , the case is considered a dropout, and the corresponding cell in the sample matrix is incremented by one unit. Similarly, if the value of $X2$ is less than or equal to λ_1 , the case is considered to have died in this time span, and the corresponding cell in the sample matrix is incremented by one unit. If neither of these conditions is satisfied, the case is entered into the next row, and the process is repeated until the case drops out, dies, or traverses the entire matrix successfully. The possibility that a case could be tagged as being both a dropout and a death within the same interval is discussed below. For a sample of N cases, this process would be repeated N times, at which point we would have the actual N of cases who have entered, dropped out, or died in each interval. If the simulation is being used to compare the distributions in two groups, the process would be repeated for the second group.

Section C holds computed statistics that correspond to the cell counts for the current sample. Concretely, in this example, the initial interval in the population (section A) has an attrition rate of 0.10 and a hazard rate of 0.04. Section B shows that, of 100 cases entered into this interval, 12 are designated as dropouts and 1 is designated as a death. In section C, the N at risk [$100 - (12/2)$] is shown as 94.0, the hazard rate ($1/94.0$) is shown as 0.011, and the cumulative survival through this interval ($1.0 - 0.011$) is shown as 0.989.

The process displayed in Figure 1 represents a single sample. The researcher would typically draw a large number (1,000–100,000) of such samples and then study the distributions of test statistics or effect size indices across them.

The issue of ties. As noted, the tags $X1$ and $X2$, used to classify cases as either dropping out or dying, are assigned independently of each other. In a small propor-

Time Interval k	Section A		Section B			Section C		
	Population Rates		Observed N of Cases in Sample			Rates Corresponding to Observed N in Section B		
	Attrition Rate ω	Hazard Rate λ	N-Enter	N-Drop	N-Died	Cases at Risk	Hazard Rate	Cumulative Survival
1	0.100	0.040	100	12	1	94.0	0.011	0.989
2	0.080	0.020	87	10	3	82.0	0.037	0.953
3	0.040	0.010	74	5	0	71.5	0.000	0.953
4	0.020	0.010	69	2	0	68.0	0.000	0.953
5	0.020	0.010	67	0	1	67.0	0.015	0.939
6	0.020	0.020	66	1	1	65.5	0.015	0.925
7	0.040	0.040	64	2	2	63.0	0.032	0.895
8	0.040	0.040	60	1	2	59.5	0.034	0.865
9	0.040	0.040	57	2	2	56.0	0.036	0.834
10	0.040	0.040	53	0	3	53.0	0.057	0.787

Figure 1. Schematic of the simulation process. Section A holds the population parameters for attrition (dropout) and hazard rates for each time interval. Section B holds the cell counts for each sample: A case enters at the upper left and either lands in the Drop cell, lands in the Died cell, or continues on to the second row, where the process is repeated. Section C holds the survival rates corresponding to the cell counts in section B for the current sample. These data would be accumulated across samples (not shown).

tion of the cases (specifically, $\lambda * \omega$), a case will be tagged as both a death and a dropout during the same interval. If, in this situation, we were to ignore the first set of tags and simply draw a second pair, a bias would be introduced in that a small proportion of the "dead" cases (concretely, ω) would be systematically excluded from the sample. For this reason, a case tagged as both a dropout and a death must *not* be allowed to continue to the next interval. How we handle this type of case depends on the model that is being simulated.

When the program is being used to simulate the Cutler-Ederer model, which assumes that patients are lost across the entire span of the interval, the program includes a loop to ensure that in the event of a tie, the case in question will be classified as either a death [with probability $\lambda/(\lambda+\omega)$] or a dropout [with probability $\omega/(\lambda+\omega)$], and it will not be allowed to continue to the next time interval.

In the Kaplan-Meier method, we tend to think of time points rather than time intervals, so if a case drops out at a given time point, it is considered to have been at risk up until that time point. Accordingly, in the event of a tie, it is assumed that the case would have been observed to die, and it is assigned to that category in the simulation. To address the fact that this approach will introduce a bias by systematically reducing the number of dropouts, the simulation works with a modified value for ω when deciding whether or not a case should be tagged as a dropout. Specifically, given that a case has not been tagged as a death, it will be classified as a dropout if the $X1$ tag is less than or equal to ω' , defined as $\omega + \{[\lambda/(1-\lambda)] * \omega\}$.

Appendix B includes an annotated copy of the computer code, where the execution of these procedures is spelled out in detail. If necessary, this code can easily be modified to reflect idiosyncratic elements of any planned study.

Accuracy of the Method

To test the accuracy of this method, we generated 50,000 samples from a population and then compared the samples with the parent population. In this study, we utilized the random-number generator included as a function by Microsoft BASIC PDS 7.1, but the same results would be expected with any appropriate random-number generator. We report on four sets of simulations: Data for the Kaplan-Meier method are displayed in Table 1, and data for the Cutler-Ederer method are displayed in Table 2. In both of the tables, data displayed at the top are for samples of 60 cases per group, and those at the bottom are for samples of 200 cases per group.

All the tests are based on the same population, which was designed to mirror a particular type of clinical trial. This model would reflect a trial in which patients who had undergone surgery for removal of a tumor were treated with either a standard treatment or a more aggressive regimen in the hope of preventing a recurrence of the growth. Under the standard treatment, the hazard rate is initially high. It falls over a 3-month period, then increases over a 3-month period, and reaches a plateau 7 months following the surgery. The hazard rate will follow the same pattern under the aggressive treatment, but the hazard rate within each period is expected to be reduced by 50%. Similarly, the attrition rate is initially high because many patients had second thoughts about their partici-

pation. The rate drops over a period of time, because the patients who remained after the initial shake-out are those who were committed to the study, and then it rises again as the tedium of returning for examinations begins to take a toll. The actual hazard and attrition rates are shown in the tables.

Mean survival. If the simulation is working properly, we would expect that the sample survival curves, averaged over 50,000 replications, would closely mirror the population from which they had been drawn. This was found to be the case. At the top left corner of Table 1, for example, the cumulative proportion surviving the 1st interval is .9603 for the mean of the samples and is .9600 for the population, for a difference of only .0003 percentage points. At the 2nd interval, the numbers are .9413 and .9408, for a difference of .0005 percentage points, and so on through the 10th interval. The average absolute deviation across the 10 intervals, shown at the bottom of the column, is .0008 percentage points. Similar data are presented for the second treatment group; the mean absolute deviation in this case is .0005. The upper right corner of this table is used to display data about the difference in cumulative survival between the two groups. At the conclusion of the 1st interval, the population mean difference in cumulative survival (.9600-.9200) is .0400. Across the 50,000 samples, the mean difference for this interval was .0408—that is, within .0008 of the expected value. At the bottom of this column, the mean absolute deviation across the 10 intervals is shown as .0011. Similarly small discrepancies are shown in the tables for each of the survival curves (both within and between groups), and they were found for a wide variety of other simulations not shown here.

Cell counts. The next several sets of analyses address the dispersion in the samples. By constructing a flow chart such as that in Figure 1, it is possible to compute the number of cases expected to enter each time point and the number of cases expected to be classified as dropouts. A comparison of these numbers with the numbers observed in the simulations provides additional evidence that the simulation process is working properly. For the Kaplan-Meier method, with an initial N of 200 and the parameters shown in this figure, we would expect to see 103.44 cases (on average) entering the 10th interval and 67.35 cases (on average) dropping out over the course of the 10 intervals. The observed values over the 50,000 samples were 103.47 and 67.36. For the Cutler-Ederer method, we would expect to see 104.81 cases entering the 10th interval and a cumulative total of 67.51 dropping out. The observed numbers were 104.81 and 67.05. Thus, the simulation is functioning as expected. Similar results were observed for all 10 intervals.

Standard deviations. If the process is working properly, the standard deviation for cumulative survival should vary as an inverse function of the sample size. Concretely, simulations were run with $N = 60$ and $N = 200$, and we would expect that $SD_{60}/SD_{200} \approx \sqrt{200/60}$. This proved to be the case. For example, for the initial interval for the first treatment group in Table 1, we find

that $.0252/.0138 \approx \sqrt{200/60} \approx 1.83$. The relationship between the dispersion and sample size is also shown in Figure 2, where hi-lo lines are used to display the standard deviation of values about the mean survival for an N of 60 per group and again for an N of 200.

Confidence intervals. The dispersion is also addressed by the following set of analyses. As each of the 50,000 samples was drawn, the program applied a formula (see Appendix A) to compute the 95% confidence interval (CI) for cumulative survival. If the formula is accurate and the simulation is working properly, we would expect that in precisely 95% of all simulations these CI would include the population parameter. On average (across both methods and treatment groups), the computed CI *within* groups was found to include the population parameter in 95.37% of the samples. The CI for the difference *between* groups was found to include the population parameter in 94.87% of the samples. The findings were consistent across all the simulations, with one exception. For the earliest intervals in samples using the smaller sample size ($N = 60$) and lower hazard rate, the within-group CI included the parameter in as many as 97.6% of the simulations. Because the anomaly is limited to these particular cells, it is assumed to reflect the fact that the formula is conservative when the cumulative survival is 100%, which is a relatively frequent occurrence under these circumstances and is not seen as a problem with the simulation process.

Sample Application

Appendix B includes a listing of code for a sample application. The purpose of this application is to evaluate a formula for computing power for the difference in cumulative survival. This formula, which is given in Appendix A, follows the logic outlined by Lachin (1981) in a related context. It may be applied even when the hazard and/or attrition rates vary from one interval to the next and is intended for studies that will be tested by the Mantel-Haenszel test (Mantel, 1966; Simon, 1986). The simulation program will be used to compute the empirical power, which will then serve as a gold standard against which the formula will be judged.

The listing begins with the *Main* program, which calls a series of subroutines. The program calls *Getparms* to retrieve the parameters for the current simulation and then enters a loop that repeatedly calls a series of subprograms. *Runsimulations* includes the algorithms for drawing samples from the population. *Computestats* computes the cumulative survival for each group for the current sample as well as the statistic of interest—in this case, the p value. *Savedata* uses a method described by Cook, Craven, and Clarke (1990) to accumulate the mean and standard deviation of cumulative survival across samples. This subprogram also tracks the proportion of cases in which the difference yields a statistically significant p value ($\leq .05$). When the final sample has been drawn, this subprogram writes the summary data to disk.

The results of this simulation are included in Tables 1 and 2. The columns at the extreme right of the tables show

Table 1
Population Parameters and Sample Statistics (50,000 Replications Using the Kaplan-Meier Method) for Treatment 1, Treatment 2, and the Between-Groups Difference

Kaplan-Meier Method																
Time	Treatment — 1						Treatment — 2						Difference			
	Population			Samples Cumulative Survival			Population			Samples Cumulative Survival			Pop'n	Samples Difference in. Surv		Power
	Hazard Rate	Drop Rate	Cumul Surv	Mean	SD	CI *	Hazard Rate	Drop Rate	Cumul Surv	Mean	SD	CI *		Mean	SD	CI *
1	.0400	.1000	.9600	.9603	.0252	.9667	.0800	.1000	.9200	.9195	.0349	.9459	.0400	.0408	.0432	.9535
2	.0200	.0800	.9408	.9413	.0309	.9621	.0400	.0800	.8832	.8825	.0424	.9494	.0576	.0588	.0523	.9530
3	.0100	.0400	.9314	.9319	.0337	.9571	.0200	.0400	.8655	.8647	.0457	.9510	.0659	.0671	.0569	.9507
4	.0100	.0200	.9221	.9228	.0360	.9600	.0200	.0200	.8482	.8476	.0486	.9522	.0739	.0752	.0605	.9511
5	.0100	.0200	.9129	.9137	.0383	.9586	.0200	.0200	.8313	.8304	.0517	.9506	.0816	.0833	.0642	.9462
6	.0200	.0200	.8946	.8957	.0424	.9537	.0400	.0200	.7980	.7975	.0563	.9494	.0966	.0981	.0703	.9457
7	.0400	.0400	.8588	.8599	.0495	.9528	.0800	.0400	.7342	.7340	.0632	.9503	.1246	.1258	.0801	.9488
8	.0400	.0400	.8245	.8256	.0547	.9516	.0800	.0400	.6754	.6756	.0683	.9491	.1490	.1500	.0877	.9449
9	.0400	.0400	.7915	.7924	.0592	.9523	.0800	.0400	.6214	.6217	.0717	.9519	.1701	.1707	.0939	.9448
10	.0400	.0400	.7598	.7604	.0642	.9498	.0800	.0400	.5717	.5716	.0746	.9523	.1881	.1889	.0991	.9417
Mean Absolute Deviation or Mean CI				.0008		.9565					.0005	.9502			.0011	.9480
1	.0400	.1000	.9600	.9602	.0138	.9616	.0800	.1000	.9200	.9196	.0185	.9618	.0400	.0405	.0229	.9566
2	.0200	.0800	.9408	.9408	.0171	.9488	.0400	.0800	.8832	.8830	.0229	.9532	.0576	.0579	.0285	.9480
3	.0100	.0400	.9314	.9315	.0187	.9485	.0200	.0400	.8655	.8653	.0243	.9557	.0659	.0662	.0304	.9506
4	.0100	.0200	.9221	.9222	.0200	.9514	.0200	.0200	.8482	.8480	.0260	.9539	.0739	.0741	.0329	.9484
5	.0100	.0200	.9129	.9131	.0212	.9520	.0200	.0200	.8313	.8309	.0277	.9544	.0816	.0822	.0347	.9448
6	.0200	.0200	.8946	.8947	.0236	.9459	.0400	.0200	.7980	.7983	.0299	.9538	.0966	.0964	.0386	.9467
7	.0400	.0400	.8588	.8594	.0271	.9447	.0800	.0400	.7342	.7342	.0340	.9556	.1246	.1252	.0442	.9537
8	.0400	.0400	.8245	.8247	.0303	.9480	.0800	.0400	.6754	.6755	.0369	.9509	.1490	.1492	.0481	.9481
9	.0400	.0400	.7915	.7917	.0325	.9538	.0800	.0400	.6214	.6210	.0388	.9548	.1701	.1707	.0511	.9525
10	.0400	.0400	.7598	.7598	.0347	.9502	.0800	.0400	.5717	.5711	.0405	.9473	.1881	.1888	.0532	.9518
Mean Absolute Deviation or Mean CI				.0002		.9505					.0003	.9542			.0004	.9501

Note.—Data at the top are based on 60 cases per group; data at the bottom are based on 200 cases per group. The data in this table correspond to the graphs in Figure 2. *Proportion of samples in which the 95% confidence interval includes the population parameter.

Table 2
Population Parameters and Sample Statistics (50,000 Replications Using the Cutler-Ederer Method)

Cutler-Ederer Method													
Time	Treatment — 1				Treatment — 2				Difference				
	Population		Samples Cumulative Survival		Population		Samples Cumulative Survival		Pop'n	Samples Difference in. Surv		Power	
	Hazard Rate	Drop Rate	Cumul Surv	Mean	SD	CI *	Hazard Rate	Drop Rate	Cumul Surv	Mean	SD	CI *	By Formula By Simulation
1	.0400	.1000	.9600	.9614	.0254	.9766	.0800	.1000	.9200	.9206	.0357	.9552	.146
2	.0200	.0800	.9408	.9425	.0311	.9630	.0400	.0800	.8832	.8847	.0433	.9623	.187
3	.0100	.0400	.9314	.9332	.0335	.9692	.0200	.0400	.8655	.8670	.0464	.9521	.205
4	.0100	.0200	.9221	.9239	.0360	.9622	.0200	.0200	.8482	.8494	.0495	.9491	.223
5	.0100	.0200	.9129	.9150	.0381	.9634	.0200	.0200	.8313	.8322	.0524	.9500	.240
6	.0200	.0200	.8946	.8968	.0422	.9565	.0400	.0200	.7980	.7991	.0572	.9471	.273
7	.0400	.0400	.8588	.8606	.0492	.9560	.0800	.0400	.7342	.7348	.0640	.9439	.336
8	.0400	.0400	.8245	.8262	.0548	.9567	.0800	.0400	.6754	.6753	.0690	.9518	.391
9	.0400	.0400	.7915	.7928	.0593	.9578	.0800	.0400	.6214	.6212	.0726	.9457	.439
10	.0400	.0400	.7598	.7609	.0636	.9558	.0800	.0400	.5717	.5711	.0751	.9476	.481
Mean Absolute Deviation or Mean CI				.0017		.9617					.0008	.9511	.9479
1	.0400	.1000	.9600	.9610	.0140	.9462	.0800	.1000	.9200	.9212	.0197	.9521	.374
2	.0200	.0800	.9408	.9425	.0170	.9570	.0400	.0800	.8832	.8847	.0234	.9537	.496
3	.0100	.0400	.9314	.9331	.0184	.9523	.0200	.0400	.8655	.8670	.0251	.9520	.546
4	.0100	.0200	.9221	.9237	.0200	.9532	.0200	.0200	.8482	.8497	.0267	.9527	.590
5	.0100	.0200	.9129	.9143	.0213	.9522	.0200	.0200	.8313	.8323	.0283	.9505	.630
6	.0200	.0200	.8946	.8962	.0235	.9495	.0400	.0200	.7980	.7992	.0310	.9516	.699
7	.0400	.0400	.8588	.8602	.0268	.9602	.0800	.0400	.7342	.7352	.0348	.9495	.804
8	.0400	.0400	.8245	.8261	.0300	.9563	.0800	.0400	.6754	.6761	.0371	.9489	.870
9	.0400	.0400	.7915	.7936	.0324	.9548	.0800	.0400	.6214	.6220	.0389	.9543	.912
10	.0400	.0400	.7598	.7619	.0348	.9545	.0800	.0400	.5717	.5715	.0401	.9546	.939
Mean Absolute Deviation or Mean CI				.0016		.9536					.0010	.9520	.9489

Note—Data at the top are based on 60 cases per group; data at the bottom are based on 200 cases per group *Proportion of samples in which the 95% confidence interval includes the population parameter.

Figure 2. Mean cumulative survival and standard deviation across 50,000 samples for Group 1 (top), Group 2 (center), and the between-groups difference (bottom). The mean survival in the samples falls within 0.001 percentage points of the population, which indicates that the simulation is working properly. The mean survival at right ($N = 200$) is identical to the mean survival at left ($N = 60$), but the standard deviation (denoted by the vertical bars) is smaller. These graphs are based on the Kaplan-Meier method (Table 1), but graphs based on the Cutler-Ederer method (Table 2) would be virtually identical to the ones shown here.

the proportion of samples in which the p value was less than or equal to .05, and the adjacent columns show the p value given by the formula. The values in the two columns match closely (the mean absolute deviation is .01 and no discrepancy exceeds .021), which indicates that the formula is working properly for the distributions tested here. We would systematically modify the values in

Getparm and run many sets of simulations to ensure that the formula works for other types of distributions as well.

Technical Details and Availability

The speed of the simulations will vary as a function of numerous factors. The most important factors are machine speed, sample size, number of analyses to which

the sample data are subjected, and the language in which the program is run. As a rule of thumb, microcomputers based on a 486-DX 33-MHz chip can be expected to run between 5,000 and 20,000 simulations per hour.

The code for this model is presented in Appendix B. This code is in BASIC, but is fully annotated and can be easily adapted for any other language and/or modified for any particular simulation. The author will provide copies of the program on disk free of charge.

Discussion

The Monte Carlo method is commonly used in the area of survival analysis to estimate the empirical power or precision of a specific study or to study the behavior of various statistics that may be considered for general use. Studies based on this method typically use mathematical formulas to generate the sample distributions and are therefore limited to distributions that can be readily expressed as mathematical formulas. Ironically, the distributions that do not conform to these requirements are the ones in which simulation studies are most likely to yield insights into the behavior of the statistics in question.

In this paper, I have proposed a simulation method that allows the researcher complete flexibility in describing the population. The hazard and attrition rates may be defined as constant across time points; the rate at one point in time may be linked to the rate at other time points; or the rate for any time point may be defined independently of the others. This method, in addition to being intuitively appealing, offers a practical approach. It requires relatively little computer code and yields an execution speed that is fast enough for simulation studies to be run on the current generation of microcomputers. The model has been discussed here in the context of clinical trials, but could also be applied to other applications, including processing times.

REFERENCES

- BELTANGADY, M. S., & FRANKOWSKI, R. F. (1989). Effect of unequal censoring on the size and power of the Logrank and Wilcoxon types of tests for survival data. *Statistics in Medicine*, **8**, 937-945.
- BERRY, G., KITCHIN, R. M., & MOCK, P. A. (1991). A comparison of two simple hazard ratio estimators based on the Logrank test. *Statistics in Medicine*, **10**, 749-755.
- BRESLOW, N. E., & LIANG, K. Y. (1982). The variance of the Mantel-Haenszel estimator. *Biometrics*, **38**, 943-952.
- CIAMPI, A., CHAPMAN, J., HOGG, S. A., & THIFFAULT, J. (1989). GENCOV: A FORTRAN program that generates censored survival data with covariates. *Computer Methods & Programs in Biomedicine*, **29**, 51-57.
- COOK, D., CRAVEN, A. H., & CLARKE, G. M. (1990). *Basic statistical computing* (2nd ed.). London: Edward Arnold.
- CUTLER, S. J., & EDERER, F. (1958). Maximum utilization of the life table method in analyzing survival. *Journal of Chronic Diseases*, **8**, 699-712.
- DOREY, F. J., & KORN, E. L. (1987). Effective sample sizes for confidence intervals for survival probabilities. *Statistics in Medicine*, **6**, 679-687.
- FLANDERS, W. D. (1985). A new variance estimator for the Mantel-Haenszel odds ratio. *Biometrics*, **41**, 637-642.
- HALPERN, J., & BROWN, B. W. (1987). Designing clinical trials with arbitrary specification of survival functions and for the log rank or generalized Wilcoxon test. *Controlled Clinical Trials*, **8**, 177-189.
- HANLEY, J., & LIPPMAN-HAND, A. (1983). If nothing goes wrong, is everything all right? *Journal of the American Medical Association*, **249**, 1743-1745.
- HSIEH, F. Y. (1987). A simple method of sample size calculations for unequal-sample-size designs that use the logrank or *t*-test. *Statistics in Medicine*, **6**, 577-581.
- JOHNSON, B., & MATHESON, D. (1992, August). Techniques. *SPSS Keywords: Tips & News for Statistical Software Users*, pp. 11-12.
- KAHN, A. H., & SEMPOS, C. T. (1989). *Statistical methods in epidemiology*. New York: Oxford University Press.
- KANE, J., & BORENSTEIN, M. (1985). Compliance in the long-term treatment of schizophrenia. *Psychopharmacology Bulletin*, **21**, 23-27.
- KAPLAN, E. L., & MEIER, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, **53**, 457-481.
- LACHIN, J. M. (1981). Introduction to sample size determination and power analysis for clinical trials. *Controlled Clinical Trials*, **2**, 93-113.
- LAWLESS, J. F. (1982). *Statistical models and methods for lifetime data*. New York: Wiley.
- MADIGAN, S. (1990). Simulating the binomial distribution with a software quincunx. *Behavior Research Methods, Instruments, & Computers*, **22**, 475-476.
- MANTEL, N. (1966). Evaluation of survival data and two new rank order statistics arising from its consideration. *Cancer Chemotherapy Reports*, **50**, 163-170.
- MILLER, R. G. (1981). *Survival analysis*. New York: Wiley.
- PETO, J. (1984). The calculation and interpretation of survival curves. In M. E. Buyse & R. J. Sylvester (Eds.), *Cancer clinical trials* (pp. 361-380). New York: Oxford University Press.
- ROTHMAN, K. J. (1978). Estimation of confidence limits for the cumulative probability of survival in life table analysis. *Journal of Chronic Diseases*, **31**, 557-560.
- RUBINSTEIN, L. V., GAIL, M. H., & SANTNER, T. J. (1981). Planning the duration of a comparative clinical trial with loss to follow-up and a period of continued observation. *Journal of Chronic Diseases*, **34**, 469-479.
- SATO, T. (1990). Confidence limits for the common odds ratio based on the asymptotic distribution of the Mantel-Haenszel estimator. *Biometrics*, **46**, 71-80.
- SELVIN, S. (1991). *Statistical analysis of epidemiologic data*. New York: Oxford University Press.
- SIMON, R. (1986). Confidence intervals for reporting results of clinical trials. *Annals of Internal Medicine*, **105**, 429-435.
- SIMON, R., & LEE, Y. J. (1982). Nonparametric confidence limits for survival probabilities and median survival time. *Cancer Treatment Reports*, **66**, 37-42.
- STIGLER, S. M. (1986). *The history of statistics: The measurement of uncertainty before 1890*. Cambridge, MA: Harvard University Press, Belknap Press.
- STIGLER, S. M. (1989). Francis Galton's account of the invention of correlation. *Statistical Science*, **4**, 73-86.

APPENDIX A

Computational Formulas

Confidence Intervals

Approximate lower and upper limits of the confidence interval for cumulative survival Φ through time interval k , within either group, are given by

$$\Phi_{kLL} = \begin{cases} 0 & \text{if } \Phi = 0 \\ ((1 - \text{CILEVEL})/\text{TAILS})^{1/\text{ATRISK}'_k} & \text{if } \Phi = 1 \\ \frac{\Phi_k + A/2 - Z_{(1-\text{CI})}\sqrt{\Phi_k(1-\Phi_k)/\text{ATRISK}'_k} + A/4 * \text{ATRISK}'_k}{(1+A)} & \text{otherwise} \end{cases}$$

$$\Phi_{kUL} = \begin{cases} 1 & \text{if } \Phi = 1 \\ 1 - (((1 - \text{CILEVEL})/\text{TAILS})^{1/\text{ATRISK}'_k}) & \text{if } \Phi = 0 \\ \frac{\Phi_k + A/2 + Z_{(1-\text{CI})}\sqrt{\Phi_k(1-\Phi_k)/\text{ATRISK}'_k} + A/4 * \text{ATRISK}'_k}{(1+A)} & \text{otherwise} \end{cases}$$

where

$$\text{ATRISK}' = \frac{1 - \Phi_k}{\Phi_k \sum_{i=1}^k \frac{\text{died}_i}{\text{atrisk}_i * \text{alive}_i}}$$

CILEVEL = 0.95 for 95% confidence one or two-tailed;

TAILS = 1 or 2;

$Z = 1.96$ for 95% confidence, two-tailed, or

$Z = 1.64$ for 95% confidence, one-tailed;

and

$$A = Z^2 / \text{atrisk}_k.$$

The first line in either equation is true by definition, the second is adapted from Hanley and Lippman-Hand (1983; note that half the error is apportioned to both tails, even when the interval is truncated at 0.0 or at 1.0), and the third line is given by Simon (1986). The computation of the effective N (ATRISK') is given by Simon (1986; see also Rothman, 1978). Since the cumulative survival curve is monotonic by definition, we constrain the upper limit of the CI to be no higher than the corresponding value in the preceding time period. (See Dorey & Korn, 1987, for a more complete discussion of this issue and alternative approaches.)

Similarly, 95% confidence intervals for the difference in cumulative survival were computed as

$$\begin{pmatrix} \text{UL} \\ \text{LL} \end{pmatrix} \Phi_{\text{Diff}_k} = \Phi_{\text{Diff}_k} \pm Z * SE_{\Phi_{\text{Diff}_k}},$$

where

$$SE_{\Phi_{\text{Diff}_k}} = \sqrt{\frac{\Phi_{1k} * (1 - \Phi_{1k})}{\text{ATRISK}'_{1k}} + \frac{\Phi_{2k} * (1 - \Phi_{2k})}{\text{ATRISK}'_{2k}}}$$

and ATRISK' is defined as above (Simon, 1986).

Statistical Power

The simulation was used to compute empirical power, which served as the gold standard against which we evaluated the following computational formula for power. To compute power by formula, we first find the value of the Mantel-Haenszel (MH) statistic corresponding to the alternate hypothesis

$$MH_k = \frac{(\text{DIED}_{1k} - \text{EXPECTED}_{1k})^2}{\sigma_{MH_k}^2},$$

APPENDIX A (Continued)

where the number of deaths in Group 1 is summed across k intervals as

$$\text{DIED}_{1k} = \sum_1^k \text{died}_{1k},$$

the expected number of deaths in Group 1 under the null hypothesis is given by

$$\text{EXPECTED}_{1k} = \sum_1^k \frac{\text{atrisk}_{1k} * \text{died}_{.k}}{\text{atrisk}_{.k}},$$

where the subscript $.k$ indicates summation over both groups for interval k and the variance is estimated by

$$\sigma_{\text{MH}_k}^2 = \sum_1^k \frac{\text{atrisk}_{1k} * \text{atrisk}_{2k} * \text{died}_{.k} * \text{alive}_{.k}}{\text{atrisk}_{.k}^2 * (\text{atrisk}_{.k} - 1)}.$$

After computing the MH statistic in this manner, we can compute the Z value corresponding to power, $Z_{\text{POWER}} = Z_{\text{OBS}} - Z_{\alpha/\text{TAILS}}$, where $Z_{\text{OBS}} = \sqrt{\text{MH}}$, and $Z_{\alpha/\text{TAILS}} = 1.960$ for two-tailed alpha of .05, or 1.645 for one-tailed alpha of .05. Power is equal to the area under the normal curve falling to the left of Z_{POWER} . For a two-tailed test, we compute power by using Z_{OBS} , then again compute power by using $-1 \times Z_{\text{OBS}}$, and then sum the two values.

(Continued on next page)

APPENDIX B

Code for Simulations

MAIN PROGRAM

```

RANDOMIZE TIMER

CALL GETCONSTANTS          ' GETS CONSTANTS
CALL GETPARMS              ' GETS POPULATION PARAMETERS
CALL COMPUTEPOP            ' COMPUTE SURVIVAL FOR POPULATION

FOR SAMPLE = 1 TO NUMSAMPLES
    CALL RUNSIMULATIONS    ' DRAWS THE SAMPLES
    CALL COMPUTESTATS      ' COMPUTES SURVIVAL FOR EACH SAMPLE
    CALL SAVEDATA          ' ACCUMULATES DATA ACROSS SAMPLES
NEXT SAMPLE

```

GETCONSTANTS: HOLDS CONSTANTS USED BY THE PROGRAM

```

'THE FOLLOWING CONSTANTS ARE USED IN CALCULATION OF P-VALUES

A1 = .0705230784#:      A2 = .0422820123#:      A3 = .0092705272#
A4 = .0001520143#:      A5 = .0002765672#:      A6 = .0000430638#

'THE FOLLOWING CONSTANTS ARE USED TO IDENTIFY MATRIX ELEMENTS IN POP AND OBS

NENTER = 1              ' N ENTERING INTERVAL
NDROP = 2               ' N DROPPING OUT DURING INTERVAL
DROPRATE = 3            ' DROPOUT RATE FOR INTERVAL
HAZRATE = 4             ' HAZARD RATE FOR INTERVAL IN SAMPLE
NATRISK = 5             ' N AT RISK IN INTERVAL
NDIED = 6               ' N OF TERMINAL EVENTS IN INTERVAL
NALIVE = 7              ' N ALIVE IN INTERVAL
CUMNATRISK = 8          ' CUMULATIVE N AT RISK
CUMNDIED = 9            ' CUMULATIVE N DIED
CUMNALIVE = 10          ' CUMULATIVE N ALIVE
NREMAIN = 11            ' N REMAINING AT END OF INTERVAL
SURVRATE = 12           ' SURVIVAL RATE WITHIN INTERVAL
CUMDIEDRATE = 13        ' CUMULATIVE DIED RATE
CUMSURVRATE = 14        ' CUMULATIVE SURVIVAL RATE
MHEXPECTED = 15         ' MANTEL HAENSZEL EXPECTED N DIED
MHDIEDM = 16           ' MANTEL HAENSZEL N DIED
MHVAR = 17             ' MANTEL HAENSZEL VARIANCE
CUMMHEXPECTED = 18      ' MANTEL HAENSZEL CUMULATIVE EXPECTED N
CUMMHDIEDM = 19         ' MANTEL HAENSZEL CUMULATIVE N DIED
CUMMHVAR = 20           ' MANTEL HAENSZEL CUMULATIVE VARIANCE
CHI = 21               ' MANTEL HAENSZEL CHI-SQUARE
PCHI = 22              ' MANTEL HAENSZEL P-VALUE
CPower = 23            ' MANTEL HAENSZEL STATISTICAL POWER

'THE FOLLOWING CONSTANTS ARE USED BY THE PROGRAM

KM = 1                 ' KAPLAN-MEIER METHOD
CE = 2                 ' CUTLER-EDERER METHOD
NO = 0:                YES = 1:                DIFF = 0

```

GETPARMS THE PROGRAM CALLS THIS SUBROUTINE TO GET THE DROPOUT RATE (OMEGA) AND HAZARD RATE (LAMBDA) FOR EACH INTERVAL

```

METHOD = CE: NSTART(1) = 50: NSTART(2) = 50: TAILS# = 2: ALPHA# = .05
NUMGROUPS = 2: NUMROWS = 10: NUMSAMPLES = 50000

OMEGA(1, 1) = .1: OMEGA(1, 2) = .08: OMEGA(1, 3) = .04: OMEGA(1, 4) = .02: OMEGA(1, 5) = .02:
OMEGA(1, 6) = .02: OMEGA(1, 7) = .04: OMEGA(1, 8) = .04: OMEGA(1, 9) = .04: OMEGA(1, 10) = .04:

LAMBDA(1, 1) = .04: LAMBDA(1, 2) = .02: LAMBDA(1, 3) = .01: LAMBDA(1, 4) = .01: LAMBDA(1, 5) = .01:
LAMBDA(1, 6) = .02: LAMBDA(1, 7) = .04: LAMBDA(1, 8) = .04: LAMBDA(1, 9) = .04: LAMBDA(1, 10) = .04:

'IN THIS EXAMPLE THE RATES FOR GROUP-2 ARE A FUNCTION OF THE RATES IN GROUP-1
FOR K = 1 TO 10: OMEGA(2, K) = OMEGA(1, K): LAMBDA(2, K) = LAMBDA(1, K) * 2: NEXT K

```

APPENDIX B (Continued)

COMPUTEPOP WORKING WITH VALUES GIVEN FOR N, LAMBDA AND OMEGA, THIS SUBROUTINE COMPUTES THE POPULATION CUMULATIVE SURVIVAL FOR EACH INTERVAL

```

FOR GROUP = 1 TO 2
  POP(GROUP, 0, CUMSURVRATE) = 1
  FOR K = 1 TO NUMROWS
    POP(GROUP, K, CUMSURVRATE) = POP(GROUP, K - 1, CUMSURVRATE) * (1 - LAMBDA(GROUP, K))
    POP(DIFF, K, CUMSURVRATE) = POP(1, K, CUMSURVRATE) - POP(2, K, CUMSURVRATE)
  NEXT K
NEXT GROUP

```

RUNSIMULATIONS DRAWS SAMPLES FROM THE POPULATION USING TECHNIQUES DISCUSSED IN TEXT

```

REDIM OBS(0 TO 2, 0 TO 11, 1 TO 23)
FOR J = 1 TO 2
  FOR K = 1 TO NUMROWS
    IF K = 1 THEN OBS(J, K, NENTER) = NSTART(J)
    OMEGAPRIME = OMEGA(J, K) + ((LAMBDA(J, K) / (1 - LAMBDA(J, K))) * OMEGA(J, K))
    COUNT = OBS(J, K, NENTER)
    OBS(J, K + 1, NENTER) = COUNT
    FOR L = 1 TO COUNT
      IF METHOD = KM THEN GOSUB KAPLAN
      IF METHOD = CE THEN GOSUB CUTLER
    NEXT L
  NEXT K
NEXT J
EXIT SUB

KAPLAN:
X1 = RND: X2 = RND

IF X1 <= LAMBDA(J, K) THEN
  OBS(J, K, NDIED) = OBS(J, K, NDIED) + 1
  OBS(J, K + 1, NENTER) = OBS(J, K + 1, NENTER) - 1
END IF

IF X1 > LAMBDA(J, K) THEN
  IF X2 <= OMEGAPRIME THEN
    OBS(J, K, NDROP) = OBS(J, K, NDROP) + 1
    OBS(J, K + 1, NENTER) = OBS(J, K + 1, NENTER) - 1
  END IF
END IF
RETURN

CUTLER:
TIE = 0

DO
  X1 = RND: X2 = RND
  IF X1 <= OMEGA(J, K) AND X2 <= LAMBDA(J, K) THEN TIE = 1
LOOP WHILE (X1 <= OMEGA(J, K) AND X2 <= LAMBDA(J, K))
      OR (TIE = 1 AND X1 > OMEGA(J, K) AND X2 > LAMBDA(J, K))

IF X1 <= OMEGA(J, K) THEN
  OBS(J, K, NDROP) = OBS(J, K, NDROP) + 1
  OBS(J, K + 1, NENTER) = OBS(J, K + 1, NENTER) - 1
END IF

IF X2 <= LAMBDA(J, K) THEN
  OBS(J, K, NDIED) = OBS(J, K, NDIED) + 1
  OBS(J, K + 1, NENTER) = OBS(J, K + 1, NENTER) - 1
END IF
RETURN

```

APPENDIX B (Continued)

COMPUTESTATS

COMPUTES SURVIVAL RATES AND P-VALUES FOR THE INDIVIDUAL SAMPLES

```

'THE MATRIX OBS(GROUP,INTERVAL,ITEM) IS USED TO TRACK THE DATA FOR EACH SAMPLE

'COMPUTE SURVIVAL FOR EACH SAMPLE
FOR GROUP = 1 TO 2

  OBS(GROUP, 0, CUMMDIED) = 0: OBS(GROUP, 0, CUMNATRISK) = 0: OBS(GROUP, 0, CUMNALIVE) = 0

  FOR K = 1 TO NUMROWS
    IF OBS(GROUP, K, NENTER) > 0 THEN
      OBS(GROUP, K, DROPRATE) = OBS(GROUP, K, NDROP) / OBS(GROUP, K, NENTER)
    END IF

    IF METHOD = CE THEN OBS(GROUP, K, NATRISK) =
      OBS(GROUP, K, NENTER) - (OBS(GROUP, K, NDROP)) / 2
    IF METHOD = KM THEN OBS(GROUP, K, NATRISK) = OBS(GROUP, K, NENTER)
    OBS(GROUP, K, NALIVE) = OBS(GROUP, K, NATRISK) - OBS(GROUP, K, NDIED)

    IF OBS(GROUP, K, NATRISK) > 0 THEN
      OBS(GROUP, K, HAZRATE) = OBS(GROUP, K, NDIED) / OBS(GROUP, K, NATRISK)
    END IF

    OBS(GROUP, K, NREMAIN) = OBS(GROUP, K, NENTER) - OBS(GROUP, K, NDROP) -
      OBS(GROUP, K, NDIED)
    OBS(GROUP, K, CUMMDIED) = OBS(GROUP, K - 1, CUMMDIED) + OBS(GROUP, K, NDIED)
    OBS(GROUP, K, CUMNATRISK) = OBS(GROUP, K - 1, CUMNATRISK) + OBS(GROUP, K, NATRISK)
    OBS(GROUP, K, CUMNALIVE) = OBS(GROUP, K - 1, CUMNALIVE) + OBS(GROUP, K, NALIVE)
    OBS(GROUP, 0, SURVRATE) = 1: OBS(GROUP, 0, CUMSURVRATE) = 1
    OBS(GROUP, K, SURVRATE) = 1 - OBS(GROUP, K, HAZRATE)
    OBS(GROUP, K, CUMSURVRATE) = OBS(GROUP, K - 1, CUMSURVRATE) * OBS(GROUP, K, SURVRATE)

  NEXT K
NEXT GROUP

OBS(0, 0, CUMMDIEDN) = 0: OBS(0, 0, CUMMHEXPECTED) = 0: OBS(0, 0, CUMMHVAR) = 0

FOR K = 1 TO NUMROWS
  OBS(DIFF, K, HAZRATE) = OBS(1, K, HAZRATE) - OBS(2, K, HAZRATE)
  OBS(DIFF, K, CUMSURVRATE) = OBS(1, K, CUMSURVRATE) - OBS(2, K, CUMSURVRATE)

  OBS(DIFF, K, NDIED) = OBS(1, K, NDIED) + OBS(2, K, NDIED)
  OBS(DIFF, K, NATRISK) = OBS(1, K, NATRISK) + OBS(2, K, NATRISK)
  OBS(DIFF, K, NALIVE) = OBS(1, K, NALIVE) + OBS(2, K, NALIVE)

  OBS(DIFF, K, CUMMDIED) = OBS(1, K, CUMMDIED) + OBS(2, K, CUMMDIED)
  OBS(DIFF, K, CUMNATRISK) = OBS(1, K, CUMNATRISK) + OBS(2, K, CUMNATRISK)
  OBS(DIFF, K, CUMNALIVE) = OBS(1, K, CUMNALIVE) + OBS(2, K, CUMNALIVE)

'COMPUTE P-VALUE FOR EACH SAMPLE
OK = YES: IF OBS(DIFF, K, NATRISK) <= 0 THEN OK = NO
IF OK = YES THEN
  OBS(0, K, MHDIEDN) = OBS(1, K, NDIED)
  OBS(0, K, MHEXPECTED) = (OBS(1, K, NATRISK) * OBS(0, K, NDIED)) / OBS(0, K, NATRISK)
  OBS(0, K, MHVAR) = (OBS(1, K, NATRISK) * OBS(2, K, NATRISK) * OBS(0, K, NDIED) *
    OBS(0, K, NALIVE)) / (OBS(0, K, NATRISK) ^ 2 * (OBS(0, K, NATRISK) - 1))
END IF

OBS(0, K, CUMMDIEDN) = OBS(0, K - 1, CUMMDIEDN) + OBS(0, K, MHDIEDN)
OBS(0, K, CUMMHEXPECTED) = OBS(0, K - 1, CUMMHEXPECTED) + OBS(0, K, MHEXPECTED)
OBS(0, K, CUMMHVAR) = OBS(0, K - 1, CUMMHVAR) + OBS(0, K, MHVAR)

OK = YES: IF OBS(0, K, CUMMHVAR) = 0 THEN OK = NO

IF OK = YES THEN
  OBS(0, K, CHI) = (OBS(0, K, CUMMDIEDN) - OBS(0, K, CUMMHEXPECTED)) ^ 2 /
    OBS(0, K, CUMMHVAR)
  OBS(DIFF, K, PCHI) = PROBSZ(SQR(OBS(DIFF, K, CHI)), TAILS#)
END IF
IF OK = NO THEN
  IF (OBS(0, K, CUMMDIEDN) - OBS(0, K, CUMMHEXPECTED)) = 0 THEN
    OBS(0, K, CHI) = 0: OBS(0, K, PCHI) = .5 * TAILS#: OBS(0, K, CPOWER) = ALPHA#
  END IF
END IF
NEXT K

```

APPENDIX B (Continued)

PROBZ THIS FUNCTION GIVES THE P-VALUE CORRESPONDING TO ANY GIVEN Z-VALUE

```

FUNCTION PROBZ (ZIN, TAILS)
Z = .7071067812# * ABS(ZIN): IF Z > 10 THEN Z = 10
TEMP = (A4 + Z * (A5 + Z * (A6)))
P1 = .5 * (1 + Z * (A1 + Z * (A2 + Z * (A3 + Z * TEMP)))) ^ -16
P2 = P1 * 2
IF TAILS = 1 THEN PROBZ = P1: EXIT FUNCTION
IF TAILS = 2 THEN PROBZ = P2: EXIT FUNCTION
END FUNCTION

```

SAVEDATA ACCUMULATES DATA ACROSS SAMPLES AND WRITES SUMMARY TO DISK

```

FOR K = 1 TO 10
  'W# M# AND T# ARE USED TO UPDATE THE MEAN AND VARIANCE FOR DIFFERENCE IN CUMULATIVE
  'SURVIVAL BETWEEN GROUPS. THIS METHOD DOES NOT REQUIRE TWO PASSES THROUGH THE DATA

  W#(DIFF, K) = OBS(DIFF, K, CUMSURVATE) - M#(DIFF, K)
  M#(DIFF, K) = ((SAMPLE - 1) * M#(DIFF, K) + OBS(DIFF, K, CUMSURVATE)) / SAMPLE
  T#(DIFF, K) = T#(DIFF, K) + W#(DIFF, K) * W#(DIFF, K) * (SAMPLE - 1) / SAMPLE

  'COMPUTE TOTAL NUMBER OF SAMPLES (AND PROPORTION OF SAMPLE) WITH P <= 0.05
  IF OBS(DIFF, K, PCHI) <= .05 THEN SIGTOT(K) = SIGTOT(K) + X
  SIGRATE(K) = SIGTOT(K) / SAMPLE
NEXT K

CLS : PRINT USING "NUMBER REPLICATIONS = ##### SAMPLE #####"; NUMSAMPLES; SAMPLE;

IF SAMPLE < NUMSAMPLES THEN EXIT SUB

FOR K = 1 TO 10: SD#(DIFF, K) = SQR(T#(DIFF, K) / (SAMPLE - 1)):NEXT K

OPEN "MONTE001" FOR APPEND AS #1

FMT$ = "## #.### #.### #.### #.###"

PRINT #1, USING "NUMBER REPLICATIONS = ##### SAMPLE ##### METHOD # N1 ### N2 #####";
NUMSAMPLES; SAMPLE; METHOD; NSTART(1); NSTART(2)

PRINT #1, " POPN SAMPLE SAMPLE POWER "
PRINT #1, " K DIFF MEAN SD "

FOR K = 1 TO 10
PRINT #1, USING FMT$; K; POP(DIFF, K, CUMSURVATE); M#(DIFF, K); SD#(DIFF, K); SIGRATE(K)
NEXT K

```

NOTES:

ALL VARIABLES ARE PUBLIC, I.E. ACCESSIBLE BY ALL SUBROUTINES.

THE FOLLOWING MATRICES ARE DIMENSIONED 0 TO 2 (GROUPS) 1 TO 10 (TIME INTERVALS)
W#, M#, T#, SD# ARE USED TO ACCUMULATE DATA FOR THE MEAN, VARIANCE AND SD OF SURVIVAL RATES
OMEGA, LAMBDA HOLD THE DROPOUT RATE AND HAZARD RATE

THE FOLLOWING MATRICES ARE DIMENSIONED 0 TO 2 (GROUPS) 0 TO 10 (TIME INTERVALS) 1 TO 23 (VARIABLES)
POP HOLDS DATA FOR THE POPULATION. OBS HOLDS DATA FOR THE OBSERVED SAMPLES

THE FOLLOWING MATRICES ARE DIMENSIONED 1 TO 10 (TIMEPOINTS):
SIGRATE HOLDS NUMBER OF STATISTICALLY SIGNIFICANT SAMPLES