

Using the Computer for Library Random Sample Selection

Random sample selection by manual methods is tedious and time-consuming. Fortunately, it is an operation that lends itself well to computerization. A FORTRAN selection program that is appropriate for a wide range of typical library sampling problems is described and made available to the profession.

THERE IS AN INCREASING AWARENESS by librarians of the value of random sampling as a data-gathering tool for library managers. The basic ideas of such sampling are not difficult to grasp, and they are adequately explained in professional library literature.^{1,2}

However, even those librarians who understand the principles of the subject are sometimes reluctant to undergo the drudgery of manually collecting random samples from a random digit table.

For the increasing number of librarians who have access to a computer, this toil is no longer necessary. Random selection, for a variety of reasons, lends itself well to automation:

1. The very repetitiveness that makes selection boring for humans makes it relatively simple to program for a computer.

2. The computer may be programmed to generate the random numbers from which the sample is selected, thus bypassing the need for the random digit tables used in manual procedures.

3. The computer will efficiently sort the random numbers selected into a desired order for use.

4. The computer can be instructed to arrange the printed output in such a manner that it can be used directly as a work sheet for taking the sample.

5. Sample selection programs can be written in a generalized manner so that a

single program will serve to solve a wide range of library sampling problems.

To serve as a concrete example, the author has written such a program in the FORTRAN language. It will solve selection problems requiring one, two, three, or four sampling categories. The nine-digit random blocks from which the sample is selected are generated by the well-known IBM subroutine RANDU. (RANDU relies upon word length and word overflow characteristics of IBM 360/370 machine architecture. Any machine with similar architecture [such as the RCA Spectra 70] can be relied upon to produce a valid sample of random numbers. However, certain character-oriented machines [such as the Burroughs 2500 or 3500] may not produce a valid random number set using RANDU.)

The results are sorted into ascending order and then printed. If a sample is required that is larger than anticipated, the selection may be continued with no loss in randomness by means of a saved seed-value. Four to ten simple input values are required: sample size, a seed-value to initiate RANDU, and the maximum and minimum values for each category included in the sample.

Although it is not possible to include the program itself in this article, librarians desiring to examine it may obtain a copy by writing to the author.

EXAMPLES OF PROGRAM USE

Program use is illustrated below in terms of three typical library sampling problems used as examples by Drott.³

Fred J. Heinritz is assistant director—operations, Division of Library Science and Instructional Technology, Southern Connecticut State College, New Haven, Connecticut.

Sampling Files

A library staff is concerned about the level of accuracy in the library holding records. To estimate the extent of the problem (and perhaps be able to avoid an inventory of the entire collection) the staff decides to select a sample for study from the shelflist. The required sample size at the desired levels of confidence and accuracy is 5,973.

The shelflist consists of 1,200 drawers, numbered consecutively from 1 to 1,200. Each drawer contains up to fourteen inches of cards. Cards will be selected by measuring into drawers to the nearest sixteenth of an inch.

If we let the first sampling category be drawer number (1-1,200), the second whole inches (0-13), and the third sixteenths of an inch (0-15), leaving the fourth unused, the input values are:

MAXA =	1,200
MAXB =	13
MAXC =	15
MAXD =	blank
MINA =	1
MINB =	0
MINC =	0
MIND =	blank
SS =	5,973
IX =	2,715 (randomly chosen)

The first five observations (using a sample size of 100 instead of 5,973 for illustrative purposes) are shown in table 1. The first card will be selected by measuring one inch and 12/16 inches (categories 2 and 3) into drawer 26 (category 1). And so on.

TABLE 1
FIRST FIVE OBSERVATIONS
IN A RANDOM SAMPLING OF FILES

Observation	Category			
	1	2	3	4
1	26	1	12	U
2	54	3	1	U
3	81	8	15	U
4	95	5	11	U
5	99	13	14	U

Book files can be sampled with equal ease. For example, in a typical case, the first category could be the volume, the second the page, the third the column, and the fourth the relative position of the bibliographic record in the column.

Sampling Times

A library staff is taking a survey of users' opinions about library services by handing out questionnaires to a random sample of patrons as they enter the library. The survey is designed to cover twelve days. The library is open a maximum of twelve hours per day during this period. Times are required to the nearest minute. (Drott settles for the nearest five minutes, but using the computer gives more precision with no additional effort.) The plan is to give a questionnaire to the first person (old enough to understand it) to enter the library after each sampling time. The required sample size is 271.

Let the first sampling category be the day (1-12), the second the hour of the day (1-12), the third the minute of the hour (0-59), and the fourth unused. The input values are:

MAXA =	12
MAXB =	12
MAXC =	59
MAXD =	blank
MINA =	1
MINB =	1
MINC =	0
MIND =	blank
SS =	271
IX =	18640587 (randomly chosen)

The first five sample times are shown in table 2. For example, the fifth time selected is the fifty-sixth minute of the second hour of the first day. If the library opens at 9:00 a.m., this is easily translated into the first day at 10:56 a.m. And so on. Note that this program eliminates the need for the elaborate tables of random sampling times found in various explanations of work sampling.⁴

Sampling Collections

A librarian needs to know whether significant shelf space can be saved by removing little-used books from the collection. A little-used book is considered to be one that has not circulated in the last five years. The sample is to be taken by examining the date due slips and book cards in the back of randomly selected books. If 15 percent or more of the collection are little used, the librarian will remove these books from the shelves.

The collection consists of about 19,000

TABLE 2
FIRST FIVE OBSERVATIONS
IN A RANDOM SAMPLING OF TIMES

Observation	Category			
	1	2	3	4
1	1	1	0	U
2	1	1	23	U
3	1	1	32	U
4	1	1	49	U
5	1	2	56	U

volumes. It is arranged on 234 sections of shelving, each with six shelves. Each shelf can contain up to 25 books. The required sample size is 288.

If we let the first sampling category be section of shelving (1-234), the second shelf number within the section (1-6), and the third book position on the shelf (0-25), leaving the fourth unused, the input values are:

MAXA = 234
 MAXB = 6
 MAXC = 25
 MAXD = blank
 MINA = 1
 MINB = 1
 MINC = 0
 MIND = blank
 SS = 288
 IX = 904237305 (randomly chosen)

The first five observations are shown in table 3. The first book to be checked is in the fourth section, first shelf, and eighth from the end of the shelf. And so on.

TABLE 3
FIRST FIVE OBSERVATIONS
IN A RANDOM SAMPLING
OF THE COLLECTION

Observation	Category			
	1	2	3	4
1	4	1	8	U
2	4	2	20	U
3	5	5	6	U
4	6	4	15	U
5	8	1	13	U

CONCLUSION

The program speaks for itself. Many librarians will be able to use it as it stands. Others may find it necessary or desirable to make minor modifications. In either case, it is hoped that the availability of this program will encourage librarians to make increased use of random sampling.

REFERENCES

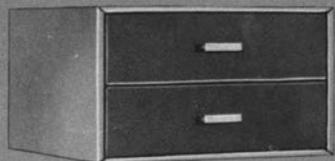
1. M. Carl Drott, "Random Sampling: A Tool for Library Research," *College & Research Libraries* 30:119-25 (March 1969).
2. Richard M. Dougherty and Fred J. Heinritz, "Sampling," in their *Scientific Management of Library Operations* (New York: Scarecrow, 1966), p.115-35.
3. Drott, "Random Sampling," p.122-25.
4. John S. Goodell, *Libraries and Work Sampling* (Littleton, Colo.: Libraries Unlimited, 1975), p.15-17.

5 Star Bonus

For as little as \$100 you can participate in the



WSI Minicat TN Microfiche Reader



Wilson 2-Drawer
Microfilm/Microfiche Cabinet



Dukane MMR 16 + 35 Microfilm Reader

Your choice! Get FREE backfile titles or take advantage of ONE-TIME-ONLY SPECIALS on top-name microfilm and microfiche readers, carrels and cabinets. Here's how—

Simply place an order before July 30th for \$100 or more worth of current UMI serials microform subscriptions. For example, a \$100 new subscription order entitles you to \$50 worth of backfile titles, FREE.

Enhance Your Library Facilities

Ordering current titles in microform has many advantages. Microforms take only 10% as much space as bound volumes. They are practically free from theft and mutilation, and last much longer than periodicals printed on today's papers. The most convenient way to order current microform titles is on the UMI Serials Subscription Service.



Xerox 350 Reader Carrel

Special

Special offer!



What's more, if you order before July 30th you are entitled to FREE backfile titles and SPECIALS on top quality microfilm and microfiche equipment—an inexpensive way to greatly enhance your library's facilities!

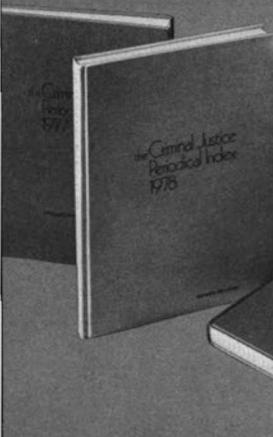
How To Order

For complete information on this UMI Serials Subscription Service special offer, send in the coupon below with no obligation.

Get more for your dollar—act now, send for further details, while this offer is still in effect!



Xerox 350 Manual
Microfilm Reader



**FREE! 1979 Edition
of Criminal Justice
Periodicals Index**

University
Microfilms
International



Don't miss this opportunity! Send in this coupon or call toll free for complete ordering information.

Please send me the 1979-80 UMI Serials on Microform Catalog.

Name _____

Title _____

Company/Institution _____

Address _____

City _____ State _____ Zip _____

Telephone _____
(area code)

MAIL TO: Serials Subscription Service,
University Microfilms International,
300 North Zeeb Road, Ann Arbor, MI 48106

C&RL-2

TOLL FREE NUMBER: 1-800-521-3042

2 Important New Books for Academic

Librarians These outstanding books from THE ORYX PRESS will provide you with important, timely information on subjects critical to contemporary library management — including current insight into budget management and justification.

Quantitative Measurement and Dynamic Library Service is designed to help and encourage those with little statistical background to utilize analytical and quantitative methods in library decision-making processes. Includes over 100 charts and tables. Edited by Ching-chih Chen. ISBN 0-912700-17-3, clothbound, 312 pages. \$16.50.

The Nature and the Future of the Catalog is a collection and analysis of the proceedings of two American Library Association Institutes sponsored by the Information Science and Automation Division. Edited by Maurice J. Freedman and S. Michael Malinconico. ISBN 0-912700-08-4. Available June, 1979. Clothbound, about 300 pages. \$16.50t.

Neal-Schuman Professional Books

Order your copies today. Please enclose 95¢ per book for postage and handling.



ORYX PRESS

3930 East Camelback Road
Phoenix, Arizona 85018 • (602) 956-6233

Visit THE ORYX PRESS at ALA in Dallas, Booth #885