

A Statistical Method for Circulation Analysis

BY WILLIAM L. LEFFLER

WHAT is the best way to distribute money allocated for the purchase of books? Should more money be spent on additional copies of the very popular works, or should a greater variety of books be bought?

Some techniques borrowed from the industrial inventory control specialists will help answer these and other questions posed by progressive library managements. One particularly applicable problem of industrial concern is the prediction of stock levels of returnable items. For example, a tool room lends many thousands of different types of tools to the mechanics it serves. Assurance of an adequate supply of each type of tool is vital. Furthermore, a limit on the amount of money to procure these tools demands a rational and exacting method of deciding what and how much to buy.

This problem is similar to those libraries where high demand calls for multiple copies of many books. With mechanical and electronic data processing,¹ the analysis below has been successfully applied to the tool crib problem above, and can be used with little modification for the corresponding library management problem.

The high turnover pattern of a particular book² can be shown graphically as in Figure 1. Each horizontal line repre-

Mr. Leffler is with the Shell Oil Company, 50 West 50th Street, New York 20.

sents the borrowing and returning of a copy of the book and the length of time it was kept out.

During the forty days of observation, fifteen loans were made, fifteen returns were made. Notice, however, that we do not need fifteen copies of our book. On the seventh day, when the sixth borrower requested a copy, the first copy was now available to be given to him. On the sixteenth day, when the seventh borrower requested a copy, four more copies had been returned. Would it not be convenient to look at a chart like this and judge visually the maximum number of copies that were in use? It would be easy to determine whether the stock of this book were being taxed to capacity, and if so, whether additional copies should be purchased.

Unfortunately, this simple analysis is an inordinately lengthy one, even for a high-speed computer. A more feasible approach is that taken in the tool crib problem—that of statistical prediction.

Let us assume some parameters of our problem. Let T = number of days of observation; N = total number of loans; A_i = length of time a copy was kept out; \bar{A} = average length of time all copies were kept out; S = (sum of all A_i)/ N .

$$S = \frac{\sqrt{\text{Sum of } (A_i - \bar{A})^2}}{N} = \frac{\sqrt{102.40}}{15} \\ = 2.58$$

¹R. W. Gibson and G. E. Randall, "Circulation Control by Computer," *Special Libraries*, LIV (July-August 1963), 333-38.

William L. Leffler, "And Now, Computer Tool Control," *Mill and Factory*, LXXIII (September 1963), 92.

²In this context, a "book" refers to a particular title, author, edition, etc., of which a library has multiple copies.

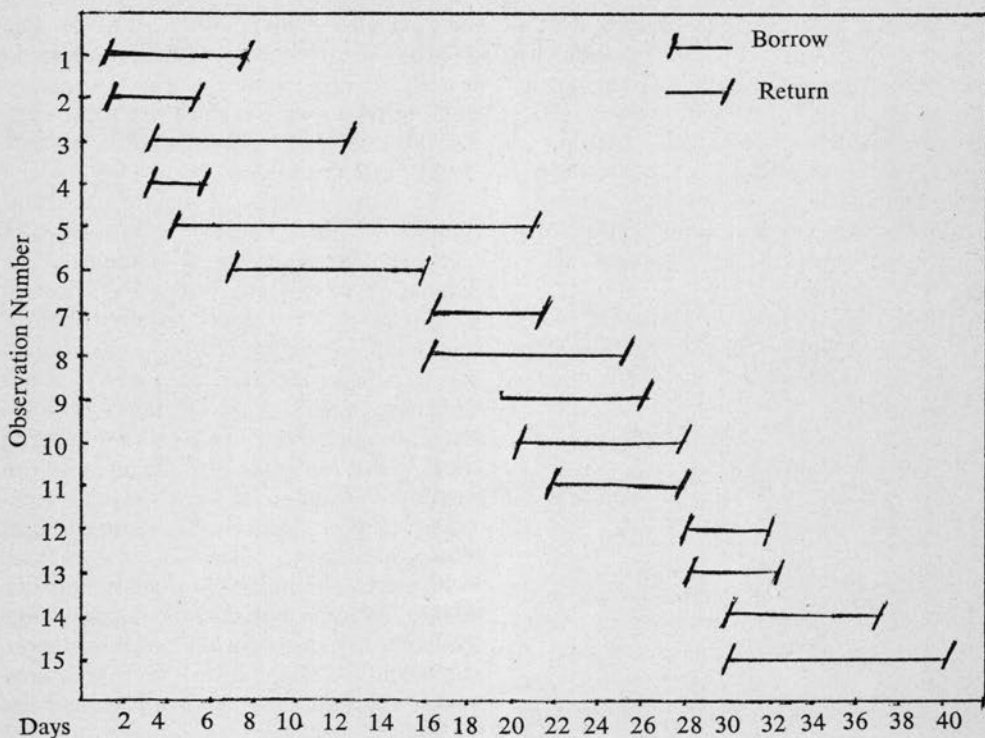


FIG. 1.—The borrow-and-return pattern for a typical multiple copy book.

$$\begin{aligned}
 N_{95\%} &= N_{av} + 3s \\
 &= 2.65 + 7.74 \\
 &= 9.39
 \end{aligned}$$

In our example graphed above in Figure 1, we find that $T =$ forty days, $N =$ fifteen loans, and A_i 's are read off the graph, and \bar{A} is the sum of all the A_i 's divided by the number of loans, or $\bar{A} = 102/15 = 6.8$.

Empirical research has shown that the average number of books that are out on loan is calculated as follows:

$$N_{av} = (N/T) \times (\bar{A})$$

In our example above,

$$\begin{aligned}
 N_{av} &= (15/40) \times (6.8) \\
 &= 2.65
 \end{aligned}$$

We must be careful, however, about interpreting this average. By its very definition an average implies that in half the cases the number will be less than the

average, and in half the cases *greater*. Since we cannot be satisfied with knowing what the usage is in only half the cases, we must make statistical adjust-

TABLE 1

CALCULATION OF THE STANDARD DEVIATION AND STOCK LEVEL OF 95 PER CENT CERTAINTY

OBSERVATION	A_i	\bar{A}	$(A_i - \bar{A})$	$(A_i - \bar{A})^2$
1	6	6.8	-0.8	0.64
2	4	6.8	-2.8	7.84
3	9	6.8	2.2	4.84
4	2	6.8	-4.8	23.04
5	11	6.8	4.2	17.64
6	9	6.8	2.2	4.84
7	5	6.8	-1.8	3.24
8	9	6.8	2.2	4.84
9	7	6.8	.2	0.04
10	8	6.8	1.2	1.44
11	6	6.8	-0.8	.64
12	4	6.8	-2.8	7.84
13	4	6.8	-2.8	7.84
14	7	6.8	.2	0.04
15	11	6.8	4.2	16.64
105	Sum of $(A_i - \bar{A})^2 =$	102.40		

ments. Just as the average gives us the upper limit 50 per cent of the time, we can predict the level of loans that covers 90 per cent, 95 per cent, or any other per cent. Since we are working with probability, we cannot give a level for 100 per cent certainty. So, as in the industrial situations, we can reasonably settle for a level which gives us 95 per cent assurance.

Table 1 shows how the statistical factor is computed. Algebraically the "standard deviation" is computed by the following formula:

$$\text{Standard Deviation} = S$$

$$S = \frac{\sqrt{\text{Sum of } (A_i - \bar{A})^2}}{N}$$

By adding three of these standard deviations to the average number of copies on loan, we arrive at the number which represents the *maximum copies that would be on loan 95 per cent of the time.*

In summary, the procedure involves the following steps:

1. Compute the average number of copies on loan.
2. Compute the average length of time on loan.
3. Compute the standard deviation.
4. Add the average number of copies on loan plus three times the standard deviation.

Now, what does this numerical manipulation do for us? First, as we pointed out above, it provides a number by which we may judge the adequacy of the number of copies we own. If we look back at our graphic example above, we would calculate that 95 per cent of the time we would have no more than ten copies on loan. (See Table 1.) If we do, in fact, have more than this number of copies, the advisability of procuring more is questionable. If we have less copies, the acquisition of additional copies is rationally justified, since we expect that we will

have occasions when we will not be able to provide a customer with his requested text. If we have a short supply of a large number of books, we can even rank them and choose those books whose shortage is the most critical.

We must, of course, mention the limitations of this analysis. Like the tool crib problem, one critical element in the evaluation of this analysis is the effort of producing it. The inputs required include the detail of the borrowings and returns for the desired period of time. Computerization makes the clerical effort feasible and must be considered an integral part of the analysis. The recent developments reported in the two articles footnoted above promise the realization of this requirement.

A second limitation is mathematical in nature. If sufficient data is not available in relation to the period of observation, the results will be distorted. For example, if the only activity of a certain book during a year were three loans on the same day for only one each, we would calculate a need for only .009 books. This is an obvious contradiction to the facts given in the analysis. We must, therefore, impose a requirement that the length of observation include that period of time in which loans and borrows actually take place. With this restriction we would calculate for three books in this last example. Other allowances for which adjustments must be made include "lost" copies (whether reported in fact or assumed by nonreturn), damaged or unusable copies, and the like.

What we have done here is outline a path toward a future in library science. The contributions to better management of libraries will not come easily or cheaply. They must be borrowed and bought from a diversity of other fields. The analysis we have developed here only indicates how a snatch of the rapid technological change we see around us can be harnessed to a seemingly unrelated field, and made to contribute to its advancement. ■■