GENERATION OF RANDOM PERMUTATIONS OF GIVEN NUMBER OF ELEMENTS USING RANDOM SAMPLING NUMBERS

By C. RADHAKRISHNA RAO

Indian Statistical Institute

SUMMARY. A general method is given for generating random permutations of integers using a table of random sampling numbers and without wasting the random numbers read. This is more convenient in practice, specially when random permutations of large numbers of elements are needed. It is suggested that even for permutations of small numbers, the method offers greater scope than consulting a table of a limited number of random permutations.

Introduction

Fisher and Yates (1957) in the introduction to their tables give several methods of obtaining random permutations of integers 1 to n. Some of these methods require random permutations of subsets of integers to build up the whole permutation. A direct method for obtaining a random permutation of integers 1 to n, (or 0 to $n-1$) is to choose k columns in a table of random digits, where $10^k \geqslant (n-1)$, and write down the numbers 0 to $(n-1)$ *in the order they occur*, omitting all those above $(n-1)$ and ignoring repetitions. Thus for obtaining a random permutation of numbers 0 to 24 we would take two columns; and the process would naturally reject on an average 75 per cent of the random numbers read. Further, if n is large there is the difficulty of remembering the integers that have not occurred up to any stage while reading the random numbers.

In this note, I give a general method of generating random permutations of integers 1 to n for any n, which does not waste any random number read, and which can be conveniently used even for large n.

One-way classification method

Choose a column in a table of random digits and read the digits from any starting point. Take the classes defined by the digits 0 to 9 and write 1 in the class of the first random number, 2 in that of the second number, 3 in that of the third number and so on up to n. For instance, if $n = 7$ and the random numbers read are 5, 3, 9, 1, 2, 3, 1 the integers 1 to 7 are recorded as shown in the first block of Table 1.

TABLE 1. ONE-WAY CLASSIFICATION METHOD

classes defined by digits									
0	1	2	3	4	5	6	7	8	9
	4	5	2		1				3
	7		6						
...	...	...	...	...	...	...	...	...	...
		4		7					
...	...	...	...	...	...	...	...	...	...
6									2

To obtain a random permutation of integers 1 to 7, first write down the groups of integers in the different classes one after the other starting from the '0' class, which in the present example leads to

(4, 7), 5, (2, 6), 1, 3.

Then randomly permute the subsets within the brackets and remove the brackets. The problem of obtaining random permutations for large numbers is thus reduced to that of obtaining for subsets with expected number in each subset being equal to 1/10 of the original number. The same process may be used for obtaining permutations of integers within subsets as shown in the last two blocks of Table 1, unless there is a ready way of obtaining these permutations such as referring to tables of random permutations. To permute the subsets continue to read the random numbers. Let the next digits be 2, 4, 9, 1, 8, 0.... Consider the subset (4, 7); 4 goes in class 2 and 7 in class 4, which completes the permutation. Similarly a permutation of (2, 6) is derived in the last block of Table 1, and the complete permutation is obtained as

4, 7, 5, 6, 2, 1, 3.

If permutations for large values of n are needed, several classes will have multiple entries, needing one or more repetitions of the process for each subset. This can be avoided to a large extent by following the two-way classification method and, if convenient, by adopting the built-in procedure of simultaneously permuting within subsets while recording the integers, as detailed in the next section.

Two-way classification method

Choose two columns in a table of random digits. Each random number would now consist of two digits which define a cell in a 10×10 two-way table as shown in Table 2.

TABLE 2. TWO-WAY CLASSIFICATION METHOD

first digit \ second digit	0	1	2	3	4	5	6	7	8	9
0		9				10				
1								2		
2				6		12				
3		1,5						7		11
4						4,16				
5				17						
6		8							15	
7					13		18			
8	3									
9									14	

Let a random permutation of numbers 1 to 18 be needed. Write 1 in the cell of the first random number, 2 in the cell of the second number, ... and so on up to 18. For instance, corresponding to the first five random numbers 31, 17, 81, 45, 31 the entries 1 to 5 would be as shown in Table 2. Proceeding further in the same manner, the first 18 random numbers are used to post all the integers 1 to 18 in the cells. In two cells there are two entries, which could have been avoided by omitting a number if its cell is already filled. Otherwise the integers within subsets (1, 5) and (4, 16) have to be randomly permuted first. Suppose (5, 1) and (4,16) are the permutations arrived at by some process. The entire permutation could be read out in any convenient manner, say from left to right row by row, the order of integers in cells with multiple entries being determined by independent permutations. In the present example the permutation is

9, 10, 2, 6, 12, 5, 1, 7, 11, 4, 16, 17, 8, 15, 13, 18, 3, 14.

The problem of multiple entries, due to repetition of random numbers, is not serious since the chance of having more than two or three entries in a cell is small if n is not very large, and it is easy to obtain a random permutation of two or three integers.

It is also possible for the random permutation of multiple entries to be built into the process of entering the numbers in the cells as follows. For instance, when 5 has to be placed in the cell (31), which already has 1, read the digit in the same row containing 31 in an adjacent column of the random number table. If this digit is even, write 5 (the second number to be placed in the cell) before 1 (the first number), and if odd, write 5 after 1. Random permutation of 1, 5 is already secured by this operation, and similarly the order of integers in (4, 16) could have been settled while posting 16 and the whole random permutation is obtained straight away from the two-way table. If a random number repeats for a third time, first read the digit in the same row in an adjacent column. If this is 0, pass on to the next column, and so on, till a non-zero digit is obtained from the row. The number to be written in the cell is put in the first position if a non-zero digit read from an adjacent column is $\equiv 0 \pmod 3$, in the second position if $\equiv 1 \pmod 3$, and in the third position if $\equiv 2 \pmod 3$. A similar device could be adopted in case a random number repeats for a fourth time, and so on. As stated earlier, the permutation of multiple entries in the cells may be undertaken at the final stage, if that is more convenient.

For larger values of n, to avoid multiple entries, one could use three columns of a random number table and consider a triple coordinate system to identify 1000 cells of a three-way table. The best way of representing would be a triple classification table involving 10 two-way sub-tables, each sub-table corresponding to a given value of the third digit. The procedure outlined before can then be followed.

It is important to note that for any value of n, one can use any arbitrary number k of columns of the random digits and proceed to distribute the integers in the cells of a k-way table. But it may be convenient to use one-way classification method for $n \leqslant 10$, two-way classification method for $n \leqslant 100$ and so on.

In this connection, the reader is referred to two papers by Lahiri (1951) and Matthai (1953) on choosing, from a list of sampling units, with unequal and equal probabilities respectively, a random sample with the help of random sampling numbers.

References

Fisher, R. A. and Yates, F. (1957) : *Statistical Tables for Biological, Agricultural and Medical Research*, 5th ed. Edinburgh, Oliver and Boyd.

Lahiri, D. B. (1951) : A method for sample selection providing unbiased ratio estimates. *Bull. Inst. Int. Stat.*, **33**, 2, 133–140.

Matthai, Abraham (1953) : On selecting random numbers for large-scale sampling. *Sankhyā*, **13**, 257–260.

Paper received : July, 1961.