Data Sets Having Integer Means and Standard Deviations

McGown and Spencer (1980) suggested that construction of tests for introductory statistics is facilitated if data sets have means and standard deviations that are integers. They presented several sets for samples of n = 5, 6, and 7. It is fairly easy to devise data sets with integer means and standard deviations for larger samples. One advantage of data sets with larger n's is that they look realistic when plotted as histograms or frequency polygons. Such frequency distributions are also typical of tabulations where data have been grouped into class intervals for computational purposes.

The procedure is to devise several sets with unit variance and zero mean, and then any combination of these also results in another set with unit variance and zero mean. A set with any desired mean (M_T) and standard deviation(s) can be achieved by employing the transformation $T = s(x) + M_T$, where T represents the reported measures or scores. Table 1

Table 1. Data Sets with Zero Mean and Unit Standard Deviation from Which Other Data Sets Having any Desired Mean and Standard Deviation Can Be Generated

x	f_A	f_B	f_C	$f_{A'}$	$f_{B'}$	$f_{C'}$	f_A+f_C	$2f_A+f_B$	T_1	T_2	T_3
2.0	1	1	1	0	1	2	2	3	44	62	45
1.5	1	1	2	1	2	1	3	3	43	59	40
1.0	2	2	3	3	1	1	5	6	42	56	35
.5	2	5	3	4	4	6	5	9	41	53	30
0	1	3	6	1	3	6	7	5	40	50	25
-.5	4	4	6	2	5	3	10	12	39	47	20
-1.0	3	1	1	2	2	3	4	7	38	44	15
-1.5	1	2	1	1	1	2	2	4	37	41	10
-2.0	0	1	2	1	1	1	2	1	36	38	5
n	15	20	25	15	20	25	40	50			
mean	0	0	0	0	0	0	0	0	40	50	25
s.d.	1	1	1	1	1	1	1	1	2	6	10

presents values of x in the first column and frequency distributions in the next three columns for data sets A, B, and C with n's = 15, 20, and 25 respectively. Each distribution has unit standard deviation and a mean of zero. From the three frequency distributions it is possible to generate many other distributions all having mean = 0 and standard deviation = 1. So, for example, in columns 5, 6 and 7 each of the original sets is simply inverted or transposed by interchanging frequencies in the highest interval with those of the lowest, those from the next to the top interval with the interval next to the bottom, etc. If a distribution where n = 30 is desired, the frequencies in distribution A can be multiplied by 2, or the frequencies in distributions A and A' can be added. To obtain a frequency distribution with n = 35, the frequencies of distribution A and distribution B can be added; or frequencies of distribution B can be added to those of A'. Shown in the seventh column of Table 1 is the distribution with n = 40 resulting when distributions A and C are combined, and in the next column a distribution of n = 50 that results from doubling the frequencies in A and then adding the frequencies from B.

Values of the measures or scores, T, can be determined by utilizing the transformation noted above to achieve any desired mean and standard deviation. The last three columns of Table 1 show three different transformations: $T_1 = 2(x) + 40$; $T_2 = 6(x) + 50$, and $T_3 = 10(x) + 25$. Any of the frequency distributions related to T_1 scores would have a mean = 40 with standard deviation = 2, while T_2 scores would result in a mean = 50 with standard deviation = 6, etc.

Reference

McGown, W. P., & Spencer, W. B. For statistics classes: Data sets with integer means and standard deviations. *Teaching of Psychology*, 1980, 7, 63.

Frank J. Dudek
Department of Psychology
University of Nebraska
Lincoln, NE 68588

Positive Side Effects of Online Information Retrieval

As online information retrieval becomes increasingly available for use by undergraduate students it becomes important to understand the appropriateness and ramifications of the use of this simple yet powerful tool for generating research literature (Atherton & Christian, 1977). In a recent article in this journal, Parr (1979) presents general arguments for the addition of training in online information searching for undergraduates. Parr argues that instead of allowing undergraduates to circumvent traditional bibliographic techniques and skills, use of online systems reinforce such bibliographic skills and adds new dimensions to library usage.

But is the influence of online retrieval of information limited to the acquisition and reinforcement of library skills? It is important to illuminate the scope of influence of these systems. Does the use of such systems influence attitudes toward assignments, class, or library usage? Does its use affect course performance? Exploration of these unintended side effects can illuminate the appropriate use of such online technology in undergraduate education.

Method. Subjects were 11 of 15 undergraduates enrolled in an environmental psychology course at a small liberal arts college. All students were invited to participate in this experiment which was introduced as an effort to evaluate a new library system. This new system was described as being relevant for a portion of their course requirement, a literature review worth 20% of their final course grade. The literature review was described in a handout during this first class session as an opportunity to actively explore research in environmental psychology in depth. It was suggested that the best approach would be to select an interesting section of the course text (Bell, Fisher, & Loomis, 1978) and review studies not included. Students were also given detailed style requirements and references of examples of literature reviews in *Psychological Bulletin*, although it was pointed out that their reviews need not be as extensive.