MICROCOMPUTER APPLICATION #3

Using SPSS for Windows To Draw Random Samples

SPSS for Windows provides a procedure that will draw random samples of any size from a file (such as the data file you constructed in Microcomputer Application #1). We can use this procedure to illustrate some points about sampling and to convince skeptics that properly selected random samples will produce sample statistics that are close approximations to population parameters. In this assignment the instructions below will provide the mean and standard deviation of five samples of various sizes drawn from this file. The sampling fractions are "roughly" 10%, 25%, 50%, 75%, and 90% of the population size, and the program selects them by a process that is similar to using a table of random numbers. Therefore, the samples constructed may be considered "simple random samples." The actual mean (and standard deviation) salary of the fifty teachers (these will be the population mean or mu and population standard deviation or sigma) can be obtained from Microcomputer Application #2. Proceed as follows:

To open an existing file (one you've saved to your disk), make sure your disk with the data file is in the appropriate drive. Follow the instructions in assignment number #1 to open SPSS for Windows. Then click on "File"; click on "Open"; and click on "Data". Under File Name, type drive and "***" (this is the name you gave your file). Finally, click "OK" and wait a few moments for you data to be displayed. Now you are ready to proceed with selecting random samples.

Click "Data"; Click "Select Cases"; Click "Random Samples". Under "Select Cases" click on "Random Sample of Cases" where it says "Sample Size" approximately, type 10%. Then click "Continue"; Then "ok". Then Click "Statistics"; then click "Summarize"; and then click "Descriptives". Under "Descriptives" with pointer arrow select salary. Then click "ok". Wait a few moments and your output will appear on the screen. Follow this same procedure for samples of 25%, 50%, 75%, and 100%.

I've included the sample means I found just to illustrate some results. When you run this task, you will, of course, draw samples that will probably be different from mine, so your results will be at least slightly different. The different sample values are due to random sampling error.

My results were as follows:

Sampling Fraction	Sample Size	Mean	Standard Deviation
.10	7	25.71	6.26
.25	12	32.08	9.22
.50	20	33.10	10.10
.75	41	32.88	10.78
.90	44	33.30	10.72

What is the population mean? ___ standard deviation? ___
Note that the small 10% sample is the only one of the estimates that is not very close in value to the population mean. In fact, it is the only one that is not within 1 standard deviation of the population mean. This is an unusual result since about 68% of all sample means will be within ±1 standard error of the population mean (or, about 68% of the area under a normal curve will fall within ±1 standard deviation of the mean). At any rate, the main thing to notice here is that some of the sample means are extremely close to the population value. This should reinforce one of the main points of statistical methodology: statistics calculated from properly drawn random samples will typically be reasonable approximations to their population counterparts.