Homework #3: Sampling and Distributions

Questions

(1) Use the provided dataset at the assignment page in Canvas.

(2) Select any two *interval* variables that you will work use for this assignment.

(3) Use SPSS to randomly select 300 cases (which you will use as your "sample") from the full sample of 3432 cases (which we will treat as the "population").

- Choose DATA then SELECT CASES
- Select "Random sample of cases" and click "Sample…"
- Tell SPSS to pick "Exactly 300 cases from the first 3432", and click continue
- Be sure to choose "Deleted" rather than "Filtered", or your file won't get any smaller
- Use FILE - SAVE AS to preserve your subsample of 300 cases.

(4) For those two variables only,

a) produce, interpret, and compare frequency distributions for these two variables;
b) examine these to determine whether there are any missing values which you need to take care of, and whether you need to do any recoding. (If you *do* need to take care of missing values or do any recoding, you will need to do part (a) again.)
c) request, interpret, and compare appropriate measures of central tendency and dispersion

(5) Using SELECT CASES again, randomly select 50% of your sample (150 cases)

(6) For the same two variables used in (4) above:

a) produce, interpret, and compare the shapes of the distributions for these two variables,
b) request, interpret, and compare appropriate measures of central tendency *and* dispersion; &
c) compare the distributions (shape, center, and spread) of this *sub*sample of 150 cases to the distributions of the same variables for your full sample of 300 cases

BONUS:

(7) In coming weeks, you'll be writing hypotheses about variables. They come in pairs – null and research – but each pair has the same hypothesized value, some number that you think the population parameter may be. For extra points on this homework assignment, find values for each of the variables you've used, for some population or some other sample – numbers that you might use in writing hypotheses – and provide the citation(s) for where you found them.

Notes

- Be as complete as possible in your answers. On this and all assignments, put all answers in prose (sentence/paragraph) form, in a way that non-statisticians will understand most of what you say; single-word or phrase answers are not sufficient.
- Copying a bunch of numbers is neither the assignment nor a reasonable college achievement; you should be *selecting* statistics that are *appropriate* for the given variables and distributions.
- Be explicit and complete in *explaining* your answers. Show all work done, including any calculations, & explain all of the steps involved and all of the parts of each calculation.