STAT8801 Case Study: Combining Motives

Group 4: Zhou, Xin, Joel, Greg, Alain
February 3, 2012

Recap of the situation:

A Psychology student has administered a questionnaire to find out people's motives for volunteering. Based on the questionnaire, she can compute six scores, corresponding to six motives for volunteering. She would like to have a single score to be used as an overall indicator of motivation for volunteering.

Questions for the client:

Motives – What are the six motives for volunteering? Are they equally distinct to each other?

Questionnaire –What kinds of questions are on the questionnaire? Can they well reflect and differentiate the six motives? Would some of the items be likert-scale type responses, or yes-no, true-false types of response?

Sampling methodology – What was the sampling methodology she used?

Objectives – What is the specific research question? This would enable us to work towards tailoring or specifying an appropriate statistical analysis.

Concerns about what she has done:

We are wondering what is her sampling frame. Selection bias would occur when some part of the target population is not in the sampled population and cause the samples under-coverage. We are concerned that she may give the questionnaire just to the students on campus. If so, the result would greatly be biased, because students are more likely to volunteer to some kinds of community service work than working people.

We are also concerned about the questionnaire design. Is it well designed so people won't have any misunderstandings? For example, certain words mean different things to different people.

Concerns about what she is going to do:

We are concerned about what she is going to do with the final score. Why she needs one score as an overall indicator of motivation for volunteering? Is this score meaningful? Is she going to use this score as a prediction? Does she intend to use this final score to make inferences in any specific populations?

Ideals for methods:

The first way to get a single score is just averaging all the scores. If she can tell us her sampling methodology, we may have different ways to do the average. She may just have used a simple random sample. She just grabbed 300 volunteers randomly from the whole population. Then we can just average all the scores.

She may have used stratified sampling, in which the population is divided into subgroups (strata). The strata might be different ethnic or age groups in her target population. Then we can do the average within each stratified group and then combine them together in some way that can give different weights to different strata. Maybe it is even better if we just combine the scores within each subgroup instead of the whole sample.

Second we can use principal component analysis (PCA) to convert these correlated variables into principle components. The shortcoming of this strategy is that there might be several principal components. The results of PCA also depend on the scaling of the variables.