Introduction

Much time has passed since you sent me the write-up on similarities and differences between the simulation studies. I don’t appear to have an electronic version of the document referred to above, but I do have a paper copy of the document. In that document, you describe some differences between the simulation studies that you conducted and the ones that I considered. Here, I want to comment on perhaps the over-riding issue underlying the differences between the simulations. It is my understanding that you used a different strategy for simulations when you were evaluating the mixed model and Scott and Smith’s model than when you evaluated the random permutation model. The difference seems to have arisen when generating the population. I will first try to summarize my understanding of how you proceeded, and then discuss the differences.

Discussion of comments in Section 3.1 on Developing the Population for the Random Permutation Model

In general, I consider a population as a free-standing entity. I don’t believe that it is necessary to specify a model to define a population. I don’t believe that it is necessary to postulate random variables to define a population. Basically, the definition of the population I think can stand on its own. To define a population, a listing of units is specified that can uniquely identify units in the population. A characteristic of the unit is the response. It can be directly observable (such as gender) or potentially observable (such as weight or pulse rate, where with repeated measures, a closer and closer approximation to the true value can be made, i.e. measurement error).

In the “Similarities and Differences” document (section 3.1), you view the population, \( y \) as one realization of a random variable \( Y \), but only specify \( Y \) in terms of the first and second moments. I think that specifying only the first two moments of \( Y \) is not adequate to make this definition equivalent to the one that I was using. Specifically, I think that if you only start by fixing the first two central moments of \( Y \), it is possible to have different distinct values of \( y \) that could serve as realizations. Thus, by considering \( Y \) as the starting point, you have defined a different problem.

I don’t think the problem you have defined in this way is a problem of practical interest when the interest concerns a finite population of units. Many problems can be defined in this way (in terms of a finite population). This is the non-stochastic definition of the problem. The stochastic version links the non-stochastic population with a model that incorporates probability (like the random permutation model). I agree that if you start with \( y \) and then give equal probability to all permutations of \( y \), you can represent this by \( Y \) with the first and second moments as you have indicated, but the two specifications are different.

You state that each realization of \( Y \) will generate the same finite population. I believe this statement is wrong. The first two moments of \( Y \) will not uniquely define \( y \).

Discussion of comments in Section 3.2 on MM and SS Models
In order to compare predictors for a problem, the comparison must be based on the same problem. I don’t believe that you consider the same problem for the RP and MM. Specifically, I don’t believe you consider the outcomes of random permutations of a finite population as defining the space over which the predictors are evaluated.

Using the random permutation model variance, the variance for a PSU is given by

\[
\begin{pmatrix}
\sigma^2_e + \sigma^2_r + \\
\sigma^2 - \frac{\sigma^2_e}{M} - \frac{\sigma^2}{N} \\
\sigma^2 - \frac{\sigma^2_e}{M} - \frac{\sigma^2}{N} \\
\end{pmatrix}
\begin{pmatrix}
\sigma^2 - \frac{\sigma^2_e}{M} - \frac{\sigma^2}{N} \\
\sigma^2 - \frac{\sigma^2_e}{M} - \frac{\sigma^2}{N} \\
\sigma^2 - \frac{\sigma^2_e}{M} - \frac{\sigma^2}{N} \\
\end{pmatrix}
\begin{pmatrix}
\sigma^2 - \frac{\sigma^2_e}{M} - \frac{\sigma^2}{N} \\
\sigma^2 - \frac{\sigma^2_e}{M} - \frac{\sigma^2}{N} \\
\sigma^2 - \frac{\sigma^2_e}{M} - \frac{\sigma^2}{N} \\
\end{pmatrix}
\] .

Obviously, the mixed model assumes that the co-variances are zero between SSUs in different PSUs. However, you can see that there are different choices for the compound symmetric terms. That’s what lead to my considering different possibilities. You say it is reasonable to assume that \(-\frac{\sigma^2_e}{M} - \frac{\sigma^2}{N}\) is zero. Why? There is no magic to what the definitions is for a “variance component parameter”. We could define a parameter to be any number of things- as for example as \(\sigma^2_e\) or \(\frac{\sigma^2_e}{M}\) or \(\frac{\sigma^2_e}{M} - \frac{\sigma^2}{N}\). With any of these definitions, the covariance for the MM will be compound symmetric, and each will approach \(\sigma^2\) as \(N\) and \(M\) get large. Thus, I don’t see your reason to limit the models.