

# A comparison of stratified simple random sampling and sampling with probability proportional to size

Edgar Bueno  
Dan Hedlin  
Per Gösta Andersson

## Abstract

This paper compares stratified and  $\pi$ ps sampling. We are interested in the estimation of the total of a right-skewed study variable in a finite population. The values of an auxiliary variable are known from the planning stage for all the elements in the population. We assume that ideal survey conditions hold.

When planning the sampling strategy, the statistician is often looking for “the most” efficient strategy. However, it has been shown that there is neither a uniformly best estimator nor a best design. Nevertheless, it is often possible to identify a set of strategies that can be considered as candidates. The task is to choose one among this set.

The objective is to use the auxiliary variable for obtaining an efficient strategy, with efficiency understood in terms of design variance. The strategy that couples  $\pi$ ps sampling with the regression estimator has been sometimes called optimal. This optimality, however, relies on a super-population model which might not exactly hold in practice. Wright (1983) proposed model-based stratification, which, making use of the same model, couples stratified sampling with the regression estimator.

The aim of this paper is to compare these two strategies and answer the following questions: How much inferior is model-based stratification compared to the optimal strategy when the model is correctly specified? Which strategy should be preferred when the model is misspecified?

There are articles focused on a particular concrete situation. Based on a simulation study we aim for more general results.

The results suggest that model-based stratification can be more efficient than the so-called optimal strategy even when the model is correctly specified.

## References

Wright, R. L. (1983). Finite population sampling with multivariate auxiliary information. *Journal of the American Statistical Association*. **78**, 879-884.