Contents

1

Introduction

Cost-effective sampling methods are of major concern in statistics, especially when the measurement of the characteristic of interest is costly and/or time-consuming. In the early 1950's, in seeking to effectively estimate the yield of pasture in Australia, McIntyre [96] proposed a sampling method which later became known as ranked set sampling (RSS). The notion of RSS provides an effective way to achieve observational economy under certain particular conditions. Although the method remained dormant for a long time, its value was rediscovered in the last 20 years or so because of its cost-effective nature. There is a surge of research interest of statisticians on RSS in recent years. There have been many new developments from the original idea of McIntyre, which made the method applicable in a much wider range of fields than originally intended. The theoretical foundation of RSS has been laid. Many variations and ramifications of the original idea have been proposed and studied. More and more applications of RSS have been cited in the literature. It is the purpose of this monograph to give a systematic account of the theory and application of RSS. In this introduction, we give a brief discussion on the notion of RSS and its applicability, a historical note on the development of RSS since it was proposed by McIntyre and, finally, an outline of the contents of the monograph.

1.1 What is ranked set sampling?

The basic premise for RSS is an infinite population under study and the assumption that a set of sampling units drawn from the population can be ranked by certain means rather cheaply *without* the actual measurement of the variable of interest which is costly and/or time-consuming. This assumption may look rather restrictive at first sight, but it turns out that there are plenty of situations in practice where this is satisfied. We shall give some examples at the end of this section.

Cycle 1

$$\begin{aligned} \boldsymbol{X}_{(1)11} \leq X_{(2)11} \leq X_{(3)11} &\Rightarrow \boldsymbol{X}_{(1)1} \\ X_{(1)21} \leq \boldsymbol{X}_{(2)21} \leq X_{(3)21} &\Rightarrow \boldsymbol{X}_{(2)1} \\ X_{(1)31} \leq X_{(2)31} \leq \boldsymbol{X}_{(3)31} &\Rightarrow \boldsymbol{X}_{(3)1} \end{aligned}$$

Cycle 2

$$\begin{aligned} \boldsymbol{X}_{(1)12} \leq X_{(2)12} \leq X_{(3)12} &\Rightarrow \boldsymbol{X}_{(1)2} \\ X_{(1)22} \leq \boldsymbol{X}_{(2)22} \leq X_{(3)22} &\Rightarrow \boldsymbol{X}_{(2)2} \\ X_{(1)32} \leq X_{(2)32} \leq \boldsymbol{X}_{(3)32} &\Rightarrow \boldsymbol{X}_{(3)2} \end{aligned}$$

$$\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots$$

Cycle m

$$\begin{aligned} \boldsymbol{X}_{(1)1m} \leq X_{(2)1m} \leq X_{(3)1m} &\Rightarrow \boldsymbol{X}_{(1)m} \\ X_{(1)2m} \leq \boldsymbol{X}_{(2)2m} \leq X_{(3)2m} &\Rightarrow \boldsymbol{X}_{(2)m} \\ X_{(1)3m} \leq X_{(2)3m} \leq \boldsymbol{X}_{(3)3m} &\Rightarrow \boldsymbol{X}_{(3)m} \end{aligned}$$

Fig. 1.1. Demonstration of the ranked set sampling procedure

The original form of RSS conceived by McIntyre can be described as follows. First, a simple random sample of size k is drawn from the population and the k sampling units are ranked with respect to the variable of interest, say X, by judgment *without* actual measurement. Then the unit with rank 1 is identified and taken for the measurement of X. The remaining units of the sample are discarded. Next, another simple random sample of size k is drawn and the units of the sample are ranked by judgment, the unit with rank 2 is taken for the measurement of X and the remaining units are discarded. This process is continued until a simple random sample of size k is taken and ranked and the unit with rank k is taken for the measurement of X. This whole process is referred to as a cycle. The cycle then repeats m times and yields a ranked set sample of size $N = mk$. For $k = 3$, the sampling procedure is illustrated in Figure 1.1.

The essence of RSS is conceptually similar to the classical stratified sampling. RSS can be considered as post-stratifying the sampling units according to their ranks in a sample. Although the mechanism is different from the stratified sampling, the effect is the same in that the population is divided into sub-populations such that the units within each sub-population are as homogeneous as possible. In fact, we can consider any mechanism, not necessarily ranking the units according to their X values, which can post-stratify the sampling units in such a way that it does not result in a random permutation of the units. The mechanism will then have similar effect to the ranking mechanism considered above. In the next chapter, we will discuss several such mechanisms. Therefore, we can extend the notion of RSS to a more general version.