**Framework for Sampling**

**Simple Random Sampling**

There are N units in the population or universe U, U = {1, 2, 3, …, N}.

Ex/ U = {1, 2, 3, 4}.

A subset, S, is chosen consisting of n units of U.

 Ex/ ***S***1 = {1, 2} ***S***3 = {1, 4} ***S***5 = {2, 4}

***S***2 = {1, 3} ***S***4 = {2, 3} ***S***6 = {3, 4}

Each possible sample ***S*** has known probability P(***S***) of being chosen, and the probabilities of possible samples sum to 1.

**Stratified Sampling**

Divide ***N*** sampling units into ***H*** strata with ***N****h* sampling units in the ***h***th stratum, where

***N***1 + ***N***2 + … + ***N***H = ***N***

We take SRSs from each stratum so that *nh* observations are randomly selected from the population in stratum ***h***.

**Cluster Sampling**

Primary Sampling Units (PSUs) or clusters are sampled first, then

Secondary Sampling Units (SSUs), often the elements in the population, are selected.

 *U* the universe, population of *N* PSUs

 *S* the sample of PSUs from the population

 *Si* the sample of SSUs from the ***i***th PSU

 *yij* measurement for the ***j***th element in the ***i***th PSU

 *N* number of PSUs in the population

 *n* number of sampled PSUs

 *Mi* number of elements in cluster ***i***

 *mi* number of elements sampled in ***i***

If the two levels of sampling are done randomly, we have two sources of sampling error that can be added.