Types of Probability Samples

- Simple Random
- Systematic Random
- Stratified Random
- Random Cluster
- Stratified Cluster
- Complex Multi-stage Random (various kinds)
Simple Random Sampling

- Each element in the population has an equal probability of selection AND each combination of elements has an equal probability of selection
- Names drawn out of a hat
- Random numbers to select elements from an ordered list
Stratified Random Sampling-1

- Divide population into groups that differ in important ways
- Basis for grouping must be known before sampling
- Select random sample from within each group
Stratified Random Sampling-2

- For a given sample size, reduces error compared to simple random sampling IF the groups are different from each other
- Tradeoff between the cost of doing the stratification and smaller sample size needed for same error
- Probabilities of selection may be different for different groups, as long as they are known
- Oversampling small groups improves inter-group comparisons
Systematic Random Sampling-1

• Each element has an equal probability of selection, but combinations of elements have different probabilities.

• Population size N, desired sample size n, sampling interval k=N/n.

• Randomly select a number j between 1 and k, sample element j and then every \( k^{th} \) element thereafter, \( j+k, j+2k, \) etc.

• Example: \( N=64, \ n=8, \ k=64/8=8. \) Random \( j=3. \)
Systematic Random Sampling-2

• Has same error rate as simple random sample if the list is in random or haphazard order
• Provides the benefits of implicit stratification if the list is grouped
Systematic Random Sampling-3

- Runs the risk of error if periodicity in the list matches the sampling interval
- This is rare.
- In this example, every 4th element is red, and red never gets sampled. If j had been 4 or 8, ONLY reds would be sampled.
Random Cluster Sampling - 1

- Done correctly, this is a form of random sampling
- Population is divided into groups, usually geographic or organizational
- Some of the groups are randomly chosen
- In pure cluster sampling, whole cluster is sampled.
- In simple multistage cluster, there is random sampling within each randomly chosen cluster
Random Cluster Sampling - 2

• Population is divided into groups
• Some of the groups are randomly selected
• For given sample size, a cluster sample has more error than a simple random sample
• Cost savings of clustering may permit larger sample
• Error is smaller if the clusters are similar to each other
Random Cluster Sampling - 3

- Cluster sampling has very high error if the clusters are different from each other
- Cluster sampling is NOT desirable if the clusters are different
- It IS random sampling: you randomly choose the clusters
- But you will tend to omit some kinds of subjects
**Stratification vs. Clustering**

**Stratification**
- Divide population into groups different from each other: sexes, races, ages
- Sample randomly from each group
- Less error compared to simple random
- More expensive to obtain stratification information before sampling

**Clustering**
- Divide population into comparable groups: schools, cities
- Randomly sample some of the groups
- More error compared to simple random
- Reduces costs to sample only some areas or organizations
Stratified Cluster Sampling

- Reduce the error in cluster sampling by creating strata of clusters
- Sample one cluster from each stratum
- The cost-savings of clustering with the error reduction of stratification
Stratified Cluster Sampling

• Combines elements of stratification and clustering
• First you define the clusters
• Then you group the clusters into strata of clusters, putting similar clusters together in a stratum
• Then you randomly pick one (or more) cluster from each of the strata of clusters
• Then you sample the subjects within the sampled clusters (either all the subjects, or a simple random sample of them)