Sample size and Sampling strategy

Dr. Abdul Sattar
Programme Officer, Assessment & Analysis
Do you agree that Sample should be a certain proportion of population???
Formula for sample

\[ N = (p) \times (1 - p) \frac{Z^2}{C^2} \]

Z is confidence level (Usually it is assumed 95% confidence level, then value of Z is 1.96)

C is confidence interval (tolerance for margin of error), usually is assumed 5%, or 0.05.

P is response distribution or variance of the indicator to be measured
Formula for sample

For standard assumptions the value of $Z^2/C^2$ is 1536.64

<table>
<thead>
<tr>
<th>Response Distribution value of P</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>138</td>
</tr>
<tr>
<td>0.2</td>
<td>246</td>
</tr>
<tr>
<td>0.3</td>
<td>323</td>
</tr>
<tr>
<td>0.4</td>
<td>369</td>
</tr>
<tr>
<td>0.5</td>
<td>384</td>
</tr>
<tr>
<td>0.6</td>
<td>369</td>
</tr>
<tr>
<td>0.7</td>
<td>323</td>
</tr>
<tr>
<td>0.8</td>
<td>246</td>
</tr>
<tr>
<td>0.9</td>
<td>138</td>
</tr>
</tbody>
</table>
Correction for Finite sample

\[ \text{New } N = \frac{N}{1 + \frac{N - 1}{\text{Population}}} \]

If Original Sample is 384 then population corrected sample is

<table>
<thead>
<tr>
<th>Population</th>
<th>New Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>80</td>
</tr>
<tr>
<td>250</td>
<td>152</td>
</tr>
<tr>
<td>500</td>
<td>217</td>
</tr>
<tr>
<td>1,000</td>
<td>278</td>
</tr>
<tr>
<td>5,000</td>
<td>357</td>
</tr>
<tr>
<td>10,000</td>
<td>370</td>
</tr>
<tr>
<td>50,000</td>
<td>381</td>
</tr>
<tr>
<td>100,000</td>
<td>383</td>
</tr>
<tr>
<td>500,000</td>
<td>384</td>
</tr>
<tr>
<td>1,000,000</td>
<td>384</td>
</tr>
<tr>
<td>2,000,000</td>
<td>384</td>
</tr>
</tbody>
</table>
Study Design

Analytical or Descriptive Depends on Comparing Groups

Yes = Analytical study
    – Case Control
    – Cross sectional study

No = Descriptive study
Buzz Words of sampling

• **Sampling Frame** (Comprehensive list of all units of population or universe, through which sample is drawn)

• **Sample** (is the subset of population or universe)

• **Sample size** (number of elements selected in a sample)

• **Sampling Unit** (are the elements into which a population is divided i.e. Villages, Health facilities, Households, individuals, etc)

• **Sampling list** (list of all element of the sample on which selected element are marked)

• **Element, or subjects** (the person/household/village/Facility, Bank)
Advantages

• Reduced cost
• Greater speed
• Greater scope
• Greater accuracy sometimes gives better results than census
Sampling

Two important things:

• Representativeness and Significance
Steps in sampling

1. Sampling strategy always and invariably follows research design
2. Sample size determination
3. Selection methods
Types of Sampling

1. Probability sampling
   - Simple random sample SRS
   - Systematic/ PPS
   - Stratified
   - Cluster
   - Multi-stage

2. Non-probability sampling
   - Purposive
   - Quota
   - Judgment/ Convenience
   - Snow ball
Simple Random Sample

- Each population element has an *equal chance* of being selected
- Selecting 1 subject does not affect selecting others
- May use random number table, lottery, Web-based applications (numerous)
Stratified Sampling

Divide population into subgroups
Mutually exclusive
Exhaustive
At least 1 common characteristic of interest
Select SRS from subgroups
Cluster Sampling

Divide population into clusters to reduce travel cost
   If managers are elements then companies are clusters
Select clusters PPS
Survey all or a random sample of elements in cluster
2. Non-probability Samples

- **Purposive**
- **Judgment**
  - Use experience to select sample
  - Example: Test markets
- **Quota**
  - Similar to stratified sampling except no random sampling
- **Convenience** (Chunk)
  - Use elements most available
- **Snowball**
  - When one cannot get list of the population who share same characteristics.
Multi-Stage Sampling

The sampling method used in national surveys is often multistage (e.g., household within cluster, themselves taken within strata)
Multi-stage Sampling

When population is Heterogeneous (different type of people) or complex population (country population), it is necessary to go for multi-stage sampling.

- **1st Stage – Stratification** (Provinces, Rural/Urban/major urban cities, income levels and other well defined social attributes)
- **2nd Stage – Clusters (Usually PPS)** Selecting clusters from each Strata (Primary Sampling Units (PSU) in rural areas and Enumeration Blocks (EB) in urban areas - divided into approx 250 households each)
- **3rd Stage – (Usually SRS/Systematics)** Households/ Individuals (Secondary Sampling Units (SSU) selected from each clusters)
Sample Size

Pakistan

Urban
  - Male
  - Female

Rural
  - Male
  - Female

2 x 384 = 768
2 x 2 x 384 = 1536

= 384
Appropriate Sample Size

Pakistan

- Punjab
  - Urban
  - Rural
- Sindh
  - Urban
  - Rural
- NWFP
  - Urban
  - Rural
- Balochistan
  - Urban
  - Rural

384 x 4 = 1536
384 x 4 x 2 = 3072
384 x 4 x 2 x 2 = 6144

M = Male
F = Female

= 384

WFP World Food Programme
VAM-Unit
## Sample Size for national Surveys

### District Representative Survey

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Districts Including ICT</td>
<td>117</td>
</tr>
<tr>
<td>Each district is divided among U-R</td>
<td>2</td>
</tr>
<tr>
<td>Total Stratums</td>
<td>234</td>
</tr>
<tr>
<td>Sample from each stratum</td>
<td>384</td>
</tr>
<tr>
<td><strong>Total Sample</strong></td>
<td><strong>89,856</strong></td>
</tr>
</tbody>
</table>

### Province Representative Survey

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major cities</td>
<td>14</td>
</tr>
<tr>
<td>Provinces</td>
<td>4</td>
</tr>
<tr>
<td>ICT</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>19</strong></td>
</tr>
<tr>
<td>U-R is taken as separate stratum</td>
<td>2</td>
</tr>
<tr>
<td><strong>Total Stratums</strong></td>
<td><strong>38</strong></td>
</tr>
<tr>
<td>Sample from each stratum</td>
<td>384</td>
</tr>
<tr>
<td><strong>Total Sample</strong></td>
<td><strong>14,592</strong></td>
</tr>
</tbody>
</table>
An example of SRS