# Sampling Methods 

Dr Wan Nor Arifin

## Outlines

- Probability Sampling
- Non-probability Sampling


## Sampling

- What is sample?
- What is sampling?


## Sample \& Sampling

Sampling

Population

## Sample

## Sampling

- Divided (Trochim, 2006) into:
- Probability
- Non-probability


## Probability Sampling

- Every subject has a chance to be selected.
- Random selection method.
- 5 Methods:

1. Simple
2. Stratified.
3. Systematic.
4. Cluster.
5. Multistage.

## 1. Simple Random

## Population N=1000



Sample n=30

7, 17, 18, 48, 71, 109, 141, 165, 214, 219, 277, 279, 288, 440, 475, 483, 576, 660, 735, 763, 764, 780, 863, 883, 888, 914, 917, 927, 993, 996

30 random number between $1-1000$ generated using computer e.g. http://www.randomizer.org/

## 2. Stratified random



## 3. Systematic Random

| 1. Interval $=\mathbf{N} / \mathbf{n}=$ | Population $\mathbf{N}=\mathbf{1 0 0}$ |  |
| :---: | :---: | :---: |
| $\mathbf{1 0 0 / 2 0}=\mathbf{5}$ | $1,2, \underline{3}, 4,5,6,7, \underline{8}, 9,10,11$, |  |
|  | $12, \underline{13}, 14,15,16,17, \underline{18}, 19,20$, | Sample $\mathbf{n}=\mathbf{2 0}$ |
| 2. Starting point $=$ | $21,22, \underline{23}, 24,25,26,27, \underline{28}, 29$, |  |
| Random number | $30,31,32, \underline{33}, 34,35,36,37, \underline{38}$, | $3,8,13,18,23,28,33$, |
| between $\mathbf{1}-\mathbf{5}$, e.g. $\mathbf{3}$ | $39,40,41,42, \underline{43}, 44,45,46,47$, | $38,43,48,53,58,63,68$, |
|  | $\underline{48}, 49,50,51,52, \underline{53}, 54,55,56$, | $73,78,83,88,93,98$ |
| 3. Then every | $57, \underline{58}, 59,60,61,62, \underline{63}, 64,65$, |  |
| interval of $\mathbf{5}$ | $66,67, \underline{68}, 69,70,71,72, \underline{73}, 74$, |  |
|  | $75,76,77, \underline{78}, 79,80,81,82, \underline{83}$, |  |
|  | $84,85,86,87, \underline{88}, 89,90,91,92$, |  |
|  | $\underline{93}, 94,95,96,97, \underline{98}, 99,100$ |  |

## 4. Cluster Sampling

- Cluster = Group of people
- Sampling Unit = Cluster e.g. House, Class, Ward etc. $\rightarrow$ Clusters to be sampled.
- Have to inflate $n$ to adjust for cluster effect (Naing, 2011)


## Inflated $\mathbf{n}=[1+($ cluster size -1$) r] \times n$

- $r$ is correlation between subjects in a cluster $\rightarrow$ unknown, can assume $\mathrm{r}=0.5$


## 4. Cluster Sampling

Population N=300
$1,2,3, \ldots, 300$

1. 50 houses in area
2. On average, 6 persons/house = cluster size

Sample n=30

$$
\begin{aligned}
& \text { Inflated } \mathrm{n}= \\
& {[1+(6-1) 0.5] \times 30} \\
& =105
\end{aligned}
$$

n of house to sample = 105/6 = 17.5 ~ 18 houses

House, N=50

$$
\begin{array}{cc}
\text { House, } \mathrm{N}=50 & \text { House, } \mathrm{n}=18 \\
\begin{array}{c}
1,2,3,4,5,6,7,8,9,10,11, \\
12,13,14,15,16,17,18,19,
\end{array} & 2,7,11,12,15,18, \\
20,21,22,23,24,25,26,27, & 20,22,25,26,30, \\
28,29,30,31,32,33,34,35, & 31,32,33,36,38, \\
36,37,38,39,40,41,42,43, & 43,44 \\
44,45,46,47,48,49,50 &
\end{array}
$$

## 5. Multistage

- Any combination of previous 4 methods.


## Non-probability Sampling

- Random selection method not used.
- Selection based on preset criteria set by researcher.
- Could be biased, not representative of population.


## Non-probability Sampling

- Among the methods:
- Convenient:
- Choose those easily available/sampled
- e.g. my friends, friends of my friends, relatives, room mates, etc.
- Purposive:
- Choose those fulfilling criteria.
- e.g. only those who come to clinic on Monday, handsome/cute persons only, etc.


## References

1.Arifin, W. N. (2012). Random sampling and allocation using SPSS. Education in Medicine Journal 4(1), 129-143.
2.Trochim, W. M. K. (2006). Research methods knowledge base. Retrieved March 27, 2012, from http://www.socialresearchmethods.net.
3.Naing, N. N. (2011). A practical guide on determination of sample size in health sciences research. Kelantan: Pustaka Aman Press.

