

Sampling

Why sample?

Impractical to measure the whole population –

- measure the height of all Malaysians
- count the orangutan in the whole of Batang Ai NP
- get biodata for all visitors to Mulu NP.

Inference

- We want to use the sample results to **infer** something about the whole population
- Sample must **represent** the whole population...
- ...problem of **bias**.

Examples of bias:



- Visitor surveys: interviewing people who look **happy**.

Examples of bias:

- Mammal trapping: bait attracts **hungry** animals.



Examples of bias:

- Forest surveys: plots and transects are **near rivers or roads**.



Likelihood

- Essential for Bayesian and frequentist methods
- Uses the laws of probability
- We need to know probability that each **unit** will be included in the sample.

Simple random sampling:

- all sample **units** have equal chance of being included
- avoids bias
- precision easily estimated.

What is the sample unit?

- the height of all Malaysians
 - a Malaysian person (adult? citizen? male?)
- number of orangutan in Batang Ai
 - number of ou in 1 km² plot
- biodata for all visitors to Mulu
 - a visitor arriving at Mulu.

Population vs sampling frame

- Population (or universe)
 - = all the units we are interested in
- Sampling frame
 - = all the units we can practically include in our sample
- need to note the difference.

Simple random sampling:

- all sample units have equal chance of being included
- ‘random’ \neq convenient
- ‘random’ \neq haphazard

- Stop here and do example of haphazard selection of plots on the whiteboard.
- Follow up later with demo of using many-sided die to select plots.

Simple random sampling:

- all sample units have equal chance of being included
- 'random' \neq convenient
- 'random' \neq haphazard
- a proper random sampling method must be used.

Sampling methods

Simple random sampling Random numbers



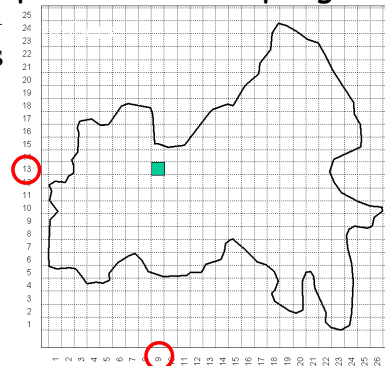
3

Simple random sampling

Random numbers

9

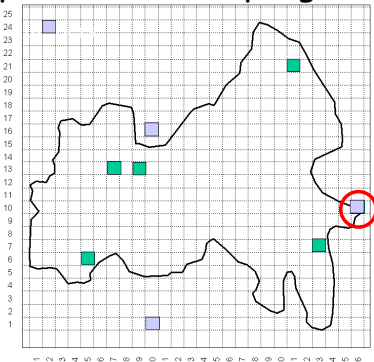
13



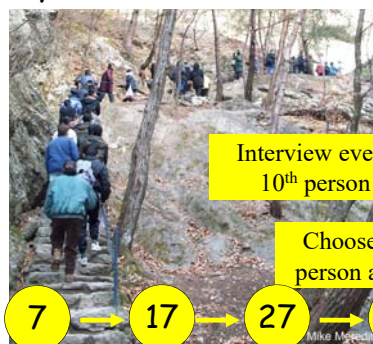
Simple random sampling

Random numbers

9, 13
10, 16
23, 7
10, 1
7, 13
2, 24
22, 21
26, 10
5, 6



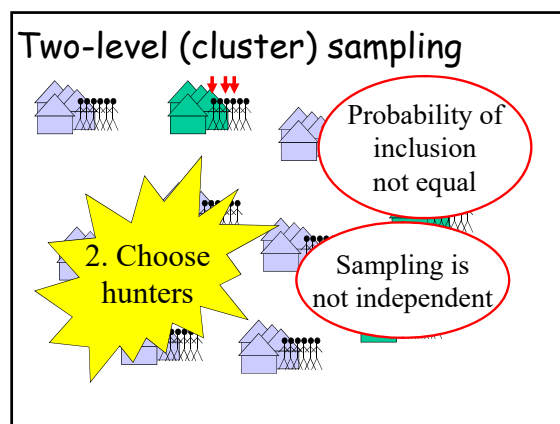
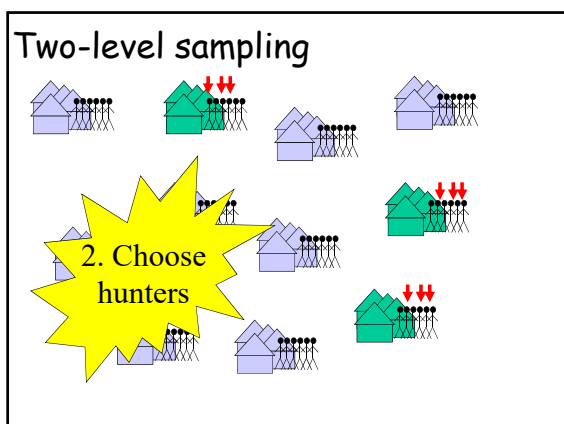
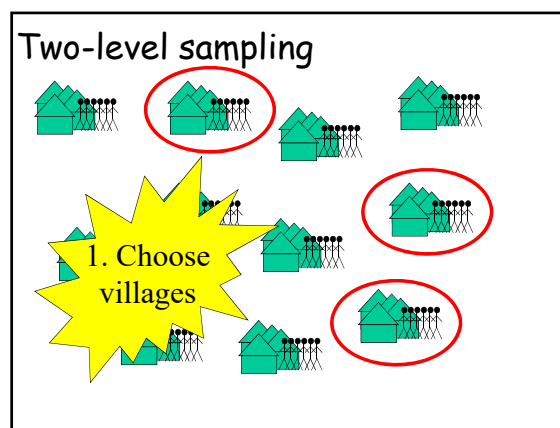
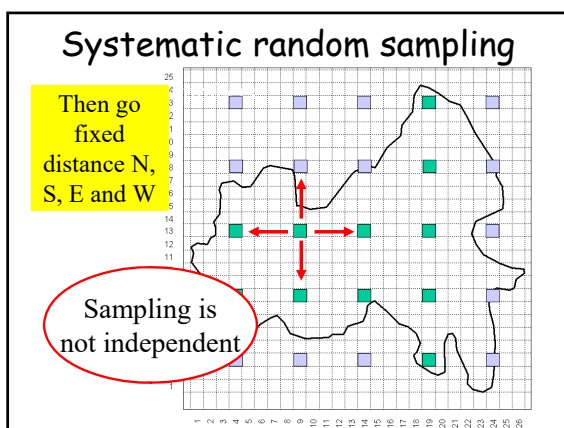
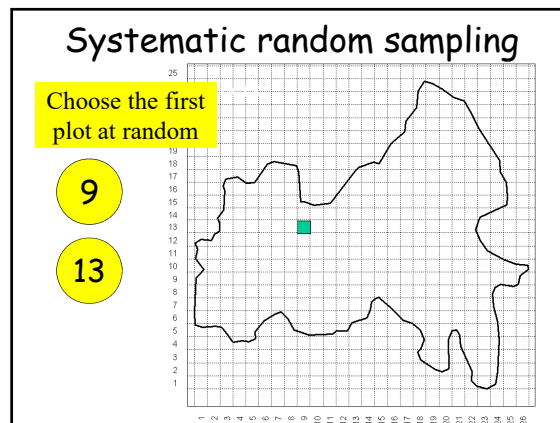
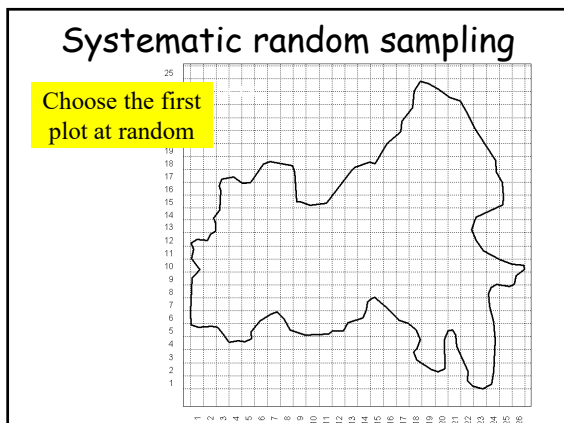
Systematic random sampling

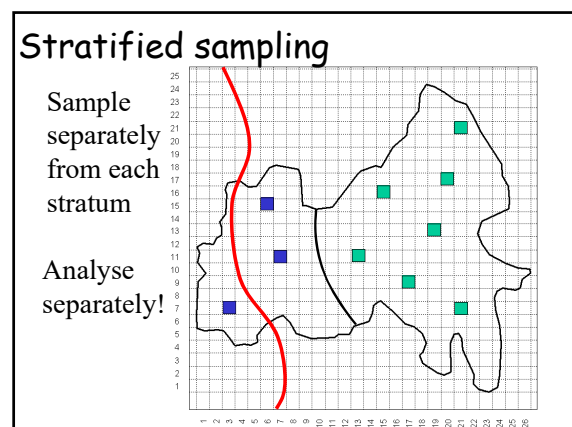
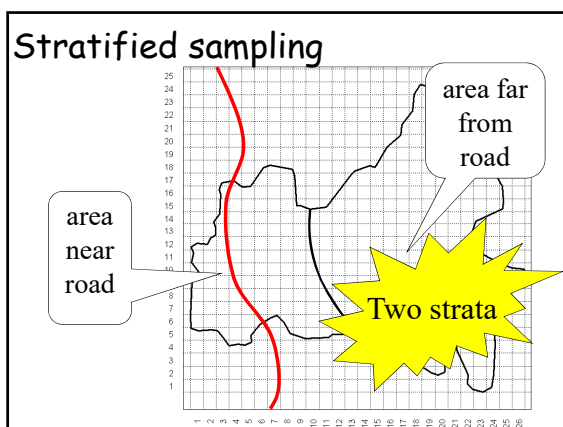
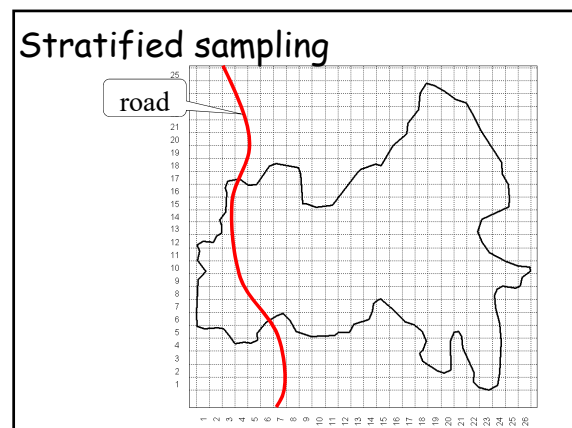
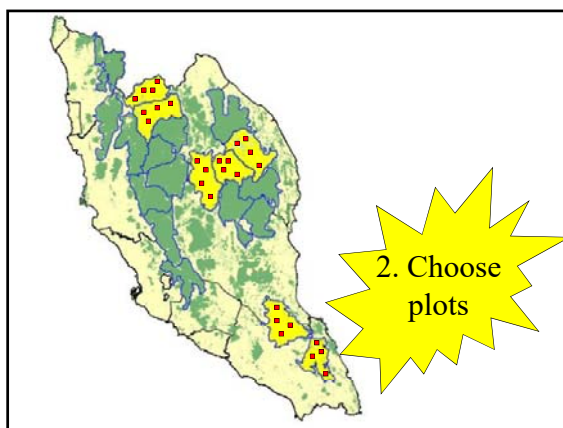
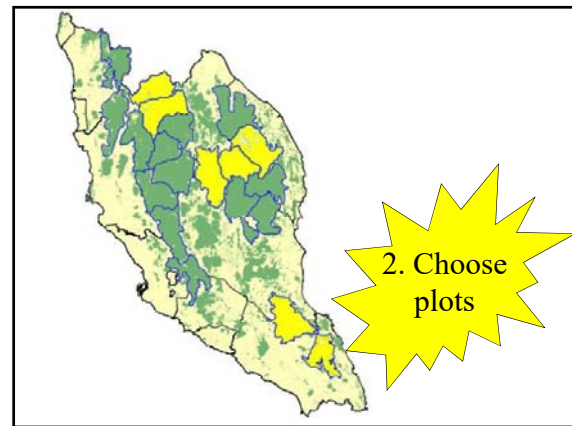
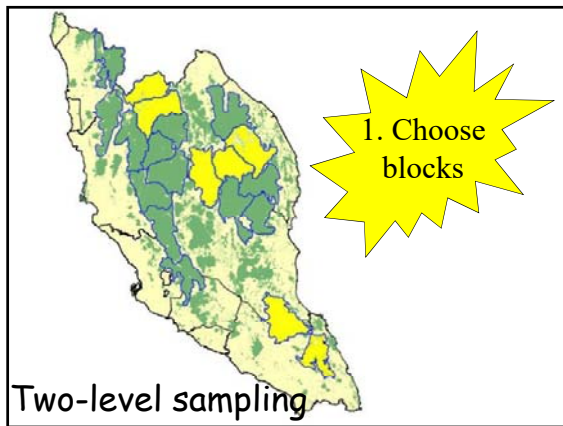


Interview every
10th person

Choose the first
person at random

7 → 17 → 27 → 37 →





Main points

- Purpose of sample is to **infer** information about whole population
- Biased samples → wrong inference
- Convenience sample → biased!
- Random samples avoid bias
- Simple vs systematic vs cluster
- Stratified sampling.

For more information:

