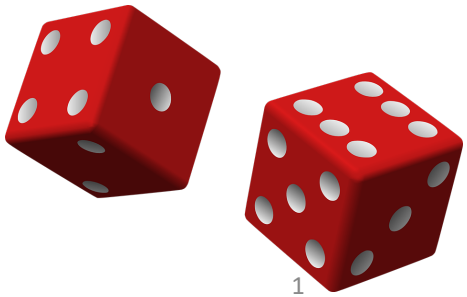
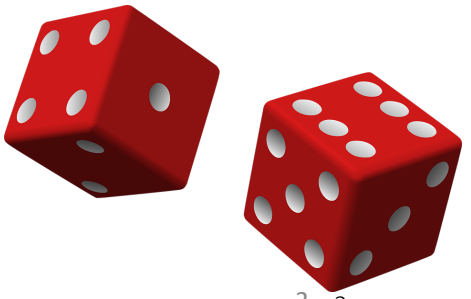
Sampling People, Records, & Networks

Jim Lepkowski, PhD
Professor & Research Professor *Emeritus*
Institute for Social Research, University of Michigan
Research Professor,
Joint Program in Survey Methodology, University of Maryland

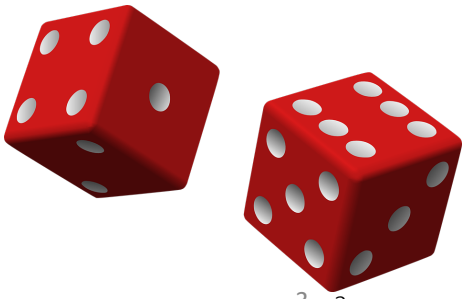


Unit 4

- 1 Forming groups
- 2 Sampling variance
- 3 More on grouping
- 4 Allocate sample
- 5 Other allocations
- 6 Weights



- Using multiple variables
- Advice on forming strata
- Multi-purpose design
- Domains



- Using multiple variables
- Advice on forming strata
- Multi-purpose design
- Domains

- We can also use more than one variable in the stratification
- For example, we can use, in addition to rank, sex:

- **Using multiple variables**
- Advice on forming strata
- Multi-purpose design
- Domains

h	Stratum	N_h	W_h
1	Female, Assistant	40	0.1000
2	Female, Associate	25	0.0625
3	Female, Full	20	0.0500
4	Male, Assistant	75	0.1875
5	Male, Associate	50	0.1250
6	Male, Full	190	0.4750
Total		400	1.0000

- Using multiple variables
- Advice on forming strata
- Multi-purpose design
- Domains

- The sample size, or allocation, needs to be determined for each stratum again
- For example, if we again select 20% of the elements in the population the sample, …

- **Using multiple variables**
- Advice on forming strata
- Multi-purpose design
- Domains

h	Stratum	N_h	W_h	n_h
1	Female, Assistant	40	0.1000	8
2	Female, Associate	25	0.0625	5
3	Female, Full	20	0.0500	4
4	Male, Assistant	75	0.1875	15
5	Male, Associate	50	0.1250	10
6	Male, Full	190	0.4750	38
Total		400	1.0000	80

- Using multiple variables
- **Advice on forming strata**
- Multi-purpose design
- Domains

- **But in general then, how should we form strata?**
- **The best advice is to make the strata internally homogeneous**

- Using multiple variables
- **Advice on forming strata**
- Multi-purpose design
- Domains

- **But in general then, how should we form strata?**
- **The best advice is to make the strata internally homogeneous**
- **That means that the strata should differ as much as possible from each other – have big differences between the means of the strata**

- Using multiple variables
- **Advice on forming strata**
- Multi-purpose design
- Domains

- **But in general then, how should we form strata?**
- **The best advice is to make the strata internally homogeneous**
- **That means that the strata should differ as much as possible from each other – have big differences between the means of the strata**
- **Another way to say this is to find background or auxiliary variables that explain as much of the variance of the variable on interest as possible**

- Using multiple variables
- **Advice on forming strata**
- Multi-purpose design
- Domains

- **Availability of data**
  - **Census**
  - **Administrative reports**
  - **Other surveys**

- Using multiple variables
- Advice on forming strata
- **Multi-purpose design**
- Domains

- **Multipurpose surveys**
  - **Survey of households in Qatar**
  - **Fixed assets, buildings, use of expatriate labor, expenditures, income, health, health care use, psychological well-being, social integration**

- Using multiple variables
- Advice on forming strata
- Multi-purpose design
- **Domains**

- **Domains of study**
  - **Subpopulations for which separate estimates are required**
  - **Geographic subdivisions such as provinces, districts, subdistricts**
  - **Socio-demographic characteristics, such as age groups, occupation, income, education**

- Using multiple variables
- Advice on forming strata
- Multi-purpose design
- **Domains**

- **When one has multiple potential stratifying variables how does one choose which to use?**

- Using multiple variables
- Advice on forming strata
- Multi-purpose design
- **Domains**

- **When one has multiple potential stratifying variables how does one choose which to use?**
- **One consideration: how large are the stratum sizes?**

- Using multiple variables
- Advice on forming strata
- Multi-purpose design
- Domains

- When one has multiple potential stratifying variables how does one choose which to use?
- One consideration: how large are the stratum sizes?
- If we are able to use only one, or a subset of variables, choose those that are going to have bigger differences in outcomes across categories.

Are there bigger differences in income between categories of rank or between categories of sex?

Unit 4

- 1 Forming groups
- 2 Sampling variance
- 3 More on grouping
- 4 Allocate sample
- 5 Other allocations
- 6 Weights

- Unit 1: Sampling as a research tool
- Unit 2: Mere randomization
- Unit 3: Saving money
- Unit 4: Being more efficient
  - Forming groups
  - Sampling variance
  - More on grouping
  - **Allocate sample**
  - Other allocations
  - Weights to combine across strata
- Unit 5: Simplifying sampling
- Unit 6: Some extensions & applications