STAT101 Worksheet: Sampling **4A**

Simple Random Sampling (SRS): Starbucks here we come!

A statistics class has 36 members. Obtain a SRS of 8 class members who will be invited to Starbucks for a free beverage of their choice. Select the sample from the following class members.

Alex	Dana	Tatiana	Ashley D.	Edosa	Victoria F.
Matt	Caitlin	Hannah	Kayla	Elyise	Danielle H1
Danielle H2	Jenna	Chris	Jason	Sabrinna	Nathaniel
Tiffini	Kyle	Stephanie N.	Alicia	Adam	Teddi-Jo
Corey	Stephanie S.	Ashley S.	Victoria S.	Deanna	Molly
Stefan	Mike	Keri	Karin	Jenna	Mary

Answers (sample) vary depending on method used (table of random numbers; computer)

- Assign number to each person (if using table of random numbers)
  - here 36 people, so each person needs a two-digit number (01 to 36).
- Using Table of Random Numbers
  - Randomly select starting point and direction (up/down; L/R; etc.); if come across same number - skip it.

* Probability occurs when you select the starting point (randomly) in table of random numbers.

Stratified Sampling: Accounting Practices

Accountants often use stratified random sampling during audits to verify a company's records of such things as accounts receivable. One company reports 5000 accounts receivable of which 200 are in accounts over $100,000, 1000 accounts are between $10,000 and $100,000 and 3800 accounts are under $10,000. The auditor decides to review 200 accounts using a stratified sample with proportional allocation. Determine the number of accounts that will be audited within each stratum.

Population Table
$N = 5000$

size	Accounts	rf
>$100k	200	.04
$10k - $100k	1000	.20
<$10k	3800	.76
Total:	5000	1.00

calculate rf; put in →

take pop. rf and mult. times sample size

rf: → .76 × 200 = 152 →

Sample Table
$n = 200$

size	Accounts f	rf
>$100k	8	.04
$10k - $100k	40	.20
<$10k	152	.76
Total	200	1.00

← calculate f given specified rf

- round partial values
- be sure you meet sample size
- if freq. sums to > required sample, reduce from largest class

* Probability occurs when you randomly select units from each stratum (e.g. randomly 8 of 200 >$100k accts.)

Systematic Sampling: Coopers

MINI builds Coopers in its Oxford England plant. Five-hundred Coopers are made per day. The management wants to select 25 Coopers daily to go through a quality control inspection consisting of 144 items. Demonstrate how you would use systematic sampling to identify the 25 Coopers.

Given
$N = 500$ cars (population)
$n = 25$ cars (sample)

$\frac{N}{n} = \frac{500}{25} = 20 = m$ = size of groups

K = random seed taken from the 1 to m group (first group)
- could obtain from a table of random numbers
- HERE ASSUME K = 4

CARS
1, 2, 3, ..., 19, 20 | 21, ..., 39, 40 | 41, ..., 480 | 481, ..., 500

first group of m size → K = 4
second group → K + m = 4 + 20 = 24
25th group → K + 24m = 4 + 24(20) = 484

* Probability occurs when you randomly select K.

Cluster Sampling: Airline Satisfaction

As part of an advertising campaign Alaska Air has hired you assess the level of customer satisfaction with its new in-flight services from Seattle to Anchorage. You elect to conduct a survey of 1,540 passengers and plan to collect data for each day of the week. The company indicates that there are 10 flights per day into Anchorage, each with an average 110 passengers. Use a cluster sampling approach to obtain the sample. Why might you select this approach?

Given
10 flights/day = 70 flights/week ← These are the clusters (here groups of 110 passengers)
n = 1540 passengers
N = 7700 passengers (70 flights × 110 pass/flight)
f = 110 passengers/flight
$\frac{n}{f} = \frac{1540}{110} = 14$ flights ← Clusters needed (take 2 of 10 flights/day)

* Probability occurs when you randomly select 14 of 70 clusters. (here randomly taking 2 of 10 daily flights)

Obtaining a Probability-based Random Sample

Suppose there are 300 students using this statistics text this semester at SUNY Oneonta. The text's publisher has commissioned you to survey 90 of these students. There are 10 course sections each with 30 students and the year level breakdown of course participants is: Fr. = 155, So. = 75, Jr. = 30, Sr. = 30, Other = 10. Describe procedures for obtaining each of the samples below.

Random: Sampling frame: List of 300 students
Randomly select 90 students

Systematic: Demonstrate how you would obtain this sample by identifying the first 4 individuals in this sample.

Given
N = 300 students
n = 90
$m = \frac{N}{n} = \frac{300}{90} = 3.33$ ← ALWAYS ROUND DOWN → Why? $\frac{300}{3} = 100$ = max. n in a sample
$\frac{300}{4} = 75$ = max. n in a sample, which does not meet required sample size
K = random from 1 to n [here 1 to 3]

Convenience:
- ask friends
- ask class members
- any approach that does not give all members (units) in the population (here 300) an equal chance of being selected.

Stratified: Determine the number of individuals for the strata.

Pop.: Stat students

class	f	nf
FR	155	.517
SO	75	.250
JR	30	.100
SR	30	.100
OTH	10	.033
Total	300	1.00

nf × n →

Sample: Stat students

class	f	nf
FR	(46.5) 47	.517
SO	(22.5) 23	.250
JR	(9) 9	.100
SR	(9) 9	.100
OTH	(2.97) 3	.033
Total	90 (91)	1.00

Rounding yields n=91. Reduce by taking away from largest class (here FR)

Cluster:
Given
10 classes
30 students/class
n = 90
$\frac{n}{f} = \frac{90}{30} = 3$ classes randomly selected from 10 classes