

# Sampling Techniques

- Simple Random Sampling
- Systematic Sampling
- Convenience Sampling
- Quota Sampling
- Stratified Sampling

## Sigma Notation Warm Up!

Ex] Evaluate the following sum  $\sum_{l=2}^4 (3l-4)$

$$(3(2)-4) + (3(3)-4) + (3(4)-4) = 2 + 5 + 8 = \boxed{15}$$

Ex] Express this series using Sigma Notation: 2 + 4 + 8 + 16 + 32

$$\sum_{n=1}^5 (2^n)$$

↗ Frequency tables can help find the sum of your set by adding a product column.

Data value (x)	Frequency (f)	Product (fx)
3	1	$1 \times 3 = 3$
4	1	$1 \times 4 = 4$
5	3	$3 \times 5 = 15$
6	7	$7 \times 6 = 42$
7	15	$15 \times 7 = 105$
8	8	$8 \times 8 = 64$
9	5	$5 \times 9 = 45$
Total	$\sum f = 40$	$\sum fx = 278$

3 4 5 5 5 6 6 6 6 6  
7 7 7 7 7 7 7 7 7 7  
8 8 8 8 8 8 8 9 9 9 9

ex) Consider the data showing the IQs of 200 people

56	62	65	68	69	70	71	71	75	77	79	79	81	81	81	83	84	85	85	85
86	86	86	87	87	87	87	87	87	87	88	88	88	88	88	89	89	89	89	89
89	89	89	89	89	89	89	91	92	92	92	93	93	93	93	93	93	94	94	94
94	94	94	94	94	95	95	95	95	95	96	96	96	96	96	96	96	97	97	97
97	98	98	98	98	98	98	99	99	99	99	99	99	99	99	99	99	99	99	99
100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
100	100	100	100	100	100	101	101	101	101	101	101	101	101	101	101	102	102	102	102
102	103	103	103	103	104	104	104	104	105	106	106	107	107	107	107	107	107	107	107
107	107	107	108	108	108	109	110	110	110	112	112	113	113	113	114	114	115	115	115
117	118	119	121	121	125	128	129	129	131	134	135	136	137	140	141	143	145	148	156

a) Find the mean ( $\bar{x}$ )



## Frequency Tables Revisited

Case 1: 2-classes

Class	Frequency ( $f_i$ )	Class Median ( $M_i$ )	$(f_i)(M_i)$
$56 \leq x < 100$	100	92	9200
$100 \leq x \leq 157$	100	104.5	10,450

$$\text{Sum } \sum_{i=1}^n (f_i)(M_i) = 19,650$$

$$\text{Actual } \bar{x} = 99.74$$

$$\bar{x} \approx \frac{\sum_{i=1}^n (f_i)(M_i)}{n} = \frac{19,650}{200} = 98.25$$

$$\text{Error: } 99.74 - 98.25 = 1.49$$

### Case 2: 5-classes

Class	Frequency ( $f_i$ )	Class Median ( $M_i$ )	$(f_i)(M_i)$
Row 1-2	40	85.5	3420
Row 3-4	40	94	3760
Row 5-6	40	99.5	3980
Row 7-8	40	102	4080
Row 9-10	40	116	4640

$$\bar{X} \approx \frac{19880}{200} = 99.4$$

Error:  $99.74 - 99.4 = 0.34$

$$\text{Sum } \sum_{i=1}^5 (f_i)(M_i) = 19880$$

### Case 3: 10-classes

Class	Frequency ( $f_i$ )	Class Median ( $M_i$ )	$(f_i)(M_i)$
Row 1	20	78	(20)(78)
Row 2	20	88	(20)(88)
Row 3	20	92	(20)(92)
Row 4	20	95	(20)(95)
Row 5	20	99	(20)(99)
Row 6	20	100	(20)(100)
Row 7	20	101	(20)(101)
Row 8	20	104.5	(20)(104.5)
Row 9	20	110	(20)(110)
Row 10	20	132.5	(20)(132.5)
Sum	200	1,000	20,000

$$\bar{X} \approx \frac{20,000}{200} = 100$$

Error:  $99.74 - 100 = 0.26$

\* The more cases you use, the more accurate  
 (If you used 200 cases, then you'd be 100% accurate)

Population - the whole group from which you collect data

Sample - a small group chosen from the population  
(a subset of the population)

ex] Sean wanted to find out how many students at the Mountain Academy enjoyed math. He surveyed the students in his Applications & Interpretations class. What is the population? What is the sample?

Population : all MA students

Sample : A&I Students

ex] Steve wants to know whether the food he serves in his restaurant is within a safe range of temperatures. He selects 70 entrees and measures their temperature just before he serves them to his customers. What is the population? What is the sample?

Population : all entrees he serves in his restaurant

Sample : 70 selected entrees

## Sampling

- the selection of a subset of individuals from within a statistical population to estimate characteristics of the whole group

Non-Random/Bias : Data is selected based on the subjective judgment of the researcher

ex: researching spending habits on cars and only garage interviewing people exiting a mechanic

- Convenience Sampling - gathering a sample by selecting data based on its convenient accessibility and proximity.  
ex: people at your school, club, team
- Quota Sampling - biased version of stratified sampling. Population is segmented into mutually exclusive sub-groups, then researcher uses judgment to select data values based on predetermined proportion.  
ex: selecting a sample of 8 girls and 8 boys,  
9 blue eyed people and 9 brown eyed people

For example, the school canteen is considering introducing a new lunch menu and would like feedback from the students. The school has 250 boys and 300 girls and so the canteen manager decides to interview 25 boys and 30 girls to find their opinion of the new menu. He stands at the entrance to the canteen and interviews only 25 boys and only 30 girls who come into the canteen

Random / Unbiased : Every data value has an equal probability of being selected.

- Simple Random Sampling - Selecting a sample completely at random. (ex: picking #'s from a hat, using a random number generator)
- Systematic Sampling - Selecting a sample using fixed, periodic intervals starting from a random place  
(ex: taking every fifth entry starting at a random place)
- Stratified Sampling - Random sampling that divides a population into strata (characteristics)  
(ex: divide a population by the state they live in, income, education level, gender, etc.)

\* To determine equitable representation of strata:  $\frac{\text{Strata Frequency}}{\text{Population size}} \cdot \text{sample size}$

## Stratified Ex]

Mandy asks all the students in her school to take a memory test. The students have to remember as many objects as they can from the 20 that Mandy shows them. The results are shown in the table.

Class 7 [20 students]	16, 15, 13, 15, 12, 8, 18, 16, 12, 11, 14, 17, 16, 9, 11, 10, 17, 13, 14, 13
Class 8 [27 students]	19, 15, 16, 14, 11, 16, 18, 15, 13, 12, 10, 8, 20, 14, 17, 12, 10, 7, 19, 20, 13, 17, 16, 16, 15, 11
Class 9 [23 students]	17, 14, 15, 8, 7, 13, 15, 19, 16, 13, 11, 10, 17, 17, 20, 15, 11, 10, 7, 13, 16, 15, 15,
Class 10 [26 students]	9, 10, 10, 12, 18, 16, 17, 15, 11, 11, 14, 16, 19, 19, 11, 15, 17, 13, 13, 14, 13, 13, 9, 10, 8, 15
Class 11 [30 students]	16, 15, 15, 16, 16, 18, 11, 12, 13, 9, 10, 11, 16, 17, 15, 12, 12, 15, 15, 15, 18, 20, 16, 17, 17, 15, 14, 14, 14, 14
Class 12 [24 students]	9, 11, 16, 14, 13, 13, 18, 19, 12, 10, 11, 9, 16, 16, 18, 14, 15, 15, 16, 13, 13, 12, 18, 19

40 students' scores are sampled out of 150

(a) What percentage of the population will be sampled?

$$\frac{40}{150} = 26.6\%$$

(b) What percentage of the population is in Class 7?

$$\frac{20}{150} = 13.3\%$$

(c) Using stratified sampling, how many students should be chosen from Class 7 to ensure equity?

$$\left[ \frac{20}{150} \cdot 40 \right] = 5.33 \text{ so } 5 \text{ students}$$

How many students should be chosen from class 11?

$$\frac{30}{150} \cdot 40 = 8 \text{ students}$$

13.3% of the population is in class 7  
To ensure an equitable representation of each strata, you need class 7 to make up 13.3% of the 40 students that are chosen.  
Thus,  $(0.133)(40)$ .

Systematic Ex | Consider the list below. We want to take a 1 in 4 systematic sample.

A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T

What is the probability that A is selected? What is the probability that F is selected?

### Practice Problems

Pg 118 - 119. Q 1-3