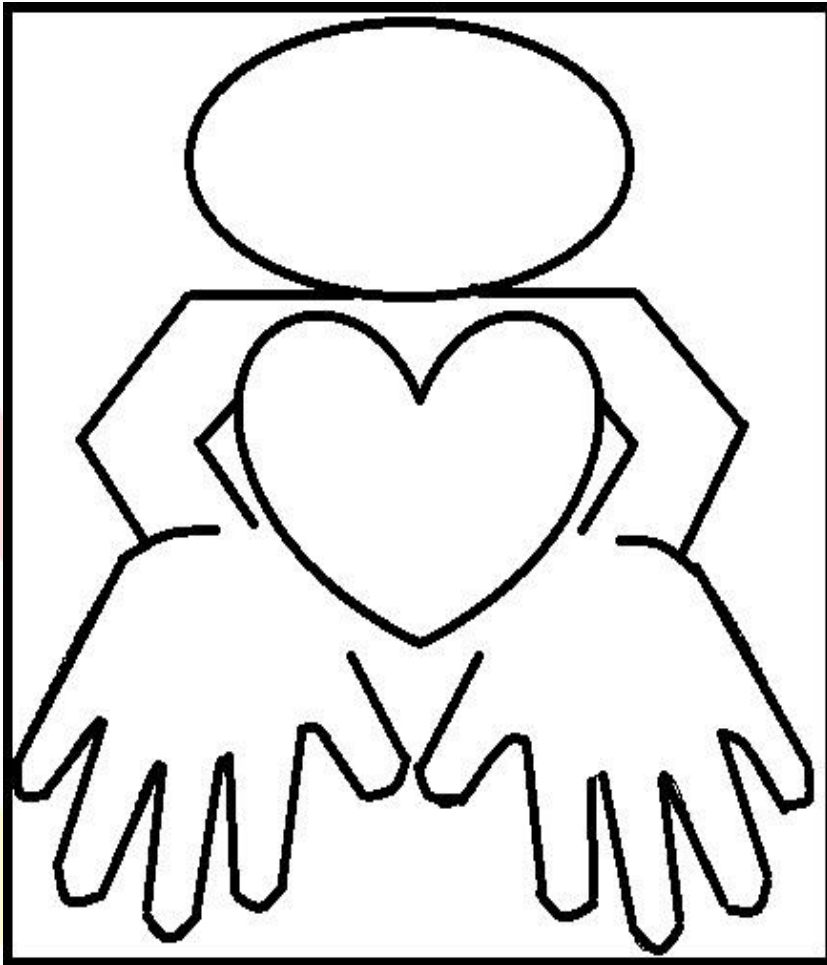


ESSENTIALS OF STATISTICS

DR. ROGER L. NUQUI

WHERE I AM?

My Think, Feel and Do **about Statistics**



Group Work

1. Print your name, on the top, left side of the *My Think, Feel and Do Profile*.
2. **On the circle**, Describe the nature of statistics.
3. **On the heart**, With your knowledge of statistics in this sense, how do you apply it in a real life situation? Illustrate your answer.
4. **On the two hands**, Why is statistics considered a science? Is interpretation possible without the analysis of the data first? Why? What is the use of the test or summary table? Is this necessary? Why?

The Nature of Statistics

Through the centuries, man has been interested in determining the number of birth rate, mortality rate, occurrence of certain events at certain periods of time, crop yield and the frequency of failures in school entrance test among many others. In these activities, he deals with the counts or numeral measures of activities, events and things, which are called statistics in a limited sense.

In counting activities, events, and things the measurement that are collected from the original information are called **raw data**. These data may be treated by statistical methods that are used to describe, to relate or associate, and to make inferences. In the method of description, the measure to apply is known as descriptive statistics. In the method to relate or associate two variables, correlational statistics is employed. And in making generalizations about the population

- In the method of description, the measure to apply is known as descriptive statistics.
- In the method to relate or associate two variables, correlational statistics is employed.
- And in making generalizations about the population from which our sample has been drawn, the measure to use is inferential statistics.

Statistics as defined scientifically

From the research point of view, statistics is a science which deals with methods in collection, gathering, presentation, analysis, and interpretation of data. Data gathering involves getting information through interviews, questionnaires, objective observations, experimentations, psychological tests and other methods.

Usually the information is translated into numerical or quantitative data. By using graphs, figures, and tables the data collected can be exhibited. This becomes part of data presentation. Data presentations in tabular form are of two types, the text or summary table which is found in the body of the research work, and the reference table which is usually found in appendices.

If your interest is on the few members of the population to represent their characteristics or traits, these members constitute a sample. The measures of the population are called parameters, while those of the sample are called estimates or statistics.

Use of Statistics

Statistics is a branch of mathematics that can be used for many purposes. Some of these are briefly described below.

1. *It can give a precise description of data.* This is a function of statistics which enables us to make accurate statements or judgments about averages, variability, and relationship. An example of this purpose is when you describe the academic performance of a group of pupils according to the computed mean, standard deviation, and correlation with another factor.

2. It can predict the behavior of individuals. In school, the grades of students can be predicted through a scholastic aptitude test. In industry, the work performance is usually predicted by an aptitude test related to that particular type of work. A teacher's performance may also be predicted by his performance in an instrument like a teacher aptitude test. Of course other sets of data or psychological tests can be used to predict the success of a teacher in his work. To measure the success of a pupil, teacher, or a worker, we may have to compute measures like the mean, standard scores, percentiles, stanines and other statistical methods.

3. It can be used to test a hypothesis. We can determine whether a variable is related or not to another variable through a test of inference such as in correlation. Other statistical measures we can apply for inferential purposes are the t-test, chi-square test, F-test, and others. It is wise to remember that your choice of the statistics to use in testing hypothesis depends upon the nature of your data. This includes the scale of measurement used such as nominal, ordinal, interval, and ratio; and its distribution whether normally distributed or not; and other considerations such as your purpose.

Grouped Frequency Distribution of Interval Data

To organize the interval-level scores into a grouped frequency distributions, we condense the separate scores into a fewer number of categories or groups. Each group contains more than one score value, called the class interval. The class interval contains the number of score values.

- Examples: Given the following scores is a psychology test, make a frequency table.

Raw Scores:

47	56	42	28	56	41	56	55	82	78	50	55
57	38	62	52	66	65	72	37	47	42	68	62
54	48	52	47	44	39	53	48	42	54	62	68
48	56	42	60								

- Step 1. *Compute the range.* In the given data, the highest score is 82 and the lowest score is 28. The range ($H - L$) is $82 - 28 = 54$
- Step 2. *Divide the range by 10 to 15 to determine the acceptable size of the interval.* Let the lowest interval begin with a number which is a multiple of the interval size. In applying this step, we should decide how large each of the interval is going to be. Generally, the most acceptable number of intervals is as few as 5 to as many as 20. In estimating the number of intervals, a trial and error fashion is suggested. If we try 2 as the size of our class interval, we reject this because it results in too many categories of intervals.

If we try 12 as the size of our class interval, we also reject this because there are only 4 categories of intervals. The number 4 is not within the acceptable range of the number of class intervals or categories of class intervals. Since we need between 5 and 20 categories of intervals in our frequency distribution, we can take 10 to 15 and divide this into the range. In this case we may divide 54 by 10, 11, 12, 13, 14, and 15. If we select 10, it results in a value of 5.4 or 5. This value could be used as our interval size.

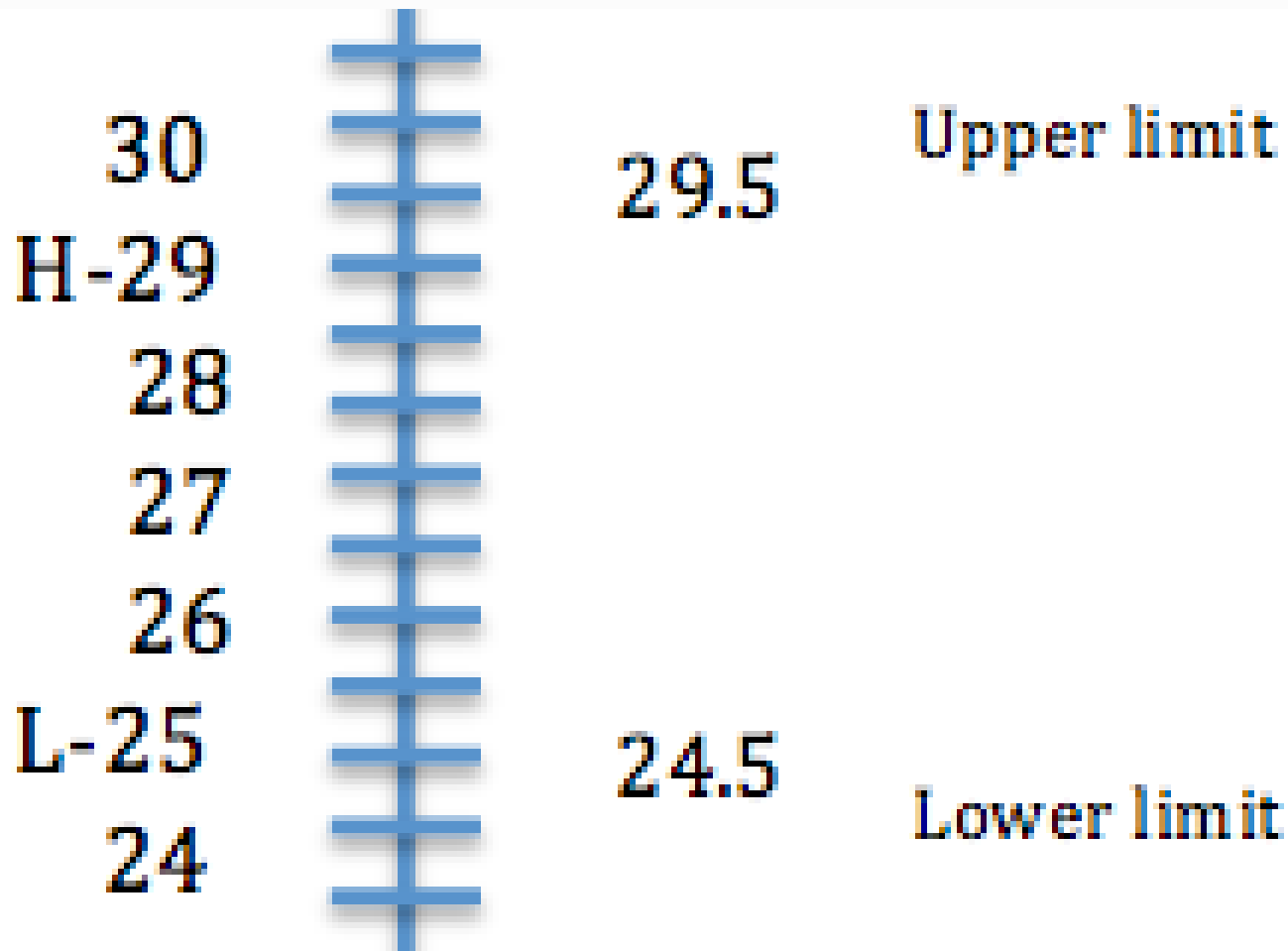
As we may have observed, most frequency distributions have odd numbers as the size of the interval. The advantage of using odd numbers is that, the midpoints of the intervals will be whole numbers.

Step 3. *Organize the class interval.* See to it that the lowest interval begins with a number that is a multiple of the interval size. Since the lowest score is 28 and the size of your interval is 5, the lowest interval would begin with 25 and end at 29. These are the interval limits. You take note that the upper and lower limits (the exact or real limits)

here are 24.5 and 29.5 respectively. These are sometimes spoken of as class boundaries.

After deciding upon the limits of the first class interval category, you determine the rest of the intervals by increasing each interval limits by 5 until you reach 80 – 84 which contains the highest score in the distribution.

Tabular Graphic Presentation of Data



- **Step 4.** *Tally each score to the category of class interval it belongs to.*

Class interval	Tally (T)
25 - 29	1
30 - 34	-
35 - 39	111
40 - 44	1111-1
45 - 49	1111-1
50 - 54	1111-1
55 - 59	1111-11
60 - 64	1111
65 - 69	1111
70 - 74	1
75 - 79	1
80 - 84	1

- Step 5. *Count the tally column and summarize it under column f.* Then add your frequency which is the total number of cases (N)

Class interval	Frequency (f)
25 - 29	1
30 - 34	0
35 - 39	3
40 - 44	6
45 - 49	6
50 - 54	6
55 - 59	7
60 - 64	4
65 - 69	4
70 - 74	1
75 - 79	1
80 - 84	1 N= 40

- Step 6. *Compute the midpoint (M) for each class interval and put it under column M.*
You can obtain the midpoint by the formula below.

$$M = \frac{LS + HS}{2}$$

Where:

M = the midpoint

LS = the lowest score in the class interval

HS = the highest score in the class interval

Class interval	Frequency (f)	M
25 - 29	1	27
30 - 34	0	32
35 - 39	3	37
40 - 44	6	42
45 - 49	6	47
50 - 54	6	52
55 - 59	7	57
60 - 64	4	62
65 - 69	4	67
70 - 74	1	72
75 - 79	1	77
80 - 84	1	82

Illustrative example for the first class interval:

$$M = \frac{25 + 29}{2} = \frac{54}{2} = 27$$

- Step 7. *Compute cumulative distributions for “less than” and “greater than”.* Then put them under column “less than” cumulative and “more than” cumulative distribution. Cumulative frequencies can be obtained by adding the frequency for any class interval or category to the total frequency for all categories above and below it.

- Tabular and Graphic Presentation of Data

Class interval	Frequency (f)	CF <	CF >
25 - 29	1	1	40
30 - 34	0	1	39
35 - 39	3	4	39
40 - 44	6	10	36
45 - 49	6	16	30
50 - 54	6	22	24
55 - 59	7	29	18
60 - 64	4	33	11
65 - 69	4	37	7
70 - 74	1	38	3
75 - 79	1	39	2
80 - 84	1 = 40	40	1

From the given less than ($CF<$) and greater than ($CF>$) cumulative frequency distribution, we can interpret that there is only 1 student whose score is less than 30, just as there are 4 students who scored less than 40. Those who scored less than a given score is evident in the CF less than column. Also, you can interpret this as that only 1 student has a score which is greater than 79, 2 students obtained a score that is greater than 74, and so forth.

- Step 8. *Compute the relative frequency distribution.* This can be obtained by the given formula below.

$$RF(\%) = (100) \frac{CF}{TF}$$

Where: RF = the relative frequency

CF = the class frequency

TF = the total frequency

Class interval	Frequency (f)	RF (%)
25 - 29	1	2.50
30 - 34	0	-
35 - 39	3	7.50
40 - 44	6	15.00
45 - 49	6	15.00
50 - 54	6	15.00
55 - 59	7	17.50
60 - 64	4	10.00
65 - 69	4	10.00
70 - 74	1	2.50
75 - 79	1	2.50
80 - 84	1	2.50
	= 40	= 100.00

- Illustrative example for the first class interval

$$RF(\%) = (100) \frac{CF}{TF}$$

$$= 100/40$$

$$= 2.50$$

Exercise

Below are scores in Statistics examination. Convert the following distribution of scores into a grouped frequency table and (a) determine the size of class intervals, (b) indicate the upper and lower limits of each class interval, (c) identify the midpoints of each class interval, (d) find the cumulative frequency “less than” and “greater than” for each class interval and (e) find the relative frequency for each class interval.

47	31	49	43	59	36	45	38	41	55
63	50	21	40	49	47	24	33	56	50
50	37	51	32	41	62	38	42	46	37
41	52	40	48	58	34	42	54	38	51
54	20	42	47	26	46	38	42	58	40
60	28	34	55	44	56	43	58	50	44
43	39	32	57	41	53	46	33	57	39
49	25	42	36	38	44	51	49	55	45
35	42	28	49	44	40	55	46	30	33

Application

Research Question: How may the length of service of the respondents be described?

Length of Service	School Heads		Teachers	
	Frequency	Percentage	Frequency	Percentage
31-40	5	20.83		
21-30	11	45.83	9	13.24
11-20	5	20.83	27	39.71
1-10	3	12.5	32	47.05
Total	24	100	68	100
Mean	23 years		12.12 or 12 years	

Presentation and Analysis

- The table reveals the respondents' length of service. From among the school heads, 5 or 20.83% had served the system from 31-40 years; likewise, 11 or 45.83 % served for 21-30 years, and another 5 Or 20.83% served from 11-20 years while 3 or 12.5% served from 1-10 years.
- From the teacher-participants, Nine (9) or 13.24 % had served for 21-30 years while twenty-seven (27) or 39.71% had served from 11-20 years and the remaining thirty-two (32) or 47.05 had served from 1-10 years.
- In general, school heads had a mean of 23 years of service while teachers had a mean of 12 years.

Interpretation

- It could be gleaned that if compared with the length of service of the school head-participants, others could have made it in the school heads' positions. But it should be understood that length of service or seniority does not necessarily mean an automatic qualification for headship position.

- It could be said that the trend is to place qualified teachers as school heads regardless of the length of service. It could be said also that other factors other than length of service or eligibility or educational attainment are necessary to weigh the credentials of those who would like to assume higher positions than a classroom teacher.