SELF-LEARNING HOME TASK (SLHT)

Subject: **Practical Research 2** Grade Level: **12** Quarter: **2** Week: **6**

MELC: ***Uses statistical techniques to analyze data – study of differences and relationships limited for bivariate analysis***

Competency Code: ***CS_RS12-IId-g-3***

Name ____________________________ Section _________ Date _________

School ____________________________ District ____________________________

A. Readings/Discussions

Statistical Techniques

1. Percentage is any proportion from the whole.

Formula: $PERCENTAGE(\%) = \left(\frac{PART}{WHOLE}\right)X100$

Example:

Here's a data gathered by AAA National High School administration regarding the number of Grade 7 parents who opted to receive digital copies of the learning modules.

Table 1: Percentage of Parents who Opted to Receive Digital Copies of Learning Modules

Sections	Total Number of Parents	Number of Parents who opted to received digital copies of learning modules	Percentage (%)
7-A	30	24	(24÷30)X100 = 80%
7-B	25	25	(25÷25)X100 = 100%
7-C	32	16	(16÷32)X100 = 50%
7-D	30	11	(12÷30)X100 = 40%
TOTAL	117	76	(76÷117)X100 = 64.96%

2. Mean or average is the middlemost value of your list of values and this can be obtained by adding all the values and divide the obtained sum to the number of values.

Formula: $MEAN(\bar{X}) = \frac{SUM\ OF\ ALL\ VALUES}{NUMBER\ OF\ VALUES}$

Example:

1. **Ungrouped Data**

Refer to Table 1 above, to get the mean or average number of parents who opted to receive digital copies of learning modules, do the following:

$$MEAN(\bar{X}) = \frac{24+25+16+11}{4} = \frac{76}{4} = 19$$

2. **Grouped Data**

Here's the data gathered from the survey on Study Habits conducted by the Grade 12 students to the 150 Grade 7 students of AAA National High School.

Table 2: Mean Distribution of the Study Habits of Students

A Questionnaire to Review Your Study Habits							
	Strongly Agree (5)	Agree (4)	Undecided (3)	Disagree (2)	Strongly Disagree (1)	Mean ($\bar{X}$)	Verbal Description
I study where there is good lighting.	120x5 =600	10x4 =40	0x3 =0	15x2 =30	5x1 =5	$\frac{600+40+0+30+5}{150}$ =4.5	Strongly Agree
I study in a room where the temperature is cool.	100x5 =500	20x4 =80	5x3 =15	10x2 =20	15x1 =15	$\frac{500+80+15+20+15}{150}$ =4.2	Agree

3. **Standard Deviation** shows the spread of data around the mean.

Formula:

$$SD = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}}$$

Example:

Table 2: Mean and Standard Deviation Distribution of the Study Habits of Students

A Questionnaire to Review Your Study Habits								
	SA	A	U	D	SD	Mean ($\bar{X}$)	Mean ($\bar{X}$)	Standard Deviation (
I study where there is good lighting.	120x25 =3000	10x16 =160	0x9 =0	15x4 =60	5x1 =5	4.5	$\frac{3000+160+0+60+5}{150}$ =21.50	$=\sqrt{21.50 - 4.5}$ =4.12
I study in a room where the temperature is cool.	100x25 =2500	20x16 =320	5x9 =45	10x4 =40	15x1 =15	4.2	$\frac{2500+320+45+40+15}{150}$ =19.47	$=\sqrt{19.47 - 4.2}$ =3.91

	Abbreviation		Numerical Values
Strongly Agree	(SA)	-	5
Agree	(A)	-	4
Undecided	(U)	-	3
Disagree	(D)	-	2
Strongly Disagree	(SD)	-	1

One need to get the range from which the mean of a five-point Likert can be interpreted. There are two methods to do this, if we treat the Likert scale as interval/ratio. First, the usual way is to calculate the interval by computing the range (e.g. 5 − 1 = 4), then divided it by the maximum value (e.g. 4 ÷ 5 = 0.80). Ultimately, we get the following result:

From 1 to 1.80 represents (strongly disagree).
From 1.81 to 2.60 represents (do not agree).
From 2.61 to 3.40 represents (true to some extent).
From 3.41 to 4.20 represents (agree).
From 4.21 to 5.00 represents (strongly agree).

The other way is to treat the selection as the range themselves, and so we get these results:

From 0.01 to 1.00 is (strongly disagree);
From 1.01 to 2.00 is (disagree);
From 2.01 to 3.00 is (neutral);
From 3.01 to 4:00 is (agree);
From 4.01 to 5.00 is (strongly agree)

Here is how it will appear in your research paper.

Study Habit	Mean ($\bar{X}$)	Standard Deviation (SD)	Verbal Interpretation
1. I study where there is good lighting.	4.5	4.12	Strongly Agree
2. I study in a room where the temperature is cool.	4.2	3.91	Agree

4. Correlation Analysis (Pearson's r) is a statistical method used to estimate the strength of relationship between two quantitative variables.

Formula: $r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n(\sum x^2) - (\sum x)^2][n(\sum y^2) - (\sum y)^2]}}$

Example:

Here's a data of five students with their corresponding grade in Math (Independent Variable) and grade in English (Dependent Variable). Is there a significant relationship between the grade in Math and the grade in English?

Table 3. Grade in Math and Grade in English of Five Students

Student	Grade in Mathematics (x)	Grade in English (y)	x^2	y^2	xy
A	96	97	9216	9409	9312
B	90	92	8100	8464	8280
C	93	96	8649	9216	8928
D	94	95	8836	9025	8930
E	92	90	8464	8100	8280
Sum	465	470	43265	44214	43730

Step 1. Compute the value of r using the Pearson's r formula.

$$r = \frac{5(43730) - (465)(470)}{\sqrt{[5(43265)-(465)^2][5(44214)-(470)^2]}} = 0.77$$

Step 2. From the table of values, there is a strong positive correlation between the grade in Math and the grade in English.

5. **Regression Analysis** is can be used to explain the relationship between dependent and independent variables.

Three major uses:

a. **Causal analysis** -shows you the possible causation of changes in Y by changes X.
b. **Forecasting an Effect-** allows you estimate and predict the value of Y given the value of X.
c. **Linear Trend Forecasting-** helps you trace the line best fit to tine series

Formula: $Y = mX + b$

$$b = \frac{(\sum y)(\sum x^2)-(\sum x)(\sum xy)}{n(\sum x^2)-(\sum x)^2} \qquad m = \frac{n(\sum xy)-(\sum x)(\sum y)}{n(\sum x^2)-(\sum x)^2}$$

Example:

Using the same data from Table 3, answer the following questions:

a. What linear equation best predicts the grade in English given the grade in Math?

Step 1: Compute the b and m.

$$b = \frac{(\sum y)(\sum x^2)-(\sum x)(\sum xy)}{n(\sum x^2)-(\sum x)^2} \qquad m = \frac{n(\sum xy)-(\sum x)(\sum y)}{n(\sum x^2)-(\sum x)^2}$$

$$b = \frac{(470)(43265)-(465)(43730)}{5(43265)-216225} \qquad m = \frac{5(43730)-(465)(470)}{5(43265)-(216225)}$$

$$m = 1$$

$$b = 1$$

Step 2: Substitute the value of m and b to the regression formula.

The regression equation for predicting the grade in English given the grade in Math is $Y = X + 1$.

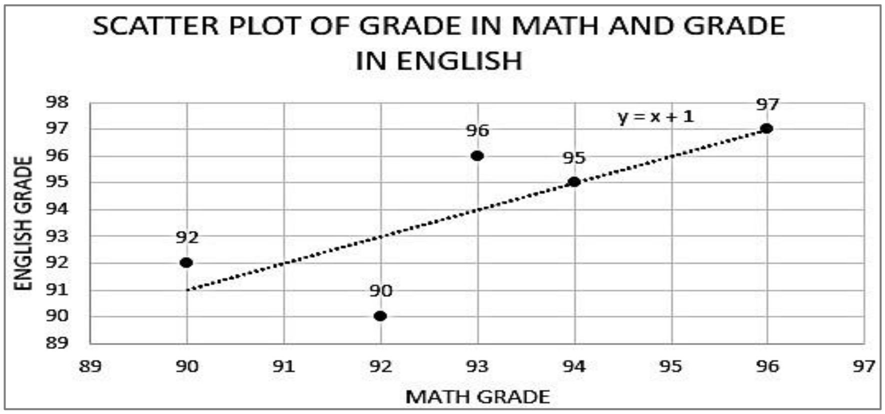
b. If a student made a grade of 91 in Math, what grade would you expect the student to obtain in English?

Using the obtain equation $Y = X + 1$, substitute 91 in X.

$Y = 91 + 1$ = 92 (Grade in English)

According to this model, for every 1point increase in the Math grade, there is a corresponding average increase of 1 point in the English grade.

c. How well does the regression equation fit the data?



Interpretation:

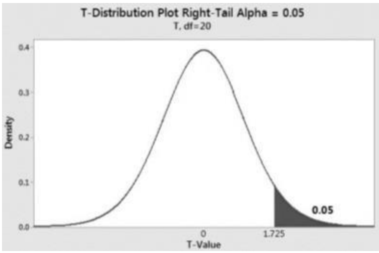
The Math grade is directly proportional to the English grade because the slope is positive.

6. Hypothesis testing. A hypothesis test helps you determine some quantity under a given assumption. The outcome of the test tells you whether the assumption holds or whether the assumption has been violated.

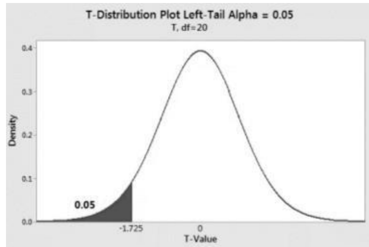
From Module 3, you were exposed to creating your **Null hypothesis (H_0)** which states that there is no difference between the two values or variables and the **Alternative hypothesis (H_1)** which states that there is a difference between two values or variables.

The **statistical test** uses the data obtained from a sample to decide about whether the null hypothesis should be rejected. In a **one-tailed test (left-tailed or right-tailed test),** when the test value falls in the critical region on one side of the mean, the null hypothesis should be rejected.

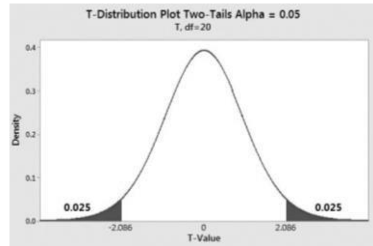
On the other hand, in a **two-tailed test**, the null hypothesis should be rejected when the test value falls in either of the two critical regions.



One-tailed, right-tailed test



One-tailed, left-tailed test



Two-tailed test

To perform hypothesis testing, you compute the mean from the sample and compare it with the mean from the population. Then, you decide whether to reject or not reject the null hypothesis. If the difference is significant, the null hypothesis is rejected. If the difference is not significant, then the null hypothesis is not rejected. In the hypothesis testing, there are four possible results.

	H_0true	$H_0 false$
Reject H_0	Error Type I	Correct decision
Do not reject H_0	Correct decision	Error Type II

The four possibilities are as follows:

1. It would be an incorrect decision and would result in a **Type I error** when you reject the null hypothesis when it is true.
2. It would be a correct decision when you reject the null hypothesis when it is false.
3. It would be a correct decision when you do not reject the null hypothesis when it is true.
4. It would be an incorrect decision and would result in a **Type II error** when you do not reject the null hypothesis when it is false.

The basic format for hypothesis testing:

1. State the hypotheses and identify them.
2. Find the critical value(s).
3. Compute the test value.
4. Make the decision.
5. Summarize the result.

Hypothesis testing can be done using the following t-value approach or critical value approach and p-value approach.

1. **The Critical Value Approach** is used to determine whether the observed test statistic is more extreme than a defined critical value. Hence, the observed test statistic (calculated based on sample data) is compared to the critical value, from t-table. If the test statistic (t^*) is more extreme than the critical value (t), the null hypothesis is rejected. If the test statistic is not as extreme as the critical value, the null hypothesis is not rejected.

One-Sample t-test Formula:

$$t^* = \frac{\bar{X}-\mu_0}{\frac{s}{\sqrt{n}}} \qquad \sigma = \sqrt{\frac{\sum(X-\bar{X})^2}{n}}$$

Example:

A random sample of 10 Grade 7 students has grades in Math, where marks range from 90 (Good) to 98 (Excellent). The general average grade (Gen. Ave.) of all Grade 7 students as of the last 5 years is 93. Is the Gen. Ave. of the 10 Grade 7 students different from the population's Gen. Ave? Use 0.05 level of significance.

Student	1	2	3	4	5	6	7	8	9	10
Math Grade	90	98	97	93	94	91	97	93	93	94

Given: n=10 α=0.05 μ_0=93 $\bar{X}$=94 sd= 2.68

Computational Procedure:

1. Define the Null and Alternative Hypothesis

   H_0: There is no significant difference between the gen. ave. of 10 Grade 7 students from the population's gen. average of 93.

   $H_0: \mu = 93$

   H_1: There is a significant difference between the gen. ave. of 10 Grade 7 students from the population's gen. average of 93.

   $H_1: \mu \neq 93$

2. State the alpha and the degree of freedom.

   $\alpha = 0.05$

   $Df = n - 1 = 10 - 1 = 9$

3. State the decision rule.

   One-tailed Test: $|t| > z_a$; Reject H_0

   Two-tailed Test: $|t| > \frac{z_a}{2}$; Reject H_0

4. Calculate the Test Statistic.

$$t = \frac{\bar{X}-\mu_0}{\frac{s}{\sqrt{n}}} = \frac{94-93}{\frac{2.68}{\sqrt{10}}}$$ =1.18

5. State results (**use t table to get the critical value, see procedure below)**.

$\frac{t_{\frac{\alpha}{2}}}{n-1} = \frac{t_{0.05}}{10-1} = t_{0.0025}$ =2.263

$|1.18| < 2.262$

6. Decision: Accept H_0
7. Conclusion: Therefore, the average grade of 10 Grade 7 students is not different from the population's average grade in Math which is 93.

2. P-value Approach involves determining the probability (assuming the null hypothesis were true) of observing a more extreme test statistic in the direction of the alternative hypothesis than the one observed. If the *P*-value is less than (or equal to) α then the null hypothesis is rejected in favor of the alternative hypothesis. And, if the *P*-value is greater than α, then the null hypothesis is not rejected.

Example:

Use the same data from Example 1 of Critical value approach:

Computational Procedure:

1. Define the Null and Alternative Hypothesis

H_0: There is no significant difference between the gen. ave. of 10 Grade 7 students from the population's gen. average of 93.

$H_0: \mu = 93$

H_1: There is a significant difference between the gen. ave. of 10 Grade 7 students from the population's gen. average of 93.

$H_1: \mu \neq 93$

2. State the alpha and the degree of freedom.

$\alpha = 0.05$

$Df = n - 1 = 10 - 1 = 9$

3. State the decision rule.

One-tailed Test: $|t| > z_a$; Reject H_0

Two-tailed Test: $|t| > \frac{z_a}{2}$; Reject H_0

4. Calculate the Test Statistic.

$$t = \frac{\bar{X}-\mu_0}{\frac{s}{\sqrt{n}}} = \frac{94-93}{\frac{2.68}{\sqrt{10}}}$$ =1.18

5. Use statistical software or an online calculator (https://www.statology.org/t-score-p-valuecalculator/) to find the corresponding p-value.
One-tailed P-value: 0.13412
Two-tailed P-value: 0.26825

6. State results.
One-tailed $|0.13412| > 0.05$
Two-tailed $|0.26825| > 0.05$

7. Decision: Accept H_0
Since this p-value is *not* less than our chosen alpha level of 0.05, we can't reject the null hypothesis.

8. Conclusion: Therefore, the average grade of 10 Grade 7 students is not different from the population's average grade in Math which is 93.

Here are the steps in finding the t-value or critical value at the t-table:

1. Locate your confidence level (alpha level) at the top row of the **t-table found below** (this tells you which column you need).
2. Intersect this column with the row for your df (degrees of freedom). The number you see is the **critical value** (or the **t-value**) for your confidence interval. Table of T-Values

***t* Table**

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221

Examples:

Given	t-value (critical value)
1. df=5, α=0.05, two-tailed test	2.571
2. df=12, α=0.05, one-tailed test	1.782

B. Exercises

Exercise 1

Test I. Direction: Read and analyze the situation below. Complete the table and write a brief interpretation of the table.

Situation: The parents of the senior high school students at AAA High School chose modular learning over online learning approach. In modular learning, the parents can have either a printed copy or a digital copy of the modules. The school administration needs to find out how many printed modules they need to reproduce for the parents. The data were summarized in the table below.

Sections	**Total number of Parents**	**Number Parents of chooses printed modules**	**Percentage**
Grade 12 - A	35	25	
Grade 12 - B	40	20	
Grade 12 - C	38	15	
Grade 12 - D	45	35	
Grade 12 - E	30	18	
Total			

Interpretation:

____
____
____
____
____

Test II. Directions: Complete the table below.

Situation: Here is the data gathered from the survey on Study Habits of the Grade 12 students at AAA High School.

A Review Your Study Habits								
	Strongly Agree (5)	Agree (4)	Undecided (3)	Disagree (2)	Strongly Disagree (1)	Mean ($\bar{X}$)	Standard Deviation (SD)	Verbal Interpretation
The desk where I study is always clear from distractions.	90	30	10	5	15			
I use earplugs to minimize distracting sounds.	10	50	30	20	40			
I study facing a wall.	15	35	30	20	50			

Exercise 2

Test I. Direction: Study carefully the table below. Is there a significant relationship in the pre-test and post-test score in General Mathematics?

Student	Pretest	Post Test
1	20	30
2	23	28
3	35	33
4	30	33
5	26	38
6	19	24
7	20	31
8	18	38
9	18	30
10	23	25

Answer the following:

1. Compute for the Pearson’s r.
2. Write a brief interpretation.
3. What linear equation best predicts the posttest given the pretest in Math?
4. Show the line of best fit and its interpretation.

Test II. Direction: A random sample of 10 Grade 7 students has grades in MAPEH, where marks range from 90 (Good) to 98 (Excellent). The general average grade (Gen. Ave.) of all Grade 7 students as of the last 5 years is 95. Is the Gen. Ave. of the 10 Grade 7 students different from the population's Gen. Ave? Use 0.05 level of significance.

Student	1	2	3	4	5	6	7	8	9	10
MAPEH Grade	92	95	95	96	97	98	95	94	98	92

Given: n=10 α=0.05 μ_0=95 $\bar{X}$=___ sd= ___

1. Perform hypothesis testing using the Critical Value Approach.
2. Perform hypothesis testing using the P-Value Approach.

C. Assessment/Application/Outputs (Please refer to DepEd Order No. 31, s. 2020)

Direction: Read and carefully analyze each item below. Choose the letter of the best answer.

1. Which of the following is a statistic?
   A) the median family income in the United States
   B) the mean number of absences per day in one school district during one school year
   C) the range of running times at the Boston Marathon
   D) the percentage of defective units in a random sample of units produced at a manufacturing plant
2. Which of the following is an example of categorical data?
   A) ages of volunteers at a community service center
   B) zip codes of shoppers at an appliance store
   C) SAT scores of students who took the SAT in March
   D) square footage of high school gymnasiums in Nebraska
3. The median is a better index of central tendency than the mean when the distribution is
   A) symmetrical.
   B) skewed.
   C) normal.
   D) bimodal.

4. Which of the following correlations would be the most likely value of Pearson's r for the relationship of the time spent practicing typing to the number of errors made on a typing test?
   A) -0.55
   B) 0
   C) 0.82
   D) 1.47
5. Normal distribution is best described when
   A) researchers draw a smooth curve instead of the series of straight lines in a frequency polygon.
   B) the tail of the distribution trails off to the left, in the direction of the lower, more negative, score values.
   C) the tail of the distribution trails off to the right, in the direction of the higher, more positive, score values.
   D) most scores are concentrated in the middle of the distribution, and the scores decrease in frequency the father away from the middle they are.
6. Which of the following is an example of a null hypothesis?
   A) Motivation is not related to achievement.
   B) Highly motivated people are more likely to be successful than those with low motivation.
   C) Highly motivated people are less likely to be successful than those with low motivation.
   D) Motivation is related to achievement.
7. The standard deviation of a statistic is called
   A) the standard error.
   B) the sampling error.
   C) a quartile.
   D) a confidence interval.
8. Which type of hypothesis specifies that there is no relationship in the population?
   A) research hypothesis
   B) descriptive hypothesis
   C) null hypothesis
   D) inferential hypothesis
9. The probabilities used for hypothesis testing are accurate for generalizing to a population when
   A) the sample size is large.
   B) confidence intervals cannot be constructed.
   C) the significance level is small.
   D) samples are random.

10. Which of the following sequences for analyzing data is in the best order?
    A) calculate descriptive statistics, construct graphs, calculate inferential statistics
    B) calculate inferential statistics, construct graphs, calculate descriptive statistics
    C) construct graphs, calculate descriptive statistics, calculate inferential statistics
    D) calculate descriptive statistics, calculate inferential statistics, construct graphs

D. Suggested Enrichment/Reinforcement Activity/ies

Direction. Perform the following tasks. You may write or encode your answer in a long bond paper. Submit your output to your teacher for checking.

Tasks:

1. Based on your methodology (of your current study), decide what statistical technique/s you will use to analyze deeply your data.
2. Why will you use this tool?
3. Indicate your data analysis.

If your study involves significant relationship and hypothesis testing, answer these questions further.

1. Use the statistical tool that you have decided upon to compute the significance of your study with relevance to the null and the alternative hypothesis.
2. Conduct hypothesis testing.

References:

Books

Calmorin, Laurentina Paler and Melchor A. Calmorin. *Research Methods and Thesis Writing, 2nd Edition.* Manila, Philippines: Rex Book Store, Inc, 2007.

Faltado, Ruben E. III, Medardo B. Bombita, Helen B. Boholano, and Angeline M. Pogoy. *Practical Research 2: Quantitative Research*. Quezon City: Lorimar Publishing, 2016.

Ragma, Feljone. *Practical Research 2: Quantitative Research*. Intramuros Manila: Mindshapers Co., Inc., 2019.

Online Sources

Esther, Baraceros. *Practical Research 2 e-Book*. Quezon City: Rex Bookstore, Inc, 2016.

https://bit.ly/3c6evv2

https://bit.ly/36EfqBZ

https://bit.ly/3gmMbYS

Prepared by:

Mrs. Florie Ann F. Sabio
Ms. Keisha Marie P. Roldan

Edited by:

Reviewed by:

GUIDE

For the Teacher

For the Learner

For the Parent/Home Tutor