SHS

AIRs - LM in

Statistics and Probability

Module 8:

T-distribution and Confidence Interval

Statistics and Probability
Module 8: T-distribution and Confidence Interval
First Edition, 2021

Development Team of the Module

Author: Abelaine Joy B. Abaquita, *TIII*

Editor: SDO La Union, Learning Resource Quality Assurance Team

Illustrator: Ernesto F. Ramos Jr., *P II*

Management Team:

Target

Suppose that the variance of the population from which we select our random sample is unknown. If the sample size n is at least 30, then the sample variance s^2 can be used to estimate σ^2 and the Central Limit Theorem still applies. However, if the sample size is small and the variance in unknown, the sampling distribution for the mean can no longer be approximated by the normal distribution and the t – distribution can be used.

Before we start, let us consider first the objectives that needs to be attained after going through this module.

Learning Competencies:

1. illustrates the t-distribution **(M11/12SP-IIIg-2)**,
2. identifies percentiles using the t-table **(M11/12SP-IIIg-5)**,
3. identifies the length of a confidence interval **(M11/12SP-IIIj-1)**,
4. computes for the length of the confidence interval **(M11/12SP-IIIj-2)**,
5. computes for an appropriate sample size using the length of the interval **(M11/12SP-IIIj-3)** and
6. solves problems involving sample size determination **(M11/12SP-IIIj-4)**.

Subtasks

1. define t-distribution
2. determine the steps to be followed in using the t-table.
3. define confidence interval
4. determine the formula and steps in solving the length of the confidence interval.
5. solve sample size

Before going on, check how much you know about this topic. Answer the pretest on the next page in a separate sheet of paper

Pretest

Directions: Read the following questions carefully and choose the letter of your answer. Use the t-table at the appendix as needed.

1. Which of the following is ***NOT*** a property of the t – distribution?
   A. It is mount – shaped and symmetric about 0.
   B. It is more variable than the standard normal distribution.
   C. For large values of n, the t – distribution is approximately normal.
   D. The shape of the t – distribution does not depend on the degrees of freedom.

2. When ***NOT*** to use t – distribution to estimate?
   A. When the sample size is small
   B. When the population variance is unknown
   C. When the population standard deviation is known
   D. When the sample size is small and population variance is unknown

3. For a sample size 24 from a normal population with unknown variance, the sampling distribution of the mean has a t – distribution with 23 degrees of freedom. Is it true or false?
   A. Either B. False C. Neither D. True

4. What is the other term for t- distribution?
   A. Z – distribution C. Student's t- distribution
   B. Teacher's t- distribution D. Teacher's z- distribution

5. Which of the following is the value of $t_{0.05}$ with a sample size of 10?
   A. 1.812 B. 2.262 C. 2.860 D. 2.900

6. Which of the following is the value of $t_{0.01}$ with 15 degrees of freedom?
   A. 2.602 B. 2.624 C. 2.650 D. 2.947

7. What is the 90^{th} percentile in a t-distribution with 12 degrees of freedom?
   A. 1.356 B. 1.363 C. 1.782 D. 1.796

8. What is the 85^{th} percentile in a t-distribution with a sample size of 11?
   A. 1.093 B. 1.296 C. 1.372 D. 1.463

9. What is the 95% interval estimate of μ given $\sigma = 8.7$, n = 12, and $\bar{x} = 73.9$?
   A. $70.31 < \mu < 72.89$ B. $70.97 < \mu < 72.49$
   C. $68.37 < \mu < 79.43$ D. $69.31 < \mu < 72.43$

10. What is the 98% interval estimate of μ given $s = 3$, n=15 and $\bar{x}$ =18?
   A. $15.57 < \mu < 20.53$ B. $15.97 < \mu < 20.03$
   C. $14.97 < \mu < 21.03$ D. $14.97 < \mu < 21.03$

11. A sample of 50 learners showed a mean height of 60 inches. If it is known that the standard deviation of heights of learners is 2.5 inches, what is the 70% confidence interval estimate for the height of all the learners?

A. $59.64 < \mu < 60.36$ B. $59.32 < \mu < 60.48$

C. $58.52 < \mu < 61.48$ D. $59.52 < \mu < 62.48$

12. A college dean wishes to estimate the average number of hours students spend doing homework per week. The standard deviation from a previous study is 4.5 hours. How large a sample must be selected, if he wants to be 95% confident of finding whether the true mean differs from the sample mean by 2.1 hours?

A. 16 B. 17 C. 18 D. 20

13. Given a standard deviation of 10.4, confidence interval of mean 95% and a margin of error or maximum allowable deviation of 4.35, what is the appropriate sample size?

A. 19 B. 20 C. 21 D. 22

14. Given the margin of error 2.4, confidence level of 90% and standard deviation of 6.4, what is the value of sample?

A. 19 B. 20 C. 21 D. 22

15. Given the margin of error 4.5, confidence level of 98% and standard deviation of 5.8, what is the value of the sample?

A. 8 B. 9 C. 10 D. 11

Lesson 1	T-distribution

Jumpstart

For you to understand the lesson well, do the following activities. Have fun and good luck!

Activity 1: Convert Me!

Directions: Solve for the t – score given the following**:**

Mean (μ)	Standard Deviation (σ)	Sample Mean ($\bar{x}$)	Sample (n)	t
1. 100	16	120	10	
2. 8.5	2.2	12	15	
3. 14.4	4.6	16	20	
4. 20	9.5	21.5	25	
5. 10	11.2	18.6	18	

Activity 2: Solve Me!

Directions: Solve for the t – score of the following problems**:**

1. Suppose a population has mean 80 and standard deviation 10. Then we get a sample of 25 cases and the mean of this sample is 82. Find the t – score if the sample is 25 cases.

   Answer: ___________________________

2. The mean weight of a banana is 92 grams with a standard deviation of 6 grams. A sample of 16 is taken from a basket and the mean weight of the sample is 94 grams. Find the t- score of the sample bananas.

   Answer: ___________________________

3. The average number of liters of fresh milk that a person consumes in a month is 18 liters. Assume that the standard deviation is 4.5 liters and the mean of the sample is 19.5. Find the probability that a person selected at random consumes more than 16 liters per month.

   Answer: ___________________________

> *Very good you answered the Jumpstart with ease. Now you are ready to Discover and Learn.* ☺

Discover

THE STUDENT'S T – DISTRIBUTION

Student's t – distribution___

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

with n – 1 degrees of freedom (df)

It was the English statistician ***William Sealy Gosset*** who was one of the first to discover the Student's t – distribution. In 1908, he published a paper "The Probable Error of a Mean" in the "Biometrika" journal under the pseudonym "Student". This t – distribution is usually used when the sample size is small ($n < 30$).

Student's t-distribution which is also known as **t- distribution**, is a probability distribution which is utilized in estimating parameters of a certain population in case of the sample size is small and/or the population variance or standard deviation is unknown. It is like with the z-distribution that is bell-shaped and symmetric about the x- axis but flatter and more spread.

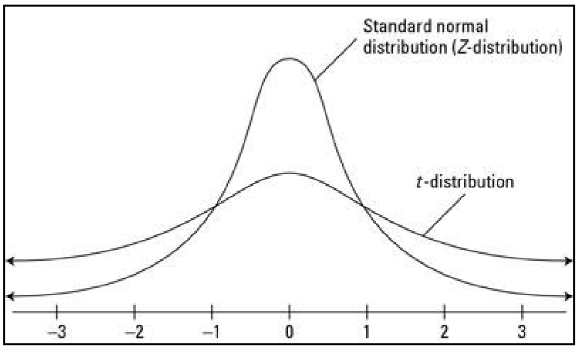


Figure 1: Comparison of a Standard Normal Distribution and a t – distribution

When defining a t – distribution, it is necessary to specify the number of *degrees of freedom* (df). **Degrees of freedom** refer to the number of independent observations in a given set of data. The number of independent observations is sample size minus one or in symbols, that is df = n-1 where df is the degree of freedom and n is the sample size. Hence, given a sample size of 9 will have a degree of freedom 8 and if given a degree of freedom of 14 will mean a sample size of 15. However, for some or other applications, degrees of freedom can be calculated in different ways.

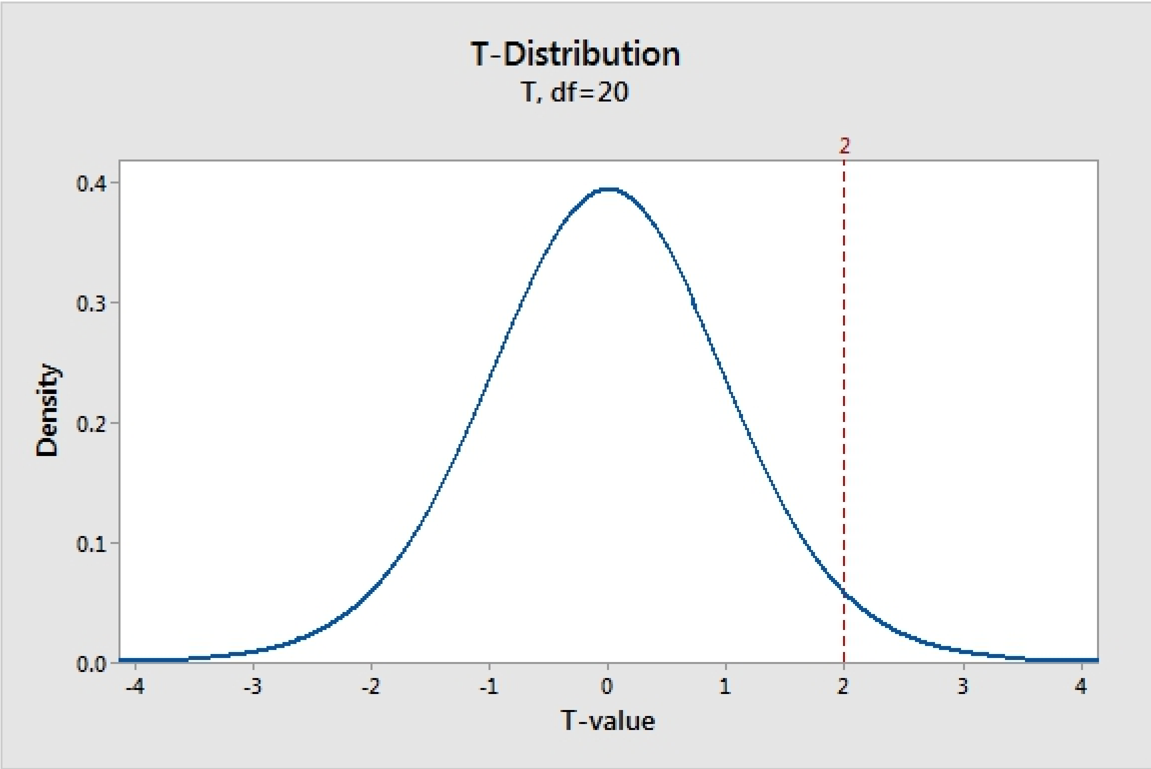
The following are the properties of t – distribution:
1. It is mound – shaped and symmetric about 0.
2. It is more variable than the standard normal distribution.
3. Its shape is dependent on the sample size n.
4. As n increases, the variability of the t – distribution decreases.
5. For large values of n, the t – distribution is approximately normal.

Also, the t-distribution can still be utilized for moderately skewed population distribution given that it is unimodal, without outliers and size is at least 40. Lastly, the t-distribution can be applied for a size greater than 40 and without outliers.

When you perform a t-test for a single study, you obtain a single t-value. However, if we drew multiple random samples of the same size from the same population and performed the same t-test, we would obtain many t-values and we could plot a distribution of all of them.

T-distributions assume that you draw repeated random samples from a population where the null hypothesis is true. You place the t-value from your study in the t-distribution to determine how consistent your results are with the null hypothesis.

The graph below shows a t-distribution that has 20 degrees of freedom, which corresponds to a sample size of 21 in a one-sample t-test. It is a symmetric, bell-shaped distribution that is similar to the normal distribution, but with thicker tails. This graph plots the probability density function (PDF), which describes the likelihood of each t-value.



The peak of the graph is right at zero, which indicates that obtaining a sample value close to the null hypothesis is the most likely. That makes sense because t-distributions assume that the null hypothesis is true. T-values become less likely as you get further away from zero in either direction. In other words, when the null hypothesis is true, you are less likely to obtain a sample that is very different from the null hypothesis.

Our t-value of 2 indicates a positive difference between our sample data and the null hypothesis. The graph shows that there is a reasonable probability of obtaining a t-value from -2 to +2 when the null hypothesis is true. Our t-value of 2 is an unusual value, but we don't know exactly *how* unusual. Our ultimate goal is to determine whether our t-value is unusual enough to warrant rejecting the null hypothesis.

In order to compute the t- value, a t-distribution table is shown below which is consist of the **degrees of freedom (df)** which are the numbers at the left most column, **"α"** which is some of the special areas at the top most row and the **t-values** which are located at the right of the degrees of freedom and below **"α"**.

The t – distribution Table

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
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
	Confidence Level										

Figure 2: t – distribution table

Steps to look for the corresponding t – value:

1. Look for the corresponding level of significance (α) which is located at the column header.
2. Look for the corresponding number of degrees of freedom (df) in the first column.

3. Look for the intersection of the level of significance and degrees of freedom to get the value of the t – value.

Example 1: Find the t – value for α = 0.05 and sample of 7

Solution:

1. Level of significance is 0.05

2. Degrees of freedom is 7 – 1 = 6

3.

ONE TAILED TEST: If you are using a significance level of .05, a one-tailed test allots all of your alpha to testing the statistical significance in the one direction of interest. This means that .05 is in one tail of the distribution of your test statistic. When using a one-tailed test, you are testing for the possibility of the relationship in one direction and completely disregarding the possibility of a relationship in the other direction.

In a one-tailed test, the critical region has just one part (the green area below). It can be a left tailed test or a right-tailed test. Left-tailed test: The critical region is in the extreme left region (tail) under the curve. Right-tailed test: The critical region is in the extreme right region (tail) under the curve.

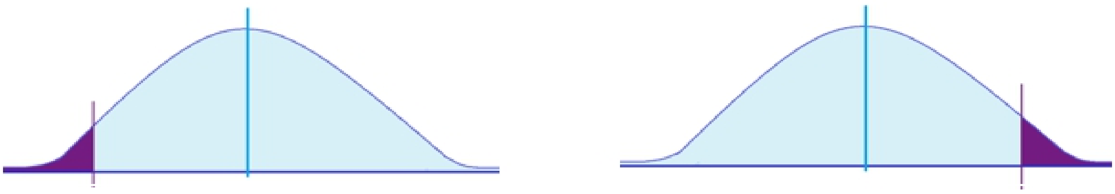


Figure 3: One Tailed Test

TWO TAILED TEST: If you are using a significance level of 0.05, a two-tailed test allots half of your alpha to testing the statistical significance in one direction and half of your alpha to testing statistical significance in the other direction. This means that .025 is in each tail of the distribution of your test statistic.

In two-tailed test, the critical region has two parts (the red areas below) which are in the two extreme regions (tails) under the curve.

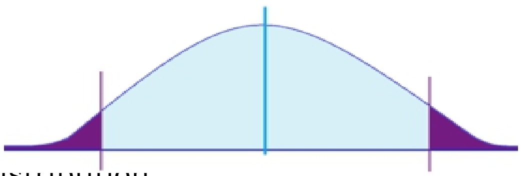


Figure 4: Two Tailed Tailed Test

Since in the problem, it did not state the why we use the two tailed test of the t – distribution.

t Table

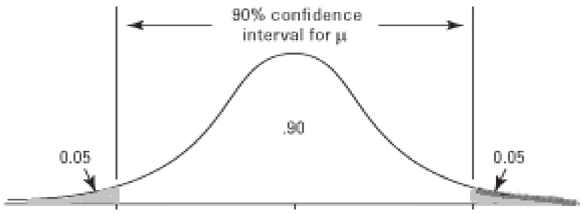
cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01
df									
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499

$t_{(0.05\,,\,6)} = 2.447$

The t – value is 2.447. Our t-value is beyond +2, hence, it is enough to warrant rejecting the null hypothesis (in our example null hypothesis not stated)

Example 2: Find the t – value for a 90% confidence interval with n = 10.

Solution: In finding the confidence level, we recall that we want the area $\frac{\alpha}{2}$ to accommodate possible mean sample values x̄ that fall below and above μ. The confidence level of 90% so the distribution is as follows:



This means that the shaded area we want has a sum of 10% or 0.1. We know that these 2 shaded areas are symmetrical about the mean thus are equal in area; hence,

$$\alpha = 0.1 \text{ thus } \frac{\alpha}{2} = 0.05$$

This means that the area in each of the shaded region is 0.05 or 5%. Thus we want to find $t_{0.05}$ with 10 – 1 = 9 degrees of freedom or simply $t_{(0.05\,,\,9)}$.

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002
df										
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144

$t_{(0.05\,,\,9)}$= 1.833

The t – value is 1.833. Our t-value is within the ± 2, hence, it is enough to warrant accepting the null hypothesis. (null hypothesis not stated)

Example 3: Find the t-value when μ = 40, $\bar{x}$ = 45, s= 6 and n = 28.

Solution: $t = \frac{(\bar{x}-\mu)}{\frac{s}{\sqrt{n}}} = \frac{(45-40)}{\frac{6}{\sqrt{36}}} = \frac{5}{\frac{6}{6}} = \frac{5}{1} = 5$

IDENTIFYING PERCENTILES USING T – DISTRIBUTION

Percentile is one of the measurements in statistics which tells the value below in which an observations' percentage in a set of observations falls. For example, you score 80 in an exam and it was mentioned that you scored at the 87^{th} percentile, it means that 87% of the scores are below you and 13% of the scores are above you.

In addition, in finding the percentile for a t- distribution, t-table can be utilized as it is a number on a statistical distribution whose less than the probability is the given percentage. So, if you are asked on the 90^{th} percentile of the t-distribution with respect to its degrees of freedom, that refers to the value whose left tail or less than probability is 90% or 0.90 and whose right tail or greater than probability is 10% or 0.10.

Example 4: Let T be a random variable having a t – distribution with a sample of 14. Find the 90^{th} percentile of T.

Solution: To find the value of 90^{th} percentile, identify first the degrees of freedom.

df = n-1 = 14-1 = 13

To solve for 90^{th} percentile, we need to understand first its implication which is to get the t-value that is less than the probability 0.90 and the right tail probability that is 0.10. Based on the table below (one – tail test), the 90^{th} percentile is 1.350.

df \ α	0.10
1	3.078
2	1.886
⋮	⋮
13	1.350

$t_{(\alpha,df)}$ = 1.350

The t – value is 1.350, null hypothesis accepted

Example 5: Let T be a random variable having a t – distribution with a sample of 20. Find the 85th percentile of T.

Solution: To find the value of 85th percentile, identify first the degrees of freedom.

df = n-1 = 20-1 = 19

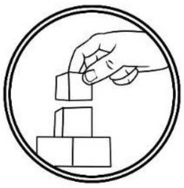
To solve for 85th percentile, we need to understand first its implication which is to get the t-value that is less than the probability 0.85 and the right tail probability that is 0.15. Based on the table below (One tail test), the 85th percentile is 1.066.

df \ α	0.15
1	1.963
2	1.386
3	1.250
4	1.190
5	1.156
6	1.134
⋮	⋮
19	1.066

$t_{(0.15,\,19)} = 1.066$

The t – value is 1.066. null hypothesis accepted.

Did you learn something from your discovery? Now let us apply these skills to explore . Enjoy ☺



Explore

Activity 1: Complete Me!

Directions: Complete the table below by solving the t – value.

Sample (n)	Level of Confidence (1 – α)100%	Level of Significance (α)	t - value
1. 14	90%		
2. 11	95%		
3. 22	98%		
4. 27	80%		
5. 16	70%		
6. 18	99%		
7. 20	90%		
8. 9	95%		
9. 7	80%		
10. 17	60%		

Activity 2: Find Me!

Directions: Solve the following problems and find your answer inside the box. Write the letter of the correct answer.

A. 1.753	D. 4.289
B. 1.383	E. 1.064
C. 3.499	F. 1150

______ 1. What is the value of $t_{0.05}$ with 15 degrees of freedom?

______ 2. Solve for the of $t_{0.005}$ with a sample of 8 in one-tailed.

______ 3. What is the 90th percentile of a t distribution with a sample of 10?

______ 4. What is the 70th percentile of a t-distribution with a sample size of 21?

______ 5. In a factory, the weight of the concrete poured into a mold by a machine follows a normal distribution with a mean of 1150 pounds and a standard deviation of 22 pounds. What weight falls at the 50th percentile?

Deepen

At this point, you are going to apply what you have learn about the t – distribution and percentile using the t – table.

What you need:

A sheet of paper
Ballpen or any writing material

What you have to do:

Read, analyze, and solve the problem below. You will be scored based on the given rubric found on the next page.

Problem:

A machine produces cylindrical metal pieces. A sample of pieces is taken, and the diameter are found to be 1.01, 0.97, 1.03, 1.04, 0.99, 0.98, 0.99, 1.01 and 1.03 cm. Find the t-value for a 98% confidence interval for the diameters of the cylindrical metal pieces. Assume that the diameters are approximately normally distributed.

RUBRICS FOR SCORING THE PROBLEM

CRITERIA	4	3	2	1
Solution	90 – 100% of the steps and solutions have no mathematical errors	Almost all (85 – 89%) of the steps and solutions have no mathematical errors	Most (75 – 84%) of the steps and solutions have no mathematical errors	More than 75% of the steps and solutions have mathematics errors
Mathematical Work and Notation	Correct terminology and notation are always used, making it easy to understand what was done	Correct terminology and notation are usually used, making it fairly easy to understand what was done	Correct terminology and notations are used, but it is sometimes not easy to understand what was done	There is little use, or a lot of inappropriate use of terminology and notations
Neatness and Organization	The work is presented in a neat, clear, organized fashion but is easy to read	The work is presented in a neat and organized fashion that is usually easy to read	The work is presented in an organized fashion but may be hard to read at times	The work appears sloppy and unorganized. It is hard to know what information goes together

Lesson 2 Confidence Interval

Jumpstart

When you make an estimate in statistics, whether it is a summary statistic or a test statistic, there is always uncertainty around that estimate because the number is based on a sample of the population you are studying.

The **confidence interval** is the range of values that you expect your estimate to fall between a certain percentage.

To be ready for the next lesson, let us enjoy the following activity… ☺

Activity 1: FACT OR BLUFF

Directions: Write **FACT** if the statement is correct otherwise write **BLUFF.**

______ 1.	The confidence level is the percentage of times you expect to reproduce an estimate between the upper and lower bounds of the confidence interval, and is set by the alpha value.
______ 2.	A confidence interval is the standard deviation of your estimate plus and minus the variation in that estimate. This is the range of values you expect your estimate to fall between if you redo your test, within a certain level of confidence.
______ 3.	Confidence, in statistics, is another way to describe probability
______ 4.	You can calculate confidence intervals for many kinds of statistical estimates, including proportions and population means.
______ 5.	In the confidence interval there is an exact value that you expect your estimate to fall

Discover

CONFIDENCE INTERVAL

Estimation is one aspect of inferential statistics; it is the process of estimating the value of a parameter from information drawn from a sample. The objective of estimation is to determine the approximate value of a population parameter on the basis of a sample statistic. We refer to the sample statistic as the ***estimator*** of the population estimator. The computed sample statistic is called the ***estimate***.

An ***estimate*** may be a point estimate or an interval estimate. A point estimate is the value of a sample statistic that is used to estimate a population parameter. Generally, whenever we utilize point estimate, we calculate the ***margin of error***. The margin of error denoted as E is calculated as follows:

E = $\pm\, Z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$ for the margin of error of population mean when variance is known or sample size is greater than 30

E = $\pm\, t_{\left(\frac{\alpha}{2},\, n-1\right)} x \left(\frac{s}{\sqrt{n}}\right)$ for margin of error of population mean when variance is unknown and sample size is less than 30

An interval estimate of a parameter on the other hand is an interval or a range of values used to estimate the parameter. This estimate may or may not contain the value of the parameter being estimated. A degree of confidence (generally a percent) can be assigned before an interval estimate is prepared. Each interval is constructed with regard to a given confidence level and is called a ***confidence interval***. The ***confidence level*** is associated with a confidence interval, states how much confidence we have that is interval contains the true population parameter. The ***confidence level*** is denoted by (1 – α)100%, where α is the Greek letter alpha. When an interval estimate has an attached confidence coefficient, it will be called **confidence interval.** Confidence interval is a range with lower limit and upper limit used to estimate population parameter. The lower and the upper limit of the interval is within the certain level of confidence.

Confidence Interval for the population mean when variance is known or sample size is greater than 30

A (1 – α) 100% confidence interval for μ is given by

$$\bar{x} \pm Z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right) \qquad or \qquad \bar{x} - Z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right) < \mu < \bar{x} + Z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

where: $\bar{x}$ = estimated mean based from the sample

$Z_{\alpha/2}$ = z – value that leaves an area of $\frac{\alpha}{2}$ to the right;

σ = sample standard deviation; and

n = sample size

Confidence Interval for the population mean when variance is unknown and sample size is less than 30

A (1 – α) 100% confidence interval for μ is given by

$$\bar{x} \pm \left(t_{\left(\frac{\alpha}{2},n-1\right)} x \left(\frac{s}{\sqrt{n}}\right)\right) \quad \text{or} \quad \bar{x} - \left(t_{\left(\frac{\alpha}{2},n-1\right)} x \left(\frac{s}{\sqrt{n}}\right)\right) < \mu < \bar{x} + \left(t_{\left(\frac{\alpha}{2},n-1\right)} x \left(\frac{s}{\sqrt{n}}\right)\right)$$

where: $\bar{x}$ = estimated mean based from the sample

$t_{\left(\frac{\alpha}{2},n-1\right)}$ = t – value with n – 1 degrees of freedom that leaves an area of $\frac{\alpha}{2}$ to the right;

s = sample standard deviation; and

n = sample size

Example 1: Solve for the confidence interval of a distribution given a confidence level of 95%, standard deviation (σ) of 9, mean of 58 and sample of 50.

Solution:

Step 1: Note that n = 50, which means that n > 30.

Find $\frac{\alpha}{2}$, 95% means α is:

1 – 0.95 = 0.05

α = 0.05

$\frac{\alpha}{2}$ = 0.025

Step 2: To locate $z_{\frac{\alpha}{2}}$, use the t – distribution table search for α = 0.05 (for two tailed test) and the degrees of freedom look for ∞

df \ α	
One tail	**0.025**
Two tail	**0.05**
df	
1	12.71
2	4.303
⋮	⋮
∞	1.960

Step 3: Substitute to the confidence interval formula:

$$\bar{x} - \left(Z_{\alpha/2} x \left(\frac{\sigma}{\sqrt{n}}\right)\right) < \mu < \bar{x} + \left(Z_{\alpha/2} x \left(\frac{\sigma}{\sqrt{n}}\right)\right)$$

$$58 - \left(1.960\left(\frac{9}{\sqrt{50}}\right)\right) < \mu < 58 + \left(1.960\left(\frac{9}{\sqrt{50}}\right)\right)$$

$$58 - (2.4946) < \mu < 58 + (2.4946)$$

$$55.5053 < \mu < 60.4947$$

Thus, the confidence interval is 55.51 < μ < 61.49 or μ = (55.51, 60.49)

Example 2: Solve for the confidence interval of a distribution given a confidence level of 90%, standard deviation (σ) of 9, mean of 58 and sample of 20.

Solution:

Step 1: Note that n = 20, which means that n < 30. Hence, we will use the t – distribution.

Find $\frac{\alpha}{2}$, 90% means α is:

1 – 0.90 = 0.10

α = 0.10

$\frac{\alpha}{2}$ = 0.05

Step 2: Locate $t_{\frac{\alpha}{2}}$ or $t_{0.05}$ with df = n – 1 = 20 – 1 = 19 on the t – table.

Note: Look for the two tailed value of 0.05 for α.

df \ α	
One tail	**0.05**
Two tail	**0.10**
df	
1	6.314
2	2.920
3	2.353
4	2.132
5	2.015
6	1.943
7	1.895
⋮	⋮
19	1.729

Step 3: Substitute to the confidence interval formula:

$$\bar{x} - \left(t_{\left(\frac{\alpha}{2}, n-1\right)} x \left(\frac{s}{\sqrt{n}}\right)\right) < \mu < \bar{x} + \left(t_{\left(\frac{\alpha}{2}, n-1\right)} x \left(\frac{s}{\sqrt{n}}\right)\right)$$

$$58 - \left(1.729\left(\frac{9}{\sqrt{20}}\right)\right) < \mu < 58 + \left(1.729\left(\frac{9}{\sqrt{20}}\right)\right)$$

$$58 - (3.4795) < \mu < 58 + (3.4795)$$

$$54.5204 < \mu < 61.4795$$

Thus, the confidence interval is 54.52 < μ < 61.48 or μ = (54.52, 61.48)

Example 3: The following is a list of 15 scores of an achievement test

38, 46, 50, 31, 22, 17, 56, 67, 29, 38, 37, 48, 50, 47, 41

Find the 95% confidence interval of the population mean.

Solution:

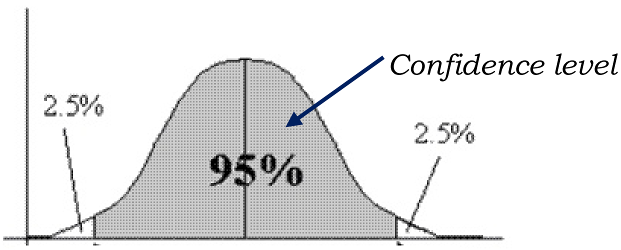
Step 1: Solve for the $\bar{x}$, where n = 15

$$\bar{x} = \frac{38+46+50+31+22+17+56+67+29+38+37+48+50+47+41}{15} = 41.13$$

Step 2: Calculate s

$$s = \frac{\sqrt{\sum (x-\bar{x})^2}}{n-1} = \frac{182.83}{15-1} = \frac{182.83}{14} = 13.06$$

Note that n = 15, which means that n < 30. Hence we will use the t – distribution. In terms of the t – distribution we have the idea as follows:



That is we want to have 95% confidence that μ falls in between the t – values according to the t – table.

Step 3: Find $\frac{\alpha}{2}$, 95% means α is:

$$1 - 0.95 = 0.05$$

$$\alpha = 0.05$$

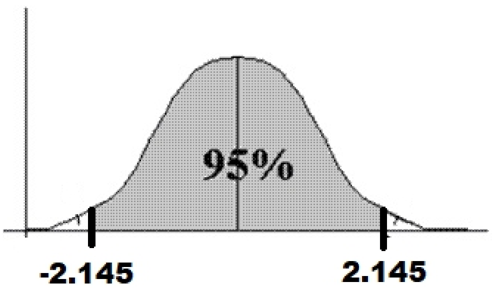
$$\frac{\alpha}{2} = 0.025$$

Step 4: Locate $t_{\frac{\alpha}{2}}$ or $t_{0.025}$ with df = n – 1 = 15 – 1 = 14 on the t – table.

Note: Look for the two tailed value of 0.05 for α.

df \ α	
One tail	**0.025**
Two tail	**0.05**
df	
1	12.71
2	4.303
3	3.182
4	2.776
5	2.571
6	2.447
7	2.365
⋮	⋮
14	2.145

This t – values of 2.145 means that in the t – distribution, the $\bar{x}$ may fall 2.145 units away from the μ with 95% confidence. This is illustrated usually as follows:



Step 5: Find the corresponding x values 2.145 away from μ, we use

$$E = \pm Z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$E = \pm\ 2.145\left(\frac{13.06}{\sqrt{15}}\right)$$

$$E = 2.145(3.37)$$

$$E = 7.2331$$

Substituting this to the confidence interval formula for the μ with unknown σ:

$$\bar{x} - \left(t_{\left(\frac{\alpha}{2},n-1\right)} x \left(\frac{s}{\sqrt{n}}\right)\right) < \mu < \bar{x} + \left(t_{\left(\frac{\alpha}{2},n-1\right)} x \left(\frac{s}{\sqrt{n}}\right)\right)$$

which is the same as

$$\bar{x} - E \ < \mu < \bar{x} + \ E$$

we have

$$\bar{x} - 7.2331 \ < \mu < \bar{x} + \ 7.2331$$

Using the computed mean $\bar{x} = 41.13$

$$41.13 - 7.2331 \ < \mu < 41.13 + \ 7.2331$$

$$33.8969 \ < \mu < \bar{x} + \ 48.3631$$

Thus, there is a 95% chance that the population means μ of the achievement test scores will fall between 33.90 and 48.40.

Sample Size Determination

One of the most important things that a researcher must consider when conducting a statistical study is the sample size. If the sample size is too small, then there is a high possibility that the estimates may be unreliable. On the other hand, if the size of the sample to be obtained is too large, it might be too costly or time – consuming for the researcher. The higher the accuracy we demand for the resulting estimate, the larger the size of the sample we need to obtain.

In estimating, the sample size needs to be considered in order to make it more valid and reliable. Especially on research, the number of respondents must be considered. The number of members or element of a population must be taken into consideration aside from being manageable. The selected samples, either through probability or non-probability sampling, must be taken at least to its minimum need to go through the study. With this, the researcher/s will be able to make a feasible conclusion about the parameter of a certain population being studied.

Below are the three factors which affect the sample size used to estimate a population parameter:

1. The ***Margin of Error*** (E). Since sampling is random, we do not expect an exact value of the population parameter when computing an estimate based from this sample. The more accurate we demand for the resulting estimate, the larger the sample size is required.
2. ***Degree of confidence*** (1 – α) of the accuracy of the estimate. This refers to the probability that the amount of error of the estimate will not exceed the margin of error (E).

3. ***Variability in the population***. The more homogeneous is the population, the smaller the sample size necessary to achieve the same level of confidence that the error of estimate will not exceed a certain amount. This variability is usually given by the standard deviation of the population and can be based on prior knowledge of

Sample Size for Estimating a Population Mean

To be (1 – α)100% confident that the estimate for the population mean is within E units of the true value, the minimum *sample size n* is given the formula below:

$$n = \left(\frac{Z_{\frac{\alpha}{2}} \bullet \sigma}{E}\right)^2$$

where: $Z_{\frac{\alpha}{2}}$ = z – value that leaves an area of $\frac{\alpha}{2}$ to the right

σ = sample standard deviation

E = maximum allowed error in the estimate

the population or through a pilot study.

Example 1: A college dean wishes to estimate the average number of hours students spend doing homework per week. The standard deviation from a previous study is 4.5 hours. How large a sample must be selected if he wants to be 95% confident of finding whether the true mean differs from the sample mean by 2.1 hours?

Solution:

Given: $\overline{x}$ = 2.1 hours

σ = 4.5 hours

level of confidence = 95%

Formula:

$$n = \left(\frac{Z_{\frac{\alpha}{2}} \bullet \sigma}{E}\right)^2$$

To locate $z_{\frac{\alpha}{2}}$, use the t – distribution table search for α = 0.05 (for two tailed test) and the degrees of freedom look for ∞

df \ α	
One tail	**0.025**
Two tail	**0.05**
df	
1	12.71
2	4.303
⋮	⋮
∞	1.960

Since α = 0.05 or (1 – 0.95 = 0.05), σ = 4.5, E = 2.1 and $Z_{\frac{\alpha}{2}}$ = 1.960, substituting in the formula:

$$n = \left(\frac{Z_{\frac{\alpha}{2}} \bullet \sigma}{E}\right)^2$$

$$n = \left(\frac{1.960 \bullet 4.5}{2.1}\right)^2 = 17.64 \approx 18$$

Therefore, to be 95% confident that the estimate is within 2.1 hours of the true mean, the college dean needs sample size of at least 18 students.

Example 2: A store manager wishes to be 90% confident that his estimate for the mean monthly family grocery expense is correct within ± 500 pesos. Based on prior information, he believes that the monthly family grocery expense follows a normal distribution, and has arrived at an estimate of Php 9, 000 for the standard deviation. Determine the minimum number of families to be taken as sample to meet his criteria.

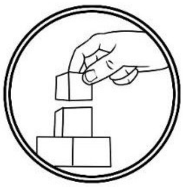
Solution: It is given that α = 0.1 so $z_{\frac{\alpha}{2}} = z_{0.05}$ = 1.645, σ = 9000, E = 500

Substituting these values into the sample size formula:

$$n = \left(\frac{z_{\frac{\alpha}{2}} \bullet \sigma}{E}\right)^2 = \left(\frac{1.645 \bullet 9000}{500}\right)^2 = 876.75 \approx 877$$

Therefore, he needs at least 877 families as sample.

Yehey, you already finished discovering. Now you are ready to Explore. Enjoy ☺



Explore

Activity 1: Decode Me!

Directions: Match Column A with the correct answer in Column B then write the letter of the correct answer at the blank before each number. Decode the secret message below using the letters of the answers.

	COLUMN A	COLUMN B
______ 1.	A random sample of size 14 is drawn from a normal population. The summary statistics are $\overline{x}$ = 933 and s = 18. Construct a 80% confidence interval for the population mean μ.	C. $68.10 < \mu < 69.10$
______ 2.	A random sample of size 28 is drawn from a normal population. The summary statistics are $\overline{x}$ = 68.6 and s = 1.28. Construct a 95% confidence interval for the population mean μ.	D. $219.86 < \mu < 258.14$
______ 3.	To estimate the number of calories in a cup of diced chicken breast meat, the number of calories in a sample of four separate cups of meat is measured. The sample mean is 211.8 calories with sample standard deviation 0.9 calorie. Assuming the caloric content of all such chicken meat is normally distributed, construct a 95% confidence interval for the mean number of calories in one cup of meat.	E. $15.97 < \mu < 20.03$
______ 4.	A college athletic program wishes to estimate the average increase in the total weight an athlete can lift in three different lifts after following a particular training program for six weeks. Twenty-five randomly selected athletes when placed on the program exhibited a mean gain of 47.3 lb with standard deviation 6.4 lb. Construct a 90% confidence interval for the mean increase in lifting capacity all athletes would experience if placed on the training program. Assume increases among all athletes are normally distributed.	G_1. $25.10 < \mu < 25.40$

______ 5.	In order to estimate the speaking vocabulary of three-year-old children in a particular socioeconomic class, a sociologist studies the speech of four children. The mean and standard deviation of the sample are $\overline{x}$ =1120 and s=215 words. Assuming that speaking vocabularies are normally distributed, construct an 80% confidence interval for the mean speaking vocabulary of all three-year-old children in this socioeconomic group.	G_2. 39.43 < μ < 44.57
______ 6.	A thread manufacturer tests a sample of eight lengths of a certain type of thread made of blended materials and obtains a mean tensile strength of 8.2 lb with standard deviation 0.06 lb. Assuming tensile strengths are normally distributed, construct a 90% confidence interval for the mean tensile strength of this thread.	L. 210.37 < μ < 213.23
______ 7.	An airline wishes to estimate the weight of the paint on a fully painted aircraft of the type it flies. In a sample of four repaintings the average weight of the paint applied was 239 pounds, with sample standard deviation 8 pounds. Assuming that weights of paint on aircraft are normally distributed, construct a 99.8% confidence interval for the mean weight of paint on all such aircraft.	0_1. 108.56 < μ < 121.44
______ 8.	A sample of 26 women's size 6 dresses had mean waist measurement 25.25 inches with sample standard deviation 0.375 inch. Construct a 95% confidence interval for the mean waist measurement of all size 6 women's dresses. Assume waist measurements are normally distributed.	0_2. 8.16 < μ < 8.24
______ 9.	In a study of dummy foal syndrome, the average time between birth and onset of noticeable symptoms in a sample of six foals was 18.6 hours, with standard deviation 1.7 hours. Assuming that the time to onset of symptoms in all foals is normally distributed, construct a 98% confidence interval for the mean time between birth and onset of noticeable symptoms.	0_3. 25.69 < μ < 30.31
____ 10.	Compute the 98% interval estimate of μ, given	0_4. 777.94 < μ <

	$\sigma = 3$, $n = 15$ and $\bar{x} = 18$	1462.07
____ 11.	Compute the 95% interval estimate of μ, given $s = 9$, $n = 10$ and $\bar{x} = 115$	R_1.926.51 < μ < 939.49
_____12.	What is the 90% interval estimate of μ given $\sigma = 12.5$, $n = 64$ and $\bar{x} = 42$?	R_2. 8.39 < μ < 11.61
_____13.	What is the 98% interval estimate of μ given $s = 15$, $n = 10$ and $\bar{x} = 21$	S. 7.62 < μ < 34.38
____ 14.	Compute the 99% interval estimate of μ, given $s = 3$, $n = 15$ and $\bar{x} = 28$.	T. 45.11 < μ < 49.49
____ 15.	Find the 70% interval estimate of μ given $s = 4.5$, $n = 7$ and $\bar{x} = 10$.	W. 16.26 < μ < 20.94

SECRET MESSAGE

8	1	11	9		2	3	6	13	10	15

4	14		12	5	7

Activity 2: Solve Me!

Directions: Analyze and Solve the following problems completely. Write your solution in a coupon bond.

1. The mean and standard deviation for the quality grade – point averages of a random sample of 36 college seniors are calculated to be 2.6 and 0.3 respectively. Find the

   a. 95% confidence interval for the mean of the entire senior class

   b. 99% confidence interval for the mean of the entire senior class

2. A sample of 10 measurements of the length of a chip gave a mean of 2.36 cm and a standard deviation of 0.06 cm. Find the

   a. 95% confidence limits

   b. 99% confidence limits

3. A researcher of a cardboard manufacturing company would like to know the estimated thickness of the cardboard a machine produces. How many cardboards should he measure if he wants to be 99% confident that the estimate is accurate to 1 mm. Study shows that the standard deviation is 3mm.

Oopss.. Before you will go to your final phase, finish first the Deepen part. Go forth and let the force be with you.

Deepen

Activity 1: Weigh to Perform Me!

Directions: Read and analyze the situation below, solve the following problem completely and accurately.

Performance Task

Just recently, you as a school biostatistician made an analysis pertaining to the heights of students. This time, your school clinic supervisor tasked you to make an analysis of the mean weight of students. You started with one class of grade 11 students, and collect their weights in kilogram (kg). To select your sample, you write each weight measurement on a slip of paper put the slips into a small bag, mix them up, pick out 5 slips at random, and write down the numbers you picked. You return the slips in the bag and repeat the process until you have 20 samples of 5 numbers. The following will be part of your analysis:

a. For each of the samples of weights obtained, compute the resulting point estimate and 95% margin of error for the mean weight of students in the class.

b. For each of the 20 samples, construct the resulting 68% confidence interval for the average weight of all the students in the class.

c. Compute the mean and standard deviation of the population weights.

The important questions you would like to answer are as follows:

- How many of the 20 confidence intervals would you expect to contain the true mean?
- Based from your results, how many actually contain the true mean?

Submit a written report. Make sure that it is accurate, detailed, neat and organized.

Rubrics for the Task:

Categories	Excellent (4)	Satisfactory (3)	Developing (2)	Beginning (1)
Representation	Shows a complete understanding of the concept of the report and the knowledge of estimation of parameters	Shows a partial understanding of the concept of the report and the knowledge of estimation of parameters	Shows limited understanding of the concept of the report and the knowledge of estimation of parameters	Not evident
Computation and Solution	Computation is correct and leads to the correct answer	Computation is correct but does not lead to the correct answer	Computation is incorrect and does not relate to the task.	Not evident
Communication	Explained the analysis of the report clearly and accurately.	Explained the analysis of the report clearly.	Explained the analysis of the report, but there some parts which are not clear.	Not evident
Neatness, Completeness and Organization of the Report	The report is very neat, complete and organized.	The report is neat, complete and organized.	The report is not that neat, complete and organized.	Not evident

Very well done indeed. But to finish this module, you must finish your assessment test. Good luck.

Gauge

Directions: Read and analyze each question carefully and write your answer in a separate sheet of paper.

1. Which of the following is not required to apply the t – distribution?
   A. $n < 30$
   B. σ is known.
   C. σ is unknown
   D. population is nearly normally distributed

2. When the population standard deviation is unknown and the sample size is greater than 30, what table value should be used in computing a confidence interval for the mean?
   A. z B. t C. x^2 D. σ

3. What is the notation of the confidence level?
   A. α B. $1 - \alpha$ C. $100\%(1 - \alpha)$ D. $100\%(\alpha)$

4. Which of the following assumptions is ***NOT*** needed to use the t – distribution to make confidence interval for μ?
   A. The sample size is at least 10.
   B. The sample size is less than 30.
   C. The population standard deviation is unknown
   D. The population from which the sample is taken s approximately normally distributed.

5. What is the t – value for a 95% confidence interval with n = 10?
   A. 1.812 B. 2.228 C. 2.262 D. 2.821

6. Which of the following is the value of $t_{0.01}$ with 15 degrees of freedom?
   A. 2.602 B. 2.624 C. 2.650 D. 2.947
7. What is the 90^{th} percentile in a t-distribution with 12 degrees of freedom?
   A. 1.356 B. 1.363 C. 1.782 D. 1.796

8. What is the t – value for a 95% confidence level with n = 14 for a two tailed test?
   A. 2.145 B. 2.160 C. 2.650 D. 2.624
9. What is the 95% interval estimate of μ given $s = 4.2$, n = 20, and $\bar{x} = 22.5$?
   A. $24.23 < \mu < 28.89$ C. $20.53 < \mu < 24.47$
   B. $19.53 < \mu < 23.47$ D. $20.63 < \mu < 27.63$

10. What is the 98% interval estimate of μ given $s = 3$, n=15 and $\bar{x}$ =18.

A. 15.57 < μ < 20.53 C. 15.97 < μ < 20.03

B. 14.97 < μ < 21.03 D. 14.97 < μ < 21.03

11. A sample of 50 learners showed a mean height of 60 inches. If it is known that the standard deviation of heights of learners is 2.5 inches, what is the 70% confidence interval estimate for the height of all the learners?

A. 59.64 < μ < 60.36 C. 59.32 < μ < 60.48

B. 58.52 < μ < 61.48 D. 59.52 < μ < 62.48

12. A random sample of 12 graduates of a certain secretarial school typed an average of 73.9 words per minute (wpm) with a standard deviation of 8.7 wpm. Assuming that the number of words typed per minute is approximately normally distributed, what is the 95% interval for the average number of words typed by all the graduates of this school.

A. 58.52 < μ < 61.48 C. 68.52 < μ < 79.48

B. 68.52 < μ < 79.48 D. 68.37 < μ < 79.43

13. Given a standard deviation of 10.4, confidence interval of mean 95% and a margin of error or maximum allowable deviation of 4.35, what is the appropriate sample size?

A. 19 B. 20 C. 21 D. 22

14. Twenty packages of mangoes are inserted in a box for shipment. To test the weight of the boxes, a few were checked. The mean weight is 9.3 kg, with a standard deviation of 0.23 pounds. How many boxes must the processor sample to be 95% confident that the sample mean does not differ from the population mean by more than 0.1 kg?

A. 19 B. 20 C. 21 D. 22

15. A college dean wishes to estimate the average number of hours students spend doing homework per week. The standard deviation from a previous study is 4.5 hours. How large a sample must be selected, if he wants to be 95% confident of finding whether the true mean differs from the sample mean by 2.1 hours?

A. 16 B. 17 C. 18 D. 20

Congratulations! You made it! You are now a Jedi Master.

References

Printed Materials:

- Melosantos, Luis Allan B. et. al. (2016). Math Connections in the Digital Age Statistics and Probability. Quezon City, Philippines: Sibs Publishing Home, Inc.
- Altares, Priscilla S. et. al. (2005). Elementary Statistics with Computer Applications. Quezon City, Philippines: Rex Bookstore. Inc.
- Sirug, Winston S. (2011). Basic Probability and Statistics: A Step by Step Approach. Manila, Philippines: Mindshapers Co., Inc.
- Chan Shio, Christian Paul O. and Marie Angeli T. Reyes. (2017). Statistics and Probability for Senior High School. Quezon City, Philippines: C & E Publishing, Inc.
- Cabero, Jonathan B. et. al. (2013). Business Statistics. Mandaluyong City, Philippines: Anvil Publishing, Inc.
- Walpole, Ronald E. (1982). Introduction to Statistics. New Jersey, USA: Prentice Hall International, Inc.

Website:

- https://www.statisticshowto.com/probability-and-statistics/t-distribution/
- https://mathworld.wolfram.com/Studentst-Distribution.html
- https://www.itl.nist.gov/div898/handbook/eda/section3/eda3672.htm
- http://www.stat.yale.edu/Courses/1997-98/101/confint.htm#:~:text=For%20a%20population%20with%20unknown%20mean%20and%20unknown%20standard%20deviation,t(n%2D1)
- https://www.statisticshowto.com/probability-and-statistics/confidence-interval/