11

SENIOR HIGH SCHOOL

DEPARTMENT OF EDUCATION
SCHOOLS DIVISION OF NEGROS ORIENTAL
REGION VII
Kagawasan Ave., Daro, Dumaguete City, Negros Oriental

STATISTICS and PROBABILITY

Quarter 3 – Module 6
Central Limit Theorem

ALTERNATIVE DELIVERY MODE
ADM

NegOr_Q3_Stat_and_Prob11_Module6_v2

Statistics & Probability – Grade 11
Alternative Delivery Mode
Quarter 3 – Module 6: Central Limit Theorem
Second Edition, 2021

Published by the Department of Education
Secretary: Leonor Magtolis Briones
Undersecretary: Diosdado M. San Antonio

Development Team of the Module

Writer: Littie Beth S. Bernadez

Editors: Didith T. Yap & Rickleoben V. Bayking

Reviewer: Mercyditha D. Enolpe

Layout Artist: Jerry Mar B. Vadil

Management Team: Senen Priscillo P. Paulin, CESO V | Elisa L. Baguio, EdD
Joelyza M. Arcilla EdD, CESE | Rosela R. Abiera
Marcelo K. Palispis JD, EdD | Maricel S. Rasid
Nilita L. Ragay, EdD | Elmar L. Cabrera

Printed in the Philippines by ________________________

Department of Education –Region VII Schools Division of Negros Oriental

Office Address: Kagawasan, Ave., Daro, Dumaguete City, Negros Oriental
Telephone #: (035) 225 2376 / 541 1117
E-mail Address: negros.oriental@deped.gov.ph

Introductory Message

This Self-Learning Module (SLM) is prepared so that you, our dear learners, can continue your studies and learn while at home. Activities, questions, directions, exercises, and discussions are carefully stated for you to understand each lesson.

Each SLM is composed of different parts. Each part shall guide you step-by-step as you discover and understand the lesson prepared for you.

Pre-tests are provided to measure your prior knowledge on lessons in each SLM. This will tell you if you need to proceed on completing this module or if you need to ask your facilitator or your teacher's assistance for better understanding of the lesson. At the end of each module, you need to answer the post-test to self-check your learning. Answer keys are provided for each activity and test. We trust that you will be honest in using these.

In addition to the material in the main text, Notes to the Teacher are also provided to our facilitators and parents for strategies and reminders on how they can best help you on your home-based learning.

Please use this module with care. Do not put unnecessary marks on any part of this SLM. Use a separate sheet of paper in answering the exercises and tests. And read the instructions carefully before performing each task.

If you have any questions in using this SLM or any difficulty in answering the tasks in this module, do not hesitate to consult your teacher or facilitator.

Thank you.

What I Need to Know

This module was designed to provide you with fun and meaningful opportunities for guided and independent learning at your own pace and time. You will be enabled to process the contents of the learning resource while being an active learner.

The module is intended for you to illustrate the Central Limit Theorem and define the sampling distribution of the sample mean using the Central Limit Theorem.

After going through this module, you are expected to solve problems involving sampling distributions of the sample mean.

What I Know

PRE-ASSESSMENT

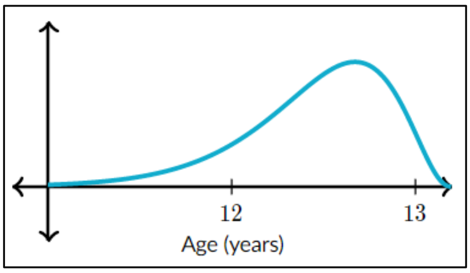
Read and understand each statement carefully. Write the letter of your answer in your activity sheets/notebook.

1. Ms. Martino always gives a very easy 10-question quiz in the first week of her classes. Over the years, the number of questions that students answer correctly on these quizzes has been strongly skewed to the left with a mean of 9 correct answers and a standard deviation of about 2.5 correct answers. Suppose we took random samples of 4 students and calculated $\bar{x}$ as the sample mean number of questions that the students answered correctly. We can assume that the students in each sample are independent. What would be the shape of the sampling distribution of $\bar{x}$?

A. Skewed to the left C. Skewed to the right

B. Approximately normal D. Unknown

2. The ages of the 175 players in a junior basketball league team have the following distribution:



Suppose that we were to take random samples of 10 players from this population and calculate the sample mean age of the players in each sample. What will be the shape of the sample mean?

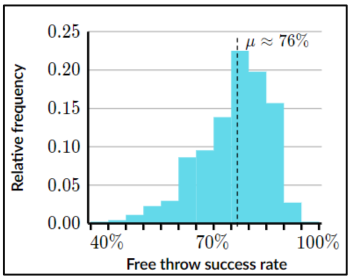
A. Skewed to the left

B. Approximately normal

C. Skewed to the right

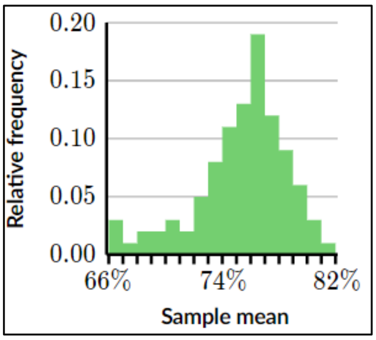
D. Unknown

3. A professional volleyball league records each player's success rate non their free throw attempts. Here is the distribution of success rates for the approximately 400 players in the league who attempted at least 20 free throws in a recent season:

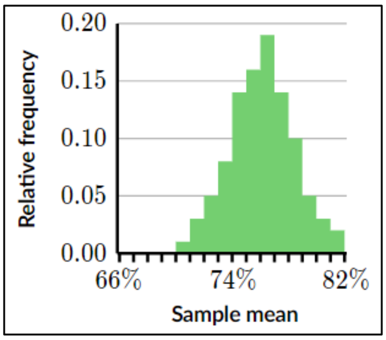


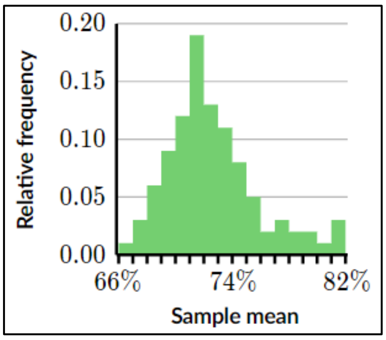
Suppose that we took random samples of 30players from this population and calculated the sample mean success rate $\overline{x}$ of the players in each sample. Which graph shows the most reasonable approximation of the sampling distribution of $\overline{x}$?

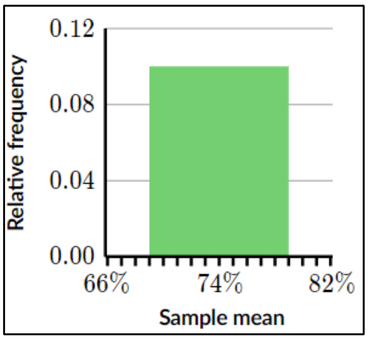
A.



C.



B. 

D. 

4. The national average NAT score (for Verbal and Math) is 1028. If we assume a normal distribution with sd = 92, what is the probability that a randomly selected score exceeds 1200?

A. 0.0257 B. 0.0307 C. 0.4693 D. 0.5678

5. An exclusive college desires to accept only the top 10% of all graduating seniors based on the results of a national placement test. This test has a mean of 500 and a standard deviation of 100. Find the cut off score for the exam. Assume the variable is normally distributed.

A. 500 B. 528 C. 628 D. 728

Lesson 1 The Central Limit Theorem

What's In

REVIEW

Let's say, we have a quantitative data set from a population with mean μ and standard deviation σ. The model for the theoretical sampling distribution of means of all random samples of size *n* has the following properties:

1. The mean of the sampling distribution of means is μ.
2. The standard deviation of the sampling distribution of means is $\sigma/\sqrt{n}\sigma/n$.
   a. Notice that as *n* grows, the standard deviation of the sampling distribution of means shrinks.

What's New

ACTIVITY 1. COMPLETE ME!

A rowing team consists of four rowers who weigh 152, 156, 160, and 164 pounds. Find all possible random samples with replacement of size two and compute the sample mean for each one.

x	$\bar{x}$	**x**	$\bar{x}$
152,152	152	160,152	156
152,156	154	160,156	
152,160	156		
152,164	158		
156,152			
156,156			
156,160			
156,164			

ACTVITY 2. CONSTRUCT ME!
Construct a probability distribution of the sample mean for the sample size two drawn from the population of four rowers. The probability distribution is:

x	152	154	156	158	160	162	164
P(x)	$\frac{1}{16}$	$\frac{2}{16}$		$\frac{4}{16}$	$\frac{\ }{16}$		

What is It

The figure is a histogram from the previous activity. It shows a side-by-side comparison of the histogram for the original population and a histogram for the distribution. Whereas, the distribution of the population is uniform, the sampling distribution of the mean has a shape approaching the shape of the familiar bell curve. This phenomenon of the sampling distribution of the mean taking on a bell shape even though the population distribution is not bell-shaped happens in general. Here is a somewhat more realistic example.

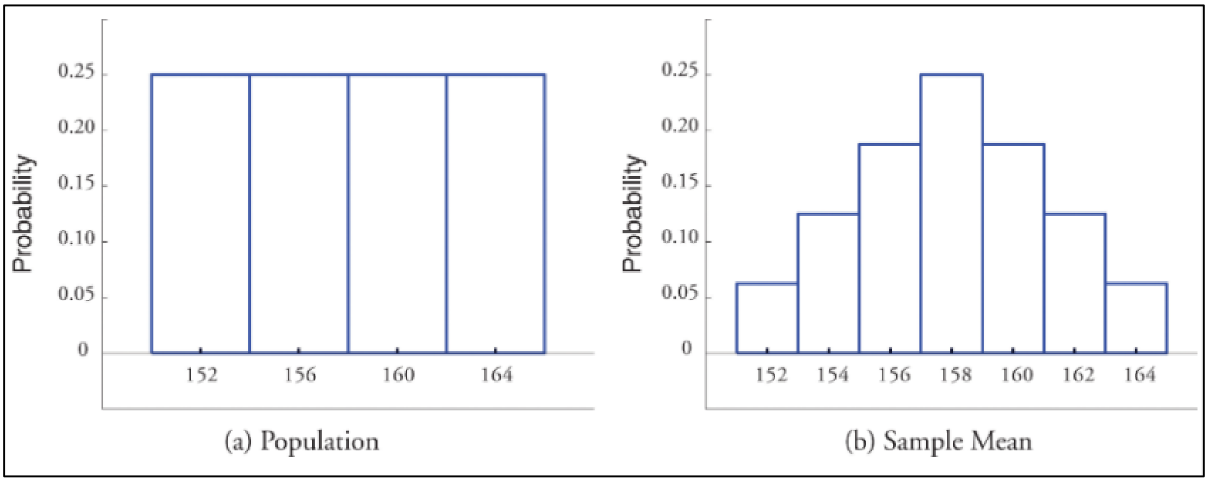


Figure 6.2.16.2.1: Distribution of a Population and a Sample Mean

Suppose we take samples of size 1, 5, 10, or 20 from a population that consists entirely of the numbers 0 and 1, half the population 0, half 1, so that the population mean is 0.5.

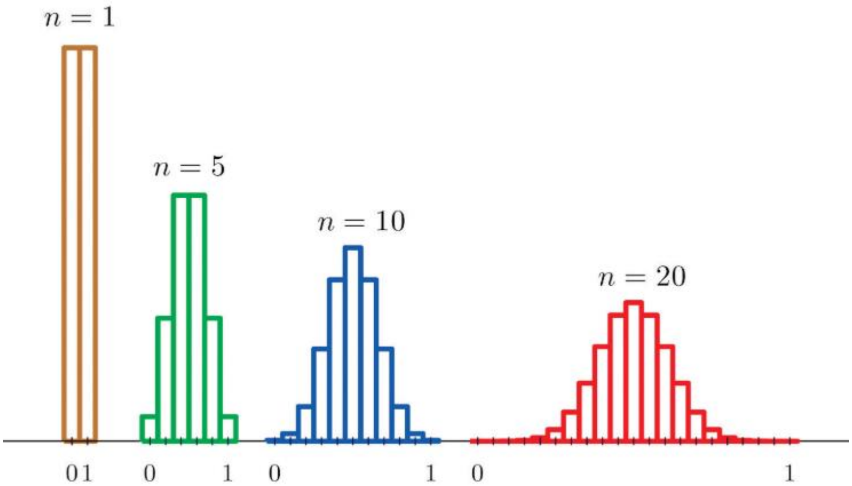


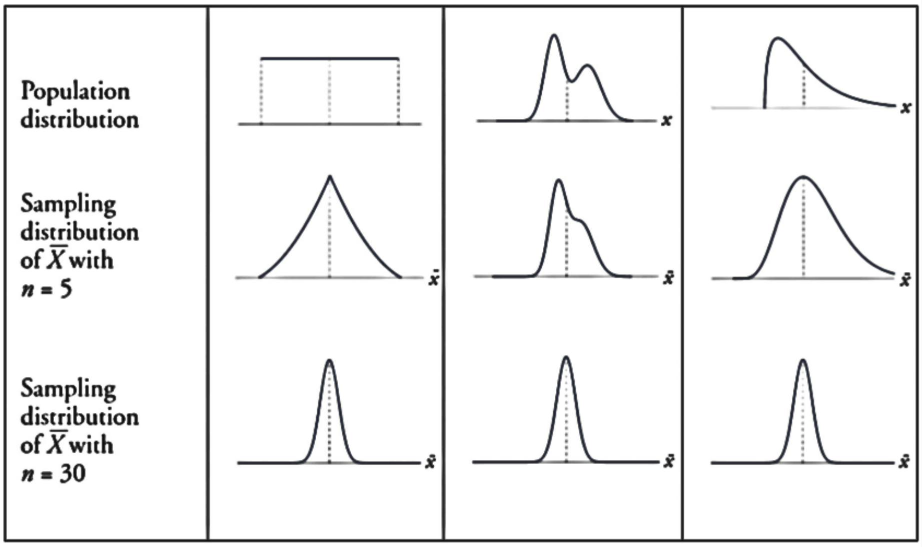
Figure 6.2.2: Distributions of the Sample Mean

(Tracyrence, 2021)

As n increases, the sampling distribution of $\bar{\boldsymbol{X}}$ evolves in an interesting way: the probabilities on the lower and the upper ends shrink and the probabilities in the middle become larger in relation to them. If we were to continue to increase, then the shape of the sampling distribution would become smoother and more bell-shaped.

The **Central Limit Theorem** tells us that as sample sizes get larger, the sampling distribution of the mean will become normally distributed, *even if the data within each sample are not normally distributed*. This implies that the concept of the standard error of the mean is very significant, since it measures the degree of accuracy of the sample mean $\bar{X}$ as an estimate of the population mean μ. It is a good estimate if the standard error is small or close to 000, and poor if the standard error is large, and this standard error is dependent on the sample size n. Specifically, as n increases, the standard error decreases. Thus, in order to attain a relatively good estimate of μ, n must be sufficiently large.

As the sample size n increases, the shape of the distribution of the sample means taken from a population with mean m and standard deviation of s will approach a normal distribution. Thus, the distribution will have a mean m and standard deviation $\frac{s}{\sqrt{n}}$.



CENTRAL LIMIT THEOREM. *If a random samples of size n are drawn from a population (finite or infinite), then as n becomes larger, the sampling distribution of the mean approaches the normal distribution, regardless of the form of the population distribution.*

The Central Limit Theorem assures that no matter what the shape of the population distribution is, the sampling distribution of the mean is closely normally distributed whenever n is large. Consequently, it justifies the use of the formula for the z-values

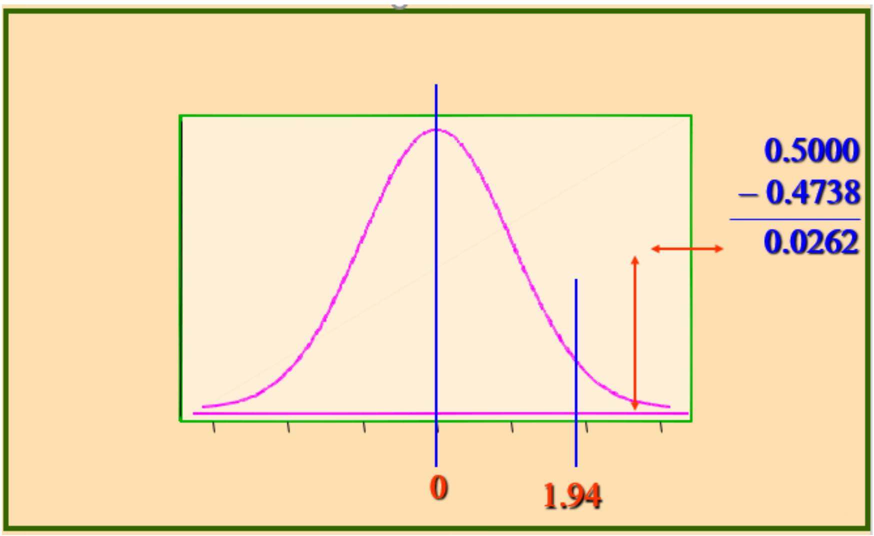
$$z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

when computing for the probability that $\bar{X}$ will take n a value within a given range in the sampling distribution of $\bar{X}$. Consider a sample that is sufficiently large and hence approximately normal whenever $n \geq 30$. It must be remembered, however, that if the population distribution is a normal distribution, the sampling distribution will always have a normal curve, no matter how small n is.

Example 1. A.C. Neilsen reported that children between the ages of 2 and 5 watch an average of 25 hours of TV per week. Assume the variable is normally distributed and the standard deviation is 3 hours. If 20 children between the ages of 2 and 5 are randomly selected, find the probability that the mean of the number of hours they watch TV is greater than 26.3 hours.

The standard deviation of the sample means is $\frac{s}{\sqrt{n}} = \frac{3}{\sqrt{20}} = 0.671$.

- The z-value is z = (26.3 - 25)/0.671= 1.94.
- Thus P(z > 1.94) = 0.5 – 0.4738 = 0.0262. That is, the probability of obtaining a sample mean greater than 26.3 is 0.0262 = 2.62%.



Example 2. Let $\bar{X}$ be the mean of a random sample of size 50 drawn from the population with mean 112 and standard deviation of 40.

a. Find the mean standard deviation of $\bar{X}$.
b. Find the probability that $\bar{X}$ assumes a value between 110 and 114.
c. Find the probability that $\bar{X}$ assumes a value greater than 113.

Solution.

a. $\mu_{\bar{X}} = \mu = 112$ and $\sigma_X = \frac{\sigma}{\sqrt{n}} = \frac{40}{\sqrt{50}} = 5.65685$

b. Since the sample size is at least 30, the Central Limit Theorem applies: $\overline{X}$ is approximately normally distributed. We compute probabilities.

$$P(110 < \overline{X} < 114) = P\left(\frac{110 - \mu_{\overline{X}}}{\sigma_{\overline{X}}} < \overline{Z} < \frac{114 - \mu_{\overline{X}}}{\sigma_{\overline{X}}}\right)$$

$$= P\left(\frac{110 - 112}{5.65685} < \overline{Z} < \frac{114 - 112}{5.65685}\right)$$

$$= P(-0.35 < \overline{Z} < 0.35)$$
$$= 0.6368 - 0.3632$$
$$= 0.2736$$

c. $P(\overline{X} > 113)$

$$P(\overline{X} > 113) = P(Z > \frac{113 - \mu_{\overline{X}}}{\sigma_{\overline{X}}})$$

$$= P(Z > \frac{113 - 112}{5.65685})$$

$$= P(Z > 0.18)$$

$$= 1 - 0.5714$$

$$= 0.4286$$

What's More

Enrichment Activity:

Answer the following.

1. The average daily jail population in the United States is 706,242. If the distribution is normal and the standard deviation is 52,145, find the probability that on a randomly selected day, the jail population is

   a. Greater than 750,000 *b.* Between 600,000 and 700,000

What I Have Learned

The Central Limit Theorem states that the sampling distribution of the sample means approaches a normal distribution as the sample size gets larger — *no matter what the shape of the* population *distribution*. This fact holds especially true for sample sizes over 30.

The Central Limit Theorem is important for statistics because it allows us to safely assume that the sampling distribution of the mean will be normal in most cases. This means that we can take advantage of statistical techniques that assume a normal distribution

What I Can Do

Read and understand each statement carefully. Write your answer in your activity sheets/notebook. Show all the necessary solutions. Be guided with the rubrics provided in scoring your answers.

1. A survey found that the American family generates an average of 17.2 pounds of glass garbage each year. Assume the standard deviation of the distribution is 2.5 pounds. Find the probability that the mean of a sample of 55 families will be between 17 and 18 pounds.

2. The mean serum cholesterol level of a large population of overweight children is 220 milligrams per deciliter (mg/dl), and the standard deviation is 16.3 mg/dl. If a random sample of 35 overweight children is selected, find the probability that the mean will be between 220 and 222 mg/dl. Assume the serum cholesterol level variable is normally distributed.

3. The average fuel efficiency of U.S. light vehicles (cars, SUVs, minivans, vans, and light trucks) for 2005 was 21 mpg. If the standard deviation of the population was 2.9 and the gas ratings were normally distributed, what is the probability that the mean mpg for a random sample of 25 light vehicles is under 20?

RUBRICS

	4	3	2	1	0
Understanding the Problem	Complete under-standing of the problem	Misinterprets minor part of the problem	Misinterprets major part of the problem	Completely misinterprets the problem	No attempt
Solving the Problem	A plan that could lead to a correct solution with no arithmetic errors	Substantially correct procedure with minor omission or procedural error	Partially correct procedure but with major fault	Totally inappropriate plan	No attempt
Answering the Problem			Correct solution	Copying error; computational error, partial answer for problem with multiple answers; no answer statement; answer labeled incorrectly	No answer or wrong answer based upon an inappropriate plan

Source: Sztela, Walter and Nicol, Cynthia. Evaluating Problem Solving in Mathematics. Educational Leadership, May 1992, pp. 42-45.

Assessment

Read and understand each statement carefully. Write the letter of your answer in your activity sheets/notebook.

1. The numerical population of grade point averages at a college has mean 2.61 and standard deviation 0.5. If a random sample of size 100 is taken from the population, what is the probability that the sample mean will be between 2.51 and 2.71?

A. 0.9544 B. 0.9244 C. 0.944 D. 0.9054

2. A prototype automotive tire has a design life of 38,500 miles with a standard deviation of 2,500 miles. Five such tires are manufactured and tested. On the assumption that the actual population mean is 38,500 miles and the actual population standard deviation is 2,500 miles, find the probability that the sample mean will be less than 36,000 miles. Assume that the distribution of lifetimes of such tires is normal.

A. 0.0251 B. 0.0105 C. 0.0125 D. 0.2510

3. An automobile battery manufacturer claims that its midgrade battery has a mean life of 50 months with a standard deviation of 6 months. Suppose the distribution of battery lives of this particular brand is approximately normal. On the assumption that the manufacturer's claims are true, find the probability that a randomly selected battery of this type will last less than 48 months.

A. 0.3707 B. 0.2525 C. 0.2565 D. 0.3525

4. On the same assumption in problem number 3, find the probability that the mean of a random sample of 36 such batteries will be less than 48 months.

A. 0.0028 B. 0.2828 C. 0.0208 D. 0.0228

5. A population of 29 year-old males has a mean salary of ₱29,321 with a standard deviation of ₱2,120. If a sample of 100 men is taken, what is the probability their mean salaries will be less than ₱29,000?

A. 0.95 B. 0.10 C. 0.08 D. 0.07

Answer Key

PRE TEST

1. A
2. A
3. C
4. B
5. B

ENRICHMENT

1.a. greater than 750,000

$P(X > 750\ 000) = 0.2004$

$Z=0.84$

2.b. between 600,000 and 700,000

$P(600\ 000 < X < 700\ 000) = 0.2004$

$Z= -2.04$ $z = -0.12$

WHAT I CAN DO

1. 0.4687 or 46.87%
2. 0.2673 or 26.73%
3. 0.0427 or 4.27%

ASSESSMENT

1. A
2. C
3. A
4. D
5. D

References

Alfin, Farhan. "Probability and the Normal Distribution." *slideshare.net.* August 3, 2021. https://www.slideshare.net/FarhanAlfin/ch3-probability-and-the-normal-distribution/ (accessed January 4, 2022).

Candela, Lumen. “Sample Means”. ttps://courses.lumenlearning.com/wmopen-concepts-statistics/chapter/distribution-of-sample-means-3-of-4/ (accessed Janaury 4, 2022).

Febre, Jr.F. (199). Introduction to Statistics. Phoenix Press.

Gauthmath.com. June 14, 2021. https://www.gauthmath.com/solution/Construct-a-probability-distribution-of-the-sample-mean-for-the-sample-size-two--1702782641833990 (accessed January 3, 2022).

"The Central Limit Theorem." *coursehero.com.* June 14, 2021. https://www.coursehero.com/file/108733634/The-Central-Limit-Theoremdocx/ (accessed January 3, 2022).

"The Mean and Standard Deviation of the Sample Mean." *http://math.fdltcc.edu/.* n.d. http://math.fdltcc.edu/wetherbee/books/m1030/pdf/6.1.pdf (accessed January 3, 2022).

"The Sampling Distribution of the Sample Mean." *stats.libretexts.org.* February 25, 2021. https://stats.libretexts.org/Bookshelves/Introductory_Statistics/Book%3A_Introductory_Statistics_(Shafer_and_Zhang)/06%3A_Sampling_Distributions/6.02%3A_The_Sampling_Distribution_of_the_Sample_Mean (accessed January 4, 2022).

Tracyrence. *medium.com.* August 28, 2021. https://medium.com/geekculture/how-increasing-sample-size-will-improve-gaussian-distribution-ae2bf5c5abaf (accessed January 10, 2022).

For inquiries or feedback, please write or call:

Department of Education – Schools Division of Negros Oriental
Kagawasan, Avenue, Daro, Dumaguete City, Negros Oriental

Tel #: (035) 225 2376 / 541 1117
Email Address: negros.oriental@deped.gov.ph
Website: lrmds.depednodis.net