7.1: The Sampling Distribution of Sample Means

In the last unit, we used sample proportions to make estimates and test claims about population proportions. In this unit, we will focus on sample means from random samples. We begin by examining the variability of sample means that are randomly obtained from a population. We will use these ideas to learn about sampling distributions of sample means.

Exploring Samples of Acorn Weights

In Texas, oak trees are important to the overall health of the ecosystem. Acorns are the seeds of oak trees. Many birds and small mammals eat the acorns that come from oak trees. A decline in the growth of new oak trees could have a serious impact on local animal and plant life.

Botanists (scientists who study plant life) are especially interested in knowing the weights of acorns in a region. Botanists use information about the weights of acorns to help them predict the future growth of oak trees.

A group of students in Austin, Texas gathered acorns to study their weights. The weights of 400 acorns were measured in grams, and ranged from 0.7g to 6.7g.

Collect the Sample

Here is a list of 9 randomly selected acorn weights (in grams): $W = [4.2, 3.2, 3.9, 4.6, 3.5, 3.9, 4.6, 3.3, 2.6]$. In this sample, we can use desmos to compute the sample mean, $\bar{x}$, by using the $\text{mean}(W)$ function after entering the set labeled W. The sample mean weight for this sample is around 3.67 grams.

Given are 5 samples of 9 acorn weights. Pick one of the samples to compute the mean acorn weight for. Round the sample mean to two decimal places.

Sample 1. [2.5, 3.8, 4, 3.4, 4.5, 5.7, 1, 4.4, 2.6]

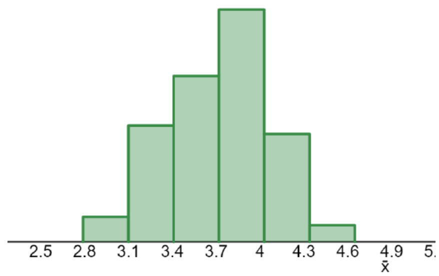
Sample 2. [3.4, 2.4, 3.2, 2, 1.4, 3.9, 0.7, 4.6, 3.8]

Sample 3. [3.8, 2.1, 1.4, 3.9, 4.6, 4.4, 3.4, 4.3, 3.9]

Sample 4. [2.8, 3.2, 3.3, 3.5, 4.8, 3.7, 1, 4.3, 6.6]

Sample 5. [5.7, 3.3, 0.7, 2.5, 4.6, 2.9, 2.6, 4.9, 2.9]

Below is a histogram of mean acorn weight from 80 random samples. Use this distribution of sample means to complete the following questions:



Images are created with the graphing calculator, used with permission from Desmos Studio PBC.

1. What is the largest average acorn weight?

2. What is the range of sample means?

3. What range of sample means occurred most frequently?

4. Describe the center, shape, and spread of the distribution of sample means.

5. Sample means are estimates of the population mean, so a distribution of sample means is a distribution of estimates. Any deviation that exists between a sample mean and the population mean is an error. This is why we call the standard deviation of all sample means the standard error. Estimate the standard error of the distribution of sample means.

Central Limit Theorem for Sample Means

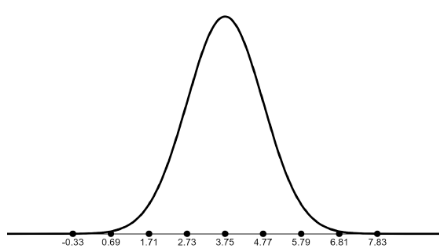
We will now shift our attention from distributions of sample means to the sampling distribution of sample means. The sampling distribution of sample means is the distribution of all possible sample means from random samples of the same size. The Central Limit Theorem for Sample Means states that: Given any population with mean μ and standard deviation σ, the sampling distribution of sample means (sampled with replacement) from random samples of size n will have a distribution that approaches normality with increasing sample size. The mean and standard error of the sampling distribution are:

$$\mu_{\bar{x}} = \mu \text{ and } \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

The criteria for the approximate normality of a sampling distribution are that either the population from which we are sampling is normal, or the sample size is greater than 30. Very non-normal populations may require samples substantially larger than 30. When a sampling distribution of sample means is approximately normal, we can use its mean and standard error to find the Z-score of any particular sample mean. The Z-score of a sample mean $\bar{x}$ from a sample of size n is found by the formula:

$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

We explored a population distribution of acorn weights from oak trees. Since we define these 400 acorns as the population, we can calculate the population mean and standard deviation. The population of acorn weights are normally distributed with a mean weight of 3.75 grams (μ) and a standard deviation of 1.02 grams (σ). The population of acorn weights is given as a desmos graph below.



Images are created with the graphing calculator, used with permission from Desmos Studio PBC.

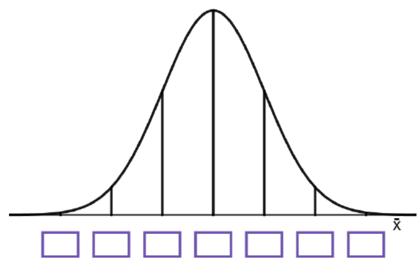
Suppose that we randomly select samples of 40 acorn weights. According to the Central Limit Theorem, since the population is normally distributed, sample means from random samples of size 40 will also be normally distributed.

6. The sample size is $n = 40$. Use the Central Limit Theorem to calculate the mean and standard error of the sampling distribution. Round the standard error to three decimal places.

   a. $\mu_{\bar{x}} = \mu =$ _____

   b. $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} =$ _____

7. We can use a normal curve to represent the sampling distribution of sample means. The boxes under the normal distribution below are one standard error apart, with the center box at the mean. Use the mean and standard error above to enter the correct values into the boxes. Use the standard error, rounded to 3 decimal places.



Images are created with the graphing calculator, used with permission from Desmos Studio PBC.

a. A particular random sample of size 40 has a mean weight of 4 grams. Plot this sample mean on the axis above. Find the Z-score for the sample mean, $\bar{x} = 4$. Round the Z-score to two decimal places.

b. Would you consider this sample mean to be unusual? Explain your answer.

c. If a sample of 40 acorns from this population is randomly selected, what is the probability that its mean weight $\bar{x}$ will be greater than 4 grams? Write your answer as a decimal rounded to three places. Write the desmos function you used to calculate the probability.

Check for Understanding

8. Everyday people watch 1 billion hours of videos on YouTube. That breaks down to every single person on earth watching YouTube videos for about 8.4 minutes per day. For U.S. teens, on any given day, the amount of time spent watching YouTube videos is approximately normal with mean 18.5 minutes and standard deviation 5.3 minutes.

    a. Find the probability that a randomly chosen U.S. teen watches YouTube for more than 25 minutes in a given day. Round to two decimal places. Show your thinking using a sketch and use desmos to do the calculation. Write any functions used in desmos to find your answer.

    b. Suppose we choose a simple random sample of 10 U.S. teens. Let $\bar{x} =$ the mean amount of time spent watching YouTube videos for the sample.

    c. Why is the sampling distribution of $\bar{x}$ normally distributed?

    d. What is the mean of the *sampling distribution* of $\bar{x}$?

    $\mu_{\bar{x}} =$

    e. Calculate the standard deviation of the sampling distribution of x. Round to three decimal places.

    $\sigma_{\bar{x}} =$

    f. Find the probability that the mean amount of time spent watching YouTube for the teens in the sample exceeds 25 minutes. Round to five decimal places. Write the desmos function used to do the calculation.

    $P(\bar{x} \quad 25) = P(Z > \_\_\_) =$