

## 6.2 One Sample Mean with Known Sigma

This section uses sample statistics (mean and standard deviation) to draw inference about a population parameter (mean,  $\mu$ ). This is the one-sample test for mean given that the population standard deviation (sigma,  $s$ ) is known. This hypothesis test is used when the population mean and standard deviation are well known has in the cases when there is a long established history about the population parameters like the mean of IQ is 100 and the standard deviation is 15. Often census data (statistics on the entire population will have well defined parameters as well as sample sizes that is very large and collected at random). For example, a doctor may have been recording the mean and variance of patient physical characteristic over a very long period of time as well has had accumulated hundreds or thousands of data measurements on the same patients' characteristic for many types of patients and have pooled or groups measurements with other doctors across the country. In the mid-1990s the mean scores on the Math section of the SAT was 460 and the standard deviation was 100. Educators have since been interested to know whether the mean has changed and to what value. This section will help educators make such judgment empirically or statistically.

If we wish to compare two means, we may ask the question, "Are these two means different?" Another way to evaluate this question is to subtract the value of one mean from the other and see if the difference is "0". Therefore, we can test the notion or hypothesis that mean one ( $\mu_1$ ) equals mean two ( $\mu_2$ ) or mean one minus mean two = 0, i.e., we assume, in both cases, that there is no difference in the means. This statement is called the *null hypothesis*.

Null Hypothesis Assumes,  $H_0: \mu_1 - \mu_2 = 0$  or  $\mu_1 = \mu_2$

The hypothesis that the null hypothesis is not true is called the **alternative hypothesis**.

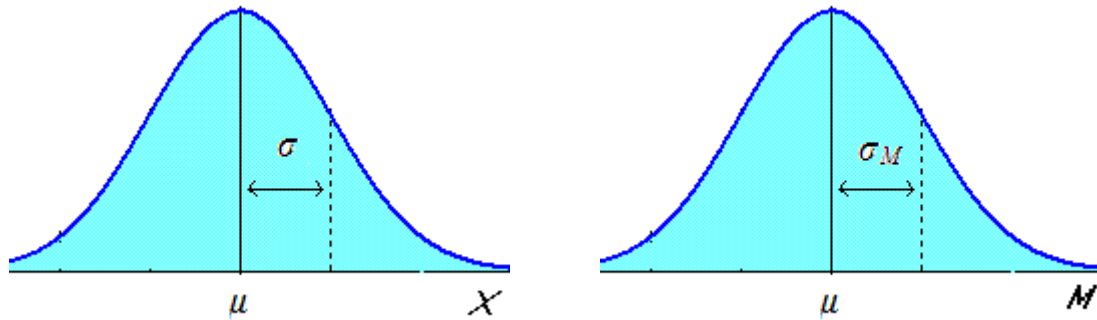
The alternative hypothesis can be stated for the above statement in one of three ways:

1.  $H_a: \mu_1 \neq \mu_2$  (non-directional)
2.  $H_a: \mu_1 > \mu_2$  (directional)
3.  $H_a: \mu_1 < \mu_2$  (directional)

### Sampling Distribution

Before we can perform hypothesis testing using sample statistics to estimate and evaluate a population parameter, we must define a few important concepts and explore some assumptions about the sample distribution being used. There are a number of ways in which we may obtain a good estimate of a population parameter from sample data. The best way is to gather and analyze a sample similar to the population that is randomly selected and is as large as the population itself. The next best way is to take many random repeated samples. If we were to take  $n$  repeated samples of a population and compute the means of each sample,  $M$ , the distribution of the sample mean would be a normal distribution with its mean very close to the population mean,  $\mu$ . The distribution of the means for these  $n$  samples will have its own standard deviation  $s_M$ , different from the population standard deviation,  $s$ . The standard deviation measures the average distance between a score  $X$  and the population mean,  $X - \mu$ . The **standard error** is the standard deviation of the sampling distribution, and it measures the average distance between a sample mean and the population mean,  $M - \mu$ . Figure 6.2.1 shows a comparison between the standard deviation,  $s$  for the distribution of scores,  $X$  and the standard error,  $s_M$  for the distribution of sample means,  $M$ .

A **sampling distribution** of the mean is the distribution of sample means from repeated sampling of a population.  
 The **standard error** is the standard deviation of a sampling distribution.



(a) Distribution of Scores

(b) Sampling Distribution of Means

Figure 6.2.1 Distributions involving standard deviation and standard error

The sample data rarely provides a perfect estimate of the population parameter. From Figure 6.2.1 we observe that values of the sample mean,  $M$  will be close to and some distance from the population mean,  $\mu$ . Therefore, a researcher who uses just one sample to estimate a population mean will have some error associated with such estimate. The standard error provides a way to measure the average distance or error between a sample mean and the population mean. The larger the sample size (assume a random sample), the smaller will be the error associated with the estimation of the population mean. The formula for computing the standard error of a sampling distribution is:

$$\text{standard error} = \sigma_M = \frac{\sigma}{\sqrt{n}}, \text{ where } n \text{ is sample size and } \sigma \text{ the standard deviation}$$

A sample of  $n = 100$  would provide a much better estimate (smaller standard error,  $s_M$ ) than a sample of  $n = 9$ . In practical terms, the sample standard deviation,  $s$ , will be used to estimate the population standard deviation. This scenario or case will be presented in the next section. Therefore larger sample size leads to smaller standard error.

$$\sigma_M = \frac{27}{\sqrt{100}} = \frac{27}{10} = 2.7 \text{ and } \sigma_M = \frac{27}{\sqrt{9}} = \frac{27}{3} = 9; \text{ for } \sigma = 27$$

When the population standard deviation is known by the researcher, the underlying sampling distribution will be the normal distribution and so the  $z$ -score statistical formula is used to test the null hypothesis. After stating the hypotheses, the researcher uses the following formula to compute the test statistics:

$$z = \frac{M - \mu}{\sigma_M} = \frac{\text{sample mean} - \text{hypothesized pop mean}}{\text{standard error between } M \text{ and } \mu} = \frac{(M - \mu)}{\sigma} \cdot \sqrt{n}$$

**The Central Limit Theorem** states that for any population with mean  $\mu$  and standard deviation  $s$ , the distribution of sample means for sample size  $n$  will have a mean of  $\mu$  and a standard deviation  $s/\sqrt{n}$  and will approach a normal distribution as  $n$  becomes very large.

We measured the height of 100 men 18 years and older, and calculated a mean height of 69.2 inches. Let us assume that the height of similar men in our population is some hypothesized mean with a standard deviation the 3.0 inches. We may argue that if we measured a large number of men our average would approach that of the population's. We know based on the Central Limit Theorem, that the true average height of all men is normally distributed with standard deviation equal to the standard error.

Let us test if our sample height of 69.2 is close to a hypothesized true height of 69 for men 18 years and older. The process of formulating and testing the hypothesis is shown below. Let us examine a few hypotheses based on the sample above.

**Problem statement 1:** Is the mean height of 69.2 inches obtained from a sample of  $n = 100$  with a hypothesized population mean of 69 inches? The population standard deviation is 3.0 inches.

**Step 1: State the hypotheses**

$$H_0: \mu = 69$$

$$H_a: \mu \neq 69 \text{ (non-directional)}$$

**Step 2: Select the significance level**

$$\alpha = 0.05 \text{ (5\%, two-tailed with 2.5\% in each tail)}$$

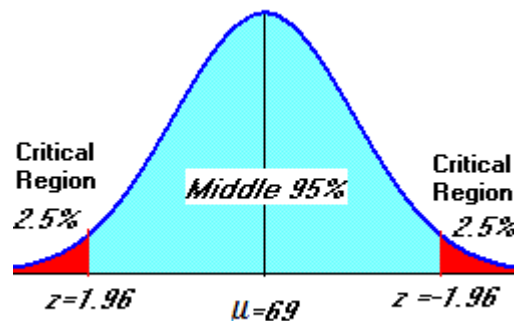
**Step 3: Compute the test statistics**

$$\text{standard error} = \sigma_M = \frac{\sigma}{\sqrt{n}} = \frac{3}{\sqrt{100}} = \frac{3}{10} = 0.3$$

$$\text{test statistics} = z = \frac{M - \mu}{\sigma_M} = \frac{69.2 - 69}{0.3} = 0.67$$

**Step 4: Determine the critical region**

The critical region separates the extreme 5% ( $\alpha = .05$ ) of data in the tails of the distribution; so the z-score lower boundary is  $z = -1.96$ ,  $\Pr(z) = 0.025$ . The upper z-score boundary is  $z = 1.96$ ,  $\Pr(z) = 0.975$ . The middle 95% is between -1.96 and 1.96 ( $0.975 - 0.025 = 0.95$  or 95%)



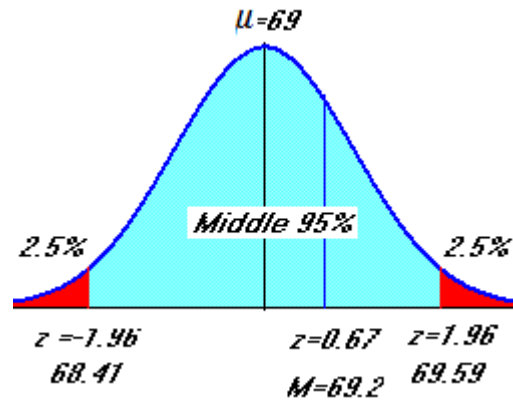
**Step 5: Make a decision**

We **don't reject** the null hypothesis:

1. The test statistics  $z = 0.67$  is within the boundaries of the 95% acceptance region of  $z = \pm 1.96$  [68.41 to 69.59 =  $69 \pm 1.96(0.3)$ ]
2. The probability of the test statistics is  $0.251 = 1 - 0.749$   
 $p\text{-value} = 0.502 > \alpha = 0.05$

The effect size is small,  $d = 0.07$

$$d = \frac{M - \mu}{\sigma} = \frac{69.2 - 69}{3} = 0.07$$



$$p\text{-value} = 2(0.251) = 0.502 > 0.05$$

since the  $\Pr(z_{test} \geq 0.67) = 0.251$

Analysis of the hypothesis testing process for problem statement 1 shows the following:

1. The null hypothesis states that the population mean is 69 and therefore the sample mean will be used to test this hypothesis. The alternative hypothesis is simply that the population mean is not 69; this is a non-directional statement.
2. The significance level is selected prior to evaluating the hypothesis at an alpha level of 5% ( $\alpha = 0.05$ ). Because the population standard deviation is known, the underlying distribution of the test statistics will be a standard normal distribution. Since the alternative hypothesis is non-directional we have a two-tailed sampling distribution with probability levels in each tail equal  $\alpha/2 = 0.025$  or 2.5% each. The researcher's hypothesis (alternative hypothesis) statement is that the population mean is not equal to 69.
3. The  $s = 3$  and  $n = 100$  is used to compute the standard error,  $s_M = 0.3$ . The z-score formula is then used to compute the test statistics,  $z = 0.67$ .

4. The critical region is the region where 5% of the most unlikely extreme samples would fall for the sampling distribution. A two-tailed standard normal distribution with a total of 5% of the distribution in the tails is bounded by  $z$ -score values (critical values) of  $z = \pm 1.96$ .
5. Any of the following criteria will show that there is not sufficient reason to reject the null hypothesis; therefore, we retain the null hypothesis:
  - a. The test statistics of  $z = 0.67$  is within the 95% acceptance region ( $-1.96 = z = +1.96$ ) of the sampling distribution with  $\mu = 0.69$  and  $s_M = 0.3$ .
  - b. Therefore,  $M = 69.2$  is typical for the sampling distribution ( $68.41 = \mu = 69.59$ ).
  - c. The  $p$ -value gives the degree of the failure of the test to accept the null hypothesis. It is the probability of the test statistics. The  $p$ -value is 0.502 (0.251 in each tail) and is greater than the alpha = 0.05. Since this probability is high it favors the null hypothesis because it shows that the sample  $M = 69.2$  has a high probability of occurring by chance given that the null hypothesis is true. We reject the null hypothesis if the  $p$ -value was less than expected (0.05) or less than or equal to alpha.
  - d. The population mean,  $\mu = 69$  is within the 95% confidence interval of **68.61** and **69.79** [ $CI_{95} = 69.2 \pm 1.96(0.3)$ ] around the sample mean of  $M = 69.2$

The  **$p$ -value** is the probability that the sample statistics or mean occurred by chance if the null hypothesis is true; it is the exact probability of a Type 1 error, alpha.

Similar criteria can be used to evaluate the null hypotheses of problem statement 2 below. Since this is a directional alternative hypothesis statement,  $\mu > 68.6$ , an upper one-tailed test is used in the hypothesis evaluation process with  $\alpha = 0.05$  in the upper tail.

The null hypothesis is rejected for the following reasons:

1. The test statistics  $z = 2.0$  is greater than the critical value of  $z = 1.64$
2.  $M = 69.2$  is outside the typical sampling distribution ( $\mu = 69.09$ )
3. The  $p$ -value of 0.02 is less than the alpha level of 0.05 and
4. The population  $\mu < 68.61$ ; the lower extreme 5% of the sampling distribution if the true population mean was 69.2  $[(69.2 - 1.64(0.3))]$

Table 6.2.1 shows the boundaries or critical regions of the standard normal sampling distribution under various conditions. The two-tailed critical value is used to construct the confidence interval. Since the confidence interval is dependent upon both the values of the critical value and the standard error; large significance levels and large sample size will reduced the width of the confidence interval.

Table 6.2.1 *z Critical Values for Sampling Distribution (Sigma Known)*

Types of Test	Alpha Levels	
	0.05	0.01
One-Tailed	1.64	2.33
Two-Tailed	$\pm 1.96$	$\pm 2.58$



**Problem statement 2:** Is the mean height of 69.2 inches obtained from a sample of  $n = 100$  **greater than** a hypothesized population mean of 68.6 inches with a standard deviation of 3.0 inches?

**Step 1: State the hypotheses**

$$H_0: \mu = 68.6$$

$$H_a: \mu > 68.6 \text{ (directional)}$$

**Step 2: Select the significance level**

$$\alpha = 0.05 \text{ (5\%, one-tailed with 5\% in upper tail)}$$

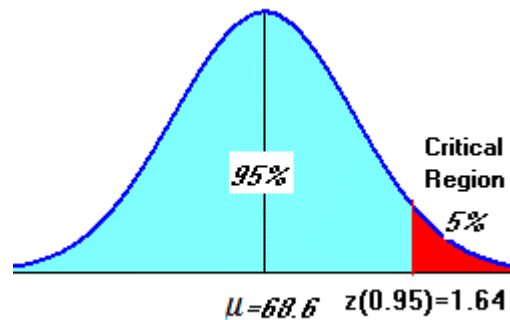
**Step 3: Compute the test statistics**

$$\text{standard error} = \sigma_M = \frac{\sigma}{\sqrt{n}} = \frac{3}{\sqrt{100}} = \frac{3}{10} = 0.3$$

$$\text{test statistics} = z = \frac{M - \mu}{\sigma_M} = \frac{69.2 - 68.6}{0.3} = \frac{0.6}{0.3} = 2$$

**Step 4: Determine the critical region**

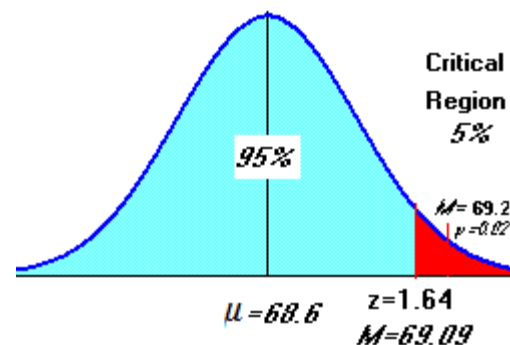
The critical region separates the extreme 5% of data in the tails of the distribution; so the z-score upper is  $z = 1.64$ ,  $\Pr(z) = 0.95$ .



**Step 5: Make a decision**

We **reject** the null hypothesis:

- The test statistics  $z = 2.0$  is outside the boundary of the 95% acceptance region of  $z = +1.64$ .  
 $M = 69.2$  is outside the 95% acceptance region  $> 69.09$   
( $> 68.6 + 1.64(0.3)$ )
- The probability of the test statistics is  $0.02 = 1 - 0.98$   
 $p\text{-value} = 0.02 < 0.05$



The effect size is small

$$d = \frac{M - \mu}{\sigma} = \frac{69.2 - 68.6}{3} = 0.2$$

$p\text{-value} = 0.02 < 0.05$   
since the  $\Pr(z_{test} \geq 2.0) = 0.02$

We reject the null hypothesis when any of the following condition is true:

1. The test statistics is greater than the critical value(s) (the boundaries of the critical region)
2. The sample mean is outside acceptance region of the sampling distribution when the sampling mean equals the population mean,  $\mu$
3. The  $p$ -value of the test statistics is less than or equal to the alpha level
4. The population mean is not within the confidence interval of the sampling distribution when  $\mu$  equals the sample mean.