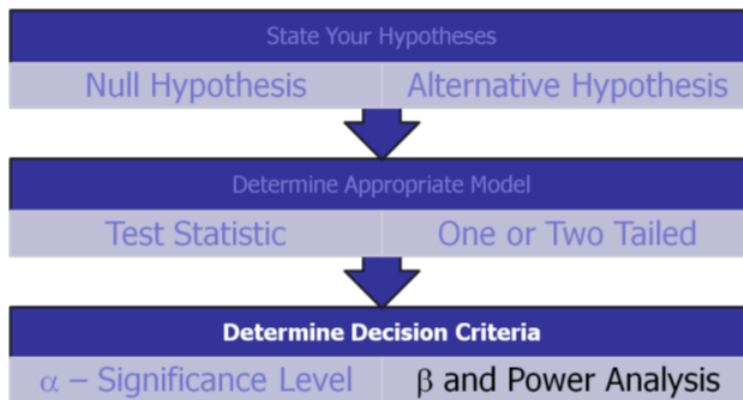


10.7: Type II Error and Statistical Power

In the prior example, the statistician failed to reject the Null Hypothesis because the probability of making Type I error (rejecting a true Null Hypothesis) exceeded the significance level of 5%. However, the statistician could have made Type II error if the machine is really operating improperly. One of the important and often overlooked tasks is to analyze the probability of making Type II error (β). Usually statisticians look at statistical power which is the complement of β .



Beta (β): The probability of failing to reject the null hypothesis when it is actually false.

Power (or Statistical Power): The probability of rejecting the null hypothesis when it is actually false.

Both beta and power are calculated for specific possible values of the Alternative Hypothesis.

	Fail to Reject H_o	Reject H_o
H_o is true	$1 - \alpha$	α Type I error
H_o is false	β Type II error	$1 - \beta$ Power

If a hypothesis test has low power, then it would be difficult to reject H_o , even if H_o were false; the research would be a waste of time and money. However, analyzing power is difficult in that there are many values of the population parameter that support H_a . For example, in the soy sauce bottling example, the Alternative Hypothesis was that the mean was not 16 ounces. This means the machine could be filling the bottles with a mean of 16.0001 ounces, making H_a technically true. So when analyzing power and Type II error, we need to choose a value for the population mean under the **Alternative Hypothesis (μ_a)** that is “**practically different**” from the **mean under the Null Hypothesis (μ_o)**. This practical difference is called the **effect size**.

Definition: Effect size

Effect Size: The “practical difference” between μ_o and $\mu_a = |\mu_o - \mu_a|$

where

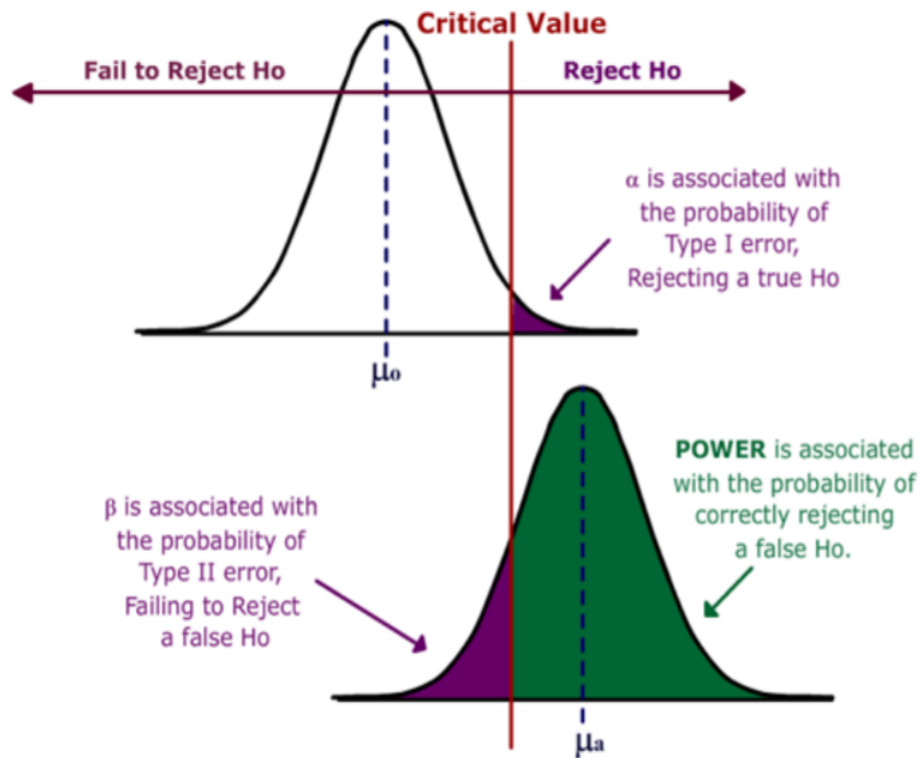
μ_o : The value of the population mean under the Null Hypothesis

μ_a : The value of the population mean under the Alternative Hypothesis

Suppose we are conducting a one-tailed test of the population mean:

$$H_o : \mu = \mu_0 \quad H_a : \mu > \mu_0$$

Consider the two graphs shown below. The top graph is the distribution of the sample mean under the Null Hypothesis, which was covered in an earlier section. The area to the right of the critical value is the rejection region.



We now add the bottom graph, which represents the distribution of the sample mean under the Alternative Hypothesis for the specific value μ_a .

We can now measure the Power of the test (the area in green) and beta (the area in purple) on the lower graph.

There are several methods of increasing Power, but they all have trade-offs:

Ways to Increase Power	Trade off
Increase Sample Size	Increased cost or unavailability of data
Increase Significance level (α)	More like to Reject a true H_o (Type I error)
Choose a value of μ_a further from μ_o	Result may be less meaningful
Redefine population to lower standard deviation	Result may be too limited to have value
Conduct as a one-tail rather than a two-tail test	May produce a biased result

Example: Bus brake pads

Bus brake pads are claimed to last on average at least 60,000 miles and the company wants to test this claim. The bus company considers a “practical” value for purposes of bus safety to be that the pads last at least 58,000 miles. If the standard deviation is 5,000 and the sample size is 50, find the power of the test when the mean is really 58,000 miles. (Assume $\alpha = .05$)



Solution

First, find the critical value of the test.

Reject H_o when $Z < -1.645$

Next, find the value of that corresponds to the critical value.

$$\bar{X} = \mu_o + \frac{Z\sigma}{\sqrt{n}} = 60000 - (1.645)(5000)/\sqrt{50} = 58837$$

H_o is rejected when $\bar{X} < 58837$

Finally, find the probability of rejecting H_o if H_a is true.

$$\begin{aligned} P(\bar{X} < 58837) &= P\left(Z < \frac{(58837 - \mu_a)}{\sigma/\sqrt{n}}\right) \\ &= P\left(Z < \frac{(58837 - 58000)}{5000/\sqrt{50}}\right) \\ &= P(Z < 1.18) \\ &= .8810 \end{aligned}$$

Therefore, this test has 88% power and β would be 12%

Power Calculation Values

Input Values

$\mu_o = 60,000$ miles

$\mu_a = 58,000$ miles

$\alpha = 0.05^*$

$n = 50$

$\sigma = 5000$ miles

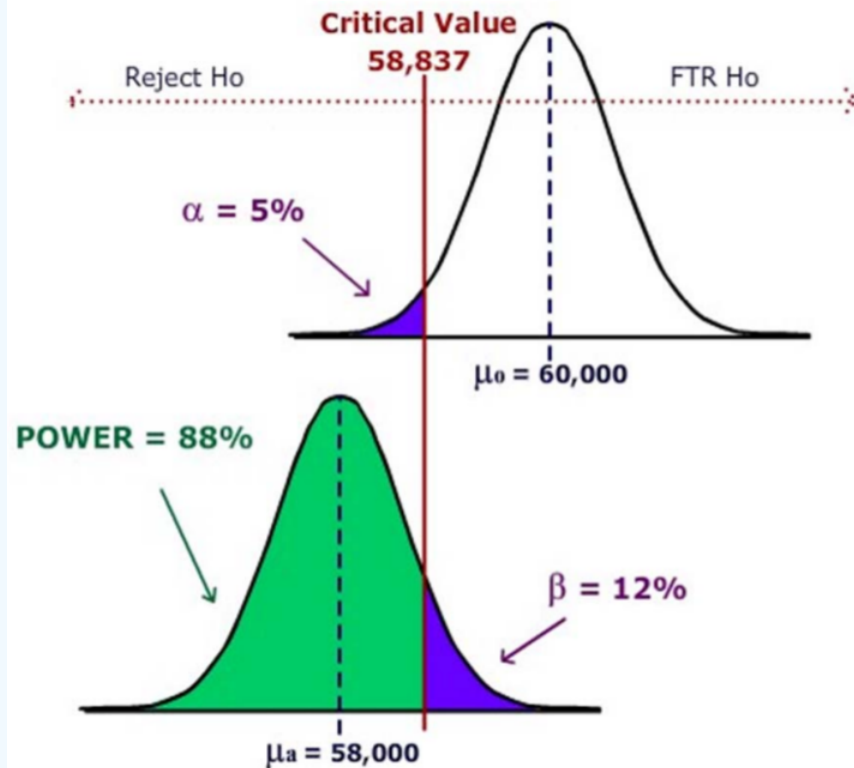
Calculated Values

Effect Size = 2000 miles

Critical Value = 58,837 miles

$\beta = 0.1190$ or about 12%

Power = 0.8810 or about 88%



This page titled [10.7: Type II Error and Statistical Power](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Maurice A. Geraghty](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.