The Data Driven Manager

5

Essentials for Effect Size Calculations

Learning Objectives

- Explain the concept of effect size and its importance in making business decisions
- Integrate business decisions and financial considerations in the selection of effect size, sample size, and error levels
- Calculate effect size for means, variances and proportions using a cost-benefit analysis
- Calculate sample size for means, variances and proportions controlling for Type 1 and Type 2 error

Effect Size Calculations

Errors & Power in Making Business Decisions

- In business and industry, research, tests and experiments which lead to the development of hypotheses, data collection, and subsequent statistical analyses are conducted to allow management to make informed, data-based decisions.

Errors & Power in Making Business Decisions

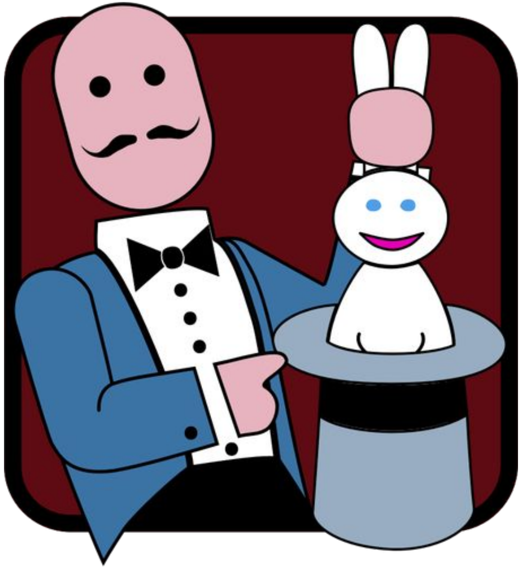
H_0 H_1

- Each test, research study, or experiment requires that management select an appropriate probability level for the **two** Types of **Errors** which may occur (Type I and Type II Errors).

- Additionally, management must determine the **Effect Size** they wish to detect in order to declare that a meaningful or beneficial change has taken place.

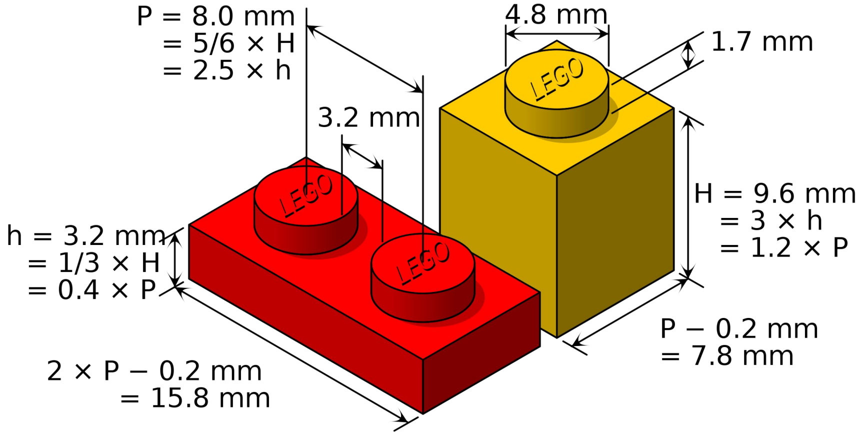
Effect Size

Frequently the Effect Size is "pulled out of thin air", but this is not the way to do it in serious, real world applications.



Effect Size

Sometimes Effect Size is determined by Engineering, Safety, Reliability, Customer, or other "Technical Requirements", which is better than the first case.



Effect Size

In business and engineering situations, Effect Size should be a computation that takes into account

1. the magnitude of the problem,
2. the cost to fix it, and
3. the requirement for the fix to pay for itself under organizational Return On Investment requirements.

Errors & Power in Making Business Decisions

H_0 H_1

- For managers, the selection of the probabilities associated with errors (α and β) should be thought of as a data-based **risk analysis**.

- **Effect size**, the shift or change we are attempting to detect, should be thought of as a data-based **cost-benefit analysis**.

Errors & Power in Making Business Decisions

H_0 H_1

Example - Oregon Saw Chain

The Oregon Saw Chain Company in Portland, Oregon manufactures a splitting maul (axe used to split wood) with a hickory (wood) handle. The safety of the user (consumer) depends on the handle being able to withstand a pressure exerted on a production line or laboratory test of at least 100 PSI without failing.

Currently, the **mean** (μ) pressure for the wood handles is **110.5,** with a **standard deviation** (σ) of **5** PSI. γ_3 and γ_4 have both been shown to be **0.00**.

Errors & Power in Making Business Decisions

Obviously, the current process is generating a number of handles which are failing the pressure test. Between end-of-line failures and scrap, plus lost processing costs, plus warranty costs, and lost customers, the current estimation is that the **process defective rate** is costing OSCD approximately **$250,000** per year.

Errors & Power in Making Business Decisions

H_0 H_1

A new manufacturer has approached the OSCD management team with a proposal to sell them fiberglass reinforced handles, ready to assemble, that they claim will have a much **higher average** resistance to breakage under pressure.

Both the prospective supplier's engineers and OSCD's engineers agree that it is **unlikely** that the new handle material could **affect** the **dispersion** or **shape** of the current pressure resistance distribution.

Errors & Power in Making Business Decisions

Unfortunately, the OSCD management team has had significant experience with vendors in the past who claimed that the use of their product would significantly enhance the performance, quality, or reliability of OSCD's products as sold. Sometimes it has been true, and sometimes not.

Errors & Power in Making Business Decisions

Further, the new handles would require OSCD to purchase a **new** handle insertion **machine** at a **cost** of **$225,000.00**, representing a significant increase in manufacturing cost. This expenditure would, naturally, be expected to be offset by a reduction in the current scrap rate and related losses.

Errors & Power in Making Business Decisions

H_0 H_1

Under current OSCD guidelines, the following product change requirements are in effect:

- The new handles must be shown to be effective in reducing the process defective rate and potential hazard to the user; and
- The resultant **reduction in losses** must have a **minimum of a 2 year Return On Investment.**

In other words, the value of the reduction in current losses per year must be equal to 50% or more of the cost of the new machine.

Errors & Power in Making Business Decisions

H_0 H_1

The OSCD management team and the prospective supplier agree to conduct a designed experiment to test a random sample of the proposed (new) handles in the OSCD production process, and subsequent laboratory test.

Using a classical approach, the hypotheses are written as:

$$H_0 : \mu = 110.5$$
$$H_1 : \mu \neq 110.5$$

Errors & Power in Making Business Decisions

H_0 : μ = 110.5

H_1 : μ ≠ 110.5

What are the possible outcomes, or results, of conducting this experiment, and testing these hypotheses, **from a business perspective**?

Errors & Power in Making Business Decisions

H_0 H_1

Experimental Outcomes

Inference About Reality	Actual Situation or Reality re: H_0	
	Handles Not Effective: H_0 is True	Handles Are Effective: H_0 is False
Accept H_0 as True	Correct Decision	Incorrect Decision
Reject H_0 as False	Incorrect Decision	Correct Decision

Errors in Making Decisions

H_0 H_1

Type I Error

- The probability of rejecting a true null hypothesis
- Probability of Type I Error is α (Alpha)
- α is the level of significance

Confidence Level

- Not Rejecting a true null hypothesis
- Probability is $\{1-\alpha\}$

Where Are These Probabilities on the Table ?

Errors & Power in Making Business Decisions

H_0 H_1

Experimental Outcomes

Inference About Reality	Actual Situation or Reality re: H_0)	
	Handles Not Effective: H_0 is True	Handles Are Effective: H_0 is False
Accept H_0 as True	Correct Decision { **1-α** } : Confidence	Incorrect Decision
Reject H_0 as False	Incorrect Decision **α** : Type I Error	Correct Decision

Errors in Making Decisions

H_0 H_1

Type II Error

- When we "Accept" a false null hypothesis
- Probability of Type II Error is **β** (Beta)

Power

- The probability that we will correctly reject a false null hypothesis
- The probability that we will detect a true difference or change when it is present.
- Power is {1-**β**}
- Power is a function of the sample size employed, and the Effect Size

Where Are These Probabilities on the Table?

Errors & Power in Making Business Decisions

H_0 H_1

Experimental Outcomes

Inference About Reality	Actual Situation or Reality re: H_0)	
	Handles Not Effective: H_0 is True	Handles Are Effective: H_0 is False
Accept H_0 as True	Correct Decision { 1-**α** } : Confidence	Incorrect Decision **β**: Type II Error
Reject H_0 as False	Incorrect Decision **α** : Type I Error	Correct Decision { 1- **β** } : Power

Errors in Making Decisions

H_0 H_1

Now, let's identify and describe the **business implications** of each type of error and the correct decision associated with this experiment and hypothesis test.

Errors & Power in Making Business Decisions

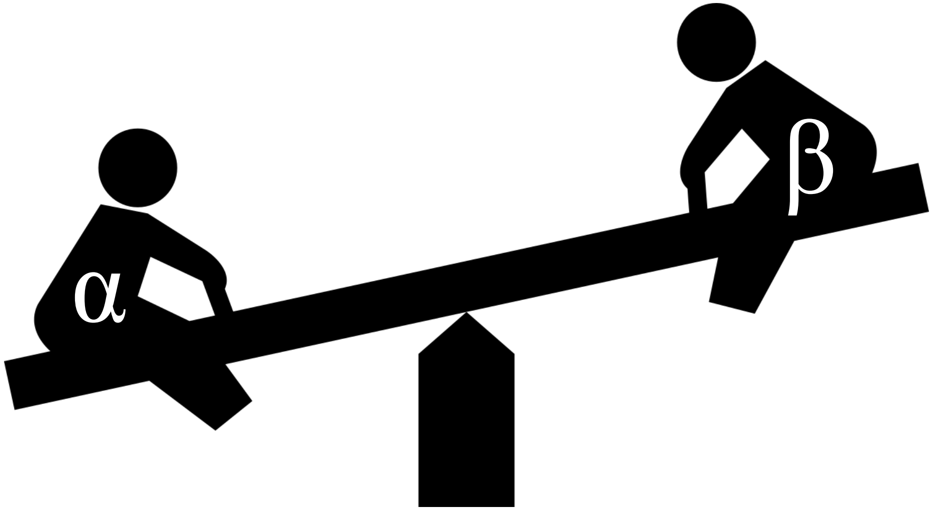
H_0 H_1

Experimental Outcomes

Inference About Reality	Actual Situation or Reality re: H_0)	
	Handles Not Effective: H_0 is True	Handles Are Effective: H_0 is False
Accept H_0 as True	Correct Decision Avoid Spending $225,000 For No Benefit [1-**α**] : Confidence	Incorrect Decision Lose Opportunity To Save Up To $25K/$250K **β** : Type II Error
Reject H_0 as False	Incorrect Decision Spend $225,000 For No Benefit **α** : Type I Error	Correct Decision Save Up To $25K/$250K [1- **β**] : Power

H_0 H_1

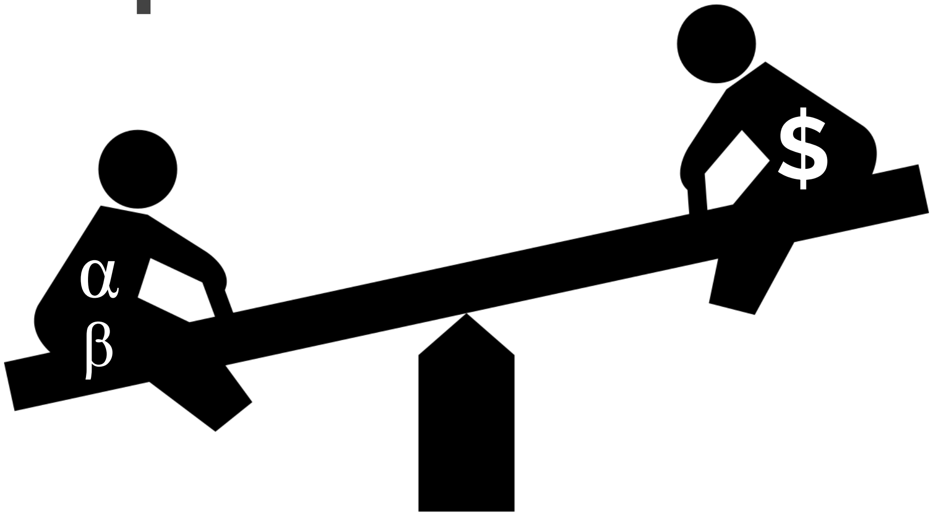
α & β Have an Inverse Relationship



Type I and Type II Errors are conditional probabilities, based upon different conditions (i.e. realities)

Note: You can NOT make both types of errors at the same time, AND the decision you make restricts the kind of error you can make.

The Smaller α & β Become, The More Expensive to Run The Experiment



Based on penalty or cost to the business, we select the appropriate levels (probabilities) for each type of error.

α & β Can Only Be Selected Meaningfully When Effect Size Has Been Correctly Identified

H_0 H_1

Effect Size (Δ_μ) :

- Effect Size is described as the ***minimum*** amount of shift or change that we want to be able **to correctly detect** as a result of conducting our test, research or experiment.
- Effect Size is a **calculated** value based upon **financial requirements** and **management direction**.

Effect Size Calculations

H_0 H_1

- We will now review the steps for effect size calculations continuing with the Oregon Saw Chain Division (OSCD) example we started earlier.

Calculating Effect Size for a Mean: The OSCD Example

H_0 H_1

Step 1: Start By Identifying the Key Information

Given:

μ = 110.5 σ = 5

LSL = 100.0 Normal Distribution

Current Loss / Year = $250,000 on N = 175,000 Units

New Handles Require One-Time Expenditure of $225,000.00

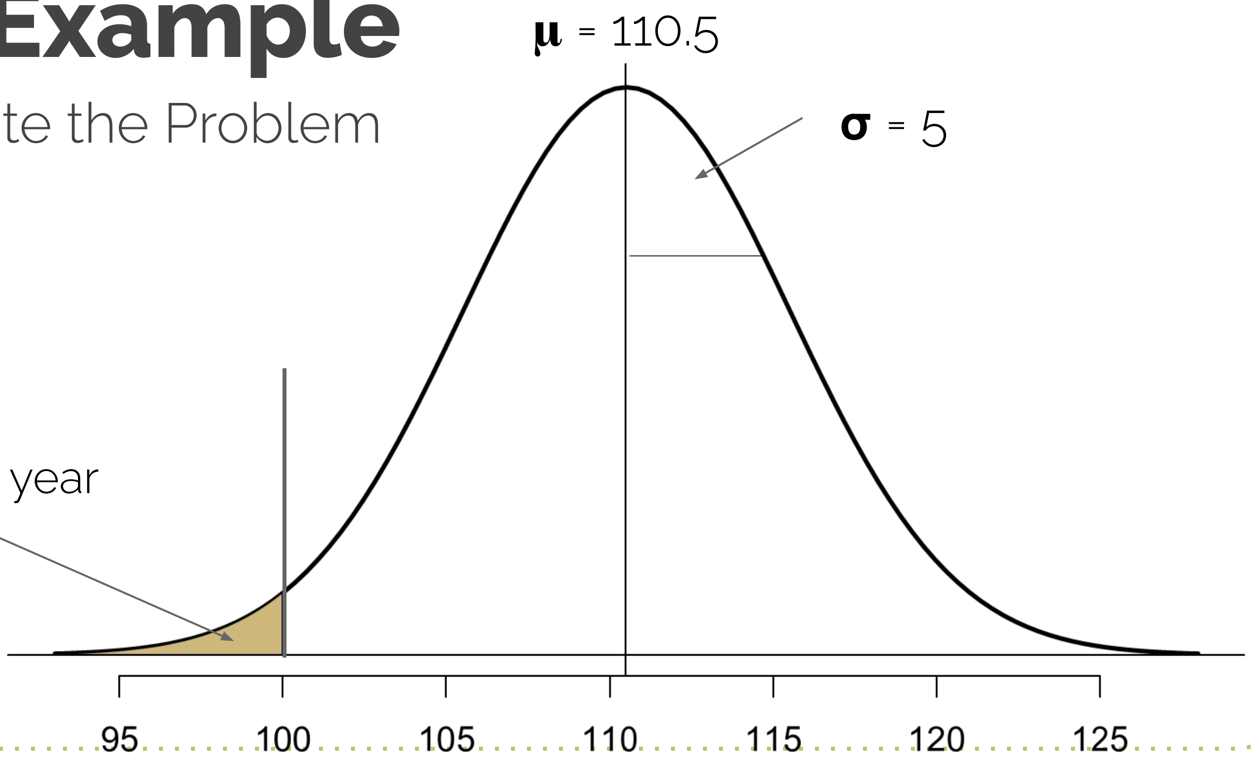
ROI Required : Savings of $112,500 / Year, or a

Maximum Allowable Scrap Related Loss of $137,500 (per year)

Calculating Effect Size: The OSCD Example

Step 2: Illustrate the Problem

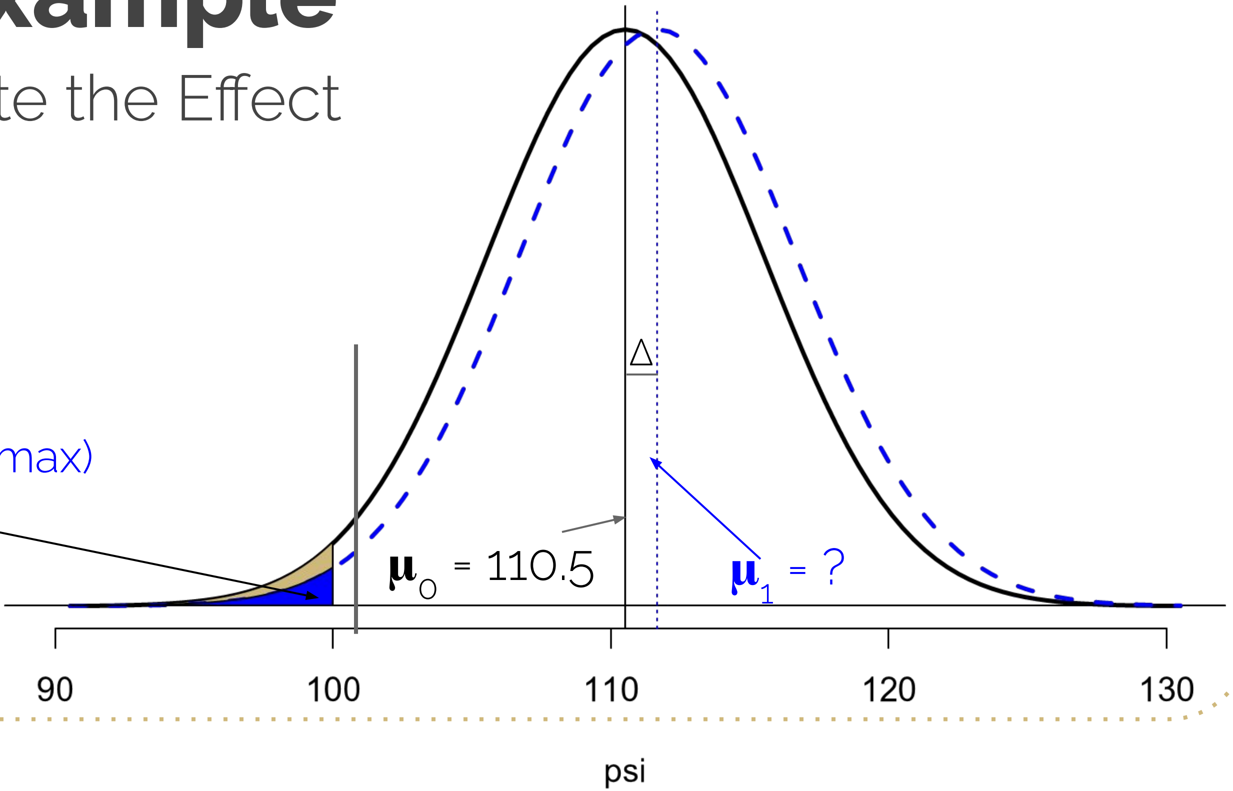
% Defective =
$250,000 loss / year



Calculating Effect Size: The OSCD Example

Step 3: Illustrate the Effect Size Required

New % Defective =
$137,500 loss / year (max)



Calculating Effect Size: The OCSD Example

H_0 H_1

Step 4 : Calculate the Effect Size Required

A. % Defective: Current Condition

Mean = 110.5
Std Dev = 5
X = 100

z-score = -2.1000

% above X = 98.2136%
% below X = 1.7864%

B. Loss / % Defective

Since 1.7864% = $250,000,
1.0000% = $139,946.26

C. New **Maximum** Defective Rate Required =

$137,500 / 139,946.26 = 0.98252

Calculating Effect Size: The OSCD Example

H_0 H_1

Step 4 : Calculate the Effect Size Required

B. Loss / % Defective

Since 1.7864% = \$250,000,
1.0000% = \$139,946.26

C. New **Maximum** Defective Rate Required =

\$137,500 / 139,946.26 = 0.98252

D. Value of μ_1 at New Defective Rate

For z p-Value = 0.0098252
Lower-tail Value = -2.333

$$-2.333 = \frac{100 - \mu_1}{5}$$

μ_1 = **111.665 so Δ = 1.165**

Step 5 : Calculating Power & β for the OSCD Example

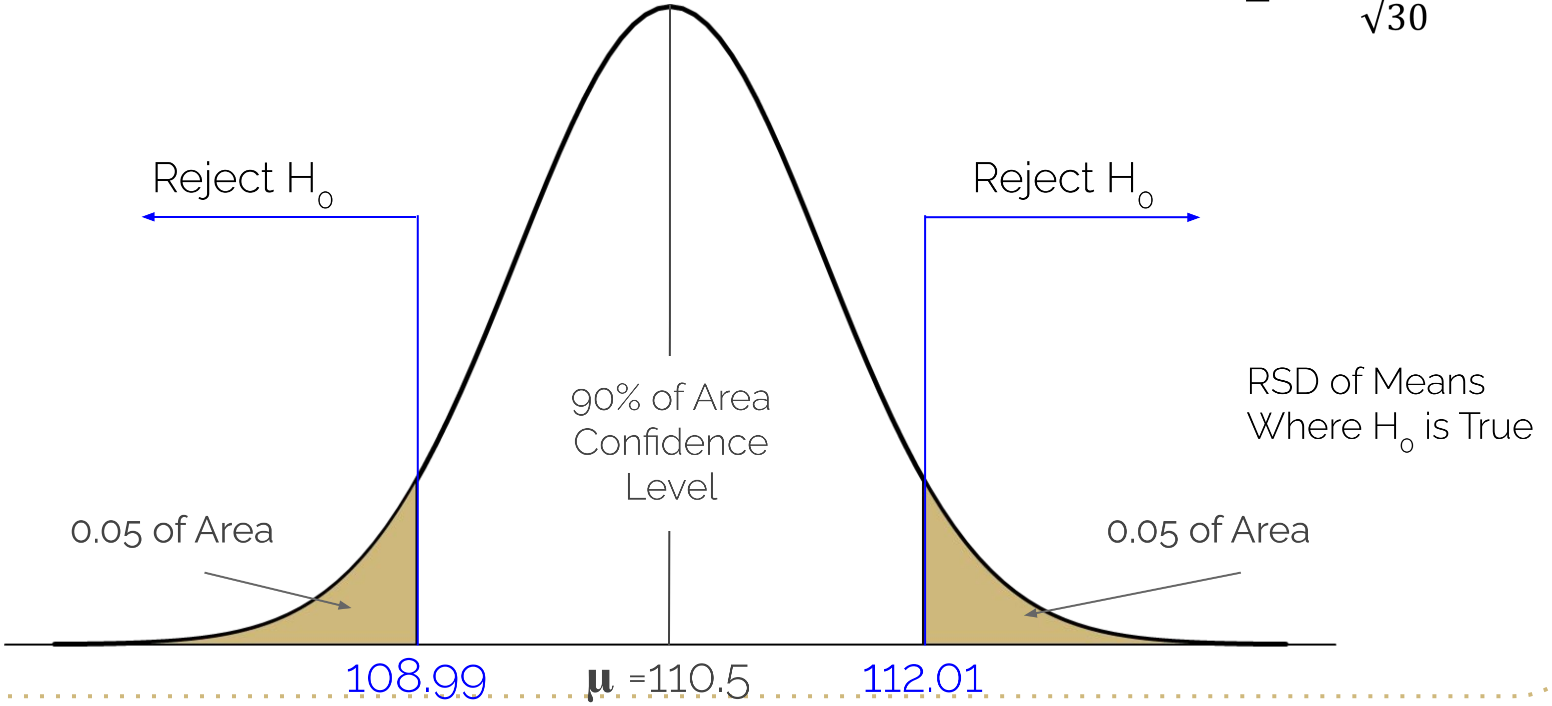
H_0 H_1

The original hypothesis test is shown below. Suppose that management has directed the researcher to maintain a maximum **α level of 10%**. Suppose further that the supplier has provided us with **30 handles** to test. What is the Power of the test ?

$H_0: \mu = 110.5$
$H_1: \mu \neq 110.5$

Step 5: Calculating Power & β for the OSCD Example

Critical Values = $110.5 \pm 1.65 \frac{5}{\sqrt{30}}$



Calculating □ and Power for Means

H_0 H_1

- https://casertamarco.shinyapps.io/power/

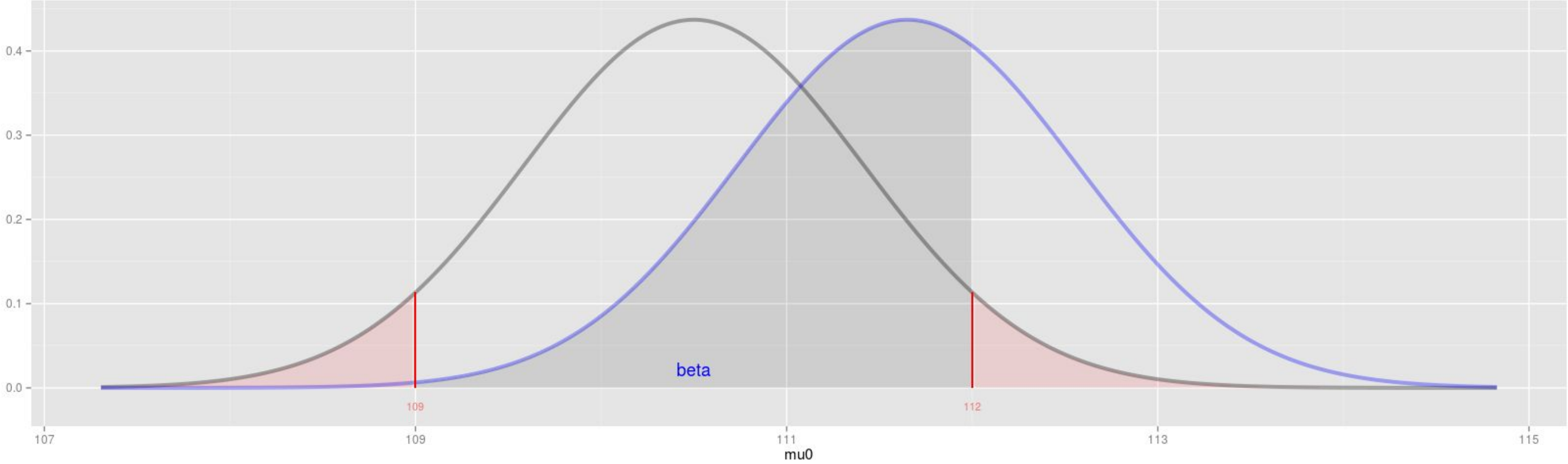
Step 5: Calculating Power & β for the OSCD Example

Why Is Power So Small?

Why Is Type II Error So Large?

Power = 0.3562

□ = 0.6438



Step 5 : Calculating Power & β for the OSCD Example

H_0 H_1

Answer: the sample size is insufficient to provide acceptable protection against a Type II error !

Solution: Select the minimum power desired (or maximum type II error required), and calculate the sample size required to conduct the experiment

Step 6 : Calculating An Appropriate Sample Size for the OSCD Example

H_0 H_1

Required: a sample size that will maintain the following conditions:

- A Maximum Type I Error Level At 10%
- A Minimum Confidence Level At 90%
- An Effect Size Equal To 1.165
- A Maximum Type II Error Level At 5%
- A Minimum Power Level of 95%

Step 6 : Calculating An Appropriate Sample Size for the OSCD Example

H_0 H_1

Solution: for a one-sample test, where the normal distribution may be used as a model for the RSD of the means; this formula is given as:

$$n = \frac{\sigma^2 (|Z_\beta| + |Z_\alpha|)^2}{\Delta^2}$$

Result: the effect of using this formula is to reduce the standard error enough so that the error levels are held at the required effect size !

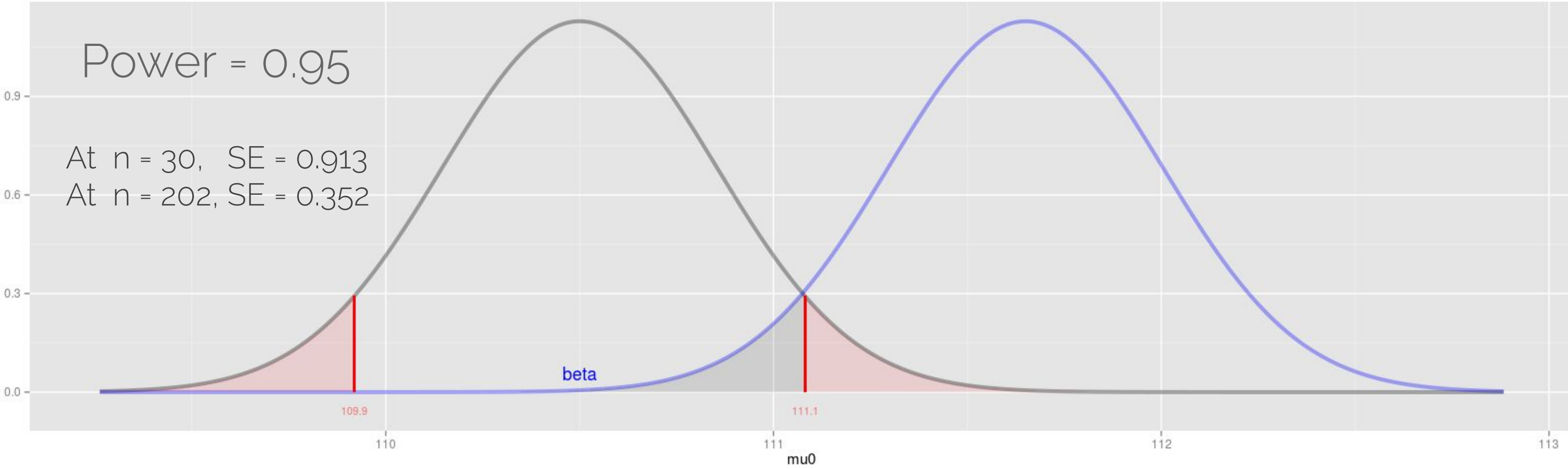
Step 6 : Calculating An Appropriate Sample Size for the OSCD Example

H_0 H_1

$$n = \frac{25\,(1.645 + 1.645)^2}{1.165^2} = 200$$

Result: the effect of using this formula is to reduce the standard error enough so that the error levels are held at the required effect size !

Step 6 : Calculating An Appropriate Sample Size for the OSCD Example



Sample Size Calculations: Means

- The function to obtain the sample size is:

```
noquote(t(sample.size.mean.t.onesample(effect.size = 1.165,  variance.est = 5^2, alpha = 0.10, beta = 0.05)))
```

- And, the results produced are:

```
test              t
type              one.sample
alternative       two.sided
sample.size       202
actual            201.1941
df                201
effect.size       1.165
variance          25
alpha             0.1
conf.level        0.9
beta              0.04890932
power             0.9510907
```