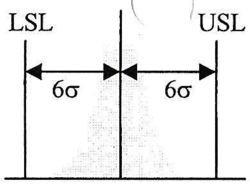
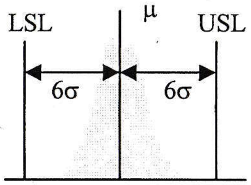
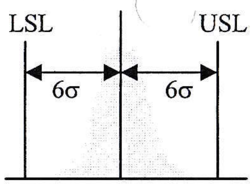


Test of Hypothesis & Power Analysis



Topics Covered

- Hypothesis testing
- Type I and Type II errors
- Alpha & Beta risk
- Delta value
- Power of the test
- Why sample size is important
- Sample size with “t” & “Z” tests
- Sample size with ANOVA & DOE

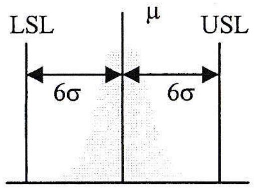


What is a Hypothesis Test ?

A ***Hypothesis Test*** is a procedure for determining if an assertion about a characteristic of a population is true. This assertion is referred to as the null hypothesis.

We then compare a sample taken from our population of interest to determine if the sample population is close enough to the population defined by our null hypothesis to be considered the same.

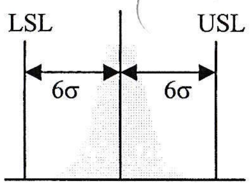
If it is, we accept the null hypothesis. If there is not enough evidence to prove it is the same, we reject the null hypothesis.



Hypothesis Testing

Every acceptance sampling test and designed experiment is a statistical <u>hypothesis test</u>.

- **Hypothesis tests separate significant effects from random chance effects.**
- **All hypothesis tests have unavoidable, but quantifiable, risks of choosing the wrong conclusion. They always involve the risk of:**
  - **Type I (alpha) error (AKA producer's risk)**
  - **Type II (beta) error (AKA consumer's risk)**



Type I & II Errors

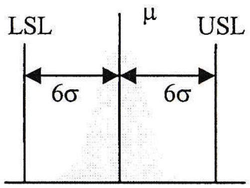
The Tech Group
WE PROVIDE INNOVATIVE SOLUTIONS

The Type I, (α) risk is the chance of deciding that a significant effect is present when it isn't.

A smoke alarm that sounds off when you make the toast too dark commits a Type I (α) error (the house is not on fire).

The Type II (β) risk is the chance of not detecting a significant effect when one exists.

A smoke alarm that does <u>not</u> sounds off when your house is on fire commits a Type II (β) error.



Type I & II Errors

- Type I: Reject the null hypothesis, when it is true.

  Say something is significant, when it is not

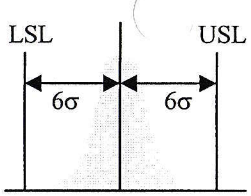
  – Known as α error

- Type II: Fail to reject the null hypothesis when it is false.

  Say something is not significant, when it is

  – Known as β error

	H_o is true	H_o is false
Do not Reject H_o	No error	Type II (β) error
Reject H_o	Type I (α) error	No error



The Null Hypothesis

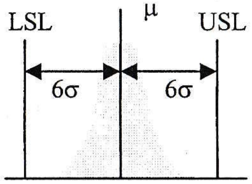
When a null hypothesis (H_o) is selected it is also necessary to select an alternative hypothesis (H_a). There are three choices for the alternative hypothesis. The correct choice depends on what you wish to prove or disprove. The alternative hypothesis can be any of the following:

1. $H_a \neq H_o$
2. $H_a > H_o$
3. $H_a < H_o$

<u>Example:</u>

<u>Null Hypothesis:</u> The average height of 6th graders is 69 inches.

<u>Alternative Hypothesis:</u> The average height or 6th graders is greater than 69 inches.



H_o: $Cpk \geq 1.33$

H_a: $Cpk < 1.33$

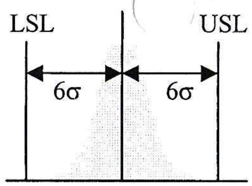
Steps to Perform a Hypothesis Test

The Tech Group
WE PROVIDE INNOVATIVE SOLUTIONS

H_o: mpg ≤ 35

H_a: mpg > 35

1. Specify the null and alternative hypothesis
2. Select a significance level (α)
3. Collect sample data from the population of interest
4. Calculate a test statistic (P-value)
5. Compare the P-value to the significance level (α)
6. Draw a conclusion:
   - Reject the null hypothesis, if P value < α
   - Fail to reject the null hypothesis, If P value > α
7. Document results and conclusions



Example #1

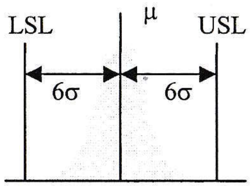
A grocer has received a truck load of 10,000 candy bars. The broker selling the bars claims the average weight of it's bars is at least 6 ounces. Before accepting the load, the grocer wants to verify this claim by weighing a sampling of 50 candy bars. The grocer would like to be 95% certain of his results (i.e. $\alpha = .05$).

A random sample of 50 bars produced the following data:

$\overline{X} = 5.96$ oz.

$\sigma = .12$ oz

The average is a bit below the minimum 6 oz specification. But this was only a sample and the average is only an estimate. What is the chance the true average is indeed at or above the minimum 6 oz. weight?



Example #1 (con't)

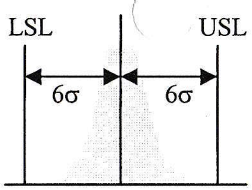
First we establish our null & alternative hypothesis:

$$H_o : \mu \geq 6 \text{ oz.}$$

$$H_a : \mu < 6 \text{ oz.}$$

<u>Null Hypothesis (H_o):</u> The average candy bar weight is at least 6 oz.

<u>Alternative Hypothesis (H_a):</u> The average candy bar weight is less than 6 oz.



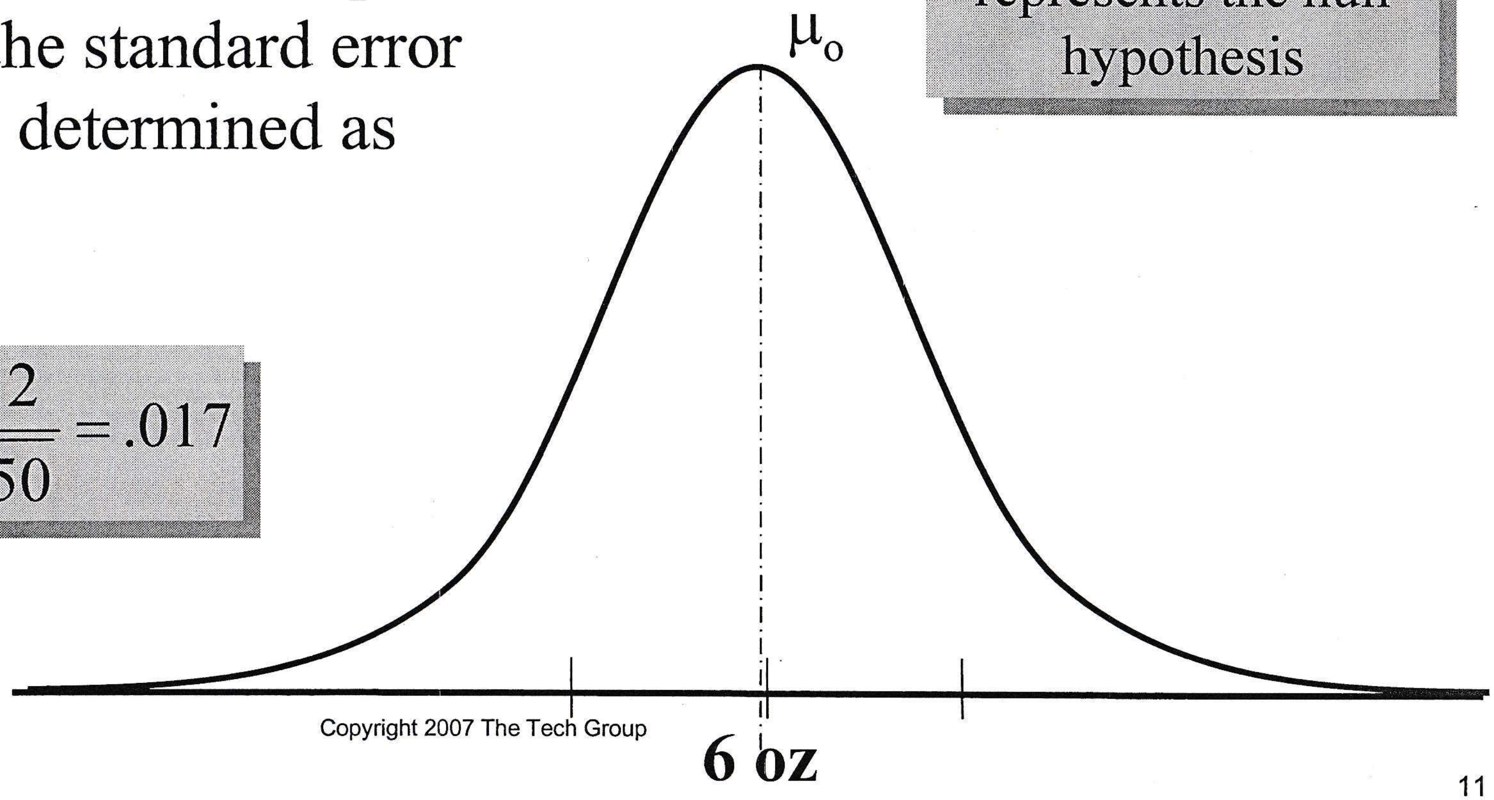
Example #1 (con't)

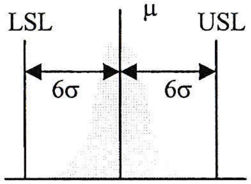
The distribution curve below represents the population of sample means that could be drawn from the candy bars in the truck.

We learned previously that the standard deviation of the sample means is called the standard error (SE(x)) which is determined as follows:

$$SE(\overline{x}) = \frac{\hat{\sigma}}{\sqrt{n}} = \frac{.12}{\sqrt{50}} = .017$$

The curve below represents the null hypothesis





Example #1 (con't)

The Tech Group
WE PROVIDE INNOVATIVE SOLUTIONS

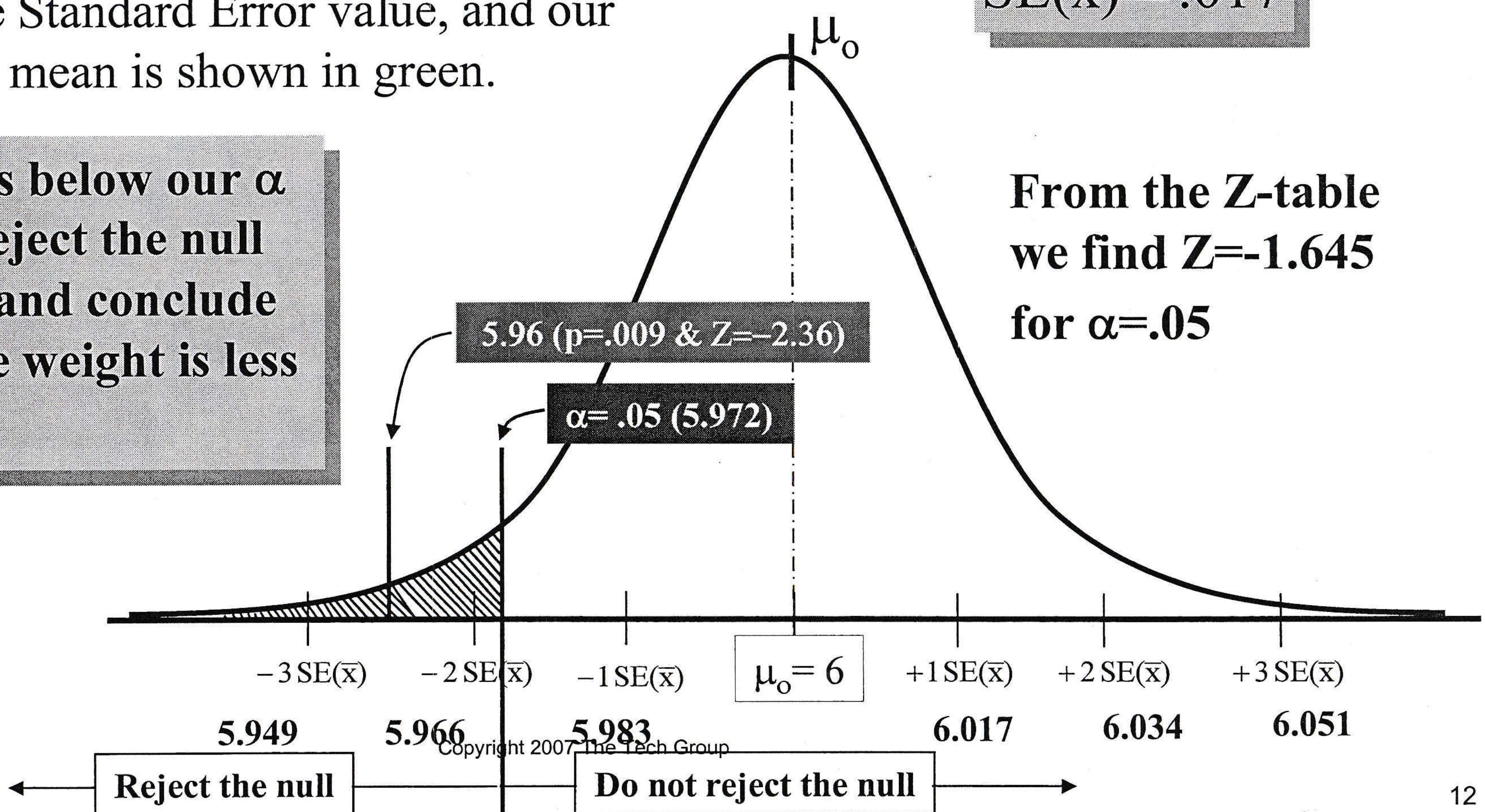
The grocer wishes to know if he can state, with 95% certainty, based on his sample mean of 5.96 oz , that the candy bars in the truck are <u>below</u> the required 6 oz average weight.

A scale has been added to the curve below based on the Standard Error value, and our 5.96 sample mean is shown in green.

$SE(\bar{x}) = .017$

Since 5.96 is below our α value, we reject the null hypothesis and conclude our average weight is less than 6 oz.

From the Z-table we find Z=-1.645 for α=.05



Ex. #1 Using Minitab (con't)

Minitab results show a p-value of .009, a Z-value of –2.36, and a SE Mean of .017, which agree with our earlier results.

As before, we reject the null hypothesis since the p-value is below our alpha of .05.

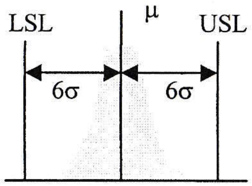
```
One-Sample Z: Mean

Test of mu = 6 vs mu < 6
The assumed sigma = 0.12

Variable           N      Mean     StDev   SE Mean
Mean              50    5.9600    0.1249    0.0170

Variable      95.0% Upper Bound        Z       P
Mean                     5.9879    -2.36   0.009
```

Based on our sample mean and our level of risk (alpha = .05) we're willing to take, the population mean on the truck is no greater than 5.9879 ozs. (The candy bars are too small!!)



Power of the Test (1-β)

In the previous problem (Ex #1) we determined there was insufficient evidence from our sample to conclude the average weight of our candy bars was at least 6 oz. So we rejected our null hypothesis. In other words, we said " There was only a .9% chance of being wrong if we concluded the average weight of the truck load of candy bars was below 6 ozs.

Does this mean there is a 99.1% (100-.009), chance of being right if we conclude the average weight of the truck load of candy bars is less than 6 ozs?

The answer is NO!

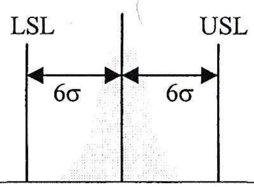
<u>The question we are now trying to answer is:</u>

"What is the chance we would reject the null and be correct?

Or

"If we randomly selected a sample of 50 candy bars, what is the chance we would conclude the average is less than 6 ounces?

This is called the <u>"Power of the Test"</u>



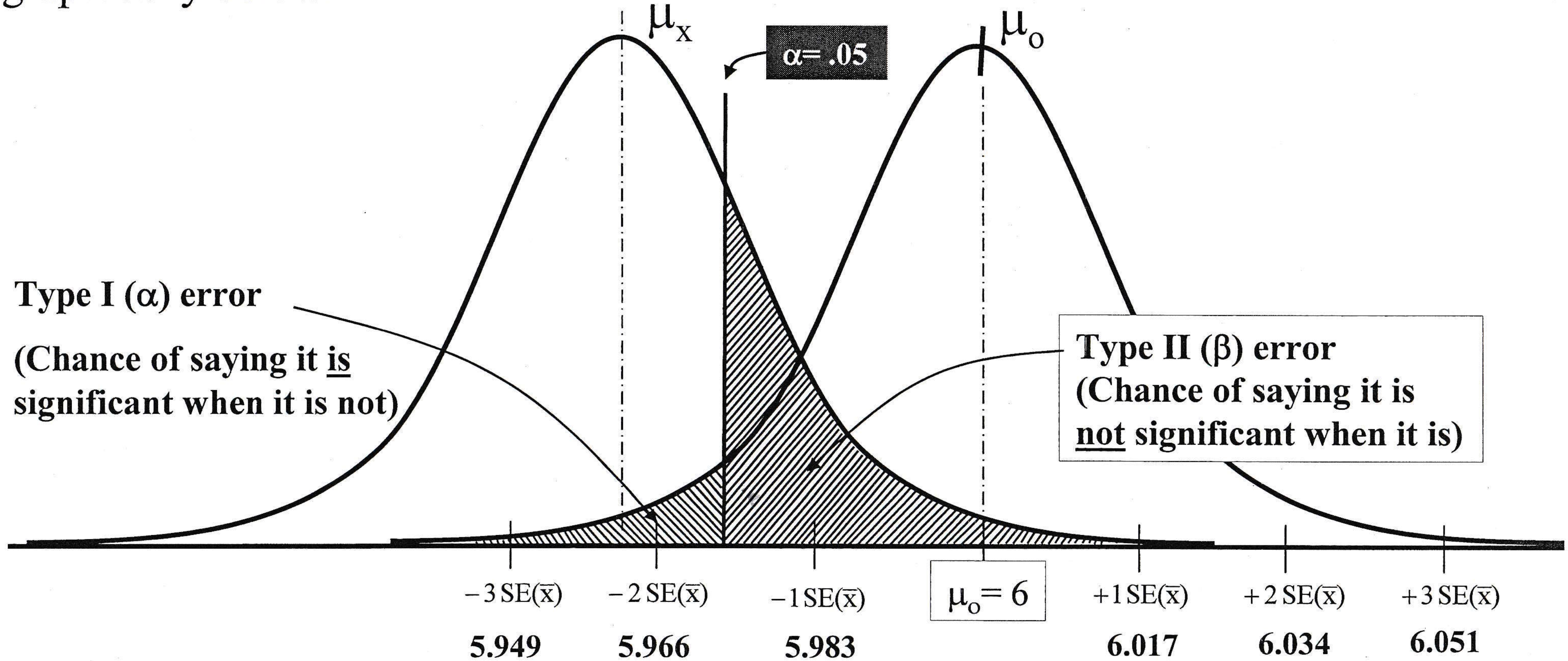
Power of the Test (cont)

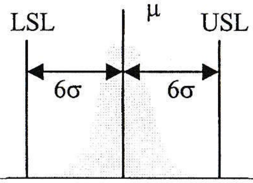
Sssssssssssssss

To answer our question we will compare the sample distribution curve from the null hypothesis (μ_o) to the sample distribution curve representing the 50 candy bar samples drawn from the truck, which we will call (μ_x).

Before we can explain "power" (1-β) we must understand β, which is shown graphically below.

S





Power of the Test (cont)

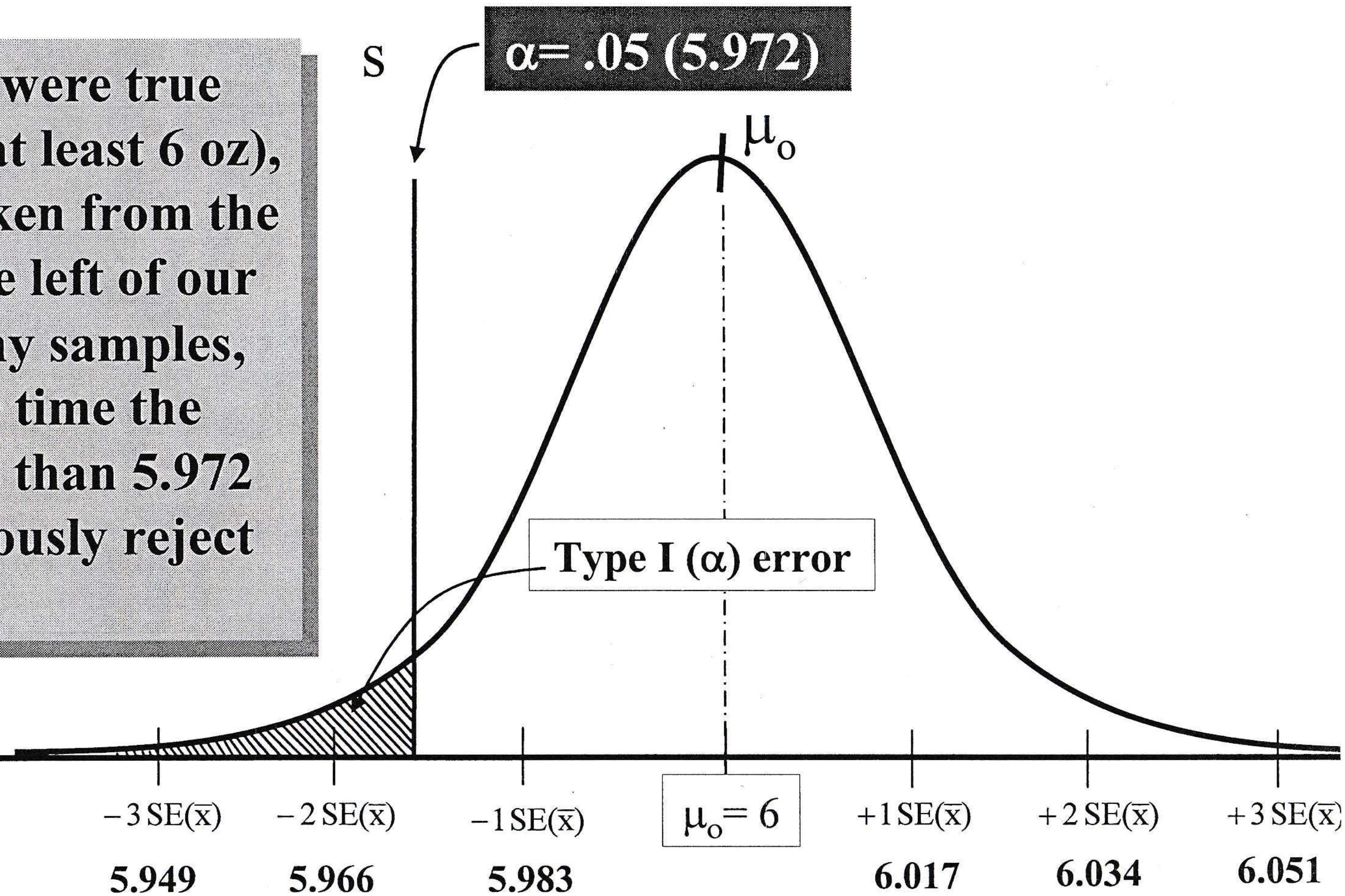
The Tech Group
WE PROVIDE INNOVATIVE SOLUTIONS

Ssssssssssssssss

Let's consider each type of error:

<u>Type I (α) error:</u> Rejecting the null hypothesis, when the null is true.

If the null hypothesis were true (candy bars average at least 6 oz), 5% of the samples taken from the truck would fall to the left of our α line. If we took many samples, (of size 50), 5% of the time the average would be less than 5.972 and we would erroneously reject the null hypothesis.



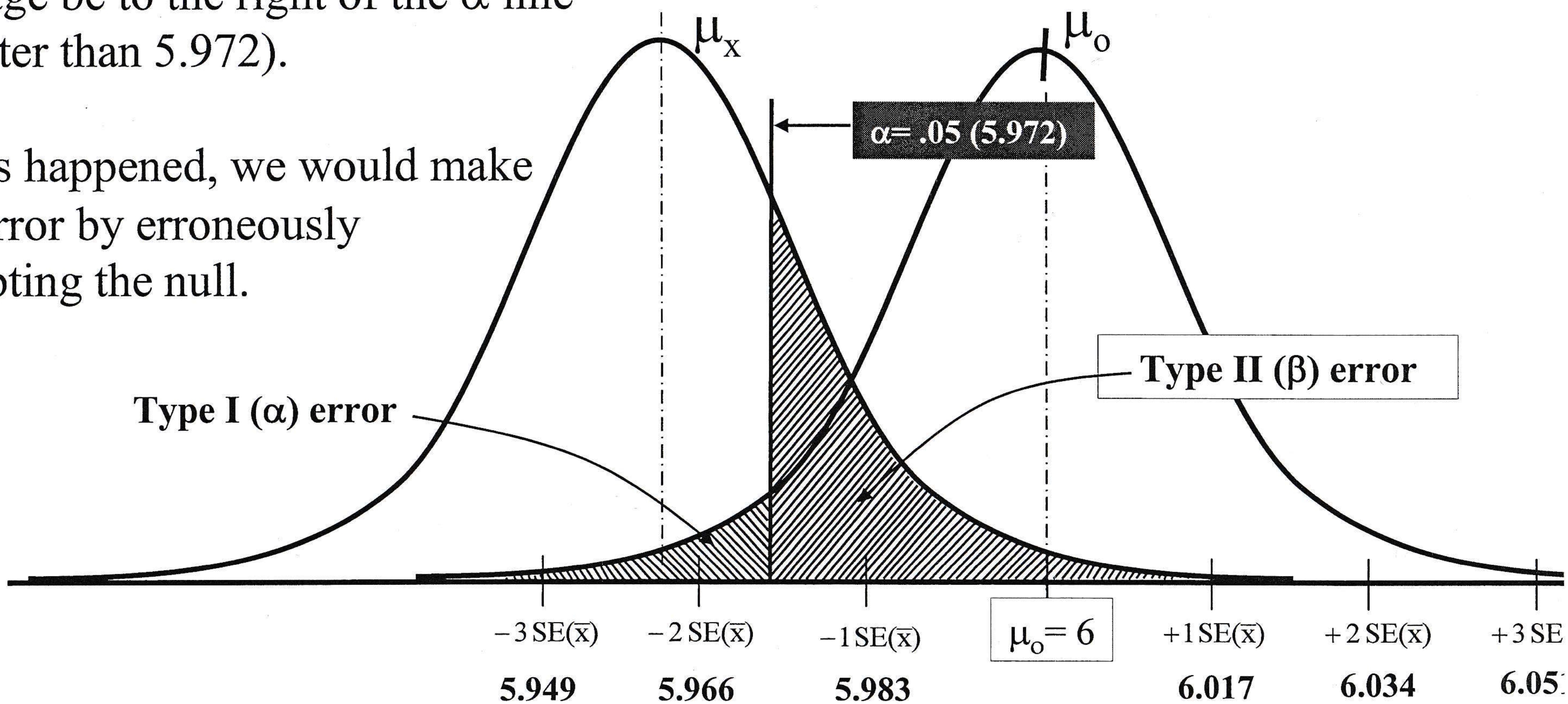
Power of the Test (cont)

We have concluded, based on our sample average, the null is not true, so let's consider our chances of making a Type II (β) error.

<u>Type II (β) error:</u> Accepting the null when the null is <u>not</u> true.

Based on our sample size of 50, the shaded area to the right of the α line represents the chance that we could draw a sample from the truck and have the average be to the right of the α line (greater than 5.972).

If this happened, we would make a β error by erroneously accepting the null.



Power of the Test (cont)

The Tech Group
WE PROVIDE INNOVATIVE SOLUTIONS

It appears there is a fairly high chance a 50 piece sample drawn from the truck will average greater than 5.972 ($\alpha = .05$). We said "the null is not true", yet, samples averaging greater than 5.972 would cause us to make a Type II (β) error.

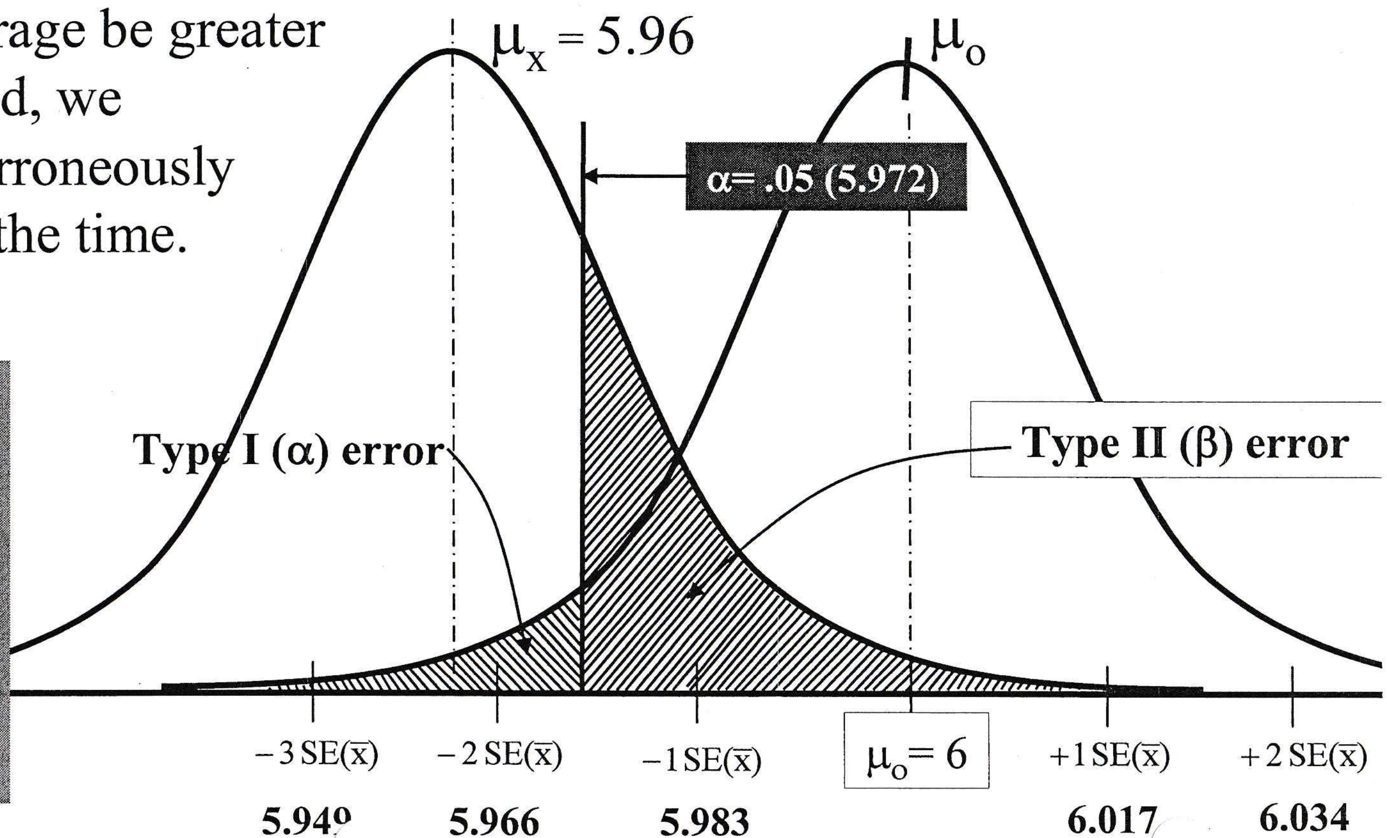
<u>Type II (β) error:</u> Accepting the null when the null is <u>not</u> true.

Based on our sample size of 50, the shaded area to the right of the α line represents the chance that we could draw a sample from the truck and have the average be greater than 5.972. If this happened, we would make a β error by erroneously accepting the null 24% of the time.

$$Z = \frac{5.972 - 5.96}{.017} = .705$$

Which corresponds to a β value of .24 (24%)

(.5 - .26 = .24)



Power of the Test (cont)

The Tech Group
WE PROVIDE INNOVATIVE SOLUTIONS

Finally, we can determine the “Power of the Test”

<u>Power of the Test:</u> Ability to detect a significant difference (1-β)

The blue shaded area to the left of the α line represents the chance that we could draw a sample from the truck and have the average be less than 5.972. This area represents the **“Power of the Test”** and is the proportion of the time we can expect to reject the null hypothesis and be right.

Since β is .24, Power = (1-β) = (1-.24) = .76

There is a 76% chance our average candy bar weight is less than 6.0 oz.