Hypothesis Testing

As was stated in the beginning of this lesson, hypothesis testing is concerned with determining if the value of a parameter is equal to a specific value. Below is an example of a hypothesis testing situation.

Many scientists have become concerned that there has been a depletion in the ozone level. The normal mean concentration level at an altitude of 19 miles is 100 parts per million (ppm). Suppose we wanted to test that there has been a depletion in the mean ozone concentration level.

Important Terms and Definitions

The **null hypothesis (H_0)** is the statement that is assumed to be true unless sufficient evidence is gathered to reject the hypothesis.

The **alternative hypothesis (H_a)** is the statement that one wishes to support as being true. This is done by gathering evidence against the null hypothesis.

A **Type I error** occurs by rejecting the null hypothesis when it is true.

A **Type II error** occurs by failing to reject the null hypothesis when it is false.

α is the probability of making a Type I error.

β is the probability of making a Type II error.

	State of Nature	
Decision	H_0 Is True	H_0 Is False
Reject H_0	Type I Error	Good Decision
Fail to Reject H_0	Good Decision	Type II Error

Many scientists have become concerned that there has been a depletion in the ozone level. The normal mean concentration level at an altitude of 19 miles is 100 parts per million (ppm). Suppose we wanted to test that there has been a depletion in the mean ozone concentration level.
What would be the null and alternative hypotheses?

What would be the Type I error for <u>this</u> problem?

What would be the Type II error for <u>this</u> problem?

If you were given a choice between the following probabilities, which would you choose? Why?

a. α=.025 and β=.10
b. α=.10 and β=.025

Suppose you were considering going bungee jumping, but would only go if the mean breaking strength was greater than 250 lbs. I would like to test if the mean breaking strength is greater than 250lbs.

What would be the null and alternative hypotheses?

What would be the Type I error?

What would be the Type II error?

If you were given a choice between the following probabilities, which would you choose? Why?

a. α=.01 and β=.10
b. α=.10 and β=.01

Conducting an Hypothesis Test

Important Terms and Definitions

The **test statistic** is a statistic, calculated from the sample data, that depends on the value of the parameter that you are testing.

The **level of significance** is the maximum chance of making a Type I Error the researcher is willing to take.

The **rejection region** is the region of the sampling distribution where the null hypothesis is rejected.

The **critical value(s)** is(are) the value(s) of the boundary of the rejection region.

The **p-value** is the probability of observing a value of the test statistic as contradictory (or more) to the null hypothesis as the computed value of the test statistic. It is also called the **observed level of significance**.

Hypothesis Testing Methods

1. Original Scale
1. Standardized Scale
2. P-value

Hypothesis Testing Procedure for Original and Standardized Scale Method

1. State the alternative hypothesis H_a.
2. State the null hypothesis H_0.
3. State the Type I and II Errors for the hypotheses.
4. State the level of significance (maximum acceptable α).
5. Determine the rejection region.
6. Compute the test statistic.
7. Compare the test statistic and rejection region and make a decision.
8. Draw a conclusion.

Hypothesis Testing Procedure for the P-value Method

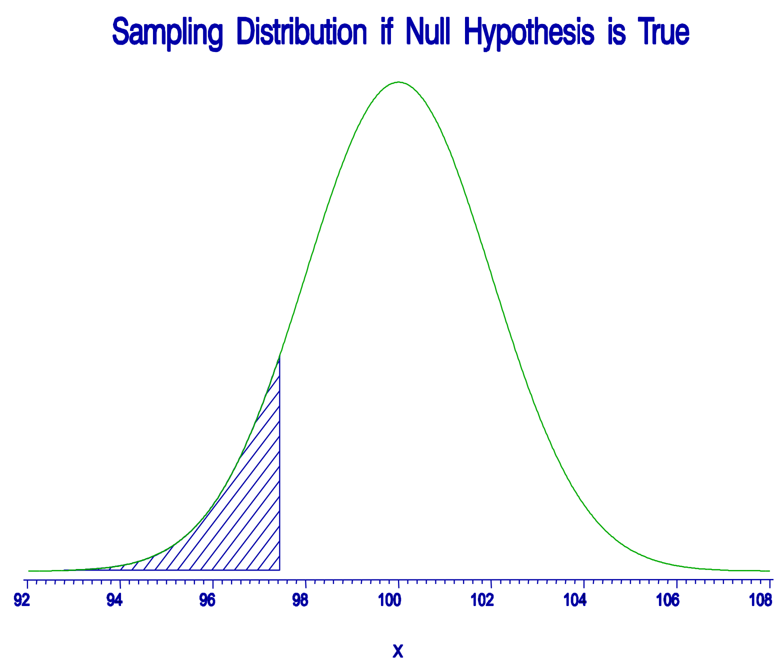
1. State the alternative hypothesis H_a.
2. State the null hypothesis H_0.
3. State the Type I and II Errors for the hypotheses.
4. State the level of significance (maximum acceptable α).
5. Compute the test statistic.
6. Calculate the p-value.
7. Compare the p-value with the level of significance and make a decision.
8. Draw a conclusion.

Original Scale Method

Many scientists have become concerned that there has been a depletion in the ozone level. The normal mean concentration level at an altitude of 19 miles is 100 parts per million (ppm). To determine if there has been a depletion, thirty-six measurements were taken, and the sample mean and sample standard deviation were 95 and 12 ppm, respectively. Do these sample results provide significant evidence that there has been a depletion in the mean ozone concentration level? (Use a level of significance of .10.)

Hypotheses

Rejection Region



Test Statistic

Decision

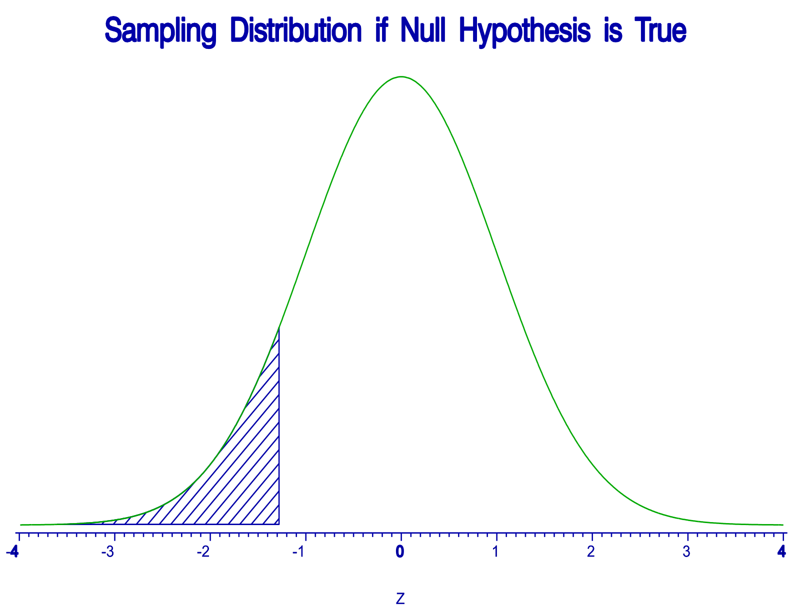
Conclusion

Standardized Scale Method

Many scientists have become concerned that there has been a depletion in the ozone level. The normal mean concentration level at an altitude of 19 miles is 100 parts per million (ppm). To determine if there has been a depletion, thirty-six measurements were taken and, the sample mean and sample standard deviation were 95 and 12 ppm, respectively. Do these sample results provide significant evidence that there has been a depletion in the mean ozone concentration level? (Use a level of significance of .10.)

Hypotheses

Rejection Region



Test Statistic

Decision

Conclusion

P-value Method

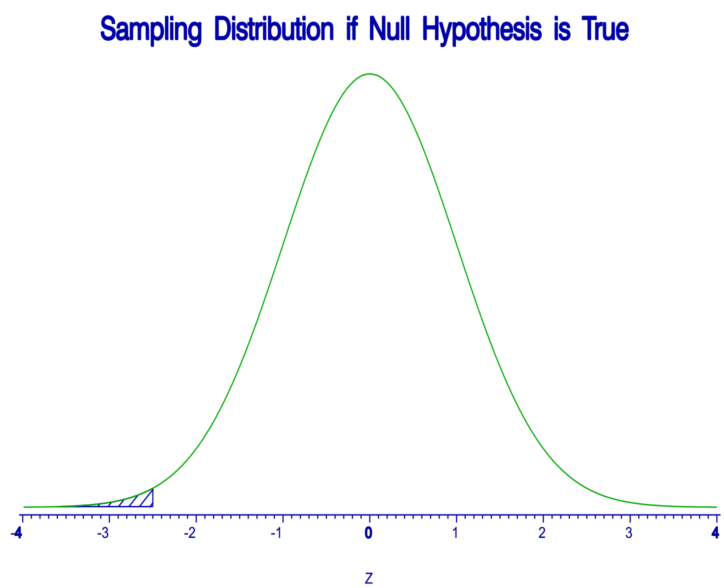
Many scientists have become concerned that there has been a depletion in the ozone level. The normal mean concentration level at an altitude of 19 miles is 100 parts per million (ppm). To determine if there has been a depletion, thirty-six measurements were taken, and the sample mean and sample standard deviation were 95 and 12 ppm, respectively. Do these sample results provide significant evidence that there has been a depletion in the mean ozone concentration level? (Use a level of significance of .10.)

<u>Hypotheses</u>

<u>Level of Significance</u>

<u>Test Statistic</u>

<u>P-value</u>



<u>Decision</u>

<u>Conclusion</u>

Decision Rule for Original Scale and Standardized Scale Hypothesis Tests

1. If H_a: $\mu > \mu_0$

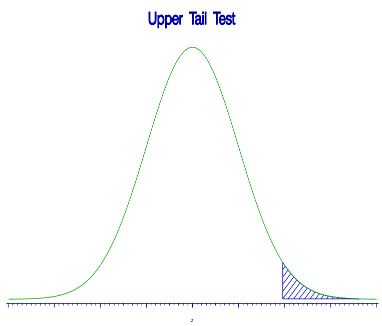
   Reject the null hypothesis if the observed value is greater than the critical value.

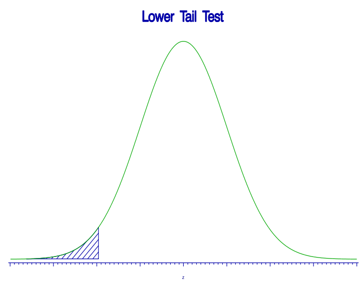
2. If H_a: $\mu < \mu_0$

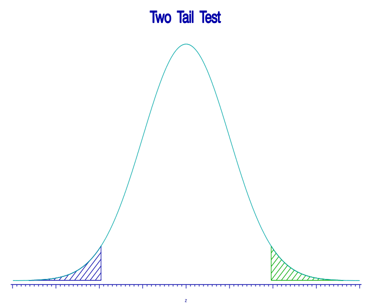
   Reject the null hypothesis if the observed value is less than the critical value.

3. If H_a: $\mu \neq \mu_0$

   Reject the null hypothesis if the absolute value of the observed value is greater than the absolute critical value.







P-value Method

Calculating the p-value

If H_a: $\mu > \mu_0$

Calculate the area to the right of the test statistic.

If H_a: $\mu < \mu_0$

Calculate the area to the left of the test statistic.

If H_a: $\mu \neq \mu_0$

Calculate the area to the right of the absolute value of the test statistic, and multiply that area by 2.

Decision Rule for P-value Method

If the p-value is less than or equal to α, then reject H_0.
If the p-value is greater than α, then fail to reject H_0.

Large Sample Hypothesis Tests for a Single Population Mean

Hypotheses

H_0: $\mu = \mu_0$	H_0: $\mu = \mu_0$	H_0: $\mu = \mu_0$
H_a: $\mu < \mu_0$	H_a: $\mu > \mu_0$	H_a: $\mu \neq \mu_0$

Test Statistic

$$Z_{obs} = \frac{\bar{y} - \mu_0}{s/\sqrt{n}}$$

Distribution for the Rejection Region

The rejection region is found using a standard normal distribution.

Small Sample Hypothesis Tests for a Single Population Mean

Hypotheses

H_0: $\mu = \mu_0$	H_0: $\mu = \mu_0$	H_0: $\mu = \mu_0$
H_a: $\mu < \mu_0$	H_a: $\mu > \mu_0$	H_a: $\mu \neq \mu_0$

Test Statistic

$$t_{obs} = \frac{\bar{y} - \mu_0}{s/\sqrt{n}}$$

Distribution for the Rejection Region

The rejection region is found using a t-distribution with df=n-1.

The Probability of a Type II Error and the Power of a Hypothesis Test

You will recall that a Type II error is failing to reject a null hypothesis that is false. The probability of making a Type II error (β) depends on how close the actual population mean (μ_a) is to the hypothesized mean (μ_0). The further μ_a is from μ_0, the smaller the probability of making a Type II error will be. The **power** of a hypothesis test is 1-β. It represents the probability of rejecting the null hypothesis when the actual population mean is equal to a specific value μ_a.

For the Ozone Problem

What is β if the true population mean were μ_a=98?

$$= \mathbf{P(Failing\ to\ reject\ H_0 \mid \mu_a = 98)}$$

$$= \mathbf{P(\bar{X} > 97.44 \mid \mu_a = 98)}$$

$$= \mathbf{P\left(\frac{\bar{X} - {}_a}{s/\sqrt{n}} > \frac{97.44 - 98}{12/\sqrt{36}} \mid \mu_a = 98 \right)}$$

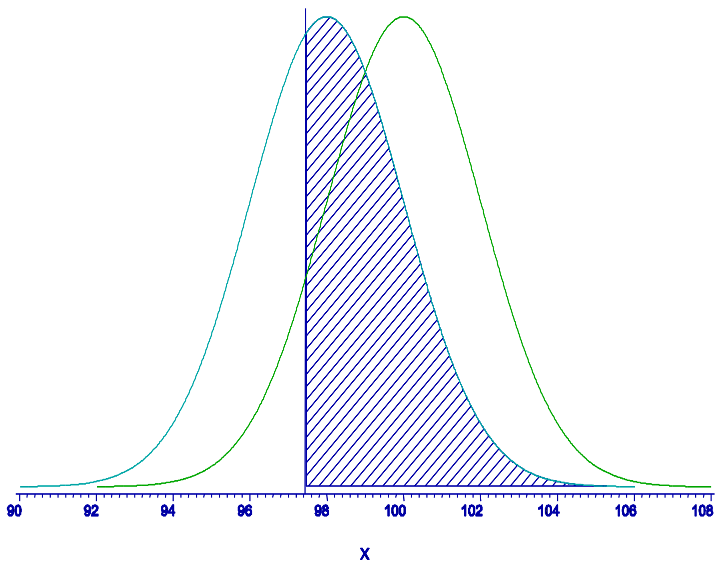
$$= \mathbf{P(Z > -.28 \mid \mu_a = 98)}$$

$$= \mathbf{.6103}$$

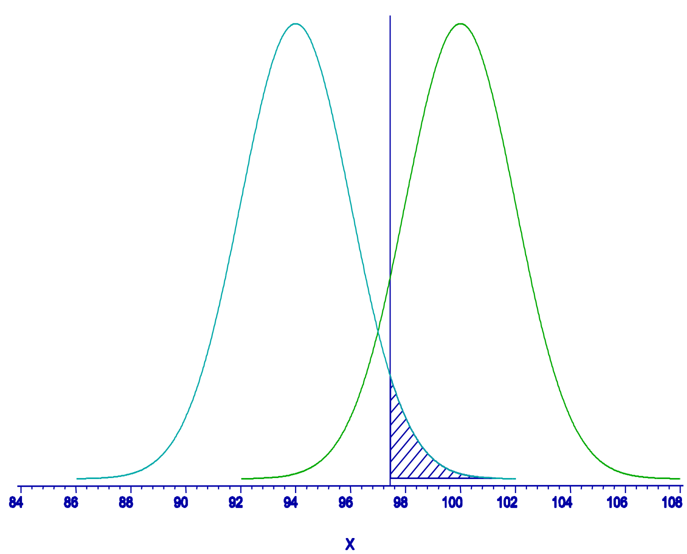
The power would be 1-.6103 = .3897.

Normal Plots Illustrating the Calculation of β

β when the true alternative is μ=98



β when the true alternative is μ=94



Choosing the Sample Size for Testing μ

The following formulae may be used for situations where the probability of a Type I error is to be α and the probability of a Type II error is to be β when the actual mean lies a distance of Δ from the hypothesized mean.

$$\text{One - Sided Test of } \mu$$

$$n = \frac{\sigma^2 \left(z_\alpha + z_\beta\right)^2}{\Delta^2}$$

$$\text{Two - Sided Test of } \mu$$

$$n = \frac{\sigma^2 \left(z_{\alpha/2} + z_\beta\right)^2}{\Delta^2}$$

Example: Ozone Problem. What sample size would be needed to detect a mean ozone concentration level of 95 ppm 90% of the time when the probability of making a Type I error is set at 10%?