

---

---

# MODULE 13

---

## HYPOTHESIS TESTS

**Objectives:**

1. Describe the relationship between the scientific method and statistical hypothesis testing.
2. Properly construct statistical hypotheses.
3. Describe the concept underlying significance testing.
4. Describe possible errors in statistical decision making.

**Contents**

13.1 Hypothesis Testing & Scientific Method . . . . .	167
13.2 Statistical Hypotheses . . . . .	167
13.3 Test Statistics and Effect Sizes . . . . .	172
13.4 Hypothesis Testing Concept Summary . . . . .	173
13.5 Errors and Power . . . . .	174

A STATISTIC IS AN IMPERFECT ESTIMATE of the unknown parameter because of sampling variability. There are two calculations using the results of a single sample that recognize this imperfectness and allow conclusions to be made about a parameter. First, a researcher may form an *a priori* hypothesis about the parameter and then use the information in the sample to make a judgment about the “correctness” of that hypothesis. Second, a researcher may form, from the information in the sample, a range of values that is likely to contain the parameter. The first method is called *hypothesis testing* and is the subject of this module. The second method consists of constructing a *confidence region*, which is introduced in Module 14. Specific applications of these two techniques are described in Modules 15-19.

## 13.1 Hypothesis Testing & Scientific Method

In its simplest form, the scientific method has four steps:

1. Observation and description of a natural phenomenon.
2. Formulation of a hypothesis to explain the phenomenon.
3. Use the hypothesis to predict new observations.
4. Experimentally test the predictions.

If the results of the experiment do not match the predictions, then the hypothesis is rejected and an alternative hypothesis is proposed. If the results of the experiment closely match the predictions, then belief in the hypothesis is gained, but the hypothesis will likely be subjected to further scrutiny.

Statistical hypothesis testing is key to using the scientific method in many fields of study and, in fact, closely follows the scientific method in concept. Statistical hypothesis testing begins by formulating two competing statistical hypotheses from a research hypothesis. One of these hypotheses (the null) is used to predict a parameter of interest. Data is then collected and statistical methods are used to determine whether the observed statistic closely matches the prediction made from the null hypothesis or not. Probability (Module 11) is used to measure the degree of matching and sampling variability is taken into account. This process and the theory underlying statistical hypothesis testing is explained in this module.

## 13.2 Statistical Hypotheses

Hypotheses in a research study are classified into two types: (1) research hypothesis and (2) statistical hypotheses. A research hypothesis is a “wordy” statement about the question or phenomenon that the researcher is testing. The research hypothesis is transferred into statistical hypotheses that are mathematical and more easily subjected to statistical methods.

**Δ Research Hypothesis:** A general statement about the question or phenomenon being tested.

**Δ Statistical Hypothesis:** Mathematical statements about the question or phenomenon being tested.

There are two types of statistical hypotheses: (1) the null hypothesis and (2) the alternative hypothesis. The **null hypothesis**, abbreviated as  $H_0$ , is a specific statement of no difference between a parameter and a specific value or between two parameters. The  $H_0$  ALWAYS contains an equals sign because it always represents “no difference.” The **alternative hypothesis**, abbreviated as  $H_A$ , always states that there is some sort of difference between a parameter and a specific value or between two parameters. The type of difference comes from the research hypothesis and will require use of a less than ( $<$ ), greater than ( $>$ ), or not equals ( $\neq$ ) sign.

**Δ Null Hypothesis:** A statistical hypothesis that states specifically that there is no difference between a parameter and a specific value or between two parameters; typically abbreviated with  $H_0$ .

**Δ Alternative Hypothesis:** A statistical hypothesis that states a specific difference between a parameter and a specific value or between two parameters; typically abbreviated with  $H_A$ .

◇ Null hypotheses always represent the “no difference” situation and, thus, always contain an equals sign.

◇ Alternative hypotheses always represent some sort of difference and, thus, always contain one of these three directional symbols ( $\neq$ ,  $>$ , and  $<$ ).

The relationships between the research, null, and alternative hypotheses are illustrated with the following examples:

1. A medical researcher is concerned that a new medicine may change the patients' mean pulse rate (from the “known” mean pulse rate of 82 bpm for individuals in the study population not using the new medicine).
  - $H_A : \mu \neq 82$  and  $H_0 : \mu = 82$  (where  $\mu$  represents the mean pulse rate for individuals in the study population that take the new medicine; thus, the alternative hypothesis represents a change from the “normal” pulse rate).
2. A chemist has invented an additive to automobile batteries that she thinks will extend the current 36 month average life of a battery.
  - $H_A : \mu > 36$  and  $H_0 : \mu = 36$  (where  $\mu$  represents mean life of batteries with the new additive; thus, this alternative hypothesis represents an extension of the current battery life).
3. An engineer wants to determine if a new type of insulation will reduce the average heating costs of a typical house (which are currently \$145 per month).
  - $H_A : \mu < 145$  and  $H_0 : \mu = 145$  (where  $\mu$  represents the mean monthly heating bill for houses that receive the new type of insulation; thus, this alternative hypothesis represents a decline in heating bills from the previous “normal” amount).

The sign used in the alternative hypothesis comes directly from the wording of the research hypothesis (Table 13.1). An alternative hypothesis that contains the  $\neq$  sign is called a **two-tailed alternative**, as the value can be “not equal” to another value in two ways; i.e., less than or greater than. Alternative hypotheses with the  $<$  or the  $>$  signs are called **one-tailed alternatives**. The null hypothesis is easily constructed from the alternative hypothesis by replacing the sign in the alternative hypothesis with an equals sign.

◇ The “not-equals” alternative is called a two-tailed alternative, whereas the other two alternative hypotheses are called one-tailed alternatives.

Table 13.1. Common words that indicate which sign to use in the alternative hypothesis.

$>$	$<$	$\neq$
is greater than	is less than	is not equal to
is more than	is below	is different from
is larger than	is lower than	has changed from
is longer than	is shorter than	is not the same as
is bigger than	is smaller than	
is better than	is reduced from	
is at least	is at most	
is not less than	is not more than	

## Review Exercises

- 13.1** A researcher is investigating the mean growth of a certain cactus under a variety of environmental conditions. Under the current environmental conditions, he hypothesizes that mean growth is no more than 4 cm. What is  $H_0$  and  $H_A$  in this situation? [Answer](#)
- 13.2** Machowiak *et al.* (1992) critically examined the belief that the mean body temperature differed from 98.6°F by measuring the body temperatures of 93 healthy humans. What is  $H_0$  and  $H_A$  in this situation? [Answer](#)
- 13.3** A study by Cheshire *et al.* (1994) reported on six patients with chronic myofascial pain syndrome. The authors were examining the hypothesis that the mean pain length was greater than 2.5 years. What is  $H_0$  and  $H_A$  in this situation? [Answer](#)

### 13.2.1 Hypothesis Testing Concept

Statistical hypothesis testing begins by using the null hypothesis to make a prediction of what value one should expect for the mean in a sample. So, for the Square Lake example (from Module ??), if  $H_0 : \mu = 105$  and  $H_A : \mu < 105$ , then one would expect, if the null hypothesis is true, that the observed sample mean would be 105. If sampling variability did not exist and the observed sample mean was NOT equal to 105, then the prediction based on the null hypothesis would not be supported and the conclusion would be that the null hypothesis is incorrect. In other words, one would conclude that the population mean was not equal to 105.

Of course, sampling variability does exist and its existence complicates matters. The simple interpretation of not supporting  $H_0$  because the observed sample mean did not equal the hypothesized population mean can NOT be made because, with sampling variability, one would not expect a statistic to exactly equal the parameter in the population from which the sample was extracted. For example, even if the null hypothesis was correct, then one would not expect, with sampling variability, the observed sample mean to exactly equal 105; rather, one would expect the observed sample mean to be **reasonably** close to 105.

Thus, hypothesis testing is a procedure for determining if the difference between the observed statistic and the expected statistic based on the null hypothesis is “large” **relative to sampling variability**. For example, the standard error of  $\bar{x}$  in samples of  $n = 50$  in the Square Lake example is  $\frac{\sigma}{\sqrt{n}} = \frac{31.5}{\sqrt{50}} = 4.45$ . Thus, with this amount of sampling variability, an observed sample mean of 103 would be considered reasonably close

to 105 and one would have more belief in  $H_0 : \mu = 105$ . However, an observed sample mean of 70 would be considered further away from 105 than one would expect based on sampling variability alone and the belief in  $H_0 : \mu = 105$  would lessen.

While the above procedure is intuitively appealing, it loses some of its objectivity when the examples chosen (i.e., samples means of 103 and 70) are not as extremely close or distant from the null hypothesized value (e.g., what would the conclusion be if the observed sample mean was 97?). A first step in creating a more objective decision criteria is to compute the “p-value.” A p-value is the probability of the observed statistic or a value of the statistic more extreme assuming that the null hypothesis is true. The p-value is described in more detail below given its centrality to making conclusions about statistical hypotheses.

Δ **p-value:** The probability of the observed statistic or a value of the statistic more extreme assuming the null hypothesis is true.

The meaning of the phrase “or more extreme” is derived from the sign in  $H_A$  (Figure 13.1). If  $H_A$  is the “less than” situation, then “or more extreme” means “less than” or “shade to the left” for the probability calculation. The “greater than” situation is defined similarly but would result in shading to the “right.” In the “not equals” situation, “or more extreme” means further into the tail AND the exact same size of tail on the other side of the distribution. It is clear from Figure 13.1 why “less than” and “greater than” are one-tailed alternatives and “not equals” is a two-tailed alternative.

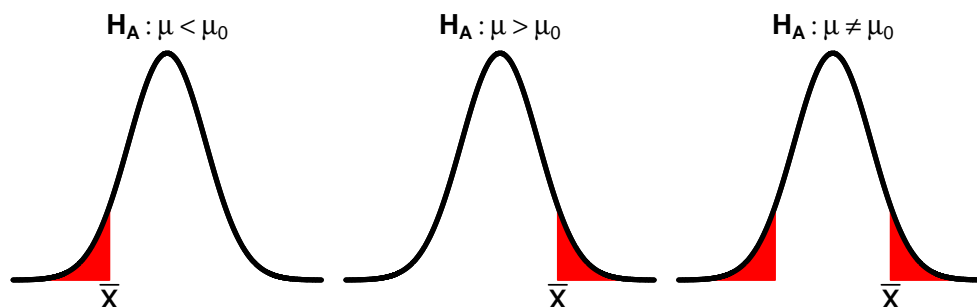


Figure 13.1. Depiction of “or more extreme” (red shaded area) in p-values for the three possible alternative hypotheses.

The “assuming that the null hypothesis is true” phrase is used to define a  $\mu$  for the sampling distribution on which the p-value will be calculated. This sampling distribution is called the **null distribution** because it depends on the value of  $\mu$  in the null hypothesis. One must remember that the null distribution represents the distribution of all possible sample means assuming that the null hypothesis is true; it does NOT represent the actual sample means.<sup>1</sup> The null distribution in the Square Lake example is thus  $\bar{x} \sim N(105, 4.45)$  because  $n = 50 > 30$ ,  $H_0 : \mu = 105$ , and  $SE = \frac{31.49}{\sqrt{50}} = 4.45$ .

The p-value is computed with a “forward” normal distribution calculation on the null sampling distribution. For example, suppose that a sample mean of 100 was observed with  $n = 50$  from Square Lake (as it was in Table 2.2). The p-value in this case would be “the probability of observing  $\bar{x} = 100$  or a smaller value assuming that  $\mu = 105$ .” This probability is computed by finding the area to the left of 100 on a  $N(105, 4.45)$  null distribution and is the exact same type of calculation that was made in Section 12.4. Thus, this p-value of  $p = 0.1308$  is computed below (Figure 13.2).

<sup>1</sup>Of course, unless the null hypothesis happens to be perfectly true.

```
> ( distrib(100,mean=105,sd=31.49/sqrt(50)) )
[1] 0.1307722
```

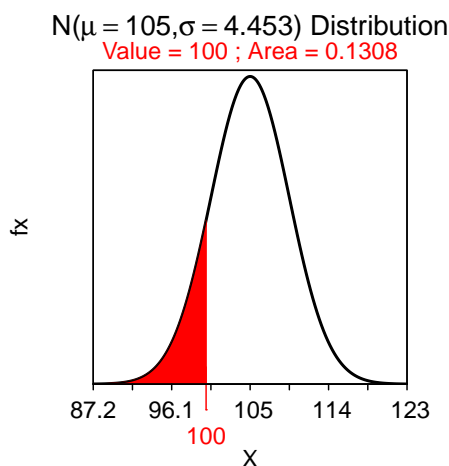


Figure 13.2. Depiction of the p-value for the Square Lake example where  $\bar{x} = 100$  and  $H_A : \mu < 105$ .

Interpreting the p-value requires critically thinking about the p-value definition and how it is calculated. Small p-values appear when the observed statistic is “far” from the value expected from the null hypothesis. In this case there is a small probability of seeing the observed statistic ASSUMING that  $H_0$  is true. Thus, the assumption is likely wrong and  $H_0$  is likely incorrect. In contrast, large p-values appear when the observed statistic is close to the null hypothesized value suggesting that the assumption about  $H_0$  may be correct.

◇ Small p-values are evidence against the null hypothesis.

The p-value serves as a numerical measure on which to base a conclusion about  $H_0$ . To do this objectively requires an objective definition of what it means to be a “small” or “large” p-value. Statisticians use a cut-off value, called the rejection criterion and symbolized with  $\alpha$ , such that p-values less than  $\alpha$  are considered small and would result in rejecting  $H_0$  as a viable hypothesis. The value of  $\alpha$  is typically small, usually set at 0.05, although  $\alpha = 0.01$  and  $\alpha = 0.10$  are also commonly used.

△  $\alpha$ : A predetermined rejection criterion value used in hypothesis testing. This value sets the “cutoff” for determining whether it was reasonable to have seen the observed statistic or not assuming the null hypothesis is true.

◇ Typical values of  $\alpha$  are 0.01, 0.05, and 0.10.

The choice of  $\alpha$  is made by the person conducting the hypothesis test and is based on how much evidence a researcher demands before rejecting  $H_0$ . Smaller values of  $\alpha$  require a larger difference between the observed statistic and the null hypothesized value and, thus, require “more evidence” of a difference for the  $H_0$  to be rejected. For example, if a rejection of the null hypothesis will be heavily scrutinized by regulatory agencies, then the researcher may want to be very sure before claiming a difference and should then set  $\alpha$  at a smaller

value, say  $\alpha = 0.01$ . The actual choice for  $\alpha$  MUST be made before collecting any data and canNOT be changed once the data has been collected. In other words, once the data are in hand, a researcher cannot lower or raise  $\alpha$  to achieve a desired outcome regarding  $H_0$ .

♦ The value of the rejection criterion ( $\alpha$ ) is set by the researcher BEFORE data is collected.

♦ Set  $\alpha$  to lower values to make it more difficult to reject  $H_0$ .

The null hypothesis in the Square Lake example is not rejected because the calculated p-value (i.e.,  $p = 0.1308$ ) is larger than any of the common values of  $\alpha$ . Thus, the conclusion in this example is that it is possible that the mean of the entire population is equal to 105 and it is not likely that the population mean is less than 105. In other words, observing a sample mean of 100 is likely to happen based on random sampling variability alone and it is unlikely that the null hypothesized value is incorrect.

## Review Exercises

- 13.4** Compute the p-value and make a decision about  $H_0$  with the following information –  $\alpha = 0.10$ ,  $H_0 : \mu = 10$ ,  $H_A : \mu > 10$ ,  $\sigma = 5$ ,  $n = 25$ , and  $\bar{x} = 12.1$ . [Answer](#)
- 13.5** Compute the p-value and make a decision about  $H_0$  with the following information –  $\alpha = 0.05$ ,  $H_0 : \mu = 50$ ,  $H_A : \mu < 50$ ,  $\sigma = 20$ ,  $n = 50$ , and  $\bar{x} = 43.8$ . [Answer](#)
- 13.6** Compute the p-value and make a decision about  $H_0$  with the following information –  $\alpha = 0.01$ ,  $H_0 : \mu = 100$ ,  $H_A : \mu \neq 100$ ,  $\sigma = 15$ ,  $n = 100$ , and  $\bar{x} = 98$ . [Answer](#)
- 13.7** Describe why we must formally go through the steps of a hypothesis test to conclude that  $\mu > 11$  when we observe  $\bar{x} = 12.1$ . [Answer](#)

## 13.3 Test Statistics and Effect Sizes

Instead of reporting the observed statistic and the resulting p-value, it may be of interest to know how “far” the observed statistic was from the hypothesized value of the parameter. This is easily calculated with

$$\text{Observed Statistic} - \text{Hypothesized Parameter}$$

where “Hypothesized Parameter” represents the specific value in  $H_0$ . However, the meaning of this value is difficult to interpret without an understanding of the standard error of the statistic. For example, a difference of 10 between the observed statistic and the hypothesized parameter seems “very different” if the standard error is 1 but does not seem “different” if the standard error is 100. Thus, it is common practice to standardize this difference by dividing by the standard error of the statistic. This measure of distance is called a *test statistic* and is generalized with

$$\text{Test Statistic} = \frac{\text{Observed Statistic} - \text{Hypothesized Parameter}}{SE_{\text{Statistic}}} \quad (13.3.1)$$

Thus, the test statistic (13.3.1) measures how many standard errors the observed statistic is away from the hypothesized parameter. Relatively large values are indicative of a difference that is likely not due to randomness (i.e., sampling variability) and suggest rejecting the null hypothesis. There are other forms for calculating test statistics, but all test statistics retain the general idea of scaling the difference between what was observed and what was expected from the null hypothesis in terms of sampling variability. Even though there is a one-to-one relationship between a test statistic and a p-value, a test statistic is often reported with a hypothesis test to give another feel for the magnitude of the difference between what was observed and what was predicted.

♦ A test statistic measures how many standard errors the observed statistic is away from the hypothesized parameter.

## 13.4 Hypothesis Testing Concept Summary

In summary, hypotheses are statistically examined with the following procedure.

1. Construct null and alternative hypotheses from the research hypothesis.
2. Construct an expected value of the statistic based on the null hypothesis (i.e., assume that the null hypothesis is true).
3. Calculate an observed statistic from the individuals in a sample.
4. Compare the difference between the observed statistic and the expected statistic based on the null hypothesis in relation to sampling variability (i.e., calculate a test statistic and p-value).
5. Use the  $p$  – value to determine if this difference is “large” or not.
  - If this difference is “large” (i.e.,  $p$  – value  $< \alpha$ ), then reject the null hypothesis.
  - If this difference is not “large” (i.e.,  $p$  – value  $> \alpha$ ), then “Do Not Reject” the null hypothesis.

Statisticians say “do not reject  $H_0$ ” rather than “accept  $H_0$  as true” when the p-value  $> \alpha$  for two reasons. First, there are several other possible values, besides the specific value in the null hypothesis, that would lead to “do not reject” conclusions. For example, if a null hypothesized value of 105 was not rejected, then values of 104.99, 104.98, etc. would also likely not be rejected<sup>2</sup> So, we don’t say that we “accept” a particular hypothesized value when we know many other values would also be “accepted.”

Second, the null hypothesis is almost always not true. Consider the null hypothesis of the Square Lake example (i.e., “that the mean length is 105”). The mean length of fish in Square Lake is undoubtedly not exactly equal to 105. It may be 104.9, 105.01, or some other more disparate value. The point is that the specific value of the hypothesis is likely never true, especially for a continuous variable. The problem is that it takes large amounts of data to be able to distinguish means that are very close to the true population mean (i.e., it is difficult to distinguish between 104.9 and 105 when sampling variability is present). Very often we will not take a sample size large enough to distinguish these subtle differences. Thus, we will say that we “do not reject  $H_0$ ” because there simply was not enough data to reject it.



<sup>2</sup>In fact, for example, the values in a 95% confidence interval – see Module ?? – represent all possible hypothesized values that would not be rejected with a two-tailed  $H_A$  using  $\alpha = 0.05$ .



## Review Exercises

- 13.8** The managers of a wastewater treatment plant monitored the amount of biological oxygen demand (BOD; lbs/day) in the effluent of the plant each month from January 1991 to October 2000. The managers would need to take corrective actions if the average BOD over this time period was significantly greater than 2200 lbs/day at a 10% rejection level. Previous studies indicated that the standard deviation was 1200 lbs/day. Summary statistics from their sample of days is given below. Use this information to answer the questions below. [Answer](#)

n	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
118	630	1600	2240	2504	3193	6023

- What are the null and alternative hypotheses?
  - What is the test statistic?
  -  Compute the p-value.
  - Use the p-value to make a decision about  $H_0$ .
  - What does this mean for the managers of the plant (i.e., will they need to take action)? Explain!
- 13.9** Admissions representatives at the University of Minnesota medical school were concerned that the average grade point average of applicants in non-science courses had dropped below 3.7. A sample of 40 of the most recent applicants indicated that the mean was 3.60. Information from the Association of American Medical Colleges suggested that the overall standard deviation was 0.35. Use this information, and an  $\alpha = 0.05$ , to answer the questions below. [Answer](#)
- What are the null and alternative hypotheses?
  - What is the test statistic?
  -  Compute the p-value.
  - Use the p-value to make a decision about  $H_0$ .
  - Was the representatives concern about the average gpa of applicants warranted? Explain!

## 13.5 Errors and Power

The goal of hypothesis testing is to make a decision about  $H_0$ . Unfortunately, because of sampling variability, there is always a risk of making an incorrect decision. Two types of incorrect decisions can be made (Table 13.2). A Type I error occurs when a true  $H_0$  is falsely rejected. In other words, even if  $H_0$  is true, there is a chance that a rare sample will occur and  $H_0$  will be deemed incorrect. The probability of making a Type I error is set when  $\alpha$  is chosen. A Type II error occurs when a false  $H_0$  is not rejected. The probability of a Type II error is denoted by  $\beta$ .

**Δ Type I error:** Rejecting  $H_0$  when  $H_0$  was actually true. Probability of Type I error is  $\alpha$ .

**Δ Type II error:** Not rejecting  $H_0$  when  $H_0$  was actually false. Probability of Type II error is  $\beta$ .

Table 13.2. Types of decisions that can be made from a hypothesis test.

		Decision from Data	
		Reject	Not Reject
Truth About Population	$H_0$	Type I	Correct
	$H_A$	Correct	Type II

The decision in the Square Lake example above produced a Type II error because  $H_0 : \mu = 105$  was not rejected even though we know that  $\mu = 98.06$  (Table 2.1). Unfortunately, in real life, it will never be known exactly when a Type I or a Type II error has been made because the true  $\mu$  is not known. However, it is known that a Type I error will be made  $100\alpha\%$  of the time. The probability of a type II error ( $\beta$ ), though, is never known because this probability depends on the true  $\mu$ . Decisions can be made, however, that affect the magnitude of  $\beta$  (discussed below with power).

A concept that is very closely related to decision-making errors is the idea of **power**. Power is the probability of correctly rejecting a false  $H_0$ . In other words, it is the probability of detecting a difference from the hypothesized value if a difference really exists. Power is used to demonstrate how sensitive a hypothesis test is for identifying a difference. High power related to a  $H_0$  that is not rejected implies that the  $H_0$  really should not have been rejected. Conversely, low power related to a  $H_0$  that was not rejected implies that the test was very unlikely to detect a difference, so not rejecting  $H_0$  is not surprising nor particularly conclusive.

**△ Power:** The probability of correctly rejecting  $H_0$  when  $H_0$  was actually false.

Power is equal to  $1 - \beta$  and, thus, like  $\beta$  it cannot be computed directly. However, a researcher can make decisions that will positively affect power (Figure 13.3). For example, a researcher can positively impact power by increasing  $\alpha$  or  $n$ . Increasing  $n$  is more beneficial because it does not result in an increase in Type I errors as would occur with increasing  $\alpha$ . In addition, power decreases as the difference between the hypothesized mean ( $\mu_0$ ) and the actual mean ( $\mu_A$ ) decreases. This means that the ability to detect increasingly smaller differences decreases. In addition, power decreases with an increasing amount of natural variability (i.e.,  $\sigma$ ). In other words, the ability to detect a difference decreases with increasing amounts of variability among individuals. A researcher cannot control the difference between  $\mu_0$  and  $\mu_A$  or the value of  $\sigma$ . However, it is important to know that if a situation with a “large” amount of variability is encountered or the difference to be detected is small, the researcher will need to increase  $n$  to gain power.

◇ **Power =  $1 - \beta$ .**

◇ **Power will increase as the difference between the actual and hypothesized value of the parameter increases.**

◇ **Power will increase as the standard error of the statistic decreases. Thus, power increases as the sample size increases.**

◇ **Power will increase as the  $\alpha$  level increases.**

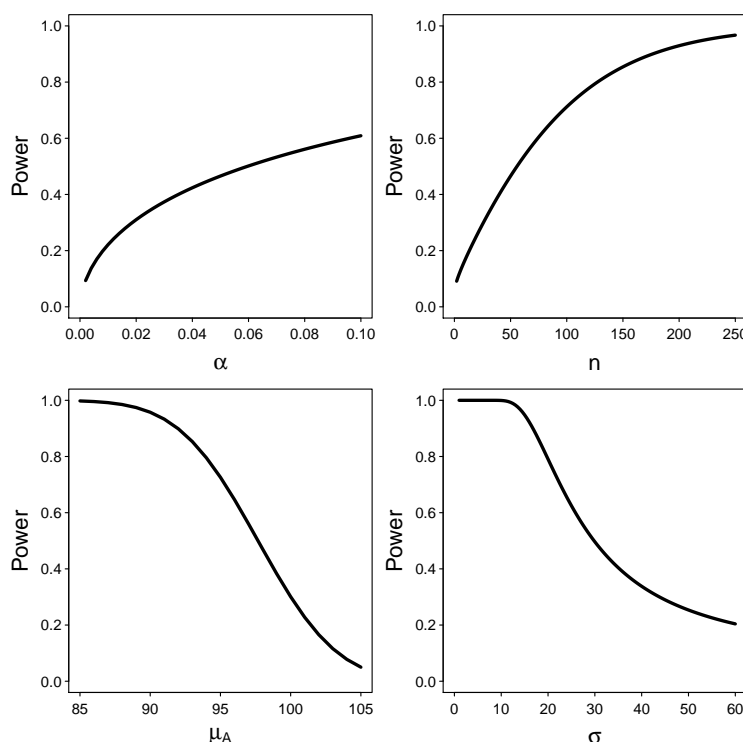


Figure 13.3. The relationship between one-tailed (lower) power and  $\alpha$ ,  $n$ , actual mean ( $\mu_A$ ), and  $\sigma$ . In all situations where the variable does not vary,  $\mu_0 = 105$ ,  $\mu_A = 98.06$ ,  $\sigma = 31.49$ ,  $n = 50$ , and  $\alpha = 0.05$ .

Power cannot usually be calculated because the actual mean ( $\mu_A$ ) is not known. However, in the Square Lake example,  $\mu_A$  is known and power can be calculated in four steps:

1. Draw the sampling distribution assuming the  $H_0$  is true (called the null distribution).
  - The null distribution is  $N(105, \frac{31.49}{\sqrt{50}})$  because  $H_0 : \mu = 105$ ,  $\sigma = 31.49$ , and  $n = 50$ .
2. Find the rejection region borders (based on  $\alpha$  and  $H_A$ ) in terms of the value of the statistic (a “reverse” calculation on the null distribution).
  - The rejection region is delineated by the  $\bar{x}$  that has  $\alpha = 0.10$  to the left (because  $H_A$  is a “less than”). The Z with 0.10 to the left of it is -1.282. Thus, the  $\bar{x}$  on the null distribution with 0.10 to the left of it is  $-1.282 \frac{31.49}{\sqrt{50}} + 105 = 99.2908$ .
3. Draw the sampling distribution corresponding to the “actual” parameter value (SE is the same as that for the null distribution).
  - The actual  $\mu$  is 98.06. Thus, the actual sampling distribution is  $N(98.06, \frac{31.49}{\sqrt{50}})$ .
4. Compute the portion of the “actual” sampling distribution in the REJECTION region of the null distribution (i.e., a “forward” calculation on the actual distribution).
  - This computation is to find the area to the left of 99.2908 on  $N(98.06, \frac{31.49}{\sqrt{50}})$ . The corresponding Z is  $\frac{99.2908 - 98.06}{\frac{31.49}{\sqrt{50}}} = 0.29$ . The area to the left of this Z is 0.6141.

Thus, the power to detect a  $\mu_A = 98.06$  was 0.6141. This means that in only about 61% of the samples will the false  $H_0 : \mu = 105$  be correctly rejected. Thus, it is not too surprising that  $H_0$  was not rejected in this

example. If  $n$  could be doubled to 100, however, the power to correctly reject  $H_0 : \mu = 105$  would increase to approximately 0.82 (Figure 13.3).

## Review Exercises

- 13.10** What is  $\beta$  if power=0.875? [Answer](#)
- 13.11** For a constant sample size,  $\sigma$ , and difference between the hypothesized and actual means, what happens to power, if  $\alpha$  is increased? [Answer](#)
- 13.12** For a constant  $\alpha$ ,  $\sigma$ , and difference between the hypothesized and actual means, what happens to power, if the sample size increases? [Answer](#)
- 13.13** For a constant  $\alpha$ ,  $\sigma$ , and sample size, what happens to power if the difference between the hypothesized and actual means increases? [Answer](#)
- 13.14** For a constant sample size,  $\sigma$ , and difference between the hypothesized and actual means, what happens to  $\beta$ , if  $\alpha$  is increased? [Answer](#)
- 13.15** For a constant  $\alpha$ ,  $\sigma$ , and difference between the hypothesized and actual means, what happens to  $\beta$ , if the sample size is increased? [Answer](#)
- 13.16** For a constant  $\alpha$ ,  $\sigma$ , and sample size, what happens to  $\beta$  if the difference between the hypothesized and actual means increases? [Answer](#)
- 13.17** Describe a real-life situation where you think that making a Type II error would be much more “costly” than making a Type I error. Completely describe the situation at hand and what Type I and a Type II errors mean in terms of the situation you describe. [Answer](#)
- 13.18** Compute power given the following information:  $\alpha = 0.05$ ,  $H_0 : \mu = 50$ ,  $H_A : \mu < 50$ ,  $\sigma = 20$ ,  $n = 50$ , and  $\mu_A = 45$ . [Answer](#)
- 13.19** Compute power given the following information:  $\alpha = 0.10$ ,  $H_0 : \mu = 10$ ,  $H_A : \mu > 10$ ,  $\sigma = 5$ ,  $n = 25$ , and  $\mu_A = 12$ . [Answer](#)
- 13.20** Compute power given the following information:  $\alpha = 0.01$ ,  $H_0 : \mu = 75$ ,  $H_A : \mu > 75$ ,  $\sigma = 15$ ,  $n = 30$ , and  $\mu_A = 82$ . [Answer](#)