



Outcomes and the Type I and Type II Errors

By:

OpenStaxCollege

When you perform a hypothesis test, there are four possible outcomes depending on the actual truth (or falseness) of the null hypothesis H_0 and the decision to reject or not. The outcomes are summarized in the following table:

ACTION	H_0 IS ACTUALLY	...
	True	False
Do not reject H_0	Correct Outcome	Type II error
Reject H_0	Type I Error	Correct Outcome

The four possible outcomes in the table are:

1. The decision is **not to reject H_0** when **H_0 is true (correct decision)**.
2. The decision is to **reject H_0** when **H_0 is true** (incorrect decision known as a Type I error).
3. The decision is **not to reject H_0** when, in fact, **H_0 is false** (incorrect decision known as a Type II error).
4. The decision is to **reject H_0** when **H_0 is false (correct decision** whose probability is called the **Power of the Test**).

Each of the errors occurs with a particular probability. The Greek letters α and β represent the probabilities.

α = probability of a Type I error = **$P(\text{Type I error})$** = probability of rejecting the null hypothesis when the null hypothesis is true.

β = probability of a Type II error = **$P(\text{Type II error})$** = probability of not rejecting the null hypothesis when the null hypothesis is false.

Outcomes and the Type I and Type II Errors

α and β should be as small as possible because they are probabilities of errors. They are rarely zero.

The Power of the Test is $1 - \beta$. Ideally, we want a high power that is as close to one as possible. Increasing the sample size can increase the Power of the Test.

The following are examples of Type I and Type II errors.

Suppose the null hypothesis, H_0 , is: Frank's rock climbing equipment is safe.

Type I error: Frank thinks that his rock climbing equipment may not be safe when, in fact, it really is safe. **Type II error:** Frank thinks that his rock climbing equipment may be safe when, in fact, it is not safe.

α = **probability** that Frank thinks his rock climbing equipment may not be safe when, in fact, it really is safe. β = **probability** that Frank thinks his rock climbing equipment may be safe when, in fact, it is not safe.

Notice that, in this case, the error with the greater consequence is the Type II error. (If Frank thinks his rock climbing equipment is safe, he will go ahead and use it.)

Try It

Suppose the null hypothesis, H_0 , is: the blood cultures contain no traces of pathogen X . State the Type I and Type II errors.

Type I error: The researcher thinks the blood cultures do contain traces of pathogen X , when in fact, they do not.

Type II error: The researcher thinks the blood cultures do not contain traces of pathogen X , when in fact, they do.

Suppose the null hypothesis, H_0 , is: The victim of an automobile accident is alive when he arrives at the emergency room of a hospital.

Type I error: The emergency crew thinks that the victim is dead when, in fact, the victim is alive. **Type II error:** The emergency crew does not know if the victim is alive when, in fact, the victim is dead.

α = **probability** that the emergency crew thinks the victim is dead when, in fact, he is really alive = $P(\text{Type I error})$. β = **probability** that the emergency crew does not know if the victim is alive when, in fact, the victim is dead = $P(\text{Type II error})$.

Outcomes and the Type I and Type II Errors

The error with the greater consequence is the Type I error. (If the emergency crew thinks the victim is dead, they will not treat him.)

Try It

Suppose the null hypothesis, H_0 , is: a patient is not sick. Which type of error has the greater consequence, Type I or Type II?

The error with the greater consequence is the Type II error: the patient will be thought well when, in fact, he is sick, so he will not get treatment.

It's a Boy Genetic Labs claim to be able to increase the likelihood that a pregnancy will result in a boy being born. Statisticians want to test the claim. Suppose that the null hypothesis, H_0 , is: It's a Boy Genetic Labs has no effect on gender outcome.

Type I error: This results when a true null hypothesis is rejected. In the context of this scenario, we would state that we believe that It's a Boy Genetic Labs influences the gender outcome, when in fact it has no effect. The probability of this error occurring is denoted by the Greek letter alpha, α .

Type II error: This results when we fail to reject a false null hypothesis. In context, we would state that It's a Boy Genetic Labs does not influence the gender outcome of a pregnancy when, in fact, it does. The probability of this error occurring is denoted by the Greek letter beta, β .

The error of greater consequence would be the Type I error since couples would use the It's a Boy Genetic Labs product in hopes of increasing the chances of having a boy.

Try It

“Red tide” is a bloom of poison-producing algae—a few different species of a class of plankton called dinoflagellates. When the weather and water conditions cause these blooms, shellfish such as clams living in the area develop dangerous levels of a paralysis-inducing toxin. In Massachusetts, the Division of Marine Fisheries (DMF) monitors levels of the toxin in shellfish by regular sampling of shellfish along the coastline. If the mean level of toxin in clams exceeds $800 \mu\text{g}$ (micrograms) of toxin per kg of clam meat in any area, clam harvesting is banned there until the bloom is over and levels of toxin in clams subside. Describe both a Type I and a Type II error in this context, and state which error has the greater consequence.

In this scenario, an appropriate null hypothesis would be H_0 : the mean level of toxins is at most $800 \mu\text{g}$, $H_0 : \mu_0 \leq 800 \mu\text{g}$.

Outcomes and the Type I and Type II Errors

Type I error: The DMF believes that toxin levels are still too high when, in fact, toxin levels are at most $800 \mu\text{g}$. The DMF continues the harvesting ban.

Type II error: The DMF believes that toxin levels are within acceptable levels (are at least $800 \mu\text{g}$) when, in fact, toxin levels are still too high (more than $800 \mu\text{g}$). The DMF lifts the harvesting ban. This error could be the most serious. If the ban is lifted and clams are still toxic, consumers could possibly eat tainted food.

In summary, the more dangerous error would be to commit a Type II error, because this error involves the availability of tainted clams for consumption.

A certain experimental drug claims a cure rate of at least 75% for males with prostate cancer. Describe both the Type I and Type II errors in context. Which error is the more serious?

Type I: A cancer patient believes the cure rate for the drug is less than 75% when it actually is at least 75%.

Type II: A cancer patient believes the experimental drug has at least a 75% cure rate when it has a cure rate that is less than 75%.

In this scenario, the Type II error contains the more severe consequence. If a patient believes the drug works at least 75% of the time, this most likely will influence the patient's (and doctor's) choice about whether to use the drug as a treatment option.

Try It

Determine both Type I and Type II errors for the following scenario:

Assume a null hypothesis, H_0 , that states the percentage of adults with jobs is at least 88%.

Identify the Type I and Type II errors from these four statements.

1. Not to reject the null hypothesis that the percentage of adults who have jobs is at least 88% when that percentage is actually less than 88%
2. Not to reject the null hypothesis that the percentage of adults who have jobs is at least 88% when the percentage is actually at least 88%.
3. Reject the null hypothesis that the percentage of adults who have jobs is at least 88% when the percentage is actually at least 88%.
4. Reject the null hypothesis that the percentage of adults who have jobs is at least 88% when that percentage is actually less than 88%.

Type I error: c

Type I error: b

Chapter Review

In every hypothesis test, the outcomes are dependent on a correct interpretation of the data. Incorrect calculations or misunderstood summary statistics can yield errors that affect the results. A **Type I** error occurs when a true null hypothesis is rejected. A **Type II** error occurs when a false null hypothesis is not rejected.

The probabilities of these errors are denoted by the Greek letters α and β , for a Type I and a Type II error respectively. The power of the test, $1 - \beta$, quantifies the likelihood that a test will yield the correct result of a true alternative hypothesis being accepted. A high power is desirable.

Formula Review

α = probability of a Type I error = $P(\text{Type I error})$ = probability of rejecting the null hypothesis when the null hypothesis is true.

β = probability of a Type II error = $P(\text{Type II error})$ = probability of not rejecting the null hypothesis when the null hypothesis is false.

The mean price of mid-sized cars in a region is \$32,000. A test is conducted to see if the claim is true. State the Type I and Type II errors in complete sentences.

Type I: The mean price of mid-sized cars is \$32,000, but we conclude that it is not \$32,000.

Type II: The mean price of mid-sized cars is not \$32,000, but we conclude that it is \$32,000.

A sleeping bag is tested to withstand temperatures of -15°F . You think the bag cannot stand temperatures that low. State the Type I and Type II errors in complete sentences.

For [Exercise 9.12](#), what are α and β in words?

α = the probability that you think the bag cannot withstand -15 degrees F, when in fact it can

β = the probability that you think the bag can withstand -15 degrees F, when in fact it cannot

In words, describe $1 - \beta$ For [Exercise 9.12](#).

Outcomes and the Type I and Type II Errors

A group of doctors is deciding whether or not to perform an operation. Suppose the null hypothesis, H_0 , is: the surgical procedure will go well. State the Type I and Type II errors in complete sentences.

Type I: The procedure will go well, but the doctors think it will not.

Type II: The procedure will not go well, but the doctors think it will.

A group of doctors is deciding whether or not to perform an operation. Suppose the null hypothesis, H_0 , is: the surgical procedure will go well. Which is the error with the greater consequence?

The power of a test is 0.981. What is the probability of a Type II error?

0.019

A group of divers is exploring an old sunken ship. Suppose the null hypothesis, H_0 , is: the sunken ship does not contain buried treasure. State the Type I and Type II errors in complete sentences.

A microbiologist is testing a water sample for E-coli. Suppose the null hypothesis, H_0 , is: the sample does not contain E-coli. The probability that the sample does not contain E-coli, but the microbiologist thinks it does is 0.012. The probability that the sample does contain E-coli, but the microbiologist thinks it does not is 0.002. What is the power of this test?

0.998

A microbiologist is testing a water sample for E-coli. Suppose the null hypothesis, H_0 , is: the sample contains E-coli. Which is the error with the greater consequence?

Homework

State the Type I and Type II errors in complete sentences given the following statements.

1. The mean number of years Americans work before retiring is 34.
2. At most 60% of Americans vote in presidential elections.
3. The mean starting salary for San Jose State University graduates is at least \$100,000 per year.
4. Twenty-nine percent of high school seniors get drunk each month.
5. Fewer than 5% of adults ride the bus to work in Los Angeles.

6. The mean number of cars a person owns in his or her lifetime is not more than ten.
7. About half of Americans prefer to live away from cities, given the choice.
8. Europeans have a mean paid vacation each year of six weeks.
9. The chance of developing breast cancer is under 11% for women.
10. Private universities mean tuition cost is more than \$20,000 per year.
 1. Type I error: We conclude that the mean is not 34 years, when it really is 34 years. Type II error: We conclude that the mean is 34 years, when in fact it really is not 34 years.
 2. Type I error: We conclude that more than 60% of Americans vote in presidential elections, when the actual percentage is at most 60%. Type II error: We conclude that at most 60% of Americans vote in presidential elections when, in fact, more than 60% do.
 3. Type I error: We conclude that the mean starting salary is less than \$100,000, when it really is at least \$100,000. Type II error: We conclude that the mean starting salary is at least \$100,000 when, in fact, it is less than \$100,000.
 4. Type I error: We conclude that the proportion of high school seniors who get drunk each month is not 29%, when it really is 29%. Type II error: We conclude that the proportion of high school seniors who get drunk each month is 29% when, in fact, it is not 29%.
 5. Type I error: We conclude that fewer than 5% of adults ride the bus to work in Los Angeles, when the percentage that do is really 5% or more. Type II error: We conclude that 5% or more adults ride the bus to work in Los Angeles when, in fact, fewer than 5% do.
 6. Type I error: We conclude that the mean number of cars a person owns in his or her lifetime is more than 10, when in reality it is not more than 10. Type II error: We conclude that the mean number of cars a person owns in his or her lifetime is not more than 10 when, in fact, it is more than 10.
 7. Type I error: We conclude that the proportion of Americans who prefer to live away from cities is not about half, though the actual proportion is about half. Type II error: We conclude that the proportion of Americans who prefer to live away from cities is half when, in fact, it is not half.
 8. Type I error: We conclude that the duration of paid vacations each year for Europeans is not six weeks, when in fact it is six weeks. Type II error: We conclude that the duration of paid vacations each year for Europeans is six weeks when, in fact, it is not.
 9. Type I error: We conclude that the proportion is less than 11%, when it is really at least 11%. Type II error: We conclude that the proportion of women who develop breast cancer is at least 11%, when in fact it is less than 11%.
 10. Type I error: We conclude that the average tuition cost at private universities is more than \$20,000, though in reality it is at most \$20,000. Type II error: We conclude that the average tuition cost at private universities is at most \$20,000 when, in fact, it is more than \$20,000.

Outcomes and the Type I and Type II Errors

For statements a-j in [Exercise 9.109](#), answer the following in complete sentences.

1. State a consequence of committing a Type I error.
2. State a consequence of committing a Type II error.

When a new drug is created, the pharmaceutical company must subject it to testing before receiving the necessary permission from the Food and Drug Administration (FDA) to market the drug. Suppose the null hypothesis is “the drug is unsafe.” What is the Type II Error?

1. To conclude the drug is safe when in, fact, it is unsafe.
2. Not to conclude the drug is safe when, in fact, it is safe.
3. To conclude the drug is safe when, in fact, it is safe.
4. Not to conclude the drug is unsafe when, in fact, it is unsafe.

b

A statistics instructor believes that fewer than 20% of Evergreen Valley College (EVC) students attended the opening midnight showing of the latest Harry Potter movie. She surveys 84 of her students and finds that 11 of them attended the midnight showing. The Type I error is to conclude that the percent of EVC students who attended is _____.

1. at least 20%, when in fact, it is less than 20%.
2. 20%, when in fact, it is 20%.
3. less than 20%, when in fact, it is at least 20%.
4. less than 20%, when in fact, it is less than 20%.

It is believed that Lake Tahoe Community College (LTCC) Intermediate Algebra students get less than seven hours of sleep per night, on average. A survey of 22 LTCC Intermediate Algebra students generated a mean of 7.24 hours with a standard deviation of 1.93 hours. At a level of significance of 5%, do LTCC Intermediate Algebra students get less than seven hours of sleep per night, on average?

The Type II error is not to reject that the mean number of hours of sleep LTCC students get per night is at least seven when, in fact, the mean number of hours

1. is more than seven hours.
2. is at most seven hours.
3. is at least seven hours.
4. is less than seven hours.

d

Outcomes and the Type I and Type II Errors

Previously, an organization reported that teenagers spent 4.5 hours per week, on average, on the phone. The organization thinks that, currently, the mean is higher. Fifteen randomly chosen teenagers were asked how many hours per week they spend on the phone. The sample mean was 4.75 hours with a sample standard deviation of 2.0. Conduct a hypothesis test, the Type I error is:

1. to conclude that the current mean hours per week is higher than 4.5, when in fact, it is higher
2. to conclude that the current mean hours per week is higher than 4.5, when in fact, it is the same
3. to conclude that the mean hours per week currently is 4.5, when in fact, it is higher
4. to conclude that the mean hours per week currently is no higher than 4.5, when in fact, it is not higher