4.1 Hypothesis test for the difference between two population means

Comparing two populations

Hypothesis testing involving a one-sample test determines whether an observed value differs statistically from a hypothesized population value. Sometimes, differences between two populations are studied instead. Ex: Pollsters often look at issues in which the opinions of two different groups may vary wildly, such as political preferences of men compared to those of women.

Hypothesis tests involving two samples follow the same steps. A survey is conducted and statistics are calculated. The standard error for the difference between populations is the square root of the sum of the squares of the standard errors of each population. The test statistic is the difference between the observed and hypothesized value divided by the standard error. Mathematically,

$$SE \text{ for the difference} = \sqrt{SE_1^2 + SE_2^2} \text{test statistic} = \frac{\text{observed difference - hypothesized difference}}{SE \text{ for the difference}}$$

Two-sample z-test for population means

The z-test can also be used to determine whether the means of two independent populations are the same when the population standard deviations are known. When performing a hypothesis test involving the means of two independent populations, the distribution of the z-test statistic is assumed to be $N\left(0, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right)$. In practice, the standard deviation for populations are generally unknown, so either the paired or unpaired t-test is needed.

The conditions that must be satisfied are similar to those of a z-test for a population mean.

Procedure 4.1.1: Hypothesis testing for two population means.

Given two randomly selected samples each taken from an independent population where the standard deviations of each population is known,

1. Set the null and alternative hypotheses

$$H_0 : \mu_1 = \mu_2 H_a : \mu_1 < \mu_2 (\text{left-tailed}) H_a : \mu_1 > \mu_2 (\text{right-tailed}) H_a : \mu_1 \neq \mu_2 (\text{two-tailed})$$

   where μ_1 and μ_2 are means from distinct populations.

2. Use statistical software to find the test-statistic

$$z = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

where $SE = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$.

3. Use statistical software to find the p-value that corresponds to z.
4. Make a decision given a previously selected significance level α, typically 0.05.
    - If the p-value is less than the significance level, sufficient evidence exists to reject the null hypothesis H_0 in favor of the alternative hypothesis H_a.
    - If the p-value is greater than or equal to the significance level, insufficient evidence exists to reject the null hypothesis H_0.

Example 4.1.1: Candle burn times.

The mean burn time of two brands of 11-ounce candles are compared by a home safety magazine. The burn times of 100 candles of each brand are measured. The results are given in the table below.

Candle	Sample mean burn time (hours)	Population standard deviation (hours)
1	27.5	2.5
2	26	3.5

Does sufficient evidence exist supporting the claim that the mean burn time of candle 1 is greater than the mean burn time of candle 2 at the $\alpha = 0.05$ significance level?

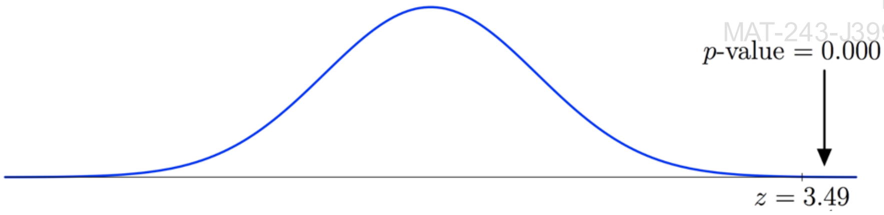
Solution

The null hypothesis is that the two mean burn times for both brands of candles are the same. Since the question asks if the mean burn time of candle 1 is greater than the mean burn time of candle 2, the hypothesis test is right-tailed. Mathematically,

$$H_0: \mu_1 = \mu_2 H_a: \mu_1 > \mu_2$$

The z-statistic is

$$z = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{27.5 - 26 - 0}{\sqrt{\frac{2.5^2}{100} + \frac{3.5^2}{100}}} \approx 3.487$$



The p-value is close to 0. Since the p-value is less than the significance level, sufficient evidence exists supporting the claim that the mean burn time of candle 1 is greater than the mean burn time of candle 2.

Example 4.1.2: Extrasensory perception.

A study is conducted to determine whether extrasensory perception (ESP) is a real phenomenon. 50 subjects claiming to have ESP answer a set of questions testing the subject's abilities. 60 subjects who do not claim to have ESP also answer the same set of questions. The results of the test are summarized in the table below.

Subject	Mean score	Population standard deviation
1 (ESP)	12	4.5
2 (non-ESP)	10	4.0

Does sufficient evidence exist supporting the claim that mean test score of subjects claiming to have ESP is different from the mean test score of subjects who don't claim to have ESP at the $\alpha = 0.05$ significance level?

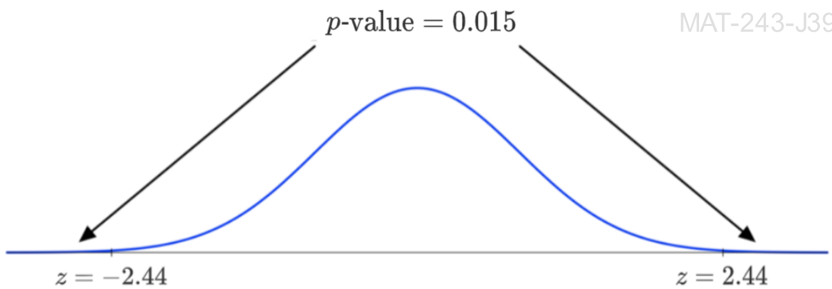
Solution

The null hypothesis is that the two mean scores for the subjects claiming to have ESP and those who do not are the same. Since the question asks if the scores are different, the hypothesis test is two-tailed. Mathematically,

$$H_0 : \mu_1 = \mu_2 H_a : \mu_1 \neq \mu_2$$

The z-statistic is

$$z = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{12 - 10 - 0}{\sqrt{\frac{4.5^2}{50} + \frac{4.0^2}{60}}} \approx 2.440$$



The p-value is $P(z \leq -2.440 \text{ or } z \geq 2.440) = 0.015$. Since the p-value is less than the significance level $(0.015 < 0.05)$, sufficient evidence exists that the mean scores of subjects claiming to have ESP and those who do not are different.

PARTICIPATION ACTIVITY 4.1.1: Commute times.

A transportation commission studies driving times between two cities to determine whether the construction of a new highway reduced commute times. Times for 40 cars driving on the old highway and times for 50 cars driving on the new highway are obtained. A summary of the data obtained from the study is given below.

Highway	Mean commute times	Population standard deviation
1 (Old)	5.35	0.5
2 (New)	4.95	0.8

1) What is the standard error? Type as: #.###

Check Show answer

2) What is the z-score? Type as: #.###

Check Show answer

3) What is the p-value? Type as: #.####

Check Show answer

4) Should the null hypothesis that the mean commute time for the old highway is equal to the mean commute time for the new highway be rejected at the $\alpha = 0.05$ significance level using the output below? Type as: yes or no

Check Show answer

Python-Function 4.1.1: ztest(x1, x2).

The `ztest(x1, x2, value = 0)` function can also be used to perform a two-sample z-test for means. However, the value parameter should be set to 0. The function requires the statsmodels.stats.weightstats library to be imported and takes two inputs. The first input x1 is an array containing sample observations from one population and the second input is also an array containing sample observations from another population.

The function returns the z-score and the two-tailed p-value.

```
from statsmodels.stats.weightstats import ztest
sample1 = [21, 28, 40, 55, 58, 60]
sample2 = [13, 29, 50, 55, 71, 90]
print(ztest(x1 = sample1, x2 = sample2))
```

```
(-0.58017208108908169, 0.56179857900464247)
```

Run example

Two-sample t-test

The t-test discussed analyzes the difference between the sample mean and the hypothesized value of the population mean. A similar method exists to compare the means of two different populations. The ***two-sample t-test*** is used to determine if a statistically significant difference exists between two population means. Two types of two-sample t-tests exist: paired and unpaired.

In a ***paired t-test*** or ***dependent t-test***, a sample taken from one population is exposed to two different treatments. The main idea is that measurements are recorded from the same group, usually before and after a treatment is applied or when each of two treatments is applied. Ex: A group of professional cycling athletes is selected for a study on the effects of caffeine dosage on exhaustion times. The populations are the cyclists for each of two dosages. The samples are the measured exhaustion times for each dosage, which implies dependence because the measurements were taken from the same group.

In an ***unpaired t-test*** or ***independent t-test***, a sample taken from one population is not related to a different sample taken from another population. In contrast to the paired t-test, measurements from an unpaired t-test are recorded from different groups when exposed to the same treatment. Ex: The effect of caffeine intake on exhaustion times is studied by measuring the exhaustion times of a randomly selected group of 9 professional cyclists taking caffeine pills and another group of 9 cyclists not taking caffeine pills. The two populations are all cyclists taking caffeine pills and those who are not taking the pills. The samples are the measured exhaustion times from the two groups, each with 9 cyclists, which implies independence because the times are for two different groups of cyclists.

PARTICIPATION ACTIVITY 4.1.2: Identifying the type of two-sample t-test.

1) A study involving children from a preschool compares the median times to recite words with two and three syllables.
   - paired
   - unpaired

2) A study on the impact of meal preparation programs compares caloric intake between treatment and control groups.
   - paired
   - unpaired

3) A study on the difficulty of a maze game involves comparing the error rates between adults and children.
   - paired
   - unpaired

4) A study involving track and field athletes compares resting heart rate to heart rate after running a race.
   - paired
   - unpaired

Paired t-test

To obtain probabilities for a paired t-test, the paired t-statistic is needed. The formula involves finding the mean and standard deviation of the differences between corresponding measurements :

$$t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}}$$

where s_d is the sample standard deviation of the differences, $\bar{d}$ is the mean difference between the samples, and n is the sample size. The most common scenario is that the hypothesized mean difference is 0. However, this scenario is not necessary. Ex: To continue the development of a new drug, a measurable improvement in the condition of the subjects must be seen. In this situation, the null hypothesis would be that the mean difference is the minimum amount of improvement set by the manufacturer in order to continue developing the drug. The differences are assumed to come from a normal distribution. Thus, the differences can be seen as a single sample following a t-distribution, which means that a paired t-test is equivalent to a one-sample t-test.

Procedure 4.1.2: Paired t-test.

Given two randomly selected samples of size n taken from each of two populations with unknown population standard deviation σ,

1. Set the null and alternative hypotheses:

$$\begin{aligned} &H_0 : \mu_d = 0 \\ &H_a : \mu_d > 0 \quad \text{(right-tailed)} \\ &H_a : \mu_d < 0 \quad \text{(left-tailed)} \\ &H_a : \mu_d \neq 0 \quad \text{(two-tailed)} \end{aligned}$$

   where μ_d is the mean difference between the populations.

2. Use statistical software to find the t-statistic and the degrees of freedom df:

$$t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}}$$

   and

$$df = n - 1$$

3. Use statistical software to find the p-value that corresponds to t.
4. Make a decision given a previously selected significance level α, typically 0.05:

- If the p-value is less than the significance level, sufficient evidence exists to reject the null hypothesis, H_0, in favor of the alternative hypothesis, H_a.
- If the p-value is greater than or equal to the significance level, insufficient evidence exists to reject the null hypothesis, H_0.

Python-Function 4.1.2: Paired t-test.

The `st.ttest_rel(x, y)` function takes two arrays or DataFrame columns x and y with the same length as inputs and returns the t-statistic and the corresponding two-tailed p-value as outputs.

```
import scipy.stats as st
import pandas as pd
df = pd.read_csv('http://data-analytics.zybooks.com/ExamScores.csv')
st.ttest_rel(df['Exam1'],df['Exam2'])
```

```
Ttest_relResult(statistic=1.4179252582484649, pvalue=0.16254101610053867)
```

Run example

Example 4.1.3: Improvement in exam scores.

In the ExamScores dataset, four exam scores of the same 50 students are recorded. The teacher believes that student exam scores improved between Exam1 and Exam2. Does statistically significant evidence exist to support the teacher's belief at the $\alpha = 0.05$ significance level? Use the output below to answer the question.

```
Ttest_relResult(statistic=1.4179252582484649, pvalue=0.16254101610053867)
```

Solution

Since both exams are taken by the same group of students, the hypothesis test is a paired t-test.

The null hypothesis is that the mean Exam1 scores and the mean Exam2 scores are the same. The alternative hypothesis is that the mean Exam1 scores is less than the mean Exam2 scores. Mathematically,

$$H_0: \mu_1 = \mu_2 H_a: \mu_1 < \mu_2$$

The p-value in the output is a two-tailed p-value. The one-tailed p-value that corresponds to the t-statistic of 1.418 is $\frac{0.163}{2} = 0.082$.

Since the p-value is greater than the significance level $(0.082 > 0.05)$, insufficient statistical evidence exists to support the teacher's claim that students showed improvement between Exam1 and Exam2.

PARTICIPATION ACTIVITY | 4.1.3: Caffeine and athletic performance.

The Caffeine dataset uses data from "The Effect of Different Dosages of Caffeine on Endurance Performance Time", *International Journal of Sports Medicine*[1] and gives the endurance times (in minutes) of 9 athletes when given a caffeine dose of 5 milligrams and 13 milligrams.

```
import scipy.stats as st
import pandas as pd
df = pd.read_csv('http://data-analytics.zybooks.com/Caffeine.csv')
st.ttest_rel(df['mg5'],df['mg13'])
```

```
Ttest_relResult(statistic=-0.12052261484527026, pvalue=0.90704115640761218)
```

Run example

1) What is the null hypothesis H_0?

- $\bar{d} = -0.4722$
- $\mu_d = 0$
- $t = -0.1205$

2) What is the alternative hypothesis H_a?

- $\mu_d \neq 0$
- $\mu_d < 0$
- $\mu_d > 0$

3) What is the p-value?

- 0.907
- 0.454

4) What is the conclusion for the t-test?

- Reject H_0
- Fail to reject H_0

Unpaired t-test

The t-test statistic for unpaired data is different from that of paired data. The formula involves subtracting the means of the two samples:

$$\frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where $\bar{x}_1$, s_1, and n_1 are the mean, standard deviation, and sample size of the sample drawn from the first population respectively; and $\bar{x}_2$, s_2, and n_2 are the mean, standard deviation, and sample size of the sample drawn from the second population. Since only sample means are subtracted and not individual observations, sample sizes do not need to be equal. The degrees of freedom are $df = n_1 + n_2 - 2$. Although $\mu_1 - \mu_2$ can be any number based on the hypothesized means for the two populations, most of the time, the accepted difference between the means of the populations is 0. Finally, the formula for the t-statistic above assumes that the variances are unequal. In most practical instances, the equality of variances should be verified using the Fisher's F-test before performing the unpaired t-test. However, this is beyond the scope of the material.

Procedure 4.1.3: Unpaired t-test.

Given two randomly selected samples taken from each of two independent populations with unknown population standard deviation σ,

1. Set the null and alternative hypotheses:

$$\begin{aligned} &H_0: \mu_1 = \mu_2 \\ &H_a: \mu_1 > \mu_2 \quad \text{(right-tailed)} \\ &H_a: \mu_1 < \mu_2 \quad \text{(left-tailed)} \\ &H_a: \mu_1 \neq \mu_2 \quad \text{(two-tailed)} \end{aligned}$$

   where $\mu_1$and μ_2 are the means of the populations.
2. Use statistical software to find the t-statistic when population variances are not equal and degrees of freedom df:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

   and

$$df = n_1 + n_2 - 2$$

3. Use statistical software to find the p-value that corresponds to t.
4. Make a decision given a previously selected significance level α, typically 0.05:
   - If the p-value is less than the significance level, sufficient evidence exists to reject the null hypothesis, H_0, in favor of the alternative hypothesis, H_a.
   - If the p-value is greater than or equal to the significance level, insufficient evidence exists to reject the null hypothesis, H_0.

Python-Function 4.1.3: Unpaired t-test.

The **`st.ttest_ind(x, y)`** command takes two arrays or DataFrame columns x and y with the same length as inputs and returns the t-statistic and the corresponding two-tailed p-value as outputs.

```
import scipy.stats as st
import pandas as pd
df = pd.read_csv('http://data-analytics.zybooks.com/Machine.csv')
st.ttest_ind(df['Old'],df['New'])
```

```
Ttest_indResult(statistic=3.397230706117603, pvalue=0.0032422494663179747)
```

Run example

Example 4.1.4: Packing machines.

To improve production capacity, a manufacturing company buys a new machine that claims to pack 50 widgets in a carton faster than the old machine can. The Machines dataset records the amount of time (in seconds) each machine can complete the packing task for 10 batches of 50 widgets. Does sufficient evidence exist at the $\alpha = 0.05$ significance level to support the claim that the new machines can pack widgets faster? Use the output below to answer the question.

```
Ttest_indResult(statistic=3.397230706117603, pvalue=0.0032422494663179747)
```


Solution

Since the times are taken from two different machines, the hypothesis test is an unpaired t-test.

The null hypothesis is that the mean packing time for the old machine μ_1 and the mean packing time for the new machine μ_2 is the same. The alternative hypothesis is that the mean packing time for the old machine is greater than the mean packing time for the new machine, because a faster machine implies a lower mean time. Mathematically,

$$H_0 : \mu_1 = \mu_2 H_a : \mu_1 > \mu_2$$

Note that the p-value in the output is a two-tailed p-value. The one-tailed p-value that corresponds to the t-statistic of 3.397 is $\frac{0.0032}{2} = 0.0016$.

Since the p-value is less than the significance level $(0.0016 < 0.05)$, sufficient statistical evidence exists to reject the null hypothesis that mean packing time for both machines is the same.

PARTICIPATION ACTIVITY 4.1.4: Memory.

The Memory dataset contains the number of errors made while completing a memory-related task by a group of 10 people taking a fictional memory enhancement drug and by another group of 10 people not taking the drug. A researcher in the study claims that the fictional drug is effective in reducing the number of errors in memory-related tasks.

Let μ_1 be the mean number of errors in the group taking the drug and μ_2 be the mean number of errors in the group not taking the drug.

```
import scipy.stats as st
import pandas as pd
df = pd.read_csv('http://data-analytics.zybooks.com/Memory.csv')
st.ttest_ind(df['nodrug'], df['drug'], equal_var=False)
```

```
Ttest_indResult(statistic=2.7992880505646385, pvalue=0.018643414767040494)
```

Run example

1) What is the null hypothesis H_0?

- $df = 18$
- $\bar{x}_1 \neq \bar{x}_2$
- $\mu_1 = \mu_2$

2) What is the alternative hypothesis H_a?

- $\mu_1 > \mu_2$
- $\mu_1 < \mu_2$
- $\mu_1 \neq \mu_2$

3) What is the p-value?

- 0.009

0.018

4) What is the conclusion for the t-test?

Reject H_0

Fail to reject H_0

CHALLENGE ACTIVITY 4.1.1: Hypothesis test for the difference between two population means.

456500.2535406.qx3zqy7

Start

An electrician wants to know whether batteries made by two manufacturers have significantly different voltages. The voltage of \(89\) batteries from each manufacturer were measured. The population standard deviations of the voltage for each manufacturer are known. The results are summarized in the following table.

Manufacturer	Sample mean voltage (millivolts)	Population standard devia
A	\(181\)	\(3\)
B	\(180\)	\(1\)

What type of hypothesis test should be performed?

What is the test statistic? Ex: 0.12

Does sufficient evidence exist to support the claim that the voltage of the batteries made by the two manufacturers is different at the \(\alpha = 0.05\) significance level?

1 2 3

Check Next

References

(*1) Pasman, WJ, et al. "The effect of different dosages of caffeine on endurance performance time." *International Journal of Sports Medicine*, 16(4):225-30, May 1995, DOI: 10.1055/s-2007-972996

4.2 Hypothesis test for the difference between two population proportions

The z-test can also be used to determine whether the proportions of two distinct populations are the same. When performing a hypothesis test involving the proportion of two distinct populations, the distribution of the z-test statistic is assumed to be

$$N\left(0, \sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}\right)$$

The conditions that must be satisfied are similar to those of the z-test for proportions involving two distinct populations. However, to satisfy the normality condition, all counts should be at least 5:

- $n_1\hat{p}_1 \geq 5$
- $n_1(1-\hat{p}_1) \geq 5$
- $n_2\hat{p}_2 \geq 5$
- $n_2(1-\hat{p}_2) \geq 5$

where $\hat{p}_1$ is the probability of success in the first sample, $\hat{p}_2$ is the probability of success in the second sample, $\hat{p}$ is the overall probability of success when two samples are combined, n_1 is the size of the first sample, and n_2 is the size of the second sample.

Procedure 4.2.1: Hypothesis testing for two population proportions.

Given two randomly selected samples each taken from a distinct population

1. Set the null and alternative hypotheses

$$H_0 : p_1 = p_2 H_a : p_1 < p_2 \text{(left-tailed)} H_a : p_1 > p_2 \text{(right-tailed)} H_a : p_1 \neq p_2 \text{(two-tailed)}$$

   where p_1 and p_2 are proportions from distinct populations.

2. Use statistical software to find the z-statistic

$$z = \frac{\hat{p}_1 - \hat{p}_2 - 0}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

3. Use statistical software to find the p-value that corresponds to z.
4. Make a decision given a previously selected significance level α, typically 0.05.

- If the p-value is less than the significance level, sufficient evidence exists to reject the null hypothesis H_0 in favor of the alternative hypothesis H_a.
- If the p-value is greater than or equal to the significance level, insufficient evidence exists to reject the null hypothesis H_0.

Example 4.2.1: Gender and voting.

The campaign team of a candidate running for statewide office is concerned about the candidate's appeal to both genders. A recent survey of voting aged adults found that if the elections were held that day, 70 of 132 men and 63 of 105 women would vote for the candidate. Does sufficient evidence exist that suggest that the percentages of men and women who would vote for the candidate are different at the $\alpha = 0.05$ significance level?

Solution

The null hypothesis is that the proportion of men and the proportion of women who would vote for the candidate are the same. Since the question asks if the the percentages of men and women voting for the candidate are different, the hypothesis test is two-tailed with an alternative hypothesis that the population proportions are different. Mathematically,

$$H_0 : p_1 = p_2 H_a : p_1 \neq p_2$$

where p_1 is the population proportion of men who would vote for the candidate and p_2 is the population proportion of women who would vote for the candidate.

The survey found that 70 of 132 men and 63 of 105 women would vote for the candidate. Thus,

$$\hat{p}_1 = \frac{\text{men who would vote for the candidate}}{\text{number of men in the survey}} = \frac{70}{132} \approx 0.530$$
$$\hat{p}_2 = \frac{\text{women who would vote for the candidate}}{\text{number of women in the survey}} = \frac{63}{105} = 0.600$$
$$\hat{p} = \frac{\text{people who would vote for the candidate}}{\text{number of people in the survey}} = \frac{70+63}{132+105} \approx 0.561$$

The test statistic is

$$z = \frac{\hat{p}_1 - \hat{p}_2 - 0}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0.530 - 0.600}{\sqrt{0.561(1-0.561)\left(\frac{1}{132} + \frac{1}{105}\right)}} \approx -1.079$$

The p-value is

$$p\text{-value} = P(z \leq -1.079 \text{ or } z \geq 1.079) \approx 0.280$$

Since the p-value is greater than the significance level $(0.280 > 0.05)$, insufficient evidence exists to suggest that population proportions for men and women who would vote for the candidate are different.

Example 4.2.2: Adverse reaction to drugs.

A medical researcher tests two types of drugs that are designed to slow the progression of a disease. adverse reaction. In a random sample of patients taking drug 2, 21 out of 30 people developed an adv showing that the proportion of patients taking drug 1 who develop an adverse reaction is greater thar

Solution

The null hypothesis is that the proportion of patients developing an adverse reaction is the same for t asks if the proportion of patients developing an adverse reaction to drug 1 is greater than that of patie Mathematically,

$$H_0 : p_1 = p_2 H_a : p_1 > p$$

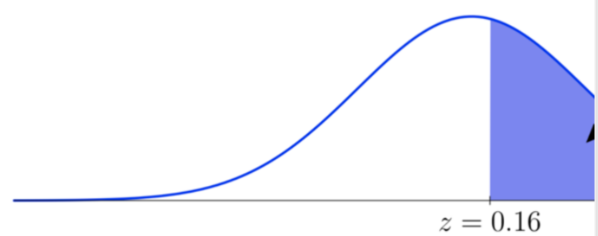
In the study, 18 of 25 people taking drug 1 and 21 of 30 people taking drug 2 developed an adverse rea

$$\hat{p}_1 = \frac{\text{people taking drug 1 who developed an adverse reaction}}{\text{number of people taking drug 1}} = \frac{18}{25} = 0.72 \hat{p}_2 = \frac{\text{people taking drug 2 who develop}}{\text{number of people tak}}$$

The test statistic is

$$z = \frac{\hat{p}_1 - \hat{p}_2 - 0}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0.72 - (}{\sqrt{0.709(1 - 0.}}$$

The p-value is $P(z \geq 0.16) \approx 0.435$.



Since the p-value is greater than the significance level $(0.435 > 0.01)$, insufficient evidence exist sugge is greater than that of patients taking drug 2.

PARTICIPATION ACTIVITY | 4.2.1: Effectiveness of a vaccine.

$10,000$ individuals are divided evenly into two groups. The treatment group is given a vaccine and the control group is given a placebo. 95 of the $5,000$ individuals in the treatment group developed a disease. 125 of the $5,000$ individuals in the control group developed a particular disease. A research team wants to determine whether the vaccine is effective in decreasing the incidence of disease. Does sufficient evidence exist to conclude that the proportion of developing a disease in individuals given the vaccine is less than that of individuals given a placebo?

1) What is the proportion of individuals in the treatment group that developed the disease? Type as: #.###

Check Show answer

2) What is the proportion of individuals in the control group that developed the disease? Type as: #.###

Check Show answer

3) What is the proportion of individuals in the overall group that developed the disease? Type as: #.###

Check Show answer

4) What is the standard error estimate? Type as: #.####

Check Show answer

5) What is the z-score? Type as: #.###

Check Show answer

6) What is the p-value? Type as: #.###

Check Show answer

7) Should the null hypothesis that the proportion of individuals taking the drug who develop the disease is the same as that of individuals not taking the drug be rejected at the $\alpha = 0.05$ significance level using the output below? Type as: yes or no

Check Show answer

Python-Function 4.2.1: proportions_ztest().

The `proportions_ztest()` function can also perform a z-test between two samples. The function requires the statsmodels.stats.proportion library to be imported, and takes two arrays, instead of two integers, as parameters. The first array is the number of individuals meeting some condition in each group, and the second array is the total number of individuals in each group.

The function returns the z-score and the two-tailed p-value.

```
from statsmodels.stats.proportion import proportions_ztest
counts = [95, 125]
n = [5000, 5000]
print(proportions_ztest(counts, n))
```

```
(-2.0452221470506315, 0.040832962004731133)
```

Run example

CHALLENGE ACTIVITY 4.2.1: Hypothesis test for the difference between two population proportions.

456500.2535406.qx3zqy7

Start

A political campaign is interested in whether a geographic difference existed in support for raising the minimum wage in a certain state. Polls were conducted in the two largest cities in the state about raising the minimum wage. In city 1, a poll of \(800\) randomly selected voters found that \(426\) supported raising the minimum wage. In city 2, a poll of \(1000\) randomly selected voters found that \(498\) supported raising the minimum wage.

What type of hypothesis test should be performed?

\(\hat{| Ex: 0.123 \(\hat{| Ex: 0.123

\(\hat Ex: 0.123

Test statistic \(= Ex: 0.12

\(p\)-value Ex: 0.123

Does sufficient evidence exist to support the claim that the level of support differs between the two cities at the \(\alpha = 0.01\) significance level?

1 2

Check Next