Testing a Claim About a Proportion

Colby Community College

Requirements

1. The sample observations are a simple random sample.

Requirements

1. The sample observations are a simple random sample.
2. The conditions of the binomial distribution are satisfied:
   - There is a fixed number of trials.
   - The trials are independent.
   - Each trial has two categories of "success" and "failure."
   - The probability of a success remains the same in all trials.

Requirements

1. The sample observations are a simple random sample.
2. The conditions of the binomial distribution are satisfied:
    - There is a fixed number of trials.
    - The trials are independent.
    - Each trial has two categories of "success" and "failure."
    - The probability of a success remains the same in all trials.
3. The conditions of $np \geq 5$ and $nq \geq 5$ are both satisfied.

Requirements

1. The sample observations are a simple random sample.
2. The conditions of the binomial distribution are satisfied:
   - There is a fixed number of trials.
   - The trials are independent.
   - Each trial has two categories of "success" and "failure."
   - The probability of a success remains the same in all trials.
3. The conditions of $np \geq 5$ and $nq \geq 5$ are both satisfied.

Note

When these requirements are met, we can approximate a binomial distribution with a normal distribution using $\mu = np$ and $\sigma = \sqrt{npq}$

Example 1

In a Pitney Bowes survey in which 1009 consumers were asked if they are comfortable with having drones deliver their purchases, and 54% of them responded with "no."

Example 1

In a Pitney Bowes survey in which 1009 consumers were asked if they are comfortable with having drones deliver their purchases, and 54% of them responded with "no."

Let us check that requirements are met for the claim that most (more than half) of consumers are uncomfortable with drone deliveries.

Example 1

In a Pitney Bowes survey in which 1009 consumers were asked if they are comfortable with having drones deliver their purchases, and 54% of them responded with "no."

Let us check that requirements are met for the claim that most (more than half) of consumers are uncomfortable with drone deliveries.

Requirements check:

1. The 1009 consumers were randomly selected.

Example 1

In a Pitney Bowes survey in which 1009 consumers were asked if they are comfortable with having drones deliver their purchases, and 54% of them responded with "no."

Let us check that requirements are met for the claim that most (more than half) of consumers are uncomfortable with drone deliveries.

Requirements check:

1. The 1009 consumers were randomly selected.
2. There are a fixed number of independent trials where the two categories are wether the subject is comfortable with drone deliveries or not.

Example 1

In a Pitney Bowes survey in which 1009 consumers were asked if they are comfortable with having drones deliver their purchases, and 54% of them responded with "no."

Let us check that requirements are met for the claim that most (more than half) of consumers are uncomfortable with drone deliveries.

Requirements check:

1. The 1009 consumers were randomly selected.
2. There are a fixed number of independent trials where the two categories are wether the subject is comfortable with drone deliveries or not.
3. We have $n = 1009$, $p = 0.5$, and $q = 0.5$ and so

$$np = 1009 \cdot 0.5 = 504.5 \geq 5$$
$$nq = 1009 \cdot 0.5 = 504.5 \geq 5$$

Test Statistic

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

Test Statistic

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

Equivalent Methods

- P-values is the most common method used.

Test Statistic

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

Equivalent Methods

- P-values is the most common method used.
- Critical values can be used when technology is unavailable.

Test Statistic

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

Equivalent Methods

- P-values is the most common method used.
- Critical values can be used when technology is unavailable.

Confidence Intervals

Confidence intervals use an estimated standard deviation based on the sample proportion, which means this method could result in a different conclusion.

Test Statistic

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

Equivalent Methods

- P-values is the most common method used.
- Critical values can be used when technology is unavailable.

Confidence Intervals

Confidence intervals use an estimated standard deviation based on the sample proportion, which means this method could result in a different conclusion.

Technology

P-values are usually provided automatically by technology. Otherwise use the standard normal distribution.

Example 2

In a Pitney Bowes survey in which 1009 consumers were asked if they are comfortable with having drones deliver their purchases, and 54% of them responded with "no."

Example 2

In a Pitney Bowes survey in which 1009 consumers were asked if they are comfortable with having drones deliver their purchases, and 54% of them responded with "no."

Let us test the claim that most (more than half) of consumers are uncomfortable with drone deliveries.

Example 2

In a Pitney Bowes survey in which 1009 consumers were asked if they are comfortable with having drones deliver their purchases, and 54% of them responded with "no."

Let us test the claim that most (more than half) of consumers are uncomfortable with drone deliveries.

The hypotheses for this example are:

$$H_0: \; p = 0.5$$
$$H_A: \; p > 0.5$$

Example 2

In a Pitney Bowes survey in which 1009 consumers were asked if they are comfortable with having drones deliver their purchases, and 54% of them responded with "no."

Let us test the claim that most (more than half) of consumers are uncomfortable with drone deliveries.

The hypotheses for this example are:

$$H_0: \ p = 0.5$$
$$H_A: \ p > 0.5$$

Using technology we get P-value of $0.005386 < 0.05$ which means we reject the null hypothesis.

Example 2

In a Pitney Bowes survey in which 1009 consumers were asked if they are comfortable with having drones deliver their purchases, and 54% of them responded with "no."

Let us test the claim that most (more than half) of consumers are uncomfortable with drone deliveries.

The hypotheses for this example are:

$$H_0: \; p = 0.5$$
$$H_A: \; p > 0.5$$

Using technology we get P-value of $0.005386 < 0.05$ which means we reject the null hypothesis.

Because we reject the null hypothesis we conclude that there is sufficient sample evidence to support the claim that more than half of consumers are uncomfortable with drone deliveries.

Caution

Be careful not to confuse the notation.

***P*-value**	The probability of a test statistic at least as extreme as the one obtained.
$\boldsymbol{p}$	The population proportion.
$\boldsymbol{\hat{p}}$	The sample proportion.

Note

When using technology, you often need to enter the sample size n and the number of successes x. But, in real applications the sample proportion $\hat{p}$ is often given instead of x.

Note

When using technology, you often need to enter the sample size n and the number of successes x. But, in real applications the sample proportion $\hat{p}$ is often given instead of x.

$$\hat{p} = \frac{x}{n}$$

Note

When using technology, you often need to enter the sample size n and the number of successes x. But, in real applications the sample proportion $\hat{p}$ is often given instead of x.

$$\hat{p} = \frac{x}{n} \quad \Rightarrow \quad x = n \cdot \hat{p}$$

Note

When using technology, you often need to enter the sample size n and the number of successes x. But, in real applications the sample proportion $\hat{p}$ is often given instead of x.

$$\hat{p} = \frac{x}{n} \quad \Rightarrow \quad x = n \cdot \hat{p}$$

Example 3

A study of sleepwalking of "nocturnal wandering" was described in *Neurology* magazine, and it included information that 29.2% of 19, 136 American adults have sleepwalked.

Note

When using technology, you often need to enter the sample size n and the number of successes x. But, in real applications the sample proportion $\hat{p}$ is often given instead of x.

$$\hat{p} = \frac{x}{n} \quad \Rightarrow \quad x = n \cdot \hat{p}$$

Example 3

A study of sleepwalking of "nocturnal wandering" was described in *Neurology* magazine, and it included information that 29.2% of 19,136 American adults have sleepwalked.

The number of adults who have sleepwalked is

$$0.292 \cdot 19136$$

Note

When using technology, you often need to enter the sample size n and the number of successes x. But, in real applications the sample proportion $\hat{p}$ is often given instead of x.

$$\hat{p} = \frac{x}{n} \quad \Rightarrow \quad x = n \cdot \hat{p}$$

Example 3

A study of sleepwalking of "nocturnal wandering" was described in *Neurology* magazine, and it included information that 29.2% of 19,136 American adults have sleepwalked.

The number of adults who have sleepwalked is

$$0.292 \cdot 19136 = 5587.712$$

Note

When using technology, you often need to enter the sample size n and the number of successes x. But, in real applications the sample proportion $\hat{p}$ is often given instead of x.

$$\hat{p} = \frac{x}{n} \quad \Rightarrow \quad x = n \cdot \hat{p}$$

Example 3

A study of sleepwalking of "nocturnal wandering" was described in *Neurology* magazine, and it included information that 29.2% of 19,136 American adults have sleepwalked.

The number of adults who have sleepwalked is

$$0.292 \cdot 19136 = 5587.712$$

But, the result must be a whole number, so we round to the nearest whole number of 5588.

Example 4

A study of sleepwalking of "nocturnal wandering" was described in *Neurology* magazine, and it included information that 29.2% of $19,136$ American adults have sleepwalked. Let us test the claim that "fewer than 30% of adults have sleepwalked" using $\alpha = 0.05$.

Example 4

A study of sleepwalking of "nocturnal wandering" was described in *Neurology* magazine, and it included information that 29.2% of $19,136$ American adults have sleepwalked. Let us test the claim that "fewer than 30% of adults have sleepwalked" using $\alpha = 0.05$.

1. The original claim is $p < 0.30$ and the opposite claim is $p \geq 0.30$.

Example 4

A study of sleepwalking of "nocturnal wandering" was described in *Neurology* magazine, and it included information that 29.2% of 19,136 American adults have sleepwalked. Let us test the claim that "fewer than 30% of adults have sleepwalked" using $\alpha = 0.05$.

1. The original claim is $p < 0.30$ and the opposite claim is $p \geq 0.30$.
2. Because $p < 0.30$ does not contain equality, we have the hypotheses:

$$H_0: \ p = 0.30 \quad \text{and} \quad H_A: \ p < 0.30$$

Example 4

A study of sleepwalking of "nocturnal wandering" was described in *Neurology* magazine, and it included information that 29.2% of 19,136 American adults have sleepwalked. Let us test the claim that "fewer than 30% of adults have sleepwalked" using $\alpha = 0.05$.

1. The original claim is $p < 0.30$ and the opposite claim is $p \geq 0.30$.
2. Because $p < 0.30$ does not contain equality, we have the hypotheses:

$$H_0: \ p = 0.30 \quad \text{and} \quad H_A: \ p < 0.30$$

3. The relevant statistic is $\hat{p}$ and the sampling distribution of sample proportions can be approximated by the normal distribution.

Example 4

A study of sleepwalking of "nocturnal wandering" was described in *Neurology* magazine, and it included information that 29.2% of 19, 136 American adults have sleepwalked. Let us test the claim that "fewer than 30% of adults have sleepwalked" using $\alpha = 0.05$.

1. The original claim is $p < 0.30$ and the opposite claim is $p \geq 0.30$.
2. Because $p < 0.30$ does not contain equality, we have the hypotheses:

$$H_0: \ p = 0.30 \quad \text{and} \quad H_A: \ p < 0.30$$

3. The relevant statistic is $\hat{p}$ and the sampling distribution of sample proportions can be approximated by the normal distribution.
4. Using technology we get the test statistic $z = -2.41039$ and P-value of 0.007968.

Example 4

A study of sleepwalking of "nocturnal wandering" was described in *Neurology* magazine, and it included information that 29.2% of 19,136 American adults have sleepwalked. Let us test the claim that "fewer than 30% of adults have sleepwalked" using $\alpha = 0.05$.

1. The original claim is $p < 0.30$ and the opposite claim is $p \geq 0.30$.
2. Because $p < 0.30$ does not contain equality, we have the hypotheses:
$$H_0: \ p = 0.30 \quad \text{and} \quad H_A: \ p < 0.30$$
3. The relevant statistic is $\hat{p}$ and the sampling distribution of sample proportions can be approximated by the normal distribution.
4. Using technology we get the test statistic $z = -2.41039$ and P-value of 0.007968.
5. Because $0.007968 \leq 0.05$ we reject the null hypothesis.

Example 4

A study of sleepwalking of "nocturnal wandering" was described in *Neurology* magazine, and it included information that 29.2% of 19, 136 American adults have sleepwalked. Let us test the claim that "fewer than 30% of adults have sleepwalked" using $\alpha = 0.05$.

1. The original claim is $p < 0.30$ and the opposite claim is $p \geq 0.30$.
2. Because $p < 0.30$ does not contain equality, we have the hypotheses:

$$H_0 : \; p = 0.30 \quad \text{and} \quad H_A : \; p < 0.30$$

3. The relevant statistic is $\hat{p}$ and the sampling distribution of sample proportions can be approximated by the normal distribution.
4. Using technology we get the test statistic $z = -2.41039$ and P-value of 0.007968.
5. Because $0.007968 \leq 0.05$ we reject the null hypothesis.

We conclude that there is sufficient evidence to support the claim that fewer than 30% of adults have sleepwalked.

Critical Value Method

If we had used critical values instead of P-values for Example 4 we would have used a critical value of $z = -1.645$.

Critical Value Method

If we had used critical values instead of P-values for Example 4 we would have used a critical value of $z = -1.645$.

We would reject the null hypothesis because $z = -2.41039$ falls in the critical region bounded by $z = -1.645$.

Critical Value Method

If we had used critical values instead of P-values for Example 4 we would have used a critical value of $z = -1.645$.

We would reject the null hypothesis because $z = -2.41039$ falls in the critical region bounded by $z = -1.645$.

Note

This is the same conclusion we reached using P-values.

Confidence Interval Method

If we were to repeat Example 4 using confidence intervals we would get a 90% confidence interval of $0.287 < p < 0.297$.

Confidence Interval Method

If we were to repeat Example 4 using confidence intervals we would get a 90% confidence interval of $0.287 < p < 0.297$.

Since the entire confidence interval is less than 0.30, there is sufficient evidence to support the claim that fewer than 30% of adults have sleepwalked.

Confidence Interval Method

If we were to repeat Example 4 using confidence intervals we would get a 90% confidence interval of $0.287 < p < 0.297$.

Since the entire confidence interval is less than 0.30, there is sufficient evidence to support the claim that fewer than 30% of adults have sleepwalked.

Note

While this is the same result, remember that we have no guarantee that confidence intervals will give the same result as P-values.