Chapter 7
HYPOTHESIS TESTING

NGUYỄN THỊ THU THỦY [1]

SCHOOL OF APPLIED MATHEMATICS AND INFORMATICS
HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY

HANOI – 2020

[1] Email: thuy.nguyenthithu2@hust.edu.vn

Content

Introduction

Introduction

Hypothesis testing was introduced by Ronald Fisher, Jerzy Neyman, Karl Pearson and Pearson's son, Egon Pearson. Hypothesis testing is a statistical method that is used in making statistical decisions using experimental data. Hypothesis Testing is basically an assumption that we make about the population parameter.

7.1.1 Key terms and concepts

Null hypothesis

- Null hypothesis is a statistical hypothesis that assumes that the observation is due to a chance factor.
- Null hypothesis is denoted by H_0.

Alternative hypothesis

- Contrary to the null hypothesis, the alternative hypothesis shows that observations are the result of a real effect.
- The alternative hypothesis, denoted by H_1.

Examples

- $H_0 : \mu_1 = \mu_2$, which shows that there is no difference between the two population means.
- $H_1 : \mu_1 \neq \mu_2$ or $H_1 : \mu_1 > \mu_2$ or $H_1 : \mu_1 < \mu_2$.

7.1.1 Key terms and concepts

Level of significance

- Refers to the degree of significance in which we accept or reject the null-hypothesis.
- 100% accuracy is not possible for accepting or rejecting a hypothesis, so we therefore select a level of significance that is usually 5%.

7.1.1 Key terms and concepts

Type I error

- When we reject the null hypothesis, although that hypothesis was true.
- P[Type I error] $= \alpha$.
- In hypothesis testing, the normal curve that shows the critical region is called the α region.

Type II error

- When we accept the null hypothesis but it is false.
- P[Type II error] $= \beta$.
- In Hypothesis testing, the normal curve that shows the acceptance region is called the β region.

Power

Usually known as the probability of correctly accepting the null hypothesis. $1-\beta$ is called power of the analysis.

7.1.1 Key terms and concepts

Note

	H_0 **is true**	H_0 **is false**
Do not reject H_0	Correct decision	Type II error
Reject H_0	Type I error	Correct decision

7.1.1 Key terms and concepts

One-tailed test

A one-tailed test is a statistical test in which the critical area of a distribution is one-sided so that it is either greater than or less than a certain value, but not both.

Two-tailed test

A two-tailed test is a method in which the critical area of a distribution is two-sided and tests whether a sample is greater than or less than a certain range of values.

Examples

1. One-tailed test:

   Left-tailed test: $H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 < \mu_2$;

   Right-tailed test: $H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 > \mu_2$.
2. Two-tailed test: $H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 \neq \mu_2$.

7.1.2 Statistical decision for hypothesis testing

Statistical decision for hypothesis testing

1. In statistical analysis, we have to make decisions about the hypothesis. These decisions include deciding
   - if we should accept the null hypothesis or;
   - if we should reject the null hypothesis.
2. The rejection rule is as follows:
   - if the standardized test statistic is not in the rejection region, then we accept the null hypothesis;
   - if the standardized test statistic is in the rejection region, then we should reject the null hypothesis.

Rejection region

The rejection region is the values of test statistic for which the null hypothesis is rejected.

7.1.2 Statistical decision for hypothesis testing

Statistical decision for hypothesis testing

1. In statistical analysis, we have to make decisions about the hypothesis. These decisions include deciding
   - if we should accept the null hypothesis or;
   - if we should reject the null hypothesis.
2. The rejection rule is as follows:
   - if the standardized test statistic is not in the rejection region, then we accept the null hypothesis;
   - if the standardized test statistic is in the rejection region, then we should reject the null hypothesis.

Rejection region

The rejection region is the values of test statistic for which the null hypothesis is rejected.

7.1.2 Statistical decision for hypothesis testing

Example 1

You wish to show that the average hourly wage of carpenters in the state of California is different from \$14, which is the national average. This is the alternative hypothesis, written as

$$H_1 : \mu \neq 14.$$

The null hypothesis is

$$H_0 : \mu = 14.$$

You would like to reject the null hypothesis, thus concluding that the California mean is not equal to \$14.

7.1.2 Statistical decision for hypothesis testing

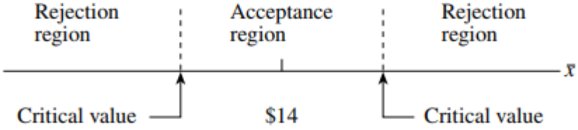

Figure: Rejection and acceptance regions for Example 1

7.1.2 Statistical decision for hypothesis testing

Example 2

A milling process currently produces an average of 3% defectives. You are interested in showing that a simple adjustment on a machine will decrease p, the proportion of defectives produced in the milling process. Thus, the alternative hypothesis is

$$H_1 : p < 0.03$$

and the null hypothesis is

$$H_0 : p = 0.03.$$

If you can reject H_0, you can conclude that the adjusted process produces fewer than 3% defectives.

7.1.2 Statistical decision for hypothesis testing

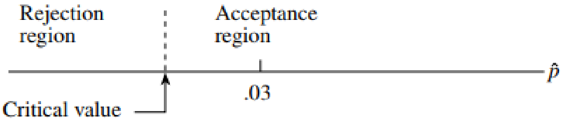

Figure: Rejection and acceptance regions for Example 2

7.1.2 Statistical decision for hypothesis testing

Approach to Hypothesis Testing with Fixed Probability of Type I Error

1. State the null and alternative hypotheses.
2. Choose a fixed significance level α.
3. Choose an appropriate test statistic and establish the critical region/rejection region based on α.
4. Reject H_0 if the computed test statistic is in the critical region/rejection region. Otherwise, do not reject.
5. Draw scientific or engineering conclusions

Content

Introduction to hypothesis testing

Steps for hypothesis testing

1. State the claim mathematically and verbally. Identify the null and alternative hypotheses (H_0 and H_1).
2. Find the standardized test statistic z.
3. Determine the rejection region W_α for H_0.
4. Make a decision to reject or fail to reject the null hypothesis. If z is in the rejection region ($z \in W_\alpha$), reject H_0. Otherwise ($z \notin W_\alpha$), fail to reject H_0.
5. Write a statement to interpret the decision in the context of the original claim.

Introduction to hypothesis testing

Note

The researcher uses the sample data to decide whether the evidence favors H_1 rather than H_0 and draws one of these two conclusions:

- Reject H_0 and conclude that H_1 is true.
- Accept (do not reject) H_0 as true.

7.2.1 Tests on a Single Mean (Variance Known)

We should first describe the assumptions on which the experiment is based. The model for the underlying situation centers around an experiment with $X_1, X_2, \ldots, X_n$ representing a random sample from a distribution with mean μ and variance $\sigma^2 > 0$. Consider first the hypothesis

$$H_0 : \mu = \mu_0, \qquad H_1 : \mu \neq \mu_0.$$

The appropriate test statistic should be based on the random variable $\overline{X}$. In Chapter 5, the Central Limit Theorem was introduced, which essentially states that despite the distribution of X, the random variable $\overline{X}$ has approximately a normal distribution with mean μ and variance σ^2/n for reasonably large sample sizes. So, $\mu_{\overline{X}} = \mu$ and $\sigma^2_{\overline{X}} = \sigma^2/n$.

7.2.1 Tests on a Single Mean (Variance Known)

It is convenient to standardize $\overline{X}$ and formally involve the standard normal random variable Z, where

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

We know that under H_0, that is, if $\mu = \mu_0$, $\dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$ follows an $\mathcal{N}(0, 1)$ distribution, and hence the expression

$$P\left[-z_{\alpha/2} < \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} < z_{\alpha/2}\right] = 1 - \alpha$$

can be used to write an appropriate non-rejection region.

7.2.1 Tests on a Single Mean (Variance Known)

Theorem 1 (Two-tailed test)

1. Identify the null and alternative hypotheses.

$$H_0 : \mu = \mu_0, \qquad H_1 : \mu \neq \mu_0.$$

2. Find the standardized test statistic z.

$$z = \frac{\overline{x} - \mu_0}{\sigma}\sqrt{n}.$$

3. Make a decision to reject or fail to reject the null hypothesis.

   If $z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$, then reject H_0;

   If $-z_{\alpha/2} < z < z_{\alpha/2}$, then fail to reject H_0.

7.2.1 Tests on a Single Mean (Variance Known)

Theorem 2 (One-tailed test)

(right-tailed test)

1. Identify the null and alternative hypotheses.

$$H_0 : \mu = \mu_0, \qquad H_1 : \mu > \mu_0.$$

2. Find the standardized test statistic z.

$$z = \frac{\overline{x} - \mu_0}{\sigma}\sqrt{n}.$$

3. Make a decision to reject or fail to reject the null hypothesis.

   If $z > z_\alpha$, then reject H_0;

   If $z < z_\alpha$, then fail to reject H_0.

7.2.1 Tests on a Single Mean (Variance Known)

Theorem 3 (One-tailed test)

(left-tailed test)

1. Identify the null and alternative hypotheses.

$$H_0 : \mu = \mu_0, \qquad H_1 : \mu < \mu_0.$$

2. Find the standardized test statistic z.

$$z = \frac{\overline{x} - \mu_0}{\sigma}\sqrt{n}.$$

3. Make a decision to reject or fail to reject the null hypothesis.

   If $z < -z_\alpha$, then reject H_0;

   If $z > -z_\alpha$, then fail to reject H_0.

7.2.1 Tests on a Single Mean (Variance Known)

Example 3

The average weekly earnings for female social workers is \$670. Do men in the same positions have average weekly earnings that are higher than those for women? A random sample of $n = 40$ male social workers showed $\overline{x} = \$725$. Assuming a population standard deviation of \$102, test the appropriate hypothesis using $\alpha = 0.01$.

Solution

You would like to show that the average weekly earnings for men are higher than \$670, the women's average. Hence, if μ is the average weekly earnings for male social workers, you can set out the formal test of hypothesis in steps.

7.2.1 Tests on a Single Mean (Variance Known)

Solution

1. Null and alternative hypotheses: $H_0 : \mu = 670$ versus $H_1 : \mu > 670$ (one-tailed test).
2. Using the sample information, calculate
$$z = \frac{(\overline{x} - \mu_0)}{\sigma}\sqrt{n} = \frac{(725 - 670)}{102}\sqrt{40} = 3.41.$$
3. Rejection region: For this one-tailed test, $z_\alpha = 2.33$ (shown in Figure 3).
4. Compare the observed value of the test statistic, $z = 3.41$, with the critical value necessary for rejection, $z_\alpha = 2.33$. Since the observed value of the test statistic falls in the rejection region, you can reject H_0.
5. Conclusion: The average weekly earnings for male social workers are higher than the average for female social workers. The probability that you have made an incorrect decision is $\alpha = 0.01$.

7.2.1 Tests on a Single Mean (Variance Known)

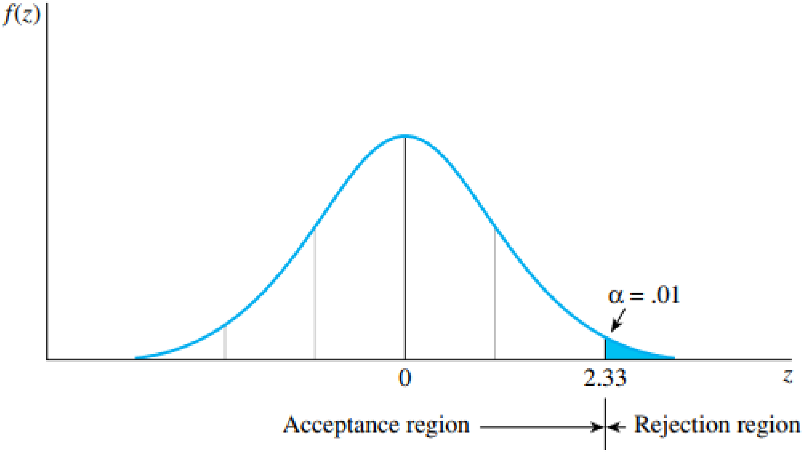

Figure: The rejection region for a right-tailed test with $\alpha = 0.01$

7.2.1 Tests on a Single Mean (Variance Known)

Note

Values of $\overline{x}$ that are either "too large" or "too small" in terms of their distance from μ_0 are placed in the rejection region. If you choose $\alpha = 0.01$, the area in the rejection region is equally divided between the two tails of the normal distribution, as shown in Figure 4. Using the standardized test statistic z, you can reject H_0 if $z > 2.58$ or $z < -2.58$. For different values of a, the critical values of z that separate the rejection and acceptance regions will change accordingly.

7.2.1 Tests on a Single Mean (Variance Known)

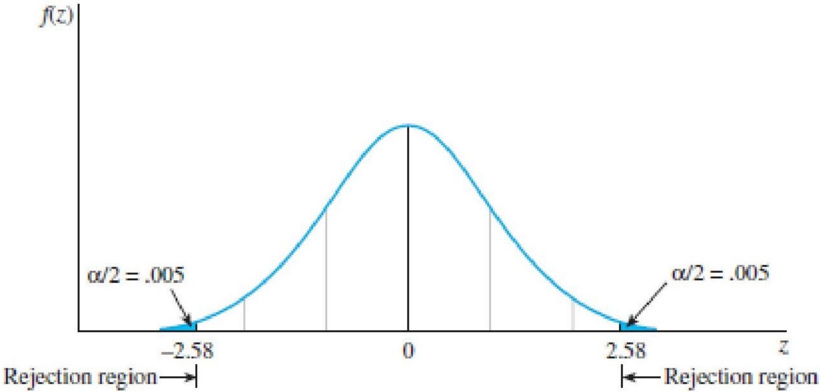

Figure: The rejection region for a two-tailed test with $\alpha = 0.01$

7.2.2 Tests on a Single Sample (Variance Unknown)

Note

The random variables $X_1, X_2, \dots, X_n$ represent a random sample from a normal distribution with unknown μ and σ^2. Then the random variable $\sqrt{n}(\overline{X} - \mu)/S$ has a t-distribution with $n-1$ degrees of freedom. The structure of the test is identical to that for the case of σ known, with the exception that the value σ in the test statistic is replaced by the computed estimate S and the standard normal distribution is replaced by a t-distribution.

7.2.2 Tests on a Single Sample (Variance Unknown)

Theorem 4 (Two-tailed test)

1. Identify the null and alternative hypotheses.

$$H_0 : \mu = \mu_0, \qquad H_1 : \mu \neq \mu_0.$$

2. Find the standardized test statistic t.

$$t = \frac{\overline{x} - \mu_0}{s}\sqrt{n}.$$

3. Make a decision to reject or fail to reject the null hypothesis.
   If $t < -t_{\alpha/2}^{(n-1)}$ or $t > t_{\alpha/2}^{(n-1)}$, then reject H_0;
   If $-t_{\alpha/2}^{(n-1)} < t < t_{\alpha/2}^{(n-1)}$, then fail to reject H_0.

7.2.2 Tests on a Single Sample (Variance Unknown)

Theorem 5 (One-tailed test)

(right-tailed test)

1. Identify the null and alternative hypotheses.

$$H_0 : \mu = \mu_0, \qquad H_1 : \mu > \mu_0.$$

2. Find the standardized test statistic t.

$$t = \frac{\overline{x} - \mu_0}{s}\sqrt{n}.$$

3. Make a decision to reject or fail to reject the null hypothesis.

   If $t > t_{\alpha}^{(n-1)}$, then reject H_0;

   If $t < t_{\alpha}^{(n-1)}$, then fail to reject H_0.

7.2.2 Tests on a Single Sample (Variance Unknown)

Theorem 6 (One-tailed test)

(left-tailed test)

1. Identify the null and alternative hypotheses.

$$H_0 : \mu = \mu_0, \qquad H_1 : \mu < \mu_0.$$

2. Find the standardized test statistic t.

$$t = \frac{\overline{x} - \mu_0}{s}\sqrt{n}.$$

3. Make a decision to reject or fail to reject the null hypothesis.
   If $t < -t_{\alpha}^{(n-1)}$, then reject H_0;
   If $t > -t_{\alpha}^{(n-1)}$, then fail to reject H_0.

7.2.2 Tests on a Single Sample (Variance Unknown)

Example 4

A local telephone company claims that the average length of a phone call is 8 minutes. In a random sample of 18 phone calls, the sample mean was 7.8 minutes and the standard deviation was 0.5 minutes. Is there enough evidence to support this claim at $\alpha = 0.05$?

7.2.2 Tests on a Single Sample (Variance Unknown)

Solution

1. Null and alternative hypotheses: $H_0 : \mu = 8$ versus $H_1 : \mu \neq 8$. The test is a two-tailed test.
2. $\alpha = 0.05$, critical region: $t < -2.110$ or $t > 2.110$, where $t = \frac{\overline{x}-\mu_0}{s}\sqrt{n}$ with 17 degrees of freedom (see Table t-distribution).
3. Computations: $\overline{x}$, $s = 0.5$, and $n = 18$. Hence, $t = \frac{(7.8-8)}{0.5}\sqrt{18} = -1.70$.
4. Decision: Do not reject H_0.
5. At the 5% level of significance, there is not enough evidence to reject the claim that the average length of a phone call is 8 minutes.

7.2.2 Tests on a Single Sample (Variance Unknown)

Note

If $n \geq 30$, $T \sim \mathcal{N}(0;1)$.

Example 5

The daily yield for a local chemical plant has averaged 880 tons for the last several years. The quality control manager would like to know whether this average has changed in recent months. She randomly selects 50 days from the computer database and computes the average and standard deviation of the $n = 50$ yields as $\overline{x} = 871$ tons and $s = 21$ tons, respectively. Test the appropriate hypothesis using $\alpha = 0.05$.

7.2.2 Tests on a Single Sample (Variance Unknown)

Solution

1. Null and alternative hypotheses: $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0,\ \mu_0 = 880$ (two-tailed test).
2. Rejection region: For this two-tailed test, you use values of z in both the right and left tails of the standard normal distribution. Using $\alpha = 0.05$, the critical values separating the rejection and acceptance regions cut off areas of $\alpha/2 = 0.025$ in the right and left tails. These values are $z_{\alpha/2} = 1.96$ and the null hypothesis will be rejected if $z > 1.96$ or $z < -1.96$.
3. Test statistic: The point estimate for μ is $\overline{x}$. Therefore, the test statistic is

$$z = \frac{(\overline{x} - \mu_0)}{s}\sqrt{n} = \frac{(871 - 880)}{21}\sqrt{50} = -3.03.$$

4. Conclusion: Since $z = -3.03$ and the calculated value of z falls in the rejection region, the manager can reject the null hypothesis that $\mu = 880$ tons and conclude that it has changed. The probability of rejecting H_0 when H_0 is true and $\alpha = 0.05$, a fairly small probability. Hence, she is reasonably confident that the decision is correct.

7.2.3 Hypothesis testing for proportions

The z-test for a population is a statistical test for a population proportion. The z-test can be used when a binomial distribution is given such that $np \geq 5$ and $n(1-p) \geq 5$. The test statistic is the sample proportion and the standardized test statistic is z.

$$z = \frac{\hat{p} - \mu_{\hat{p}}}{\sigma_{\hat{p}}} = \frac{\hat{p} - p}{\sqrt{p(1-p)}}\sqrt{n}.$$

Verify that $np \geq 5$ and $n(1-p) \geq 5$.

7.2.3 Hypothesis testing for proportions

Theorem 7 (two-tailed test)

1. Identify the null and alternative hypotheses.

$$H_0 : p = p_0, \qquad H_1 : p \neq p_0.$$

2. Find the standardized test statistic z.

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)}}\sqrt{n}.$$

3. Make a decision to reject or fail to reject the null hypothesis.

   If $z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$, then reject H_0;

   If $-z_{\alpha/2} < z < z_{\alpha/2}$, then fail to reject H_0.

7.2.3 Hypothesis testing for proportions

Theorem 8 (One-tailed test)

(right-tailed test)

1. Identify the null and alternative hypotheses.

$$H_0 : p = p_0, \qquad\qquad H_1 : p > p_0.$$

2. Find the standardized test statistic z.

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)}}\sqrt{n}.$$

3. Make a decision to reject or fail to reject the null hypothesis.

   If $z > z_\alpha$, then reject H_0;

   If $z < z_\alpha$, then fail to reject H_0.

7.2.3 Hypothesis testing for proportions

Theorem 9 (One-tailed test)

(left-tailed test)

1. Identify the null and alternative hypotheses.

$$H_0 : p = p_0, \qquad\qquad H_1 : p < p_0.$$

2. Find the standardized test statistic z.

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)}}\sqrt{n}.$$

3. Make a decision to reject or fail to reject the null hypothesis.

   If $z < -z_\alpha$, then reject H_0;

   If $z > -z_\alpha$, then fail to reject H_0.

7.2.3 Hypothesis testing for proportions

Example 6

A college claims that more than 94% of their graduates find employment within 6 months of graduation. In a sample of 500 randomly selected graduates, 475 of them were employed. Is there enough evidence to support the college's claim at a 1% level of significance?

7.2.3 Hypothesis testing for proportions

Solution

Verify $np_0 \geq 5$ and $n(1-p_0) \geq 5$: $np_0 = (500)(0.94) = 470$; $n(1-p_0) = (500)(0.06) = 30$. Normal Distribution.

1. Null and alternative hypotheses: $H_0 : p = 0.94$ versus $H_1 : p > 0.94$ (right-tailed test).
2. Computations: $\hat{p} = \frac{475}{500} = 0.95$,
$$z = \frac{(0.95 - 0.94)}{\sqrt{0.94 \times 0.06}}\sqrt{500} = 0.94.$$
3. The critical value: $z_\alpha = z_{0.01} = 2.33$ (see Table the values of standard normal CDF $\Phi(x) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{x} e^{\frac{-t^2}{2}} dt$).
4. Decision: $z = 0.94 < 2.33 = z_\alpha$, H_0 is not rejected.
5. At the 1% level of significance, there is not enough evidence to support the college's claim.

7.2.3 Hypothesis testing for proportions

Example 7

A cigarette manufacturer claims that 1/8 of the US adult population smokes cigarettes. In a random sample of 100 adults, 5 are cigarette smokers. Test the claim at $\alpha = 0.05$.

7.2.3 Hypothesis testing for proportions

Solution

Verify np_0 and $n(1-p_0)$ are at least 5. $np_0 = (100)(0.125) = 12.5$; $n(1-p_0) = (100)(0.875) = 87.5$.

1. Null and alternative hypotheses: $H_0 : p = 0.125$; $H_1 : p \neq 0.125$ (two-tailed test).
2. Computations: $\hat{p} = \frac{5}{100} = 0.05$,

$$z = \frac{(0.05 - 0.125)}{\sqrt{(0.125)(0.875)}}\sqrt{100} = -2.27.$$

3. The critical value: $z_{\alpha/2} = z_{0.025} = 1.96$ (see Table the values of standard normal CDF $\Phi(x) = \dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{x} e^{\frac{-t^2}{2}} dt$).
4. $z = -2.27 < -1.96$. H_0 is rejected.
5. At the 5% level of significance, there is enough evidence to reject the claim that one-eighth of the population smokes.

Content

Two Sample: Test on Two Mean

Null and Alternative Hypothesis

1. In a two-sample hypothesis test, two parameters from two populations are compared.
2. The null hypothesis H_0 is a statistical hypothesis that usually states there is no difference between the parameters of two populations. The null hypothesis always contains the symbol " $=$ ".
3. The alternative hypothesis H_1 is a statistical hypothesis that is true when H_0 is false. The alternative hypothesis always contains the symbol " $>, \neq, <$ ".

Null and Alternative Hypothesis

$$\begin{cases} H_0 : \mu_1 = \mu_2, \\ H_1 : \mu_1 \neq \mu_2, \end{cases} \qquad \begin{cases} H_0 : \mu_1 = \mu_2, \\ H_1 : \mu_1 > \mu_2, \end{cases} \qquad \begin{cases} H_0 : \mu_1 = \mu_2, \\ H_1 : \mu_1 < \mu_2. \end{cases}$$

Regardless of which hypotheses used, $\mu_1 = \mu_2$ is always assumed to be true.

Two Sample: Test on Two Mean

Two Sample z-Test

Three conditions are necessary to perform a u-test for the difference between two population means μ_1 and μ_2.

1. The samples must be randomly selected.
2. The samples must be independent. Two samples are independent if the sample selected from one population is not related to the sample selected from the second population.
3. Each sample size must be at least 30, or, if not, each population must have a normal distribution with a known standard deviation.

7.3.1 σ_1^2 and σ_2^2 are Known

Theorem 10 (σ_1^2 and σ_2^2 are Known)

1. Null hypothesis: $H_0 : \mu_1 - \mu_2 = D_0$, where D_0 is some specified difference that you wish to test. For many tests, you will hypothesize that there is no difference between μ_1 and μ_2; that is, $D_0 = 0$.
2. Alternative hypothesis:
   - **(a)** One-tailed test: $H_1 : \mu_1 - \mu_2 > D_0$ or $\mu_1 - \mu_2 < D_0$.
   - **(b)** Two-tailed test: $H_1 : \mu_1 - \mu_2 \neq D_0$.
3. Test statistic: $z = \dfrac{(\overline{x}_1 - \overline{x}_2) - D_0}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$.
4. Rejection region: Reject H_0 when
   - **(a)** One-tailed test: $z > z_\alpha$ (when the alternative hypothesis is $H_1 : \mu_1 - \mu_2 > D_0$) or $z < -z_\alpha$ (when the alternative hypothesis is $H_1 : \mu_1 - \mu_2 < D_0$).
   - **(b)** Two-tailed test: $z > z_{\alpha/2}$ or $z < -z_{\alpha/2}$.

7.3.2 σ_1^2 and σ_2^2 are Unknown, $n_1, n_2 \geq 30$

Theorem 11 (σ_1^2 and σ_2^2 are Unknown, $n_1, n_2 \geq 30$)

1. Null hypothesis: $H_0 : \mu_1 - \mu_2 = D_0$, where D_0 is some specified difference that you wish to test. For many tests, you will hypothesize that there is no difference between μ_1 and μ_2; that is, $D_0 = 0$.
2. Alternative hypothesis:
   - **(a)** One-tailed test: $H_1 : \mu_1 - \mu_2 > D_0$ or $\mu_1 - \mu_2 < D_0$.
   - **(b)** Two-tailed test: $H_1 : \mu_1 - \mu_2 \neq D_0$.
3. Test statistic: $z = \dfrac{(\overline{x}_1 - \overline{x}_2) - D_0}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$.
4. Rejection region: Reject H_0 when
   - **(a)** One-tailed test: $z > z_\alpha$ (when the alternative hypothesis is $H_1 : \mu_1 - \mu_2 > D_0$) or $z < -z_\alpha$ (when the alternative hypothesis is $H_1 : \mu_1 - \mu_2 < D_0$).
   - **(b)** Two-tailed test: $z > z_{\alpha/2}$ or $z < -z_{\alpha/2}$.

7.3.2 σ_1^2 and σ_2^2 are Unknown, $n_1, n_2 \geq 30$

Example 8

A high school math teacher claims that students in her class will score higher on the math portion of the ACT then students in a colleague's math class. The mean ACT math score for 49 students in her class is 22.1 and the sample standard deviation is 4.8. The mean ACT math score for 44 of the colleague's students is 19.8 and the sample standard deviation is 5.4. At $\alpha = 0.10$, can the teacher's claim be supported?

7.3.2 σ_1^2 and σ_2^2 are Unknown, $n_1, n_2 \geq 30$

Example 8 Solution

Let μ_1 and μ_2 represent the population means of the ACT math in two classes, respectively.

1. $H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 > \mu_2$.
2. The standardized error is

$$\sigma_{\overline{x}_1 - \overline{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{4.8^2}{49} + \frac{5.4^2}{44}} \simeq 1.0644.$$

   The standardized test statistic is

$$z = \frac{22.1 - 19.8}{1.0644} \simeq 2.161.$$

3. $\alpha = 0.10$, the critical value $z_\alpha = 1.28$.
4. Since $z = 2.161 > z_\alpha = 1.28$, reject H_0.
5. There is enough evidence at the 10% level to support the teacher's claim that her students score better on the ACT.

7.3.3 σ_1^2 and σ_2^2 are Unknown, $n_1, n_2 < 30$

Two Sample t-Test

If samples of size less than 30 are taken from normally-distributed populations, a t-test may be used to test the difference between the population means μ_1 and μ_2.

Three conditions are necessary to use a t-test for small independent samples.

1. The samples must be randomly selected.
2. The samples must be independent. Two samples are independent if the sample selected from one population is not related to the sample selected from the second population.
3. Each population must have a normal distribution.

7.3.3 σ_1^2 and σ_2^2 are Unknown, $n_1, n_2 < 30$

Theorem 12 (σ_1^2 and σ_2^2 are Unknown but equal variances, $n_1, n_2 < 30$)

1. Null hypothesis: $H_0 : \mu_1 - \mu_2 = D_0$, where D_0 is some specified difference that you wish to test. For many tests, you will hypothesize that there is no difference between μ_1 and μ_2; that is, $D_0 = 0$.
2. Alternative hypothesis:
   - **(a)** One-tailed test: $H_1 : \mu_1 - \mu_2 > D_0$ or $\mu_1 - \mu_2 < D_0$.
   - **(b)** Two-tailed test: $H_1 : \mu_1 - \mu_2 \neq D_0$.
3. Test statistic: $t = \dfrac{(\overline{x}_1 - \overline{x}_2) - D_0}{\sqrt{\dfrac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$.
4. Rejection region: Reject H_0 when
   - **(a)** One-tailed test: $t > t_{\alpha}^{(n_1+n_2-2)}$ (when the alternative hypothesis is $H_1 : \mu_1 - \mu_2 > D_0$) or $t < -t_{\alpha}^{(n_1+n_2-2)}$ (when the alternative hypothesis is $H_1 : \mu_1 - \mu_2 < D_0$).
   - **(b)** Two-tailed test: $t > t_{\alpha/2}^{(n_1+n_2-2)}$ or $t < -t_{\alpha/2}^{(n_1+n_2-2)}$.

7.3.3 σ_1^2 and σ_2^2 are Unknown, $n_1, n_2 < 30$

Example 9

A random sample of 17 police officers in Brownsville has a mean annual income of \$35800 and a sample standard deviation of \$7800. In Greensville, a random sample of 18 police officers has a mean annual income of \$35100 and a sample standard deviation of \$7375. Test the claim at $\alpha = 0.01$ that the mean annual incomes in the two cities are not the same. Assume the population variances are equal.

7.3.3 σ_1^2 and σ_2^2 are Unknown, $n_1, n_2 < 30$

Example 9 Solution

Let μ_1 and μ_2 represent the population means of annual incomes in Brownsville and Greensville, respectively.

1. State the claim mathematically. $H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 \neq \mu_2$.
2. The standardized error is

$$\sigma_{\overline{x}_1 - \overline{x}_2} = \hat{\sigma}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$= \sqrt{\frac{(17-1)7800^2 + (18-1)7375^2}{17+18-2}}\sqrt{\frac{1}{17} + \frac{1}{18}} = 2564.92.$$

The standardized test statistic is

$$t = \frac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)}{\sigma_{\overline{x}_1 - \overline{x}_2}} = \frac{(35800 - 35100) - 0}{2564.92} \simeq 0.273.$$

7.3.3 σ_1^2 and σ_2^2 are Unknown, $n_1, n_2 < 30$

Example 9 Solution

1. $\alpha = 0.01$, $t_{\alpha/2}^{(n_1+n_2-2)} = 2.576$.
2. Since $-2.576 < t = 0.273 < 2.576$, fail to reject H_0.
3. There is not enough evidence at the 1% level to support the claim that the mean annual incomes differ.

Content

7.4 Testing the Difference Between Proportions

Two Sample z-Test for Proportions

A z-test is used to test the difference between two population proportions, p_1 and p_2. Three conditions are required to conduct the test.

1. The samples must be randomly selected.
2. The samples must be independent.
3. The samples must be large enough to use a normal sampling distribution. That is, $n_1p_1 \geq 5$, $n_1(1-p_1) \geq 5$, $n_2p_2 \geq 5$, $n_2(1-p_2) \geq 5$.

7.4 Testing the Difference Between Proportions

Mean and Standard error

If these conditions are met, then the sampling distribution for $\hat{p}_1 - \hat{p}_2$ is a normal distribution with mean

$$\mu_{\hat{p}_1 - \hat{p}_2} = p_1 - p_2.$$

and standard error

$$\sigma_{\hat{p}_1 - \hat{p}_2} = \sqrt{\bar{p}(1-\bar{p})\Big(\frac{1}{n_1} + \frac{1}{n_2}\Big)}.$$

A weighted estimate of p_1 and p_2 can be found by using

$$\bar{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}$$

7.4 Testing the Difference Between Proportions

Standardized test statistic

A two sample z-test is used to test the difference between two population proportions p_1 and p_2 when a sample is randomly selected from each population. The test statistic is

$$\hat{p}_1 = \hat{p}_2,$$

and the standardized test statistic is

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}.$$

7.4 Testing the Difference Between Proportions

Steps for hypothesis testing

Step 1. State the claim mathematically and verbally. Identify the null and alternative hypotheses.

Null hypothesis H_0	$p_1 = p_2$	$p_1 = p_2$	$p_1 = p_2$
Alternative hypothesis H_1	$p_1 \neq p_2$	$p_1 > p_2$	$p_1 < p_2$

Step 2. Find the standardized test statistic.

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}.$$

7.4 Testing the Difference Between Proportions

Steps for hypothesis testing

Step 3. Determine the rejection region.

H_0	H_1	Rejection region W_α
$p_1 = p_2$	$p_1 \neq p_2$	$(-\infty; -z_{\alpha/2}) \cup (z_{\alpha/2}; +\infty)$
$p_1 = p_2$	$p_1 > p_2$	$(z_\alpha; +\infty)$
$p_1 = p_2$	$p_1 < p_2$	$(-\infty; -z_\alpha)$

where $z_{\alpha/2}$ and $z_{1-\alpha}$ are in Table 1.

Step 4. Make a decision to reject or fail to reject the null hypothesis.

Step 5. Interpret the decision in the context of the original claim.

7.4 Testing the Difference Between Proportions

Example 10

A recent survey stated that male college students smoke less than female college students. In a survey of 1245 male students, 361 said they smoke at least one pack of cigarettes a day. In a survey of 1065 female students, 341 said they smoke at least one pack a day. At $\alpha = 0.01$, can you support the claim that the proportion of male college students who smoke at least one pack of cigarettes a day is lower then the proportion of female college students who smoke at least one pack a day?

7.4 Testing the Difference Between Proportions

Example 10 Solution

Let p_1 and p_2 represent the population proportions of of male and female college students, respectively.

1. $H_0 : p_1 = p_2$, $H_1 : p_1 < p_2$.
2. Caculate

$$\bar{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} = \frac{361 + 341}{1245 + 1065} = \frac{702}{2310} \simeq 0.304 \text{ and } 1 - \bar{p} = 0.696.$$

Because 1245×0.304, 1245×0.696, 1065×0.304, and 1065×0.696 are all at least 5, we can use a two-sample z-test.

$$z = \frac{0.29 - 0.32}{\sqrt{0.304 \times 0.696 \times \left(\frac{1}{1245} + \frac{1}{1065}\right)}} \simeq -1.56.$$

SAMI 1956

7.4 Testing the Difference Between Proportions

Example 10 Solution (continuous)

1. $\alpha = 0.01$, the critical value $z_\alpha = 2.33$.
2. Since $z = -1.56 > -z_\alpha = -2.33$, fail to reject H_0.
3. There is not enough evidence at the 1% level to support the claim that the proportion of male college students who smoke is lower then the proportion of female college students who smoke.

Content

Problems

Problem 1

A random sample of 64 bags of white cheddar popcorn weighed, on average, 5.23 ounces with a standard deviation of 0.24 ounce. Test the hypothesis that $\mu = 5.5$ ounces against the alternative hypothesis, $\mu < 5.5$ ounces, at the 0.05 level of significance.

Problem 2

A local telephone company claims that the average length of a phone call is 8 minutes. In a random sample of 18 phone calls, the sample mean was 7.8 minutes and the standard deviation was 0.5 minutes. Is there enough evidence to support this claim at $\alpha = 0.05$?

Problems

Problem 3

According to a dietary study, high sodium intake may be related to ulcers, stomach cancer, and migraine headaches. The human requirement for salt is only 220 milligrams per day, which is surpassed in most single servings of ready-to-eat cereals. If a random sample of 20 similar servings of a certain cereal has a mean sodium content of 244 milligrams and a sample standard deviation of 24.5 milligrams, does this suggest at the 0.05 level of significance that the average sodium content for a single serving of such cereal is greater than 220 milligrams? Assume the distribution of sodium content to be normal.

Problem 4

A marketing expert for a pasta-making company believes that 40% of pasta lovers prefer lasagna. If 9 out of 20 pasta lovers choose lasagna over other pastas, what can be concluded about the expert's claim? Use a 0.05 level of significance.

Problems

Problem 5

It is believed that at least 60% of the residents in a certain area favor an annexation suit by a neighboring city. What conclusion would you draw if only 110 in a sample of 200 voters favored the suit? Use a 0.05 level of significance.

Problem 6

To determine whether car ownership affects a student's academic achievement, two random samples of 100 male students were each drawn from the student body. The grade point average for the $n_1 = 100$ non-owners of cars had an average and variance equal to $\overline{x}_1 = 2.70$ and $s_1^2 = 0.36$, while $\overline{x}_2 = 2.54$ and $s_2^2 = 0.40$ for the $n_2 = 100$ car owners. Do the data present sufficient evidence to indicate a difference in the mean achievements between car owners and nonowners of cars? Test using $\alpha = 0.05$.

Problems

Problem 7

A manufacturer claims that the average tensile strength of thread A exceeds the average tensile strength of thread B by at least 12 kilograms. To test this claim, 50 pieces of each type of thread were tested under similar conditions. Type A thread had an average tensile strength of 86.7 kilograms with a standard deviation of 6.28 kilograms, while type B thread had an average tensile strength of 77.8 kilograms with a standard deviation of 5.61 kilograms. Test the manufacturer's claim using a 0.05 level of significance.

Problem 8

Engineers at a large automobile manufacturing company are trying to decide whether to purchase brand A or brand B tires for the company's new models. To help them arrive at a decision, an experiment is conducted using 12 of each brand. The tires are run until they wear out. The results are as follows:
Brand A: $\overline{x}_A = 37,900$ kilometers, $s_A = 5100$ kilometers.
Brand B: $\overline{x}_B = 39,800$ kilometers, $s_B = 5900$ kilometers.
Test the hypothesis that there is no difference in the average wear of the two brands of tires. Assume the populations to be approximately normally distributed with equal variances. Use a 0.01 level of significance.

Problems

Problem 9

A recent survey stated that male college students smoke less than female college students. In a survey of 1245 male students, 361 said they smoke at least one pack of cigarettes a day. In a survey of 1065 female students, 341 said they smoke at least one pack a day. At $\alpha = 0.01$, can you support the claim that the proportion of male college students who smoke at least one pack of cigarettes a day is lower then the proportion of female college students who smoke at least one pack a day?

Problem 10

In a study to estimate the proportion of residents in a certain city and its suburbs who favor the construction of a nuclear power plant, it is found that 63 of 100 urban residents favor the construction while only 59 of 125 suburban residents are in favor. Is there a significant difference between the proportions of urban and suburban residents who favor the construction of the nuclear plant? Use a 0.01 level of significance.