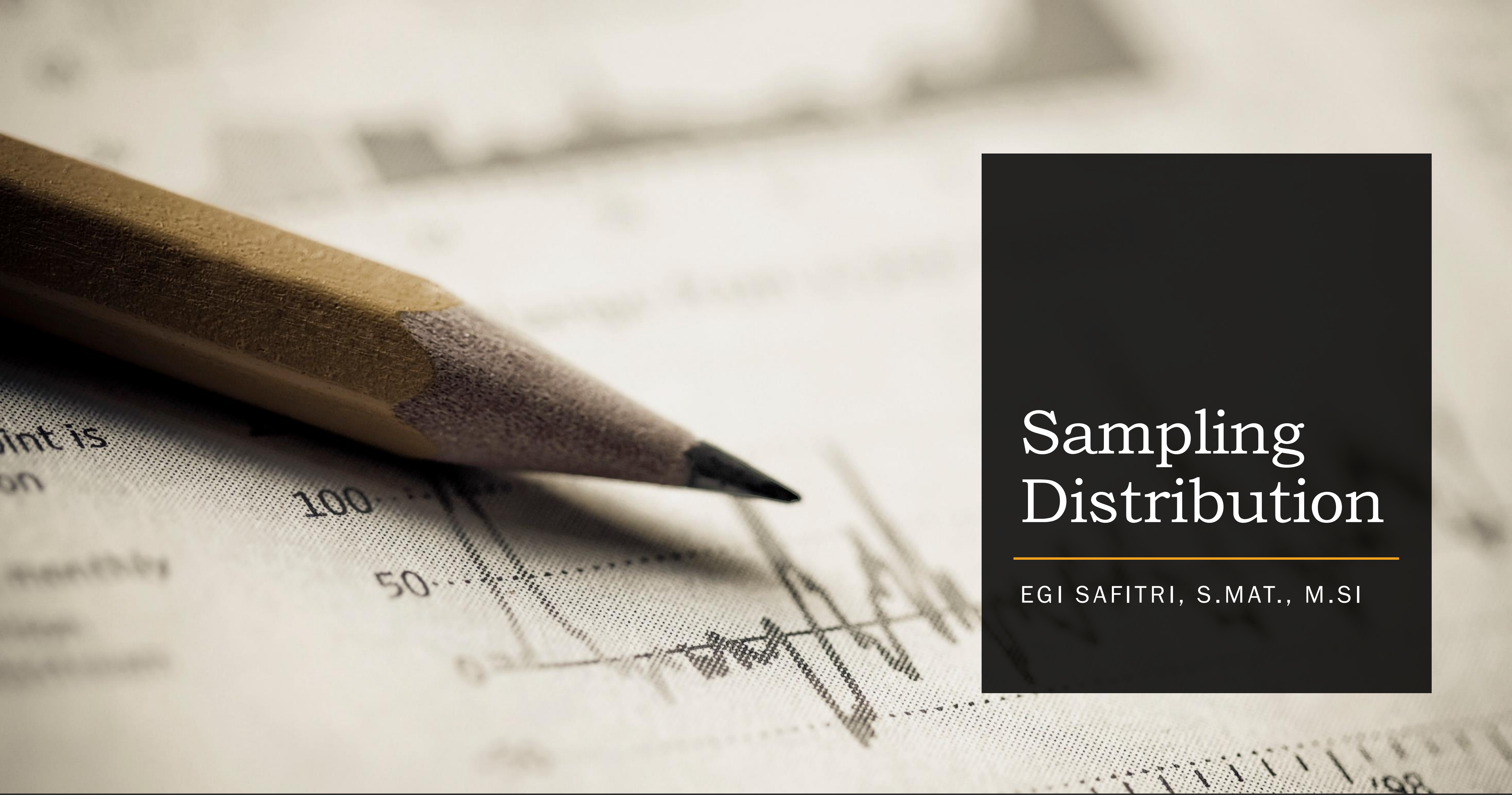


Sampling Distribution

EGI SAFITRI, S.MAT., M.SI

Hypothesis

- Hypothesis is a temporary assumption or statement that has a weak level of certainty and must be tested using specific techniques.
- Hypothesis is a theoretical or deductive answer and is temporary.
- Hypothesis is a statement about the condition of the population that will be tested for its truthfulness using data/information collected through sampling.
- If the statement is made to explain the value of a population parameter, it is called a statistical hypothesis.

Source Hypothesis

Hypotheses are formulated based on theory, research results, journals/scientific magazines, personal experiences/other people's experiences, or general phenomena.

The formulation of the hypothesis serves as a guide in the research design, data collection techniques, data analysis, and conclusions.

Hypothesis Testing

Hypothesis testing is a procedure that allows decisions to be made, i.e., a decision to reject or accept a hypothesis that is being tested.

Hypothesis Based on:

Theory: A hypothesis can be based on pre-existing theories. Theories are concepts or chains of concepts that have been scientifically explained and tested. Hypotheses derived from theories often attempt to test or complement the understanding within the theory.

Experience: Personal or collective experiences can be used as the basis for formulating hypotheses. Observations or insights from daily experiences can contribute to the development of a hypothesis.

Critical Thinking: A hypothesis can also arise from critical thinking or conceptual models, where relationships between variables or concepts are theorized and then tested in research.

Error Type

Types of Errors in Hypothesis Testing:

Type I Error: Rejecting a true null hypothesis (false positive) (α).

Type II Error: Accepting a false null hypothesis (false negative) (β).

Hypothesis Formulation

- Dinyatakan sebagai kalimat pernyataan (deklaratif)
- Melibatkan minimal dua variabel penelitian
- Mengandung suatu prediksi
- Harus dapat diuji (testable)

Hypothesis by Analysis

- Hipotesis korelatif yaitu pernyataan tentang ada atau tidak adanya hubungan antara dua variabel atau lebih
- Hipotesis komparatif yaitu pernyataan tentang ada atau tidak adanya perbedaan antara dua kelompok atau lebih

Hypothesis by Form

Dibedakan 2 jenis :

1. **Hipotesis nol** : suatu pernyataan yang akan diuji, hipotesis tersebut tidak memiliki perbedaan/ perbedaannya nol dengan hipotesis sebenarnya.
2. **Hipotesis alternatif** : segala hipotesis yang berbeda dengan hipotesis nol. Pemilihan hipotesis ini tergantung dari sifat masalah yang dihadapi

Null Hypothesis

- Dinotasikan dengan Ho
- Penulisan, Ho : μ = suatu angka numerik
- Ditulis dengan tanda =, walaupun maksudnya adalah ≤, ataupun ≥

Alternative Hypothesis

- Sebagai lawan dari hipotesis nol (komplemen)
- Mempunyai tanda ≠, atau <, atau >
- Dinotasikan dengan Ha
- Penulisan,
  - Ha : μ ≠ suatu angka → sebagai pengujian dua arah
  - Ha : μ > suatu angka → sebagai pengujian satu arah (positif/kanan)
  - Ha : μ < suatu angka → sebagai pengujian satu arah (negatif/kiri)
- Penentuan pengujian satu atau dua arah berdasarkan pernyataan hipotesis penelitian.

Determining Ho and Ha

Langkah :

1. Menyatakan hipotesis secara matematik
2. Menyatakan alternatif secara matematik
3. Pilih dan tentukan hipotesis alternatif
4. Nyatakan hipotesis nolnya

- Contoh : Apakah rata-rata lama menonton TV adalah 12 jam ?

1. $\mu = 12$
2. $\mu \neq 12$
3. Ha: $\mu \neq 12$
4. Ho: $\mu = 12$

Hypothesis Testing

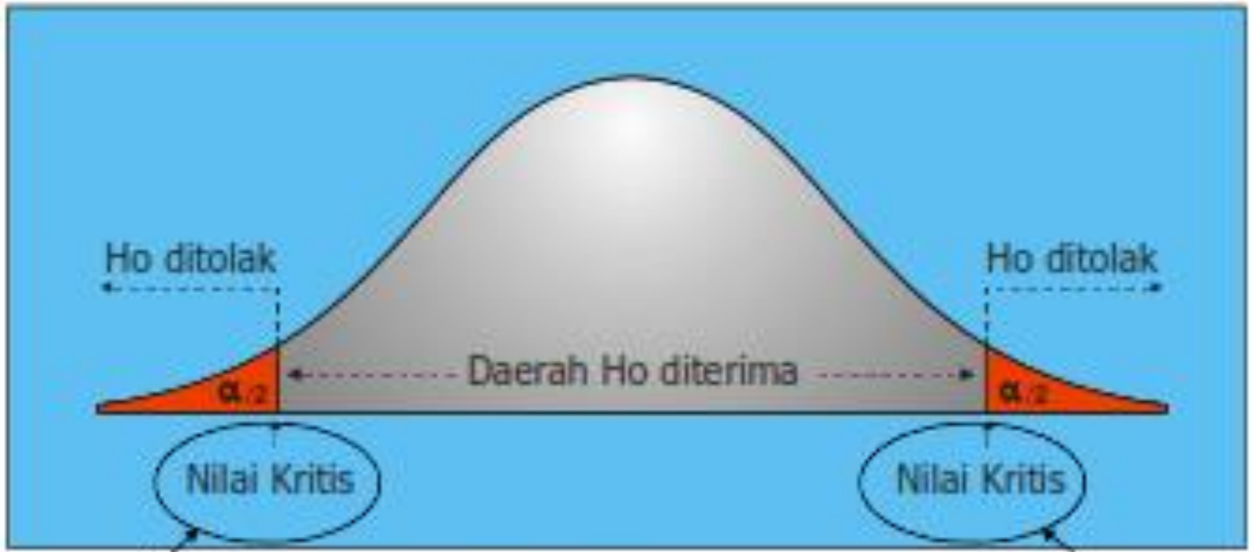
- Pengujian dua arah *(two tailed)*
  - Digunakan untuk menguji hipotesis nondirectional (belum jelas arahnya)
  - Misalnya ada perbedaan, ada korelasi
- Pengujian satu arah *(one tailed)*
  - Uji arah kanan, Misalnya:
    - IPK Mhs Wanita lebih baik daripada pria,
    - ada hubungan yang positif antara X dan Y
  - Uji arah kiri, Misalnya:
    - IPK mhs wanita lebih rendah daripada pria,
    - ada hubungan yang negatif antara X dan Y

Testing Example

Ada perbedaan prestasi belajar antara mahasiswa pria dan wanita

UJI DUA ARAH

- Ho: $\mu_1 = \mu_2$
- Ha: $\mu_1 \neq \mu_2$



Ho diterima jika:

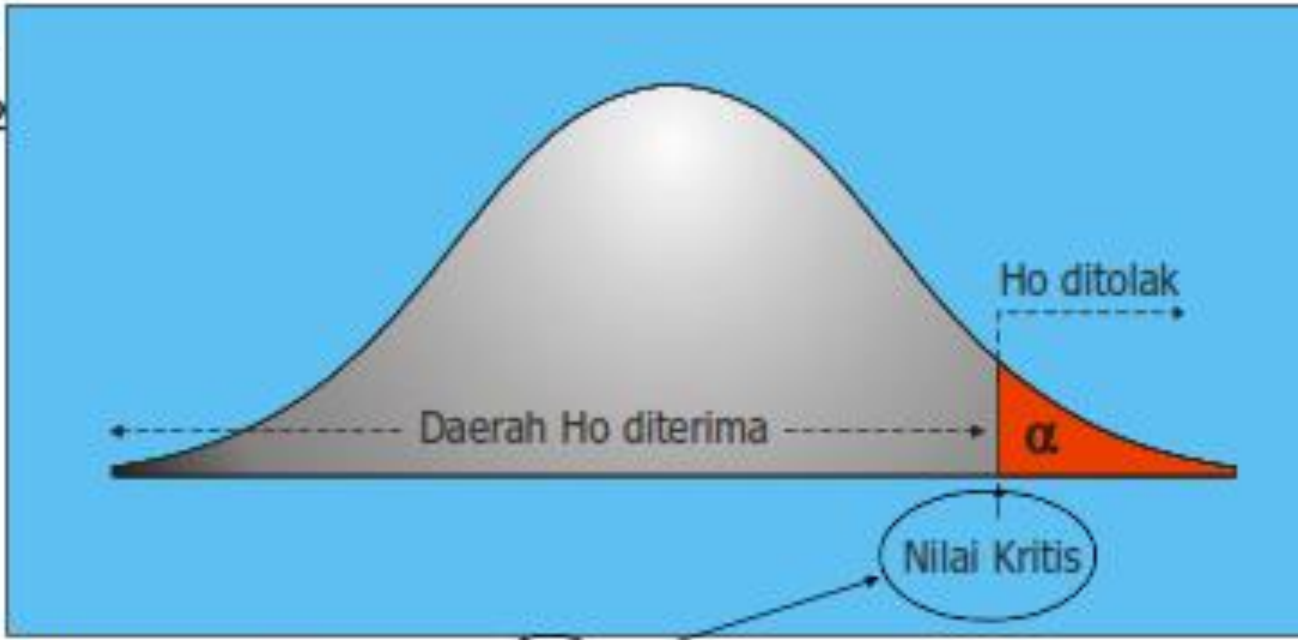
$$-Z_{(1-\alpha/2)} \leq Z_h \leq Z_{(1-\alpha/2)}$$

Testing Example

Metode Pembelajaran CTL Lebih Unggul daripada Metode Pembelajaran Tradisional

UJI SATU ARAH (KANAN)

- Ho: $\mu_1 \leq \mu_2$
- Ha: $\mu_1 > \mu_2$



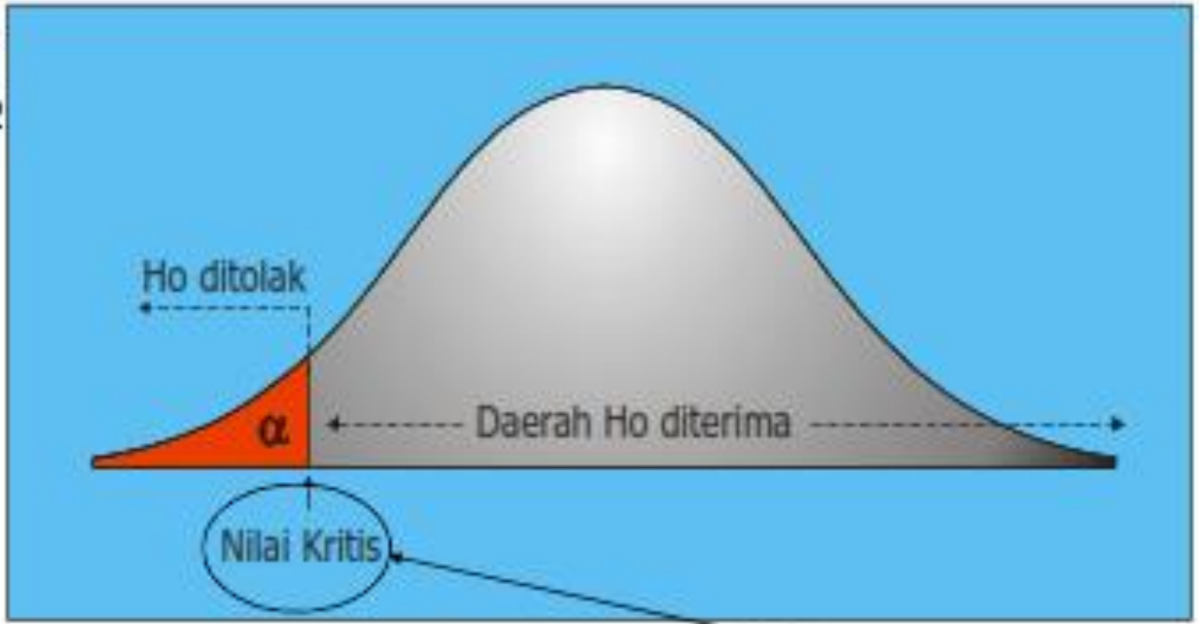
Ho diterima jika: $z_h \leq z_{1-\alpha}$

Testing Example

Masa studi lulusan wanita lebih cepat dibandingkan dengan lulusan pria

UJI SATU ARAH (KIRI)

- Ho: $\mu_1 \geq \mu_2$
- Ha: $\mu_1 < \mu_2$



Ho diterima jika: $z_h \geq -z_{1-\alpha}$

Testing Procedure

1. **Formulate the null hypothesis (H_0) and alternative hypothesis (H_1)**.
2. **Determine the significance level (α)** to be used.
3. **Calculate the test statistic** (z-test, t-test, F-test, or χ^2 test).
4. **Determine the critical value** (z, t, F, or χ^2).
5. **Compare the calculated test statistic with the critical value** to decide whether to accept or reject H_0.
6. **Draw a conclusion** based on the comparison.

Testing Averages: Two-Party Test

Two-Party Test

- Suppose a population is normally distributed with an average rate of μ and a standard deviation of σ. A sample of size n is taken, and the statistics calculated are the sample mean ($\bar{x}$) and standard deviation (s). The test is conducted as follows:

When σ is known:

- **Null Hypothesis (H_0)**: $\mu = \mu_0$ (the population mean equals a specified value).
- **Alternative Hypothesis (H_1)**: $\mu \neq \mu_0$ (the population mean is different from the specified value).

The statistic used is:

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

- If the test statistic falls within the acceptance region, we fail to reject H_0. If it falls outside, we reject H_0.

Two-Party Test

When σ is unknown:

- The sample standard deviation (s) is used instead of σ, and the t-distribution is applied:

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

Example

A light bulb manufacturer claims the bulbs last about 800 hours. However, due to recent concerns, it is suspected the average lifespan might have changed. A test is done with a sample of 50 bulbs, and the sample mean lifespan is found to be 792 hours with a standard deviation of 60 hours. The significance level is set at 0.05.

- **Null Hypothesis (H_0)**: μ = 800 hours
- **Alternative Hypothesis (H_1)**: μ ≠ 800 hours

The test statistic calculated using the z-test is:

$$z = \frac{792 - 800}{60/\sqrt{50}} = -0.94$$

- Since the calculated z-value is within the acceptance region, we conclude that the average lifespan is still 800 hours.

Answer:

1. **Hypothesis Formulation**
    - $H_0 : \mu = 800$ hours, meaning the lamp's lifetime is around 800 hours.
    - $H_1 : \mu \neq 800$ hours, meaning the lamp's quality has changed, not 800 hours anymore.
2. **Since the sample is sufficiently large**, the normal distribution can be used.
3. **Two-tailed test**
4. **Significance level** $\alpha = 0.05$, so:

    $$-z_{\alpha/2} < z < z_{\alpha/2}$$

    which gives:

    $$-1.96 < z < 1.96$$

5. **Test statistic**:

    $$z = \frac{792 - 800}{60/\sqrt{50}} = -0.94$$

6. **Conclusion**: $z_{\text{hit}} = -0.94$, which falls within the acceptance region of H_0. Therefore, at the 0.05 significance level, we fail to reject H_0, meaning the average lamp lifetime is still about 800 hours.

b. σ is unknown

For the hypothesis pair:

- $H_0 : \mu = \mu_0$
- $H_1 : \mu \neq \mu_0$

Since the standard deviation σ is unknown, it is estimated with the sample standard deviation s. The test statistic used is:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

With $dk = n - 1$. Thus, H_0 is accepted if:

$$-t_{1-\alpha/2} < t < t_{1-\alpha/2}$$

where $t_{1-\alpha/2}$ is taken from the t-distribution table with degrees of freedom $dk = n - 1$.

Example

For the example above, if the population standard deviation is unknown and the sample standard deviation is $s = 55$ hours, the test statistics would be:

1. **Hypothesis Formulation**
   - $H_0 : \mu = 800$ hours, meaning the lamp's lifetime is around 800 hours.
   - $H_1 : \mu \neq 800$ hours, meaning the lamp's quality has changed.
2. **Test statistic**: t
3. **Two-tailed test**
4. **Significance level** $\alpha = 0.05$, so:

   $$-t_{1-\alpha/2} < t < t_{1-\alpha/2}$$

   which gives:

   $$-2.011 < t < 2.011$$

5. **Test statistic**:

   $$t = \frac{792 - 800}{55/\sqrt{50}} = -1.029$$

6. **Conclusion**: $t = -1.029$, which falls within the acceptance region of H_0. Therefore, at the 0.05 significance level, we fail to reject H_0, meaning the average lamp lifetime is still about 800 hours.

Testing the Averages: One-Party Test

For the example above, if the population standard deviation is unknown and the sample standard deviation is $s = 55$ hours, the test statistics would be:

1. **Hypothesis Formulation**
   - $H_0 : \mu = 800$ hours, meaning the lamp's lifetime is around 800 hours.
   - $H_1 : \mu \neq 800$ hours, meaning the lamp's quality has changed.
2. **Test statistic**: t
3. **Two-tailed test**
4. **Significance level** $\alpha = 0.05$, so:

   $$-t_{1-\alpha/2} < t < t_{1-\alpha/2}$$

   which gives:

   $$-2.011 < t < 2.011$$

5. **Test statistic**:

   $$t = \frac{792 - 800}{55/\sqrt{50}} = -1.029$$

6. **Conclusion**: $t = -1.029$, which falls within the acceptance region of H_0. Therefore, at the 0.05 significance level, we fail to reject H_0, meaning the average lamp lifetime is still about 800 hours.

For normally distributed populations with known mean μ and standard deviation σ, the test statistic is:

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

Where H_0 is rejected if:

$$z \geq z_{0.5-\alpha}$$

This value of $z_{0.5-\alpha}$ can be obtained from the standard normal distribution table.

Example

The manufacturing process produces an average of 15.7 units per hour. The production results have a variance of 2.3. A new method is proposed to replace the old one if the average per hour produces at least 16 units. To determine whether the method should be replaced or not, the new method was tested 20 times, and the average result per hour was 16.9 units. The business owner is willing to take a 5% risk of using the new method if it produces an average greater than 16 units per hour. What should be the business owner's decision?

Solution:

1. Determine the hypothesis:

$H_0 : \mu = 16$

$H_1 : \mu > 16$

2. The test statistic: z

3. One-sided test

The significance level $\alpha = 0.05$, thus $z \geq z_{0.5-0.05}$ or $z \geq 1.64$

4. Test statistic calculation:

$$z = \frac{16.9-16}{\sqrt{\frac{2.3}{20}}} = 2.65$$

5. Conclusion: The value of $z_{hit} = 2.65$ is in the rejection region of H_0. Therefore, the null hypothesis H_0 is rejected, meaning that the new method can replace the old one.

For the second pair of hypotheses:

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu < \mu_0$$

If the standard deviation σ is not known, use the t-statistic:

The test statistic:

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

With degrees of freedom $df = n - 1$ and a significance level $(1 - \alpha)$.

Example

It is stated that by injecting a certain hormone to chickens, it will increase the average weight of their eggs by 4.5 grams. A random sample of 31 eggs from chickens that have been given the hormone injection shows an average weight of 4.9 grams and a standard deviation of $s = 0.8$ grams. Is it reasonable to accept the statement that the average increase in egg weight is at least 4.5 grams?

Solution

For the second example regarding hormone injection for egg weight increase:

Hypothesis:

$H_0 : \mu = 4.5$

$H_1 : \mu > 4.5$

Test statistic t

Significance level $\alpha = 0.01$, thus $t \geq t_{1-0.01}$ or $t \geq 2.46$

Test statistic calculation:

$t = \frac{4.9-4.5}{0.8/\sqrt{31}} = 2.78$

Conclusion: The calculated t-statistic $t_{hit} = 2.78$ falls into the rejection region for H_0. Therefore, the null hypothesis is rejected, meaning the average weight increase is at least 4.5 grams.