Mathematical Statistics and Data Analysis

Lecture 9: Hypothesis Testing

Lyu Ni

DaSE@ECNU
(lni@dase.ecnu.edu.cn)

December 11, 2019

Outlines

1. Hypothesis Testing
2. Hypothesis Testing with One Population Distribution $N(\mu, \sigma^2)$
   - Testing for μ
     - Testing for μ when σ^2 is known
     - Testing for μ when σ^2 is unknown
   - Testing for σ^2
3. Hypothesis Testing for Two Population Distributions $N(\mu_i, \sigma_i^2)$
   - Testing for $\mu_1 - \mu_2$
     - Testing for $\mu_1 - \mu_2$ when σ_1^2 and σ_2^2 are known
     - Testing for $\mu_1 - \mu_2$ when σ_1^2 and σ_2^2 are equal but unknown
     - Testing for $\mu_1 - \mu_2$ in paired sample
   - Testing for σ_1^2/σ_2^2
4. Hypothesis testing for different populations
   - Testing for $Exp(1/\theta)$

Reading Material

Textbook:

- Rice: Chapter 9;
- Mao: Chapter 7.1, 7.2, 7.3, 7.4, 6.6;

Hypothesis Testing

Example: The Lady Tasting Tea

- Lady Ottoline claimed that she was able to point out that the server had poured milk first or tea first. This means she could distinguish
  - TM: Tea first and then Milk;
  - MT: Milk first and then Tea.
- There is a hypothesis:

$$H: \text{This lady was not able to distinguish TM and MT.}$$

- The experiment was designed as follows:
  - Prepared 8 cups: 4 cups for TM and 4 cups for MT;
- Result: Lady Ottoline identified 8 out of 8 correctly.
- What can you conclude from this experiment?

Example: The Lady Tasting Tea (Con'd)

- Fisher's Idea:
  - Suppose the hypothesis is correct. The probability that Lady Ottoline correctly identified 8 out of 8 is

$$\binom{8}{4}^{-1} = \frac{1}{70} \approx 0.014$$

    which is a small probability.
  - A small probability event is considered to an event that cannot be actually occurred in an experiment.
  - However, this small probability event occurred.
  - This means the hypothesis is **not correct** and we need to **reject the hypothesis**.
  - Therefore, Lady Ottoline was deemed to be able to distinguish TM and MT.

Example: Normal Distribution

- A plant casts a type of alloy.
- The alloy intensity is thought to be distributed as $N(\theta, 16)$, where θ is required to be not less than 110 Pa.
- To guarantee the alloy quality, the plant needs to examine whether the manufacturing process goes wrong, that is, the intensity of the alloy is less than 110 Pa.
- The plant randomly selects 25 pieces of alloy and measures their intensity: $x_1, x_2, \cdots, x_{25}$.
- The sample mean is $\bar{x} = 108.2$ Pa.
- **Problem: Does the manufacturing process go wrong?**

Example: Normal Distribution (Con'd)

Let's analyze this problem as follows:

- It is a **(statistical) hypothesis testing problem**.
  - For example, we are interested in the proposition whether the alloy intensity is less than 110 Pa?
  - It is not a parameter estimation problem.
  - We need to make a decision, that is, the answer is "Yes" or "No".
- Define the hypothesis.
  - For example, the involved parameter spaces are, respectively,

$$\Theta_0 = \{\theta : \theta \geq 110\}, \quad \Theta_1 = \{\theta : \theta < 110\}.$$

  - If the hypothesis is correct, $\theta \in \Theta_0$; otherwise, $\theta \in \Theta_1$.

Example: Normal Distribution (Con'd)

Let's analyze this problem as follows:

- It is a **(statistical) hypothesis testing problem**.
    - For example, we are interested in the proposition whether the alloy intensity is less than 110 Pa?
    - It is not a parameter estimation problem.
    - We need to make a decision, that is, the answer is "Yes" or "No".
- Define the hypothesis.
    - For example, the involved parameter spaces are, respectively,

$$\Theta_0 = \{\theta : \theta \geq 110\}, \quad \Theta_1 = \{\theta : \theta < 110\}.$$

    - If the hypothesis is correct, $\theta \in \Theta_0$; otherwise, $\theta \in \Theta_1$.

Hypothesis Testing

Example: Normal Distribution (Con'd)

- Conduct a test via a statistic.
  - For example, $x_1, x_2, \cdots, x_{25} \overset{\text{iid}}{\sim} N(\theta, 16)$ and $\bar{x} = 108.2$;
  - The sample mean $\bar{x}$ is a reasonable statistic since $\bar{x}$ is a complete and sufficient statistic for θ.
  - It is known that $\bar{x} \sim N(\theta, 16/25)$.
  - If $\bar{x}$ is smaller, θ is thought to be smaller, and thus we are more likely to reject the null hypothesis: $H_0 : \theta \geq 110$.
  - Our decision is to reject H_0 if $\bar{x} \leq c$, where c is a constant.
  - Since the sample is random, the decision may be wrong.
  - We would like to minimize the probability that we reject H_0 when H_0 is true.

Example: Normal Distribution (Con'd)

- Conduct a test via a statistic.
  - Under H_0, the probability that we reject H_0 is

$$\begin{aligned} P(\bar{x} \leq c | \theta \geq 110) &= P\left(\frac{\bar{x}-\theta}{\sqrt{16/25}} \leq \frac{c-\theta}{\sqrt{16/25}} \middle| \theta \geq 110 \right) \\ &= \Phi(1.25 * (c-\theta) | \theta \geq 110) \\ &\leq \Phi(1.25 * (c-110)) \end{aligned}$$

  - Let $\Phi(1.25 * (c-110)) = 0.05$. Then, we obtain $c = \Phi^{-1}(0.05) * 0.8 + 110 \approx 108.684$.
- Make a conclusion: we will reject H_0 since $\bar{x} = 108.2 < 108.684$.

Hypothesis Testing

Remark

- this is a **parametric hypothesis testing** if the parameters are involved in the hypotheses.
- otherwise it is a **nonparametric hypothesis testing**.
    - For example, we would like to test a hypothesis that **the population is a normal distribution**.

Hypothesis Testing

Basic Step 1: Construct hypotheses

Suppose that there is a parametric distribution $\{F(x,\theta), \theta \in \Theta\}$ and the sample is $x_1, x_2, \cdots, x_n$, where Θ is a parameter space.

- Suppose that $\Theta_0 \in \Theta$ and $\Theta_0 \neq \emptyset$. The **null hypothesis** is defined as a proposition $H_0 : \theta \in \Theta_0$.
- Suppose that $\Theta_1 \in \Theta$ and $\Theta_1 \cap \Theta_0 = \emptyset$.
  - The most common choice: $\Theta_1 = \Theta - \Theta_0$.

  The **alternative hypothesis** is defined as a proposition $H_1 : \theta \in \Theta_1$.

Thus, we are interested in a pair of hypotheses that

$$H_0 : \theta \in \Theta_0 \quad \text{vs} \quad H_1 : \theta \in \Theta_1$$

Basic Step 1: Construct hypotheses (Con'd)

- **Simple & Composite**:
  - If $\Theta_0 = \{\theta : \theta = \theta_0\}$, a null hypothesis is called a **simple** null hypothesis; otherwise, a null hypothesis is called a **composite** null hypothesis.
  - The simple null hypothesis could be written as

$$H_0 : \theta = \theta_0$$

- **Two-sided & One-sided**: When $H_0 : \theta = \theta_0$,
  - H_0 vs $H_1' : \theta \neq \theta_0$ is called **two-sided** hypothesis.
  - H_0 vs $H_1'' : \theta < \theta_0$ and H_0 vs $H_1''' : \theta > \theta_0$ are called **one-sided** hypothesis.

Hypothesis Testing

Basic Step 2: Find a test statistic and give a rejection region

- Given the sample $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$, the possible outcomes of a test:
  - Reject the null hypothesis H_0;
  - Fail to reject the null hypothesis H_0;
- The sample space is divided into two disjoint parts:
  - The **rejection region** W: Reject H_0 if the sample
    $$\boldsymbol{x} = (x_1, x_2, \cdots, x_n) \in W;$$
  - The **acceptance region** $\overline{W}$: Fail to reject H_0 if
    $$\boldsymbol{x} = (x_1, x_2, \cdots, x_n) \in \overline{W};$$
- Find a **test statistic** and give a rejection region.

Basic Step 3: Choose a significance level

Since the sample is random, we may make a right or wrong decision. Two types of error are defined as follows:

- **Type I error**: $\boldsymbol{x} \in W$ when $\theta \in \Theta_0$;
- **Type II error**: $\boldsymbol{x} \in \overline{W}$ when $\theta \in \Theta_1$;

We give two notations for the probabilities:

- The probability of type I error: $\alpha = P\{\boldsymbol{x} \in W | H_0\}$;
- The probability of type II error: $\beta = P\{\boldsymbol{x} \in \overline{W} | H_1\}$;

Basic Step 3: Choose a significance level (Con'd)

Definition

Suppose that there is a testing problem

$$H_0 : \theta \in \Theta_0 \quad \text{vs} \quad H_1 : \theta \in \Theta_1$$

and the rejection region is W. The **power function** is defined as the probability that $\boldsymbol{x} \in W$, that is,

$$g(\theta) = P_\theta(\boldsymbol{x} \in W), \theta \in \Theta = \Theta_0 \cup \Theta_1.$$

Thus, the power function is defined on the parameter space Θ:

- $g(\theta) = \alpha = \alpha(\theta), \theta \in \Theta_0$;
- $g(\theta) = 1 - \beta = 1 - \beta(\theta), \theta \in \Theta_1$;

Hypothesis Testing

Basic Step 3: Choose a significance level (Con'd)

Obviously, α and β is a function of θ, that is

$$\begin{cases} \alpha(\theta) = g(\theta), & \theta \in \Theta_0, \\ \beta(\theta) = 1 - g(\theta), & \theta \in \Theta_1. \end{cases}$$

Revisit example: Normal Distribution

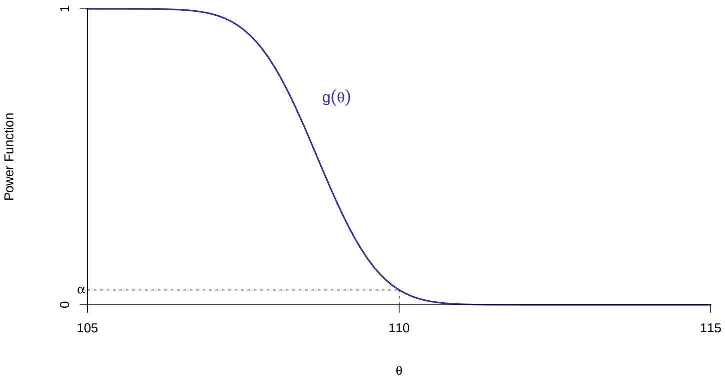
The rejection region is defined as $W = \{\bar{x} \leq c\}$. The power function is

$$\begin{aligned} g(\theta) &= P_\theta(\bar{x} \leq c) = P\left(\frac{\bar{x} - \theta}{4/5} \leq \frac{c - \theta}{4/5}\right) \\ &= \Phi\left(\frac{c - \theta}{4/5}\right). \end{aligned}$$

Hypothesis Testing

Revisit example: Normal Distribution (Con'd)

The power function is decreasing in θ shown as follows:



Revisit example: Normal Distribution (Con'd)

The probability of Type I error and Type II error are defined as follows:

$$\begin{aligned}\alpha(\theta) &= \Phi\left(\frac{c-\theta}{4/5}\right), \theta \in \Theta_0 \\ \beta(\theta) &= 1-\Phi\left(\frac{c-\theta}{4/5}\right), \theta \in \Theta_1\end{aligned}$$

Remark

- $\alpha\downarrow \Rightarrow c\downarrow \Rightarrow \beta\uparrow$;
- $\beta\downarrow \Rightarrow c\uparrow \Rightarrow \alpha\uparrow$
- There is a tradeoff between α and β.

Hypothesis Testing

Definition

Consider a testing problem

$$H_0 : \theta \in \Theta_0 \quad \text{vs} \quad H_1 : \theta \in \Theta_1.$$

If a test satisfies

$$g(\theta) \leq \alpha$$

for every $\theta \in \Theta_0$, then the test is said to be a **significance test of (significance) level α**

Thumb rule

- $\alpha = 0.05$ is the most common choice;
- Sometimes, $\alpha = 0.1$ or $\alpha = 0.01$ is also useful.

Basic step 4: Give a rejection region

After the significance level α is determined, we can give a rejection region W for the test. For example,

- Given a significance level α, for $\theta \geq 110$,

$$g(\theta) = \Phi\left(\frac{5(c-\theta)}{4}\right) \leq \alpha.$$

- $g(\theta)$ is a decreasing function of θ.
- Just let $g(110) = \alpha$, that is,

$$\Phi\left(\frac{5(c-110)}{4}\right) = \alpha$$

- The rejection region is $W = \{\bar{x} \leq 110 + 0.8 * \Phi^{-1}(\alpha)\}$.

Basic step 5: Make a decision

After the rejection region W is determined, we can make a decision. For example,

- When $\bar{x} \leq 110 + 0.8 * \Phi^{-1}(\alpha)$, we reject H_0;
- When $\bar{x} > 110 + 0.8 * \Phi^{-1}(\alpha)$, we fail to reject H_0.

Summary

Find a significance test in following steps:

- Construct a statistical hypothesis H_0 vs H_1;
- Find an appropriate test statistic $T(\boldsymbol{x})$ of which the distribution is known under H_0;
- Given a significance level α, derive the rejection region W;
- Calculate $T(\boldsymbol{x})$ from the sample $\boldsymbol{x} = (x_1, \cdots, x_n)$ and make a decision by judging whether $T(\boldsymbol{x}) \in W$.

p value

By determining different significance levels, we may make different conclusion:

Significance level α	Rejection Region W	Conclusion
$\alpha = 0.1$	$\bar{x} \leq 108.975$	Reject H_0
$\alpha = 0.05$	$\bar{x} \leq 108.684$	Reject H_0
$\alpha = 0.025$	$\bar{x} \leq 108.432$	Reject H_0
$\alpha = 0.01$	$\bar{x} \leq 108.139$	Not reject H_0
$\alpha = 0.005$	$\bar{x} \leq 107.939$	Not reject H_0

- If $\alpha = 0.05$ is chosen, H_0 could be rejected;
- If $\alpha = 0.01$ is chosen, H_0 could not be rejected.

Hypothesis Testing

p value

From a different perspective, when $\theta = 110$, the test statistic

$$u = \frac{\bar{x} - \theta}{4/5} \sim N(0, 1).$$

It is calculated that $u_0 = \theta + 0.8 * \Phi^{-1}(\alpha) = -2.25$ from the sample if the significance level $\alpha = 0.05$. The probability is

$$P(u \leq u_0) = P(u \leq -2.25) = \Phi(-2.25) \approx 0.0122$$

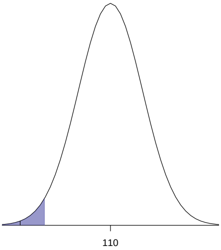
- When $\alpha \geq 0.0122$ and $u_\alpha \geq -2.25$, H_0 could be rejected since the rejection region is $W = \{u \leq u_\alpha\}$;
- When $\alpha < 0.0122$ and $u_\alpha < -2.25$, H_0 could be not rejected since the rejection region is $W = \{u \leq u_\alpha\}$;

Hypothesis Testing

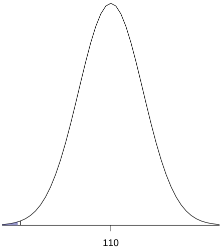
Definition

p **value** is defined as the probability under the null hypothesis of a result as or more extreme than that actually observed.

- If $\alpha \geq p$, then reject H_0 at the significance level α;
- If $\alpha < p$, then do not reject H_0 at the significance level α.



$\alpha = 0.05$



$\alpha = 0.01$

Testing for μ

- Suppose that $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$ is a sample from a normal distribution $N(\mu, \sigma^2)$.
- We are interested in making decisions about the population mean μ.
- Consider the following hypothesis testing problem:

$$\begin{array}{ll} \mathrm{I} & H_0 : \mu \leq \mu_0 \text{ vs } H_1 : \mu > \mu_0, \\ \mathrm{II} & H_0 : \mu \geq \mu_0 \text{ vs } H_1 : \mu < \mu_0, \\ \mathrm{III} & H_0 : \mu = \mu_0 \text{ vs } H_1 : \mu \neq \mu_0, \end{array}$$

  where μ_0 is a constant.
- Different test statistics will be adopted corresponding to whether σ^2 is known or not.

Testing for μ when σ^2 is known

Consider the testing problem I

$$H_0 : \mu \leq \mu_0 \text{ vs } H_1 : \mu > \mu_0.$$

Since $\bar{x}$ is an reasonable estimate of μ and is distributed as $N(\mu, \sigma^2/n)$, the test statistic

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

is proper. A natural idea is that

- if the sample mean $\bar{x}$ is not larger than the pre-determined constant μ_0, we tend to fail to reject the null hypothesis;
- if the sample mean $\bar{x}$ exceeds μ_0, we tend to reject the null hypothesis.

Testing for μ when σ^2 is known

Method 1: Derive a rejection region

Thus, the rejection region is

$$W_I = \{\boldsymbol{x} : z \geq c\}$$

where c is a critical value.
At the significance level α, the critical value c satisfies

$$P_{\mu_0}(z \geq c) = \alpha.$$

When $\mu = \mu_0$, $z \sim N(0, 1)$ and then $c = z_{1-\alpha}$, where z_α is the α quantile of a standard normal distribution.
Thus, the rejection region is

$$W_I = \{\boldsymbol{x} : z \geq z_{1-\alpha}\}.$$

Testing for μ when σ^2 is known

Method 1: Derive a rejection region

Reason: The power function is

$$\begin{aligned} g(\mu) &= P_\mu(\boldsymbol{x} \in W_I) = P_\mu(z \geq z_{1-\alpha}) \\ &= P_\mu\left(\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \geq z_{1-\alpha}\right) = P_\mu\left(\frac{\bar{x}-\mu+\mu-\mu_0}{\sigma/\sqrt{n}} \geq z_{1-\alpha}\right) \\ &= P_\mu\left(\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \geq -\frac{\mu-\mu_0}{\sigma/\sqrt{n}} + z_{1-\alpha}\right) \\ &= 1 - \Phi\left(-\frac{\mu-\mu_0}{\sigma/\sqrt{n}} + z_{1-\alpha}\right). \end{aligned}$$

Thus, the power function $g(\mu)$ is an increasing function of μ. The inequality $g(\mu) \leq \alpha$ holds for all $\mu \leq \mu_0$ as long as $g(\mu_0) \leq \alpha$. So, this is a significance test of level α.

Testing for μ when σ^2 is known

Method 2: Calculate the p-value

Given the sample $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$, it can be calculated that

$$z_0 = \frac{\sqrt{n}(\bar{x} - \mu_0)}{\sigma}$$

where $\bar{x}$ is the sample mean. Let

$$p_I = P(z \geq z_0) = I - \Phi(z_0)$$

where z is a standard normal random variable. Then, $z_{1-p} = z_0$.

Remark:

- When $p > \alpha$, $z_0 = z_{1-p} < z_{1-\alpha}$. We will not reject H_0 since $\boldsymbol{x} \notin W_I$.
- When $p \leq \alpha$, $z_0 = z_{1-p} \geq z_{1-\alpha}$. We will reject H_0 since $\boldsymbol{x} \in W_I$.

Testing for μ when σ^2 is known

Consider the testing problem:

$$\mathrm{IV} \quad H_0 : \mu = \mu_0 \text{ vs } H_1 : \mu > \mu_0$$

Remarks:

- The test statistic $z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$ can still be used;
- The rejection region is $W_{\mathrm{IV}} = \{\boldsymbol{x} : z \geq z_{1-\alpha}\}$;
- This is also a significance test of level α;
- The p value is also $p_{\mathrm{IV}} = 1 - \Phi(z_0)$, where $z_0 = \frac{\sqrt{n}(\bar{x}-\mu_0)}{\sigma}$ given the sample.

Testing for μ when σ^2 is known

Consider the testing problem:

$$\text{II} \quad H_0 : \mu \geq \mu_0 \text{ vs } H_1 : \mu < \mu_0$$

or

$$\text{V} \quad H_0 : \mu = \mu_0 \text{ vs } H_1 : \mu < \mu_0$$

- The test statistic is $z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$ and the rejection region is $W_{\mathrm{I}} = W_{\mathrm{V}} = \{\boldsymbol{x} : z \leq z_\alpha\}$;
- The p value is also $p_{\mathrm{II}} = p_{\mathrm{V}} = \Phi(z_0)$, where $z_0 = \frac{\sqrt{n}(\bar{x}-\mu_0)}{\sigma}$ given the sample.

Testing for μ when σ^2 is known

Consider the testing problem:

$$\text{III} \quad H_0 : \mu = \mu_0 \text{ vs } H_1 : \mu \neq \mu_0$$

First, the test statistic is $z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$ and the rejection region is

$$W_{\text{III}} = \{|z| \geq c\},$$

where c is a positive critical value.
It is known that $z \sim N(0, 1)$ under $H_0 : \mu = \mu_0$. Given the significance level α, $c = z_{1-\alpha/2}$ since $P_{\mu_0}(|z| \geq c) = \alpha$. Thus, the rejection region is

$$W_{\text{III}} = \{|z| \geq z_{1-\alpha/2}\}.$$

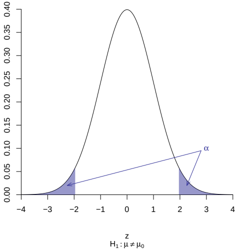
Testing for μ when σ^2 is known

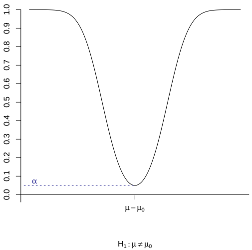
Next, we will introduce how to calculate the p-value in a two-sided hypothesis testing.
Let

$$p_{\text{III}} = P(|z| \geq |z_0|) = 2(1 - \Phi(z_0)).$$

Here $|z_0| = z_{1-p/2}$. Note that we need to use $|z_0|$ when we calculate p value since z_0 may be positive or negative.

- When $p \leq \alpha$, $z_{1-\alpha/2} \leq |z_0|$. $\Rightarrow$ Reject H_0;
- When $p > \alpha$, $z_{1-\alpha/2} > |z_0|$. $\Rightarrow$ Fail to reject H_0;





Testing for μ when σ^2 is known

Example

- Suppose that we obtain the signal at Area B and it is distributed as $N(\mu, 0.2^2)$.
- The population mean μ is the true signal.
- The same signal is sent from Area A to Area B five times.
- The obtained signals at Area B: 8.05, 8.15, 8.2, 8.1, 8.25.

We guess the true signal is 8. Is this guess right?

Testing for μ when σ^2 is known

Example (Con'd)

Solution: This is a hypothesis testing problem. Suppose that X is distributed as $N(\mu, 0.2^2)$. Construct a hypothesis. The null hypothesis and the alternative hypothesis are, respectively,

$$H_0 : \mu = 8 \quad \text{vs} \quad H_1 : \mu \neq 8.$$

This is two-sided hypothesis testing problem.
Here we will provide two solution at significance level $\alpha = 0.05$.

- One is how to construct the rejection region.
- The other is how to calculate p values.

Testing for μ when σ^2 is known

Example (Con'd)

Solution:

- Method One: The rejection region is $\{|z| \geq z_{1-\alpha/2}\}$. Since the significance level is $\alpha = 0.05$, we can know $z_{1-\alpha/2} = 1.96$. Then,

$$\bar{x} = 8.15 \quad \text{and} \quad z = 1.677 < 1.96$$

- Method Two: The p value is

$$p = 2(1 - \Phi(1.677)) = 0.0935 > 0.05 = \alpha$$

Then, we cannot reject the null hypothesis. Therefore, we think the true signal is 8.

Testing for μ when σ^2 is known

Example (Con'd)

z_p	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936

Testing for μ when σ^2 is unknown

Consider the testing problem I:

$$H_0 : \mu \leq \mu_0 \text{ vs } H_1 : \mu > \mu_0.$$

The test statistic is

$$u = \frac{\sqrt{n}(\bar{x} - \mu_0)}{\sigma}.$$

Since the population variance σ^2 is unknown, a natural idea is to replace σ with its reasonable estimate s. Then, the test statistic is

$$t = \frac{\sqrt{n}(\bar{x} - \mu_0)}{s}.$$

When $\mu = \mu_0$, it can be proved that $t \sim t(n-1)$.

Testing for μ when σ^2 is unknown

Method 1: Derive a rejection region

The rejection region is

$$W_{\mathrm{I}} = \{t \geq t_{1-\alpha}(n-1)\}.$$

Method 2: Calculated the p-value

The p-value is

$$p_{\mathrm{I}} = P(t \geq t_0)$$

where t is distributed as $t(n-1)$ and $t_0 = \frac{\sqrt{n}(\bar{x}-\mu_0)}{s}$ with the sample mean $\bar{x}$ and the sample standard deviation s.

Testing for μ when σ^2 is unknown

Consider the testing problem II:

$$H_0 : \mu \geq \mu_0 \text{ vs } H_1 : \mu < \mu_0.$$

Method 1: Derive a rejection region

The test statistic is $t = \frac{\sqrt{n}(\bar{x}-\mu_0)}{s}$ and the rejection region is

$$W_{\text{II}} = \{t \leq t_\alpha(n-1)\}.$$

Method 2: Calculated the p-value

The p-value is

$$p_{\text{II}} = P(t \leq t_0)$$

where t is distributed as $t(n-1)$ and $t_0 = \frac{\sqrt{n}(\bar{x}-\mu_0)}{s}$ with the sample mean $\bar{x}$ and the sample standard deviation s.

Testing for μ when σ^2 is unknown

Consider the testing problem III:

$$H_0 : \mu = \mu_0 \text{ vs } H_1 : \mu \neq \mu_0.$$

Method 1: Derive a rejection region

The test statistic is $t = \frac{\sqrt{n}(\bar{x}-\mu_0)}{s}$ and the rejection region is

$$W_{\mathrm{III}} = \{|t| \geq t_{1-\alpha/2}(n-1)\}.$$

Method 2: Calculated the p-value

The p-value is

$$p_{\mathrm{III}} = P(|t| \geq |t_0|)$$

where t is distributed as $t(n-1)$ and $t_0 = \frac{\sqrt{n}(\bar{x}-\mu_0)}{s}$ with the sample mean $\bar{x}$ and the sample standard deviation s.

Testing for μ when σ^2 is unknown

Example

A plant manufactures a type of aluminum steel, which the length is distributed as a normal distribution. Suppose that the population mean is 240 cm. Now five products are randomly selected and their lengths are respectively, 239.7,239.6,239,240,239.2. Is this type of aluminum steel qualified?

Solution: This is a two-sided hypothesis testing problem. The hypotheses are

$$H_0 : \mu = 240 \quad \text{vs} \quad H_1 : \mu \neq 240$$

Since σ is unknown, we use t test. Consider that the significant level is $\alpha = 0.05$. It is calculated that

$$\bar{x} = 239.5 \quad \text{and} \quad s = 0.4.$$

Testing for μ when σ^2 is unknown

Example (Con'd)

- Method 1: The rejection region is $\{|t| \geq t_{1-\alpha/2}(n-1)\}$. It is known that $t_{0.975}(4) = 2.776$. Then,

  $$t = \frac{\sqrt{5}|\bar{x} - \mu_0|}{s} = 2.975 > 2.776$$

  Then, the test statistic falls in the rejection region.

- Method 2: The p-value is calculated to be

  $$p = P(|t| \geq t_0) = P(|t| \geq 2.975) = 0.0491$$

  Then, p value is smaller than $\alpha = 0.05$.

Thus, the null hypothesis is rejected. Therefore, we do not think the aluminum steel is qualified.

Testing for μ

Method		σ known Z Test	σ unknown T Test
Test Statistic		$z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$	$t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$
H_0	H_1	Rejection Region	
$\mu \le \mu_0$	$\mu > \mu_0$	$\{z \ge z_{1-\alpha}\}$	$\{t \ge t_{1-\alpha}(n-1)\}$
$\mu \ge \mu_0$	$\mu < \mu_0$	$\{z \le z_{\alpha}\}$	$\{t \le t_{\alpha}(n-1)\}$
$\mu = \mu_0$	$\mu \ne \mu_0$	$\{\lvert z\rvert \ge z_{1-\alpha/2}\}$	$\{\lvert t\rvert \ge t_{1-\alpha/2}(n-1)\}$
Sample		$z_0 = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$	$t_0 = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$
H_0	H_1	p value	
$\mu \le \mu_0$	$\mu > \mu_0$	$1-\Phi(z_0)$	$P(t \ge t_0)$
$\mu \ge \mu_0$	$\mu < \mu_0$	$\Phi(z_0)$	$P(t \le t_0)$
$\mu = \mu_0$	$\mu \ne \mu_0$	$2(1-\Phi(\lvert z_0\rvert))$	$P(\lvert t\rvert \ge \lvert t_0\rvert)$

Testing for σ^2

Suppose that $x_1, x_2, \cdots, x_n$ is a sample from a normal distribution $N(\mu, \sigma^2)$. Consider the following testing problems:

$$\begin{array}{lccc} \text{I} & H_0 : \sigma^2 \leq \sigma_0^2 & \text{vs} & H_1 : \sigma^2 > \sigma_0^2, \\ \text{II} & H_0 : \sigma^2 \geq \sigma_0^2 & \text{vs} & H_1 : \sigma^2 < \sigma_0^2, \\ \text{III} & H_0 : \sigma^2 = \sigma_0^2 & \text{vs} & H_1 : \sigma^2 \neq \sigma_0^2, \end{array}$$

where σ_0^2 is a known constant. Suppose that the population mean μ is unknown.

The test statistic is

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}.$$

When $\sigma = \sigma_0^2$, the test statistic is distributed as $\chi^2(n-1)$.

Testing for σ^2

Suppose that $x_1, x_2, \cdots, x_n$ is a sample from a normal distribution $N(\mu, \sigma^2)$. Consider the following testing problems:

$$\begin{array}{llcl} \text{I} & H_0 : \sigma^2 \leq \sigma_0^2 & \text{vs} & H_1 : \sigma^2 > \sigma_0^2, \\ \text{II} & H_0 : \sigma^2 \geq \sigma_0^2 & \text{vs} & H_1 : \sigma^2 < \sigma_0^2, \\ \text{III} & H_0 : \sigma^2 = \sigma_0^2 & \text{vs} & H_1 : \sigma^2 \neq \sigma_0^2, \end{array}$$

where σ_0^2 is a known constant. Suppose that the population mean μ is unknown.
The test statistic is

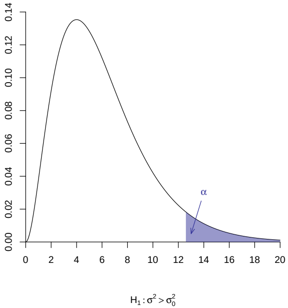
$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}.$$

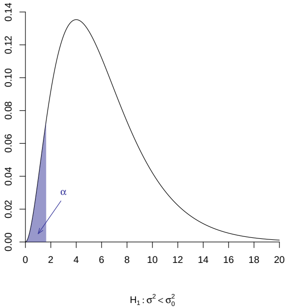
When $\sigma = \sigma_0^2$, the test statistic is distributed as $\chi^2(n-1)$.

Testing for σ^2

For the one-sided testing problems, the rejection regions are respectively

$$W_{\mathrm{I}} = \{\chi^2 \geq \chi^2_{1-\alpha}(n-1)\} \quad W_{\mathrm{II}} = \{\chi^2 \leq \chi^2_{1-\alpha}(n-1)\}$$





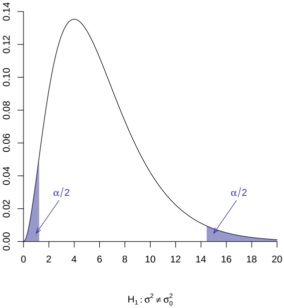
The p values are respectively

$$p_{\mathrm{I}} = P(\chi^2 \geq \chi^2_0) \quad p_{\mathrm{II}} = P(\chi^2 \leq \chi^2_0)$$

Testing for σ^2

For the two-sided testing problem, the rejection region is

$$W_{\text{III}} = \{\chi^2 \leq \chi^2_{\alpha/2}(n-1) \text{ or } \chi^2 \geq \chi^2_{1-\alpha/2}(n-1)\}$$



The p values are respectively

$$p_{\text{III}} = 2\min\{P(\chi^2 \geq \chi_0^2), P(\chi^2 \leq \chi_0^2)\}$$

Testing for σ^2

Example

Suppose that the weight of a steel plate is distributed as a normal distribution. If the variance of the weights of steel plates is not larger than 0.016, the steel plates are considered as qualified. Now a batch of steal plates are produced and 25 pieces are randomly selected and the sample variance is $s^2 = 0.025$. Is this batch of the steal plate is qualified?

Solution: This is one-sided hypothesis testing problem. The hypotheses are

$$H_0 : \sigma^2 \leq 0.016 \quad \text{vs} \quad H_0 : \sigma^2 > 0.016$$

Consider the significance level as $\alpha = 0.05$. It is known that the critical value is $\chi^2_{0.95}(24) = 36.415$

Testing for σ^2

Example (Con'd)

- The rejection region is

$$W = \{\chi^2 \geq \chi^2_{0.95}(24)\} = \{\chi^2 \geq 36.415\}.$$

  The test statistic is calculated as

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{(25-1) \times 0.025}{0.016} = 37.5 > 36.415$$

- The p value is calculated as

$$p = P(\chi^2 \geq \chi_0^2) = P(\chi^2 \geq 37.5) = 0.0390 < 0.05$$

Then, we will reject the null hypothesis. Therefore, we do not think this batch of the steel plate is qualified.

Testing for $\mu_1 - \mu_2$

Example

- Two methods, A and B, were used in a determination of the latent heat of fusion of ice.
- The data shows the change in total heat from ice at -0.72°C to water 0°C in calories per gram of mass.
- Method A (13 Obs): 79.98, 80.04, 80.02, 80.04, 80.03, 80.03, 80.04, 79.97, 80.05, 80.03, 80.02, 80.00, 80.02
- Method B (8 Obs): 80.02, 79.94, 79.98, 79.97, 79.97, 80.03, 79.95, 79.97
- Are these two methods different?

Testing for $\mu_1 - \mu_2$

Introduction

- The first sample: $x_1, x_2, \cdots, x_m$ is a sample from $N(\mu_1, \sigma_1^2)$;
- The second sample: $y_1, y_2, \cdots, y_n$ is a sample from $N(\mu_2, \sigma_2^2)$;
- Two samples are independent;
- Consider the following hypothesis testing problems:

$$
\begin{array}{llcl}
\mathrm{I} & H_0: \mu_1 - \mu_2 \leq 0 & \text{vs} & H_1: \mu_1 - \mu_2 > 0 \\
\mathrm{II} & H_0: \mu_1 - \mu_2 \geq 0 & \text{vs} & H_1: \mu_1 - \mu_2 < 0 \\
\mathrm{III} & H_0: \mu_1 - \mu_2 = 0 & \text{vs} & H_1: \mu_1 - \mu_2 \neq 0
\end{array}
$$

Testing for $\mu_1 - \mu_2$ when σ_i^2's are known

Method One: Derive the rejection region

- $\bar{x} - \bar{y}$ is an estimate of $\mu_1 - \mu_2$;
- The distribution of $\bar{x} - \bar{y}$ is

$$N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}\right);$$

- The test statistic is

$$z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} \overset{\mu_1 = \mu_2}{\sim} N(0, 1);$$

- In the example, the rejection region is

$$W_{\text{III}} = \{|z| \geq z_{1-\alpha/2}\}.$$

Testing for $\mu_1 - \mu_2$ when σ_i^2's are known

Method Two: Calculate the p value

- Calculate the test statistic

$$z_0 = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}},$$

  where $\bar{x}$ and $\bar{y}$ are respectively two sample means.

- In the example, the p value is

$$\begin{aligned} p &= P(|z| \geq |z_0|) \\ &= 2(1 - \Phi(|z_0|)) \end{aligned}$$

Testing for $\mu_1 - \mu_2$ when σ_i^2's are known

Example

- It is supposed that $\sigma_1 = 0.025$ and $\sigma_2 = 0.03$.
- Let the significance level $\alpha = 0.05$
- The hypotheses are

$$H_0 : \mu_1 - \mu_2 = 0 \quad \text{vs} \quad \mu_1 - \mu_2 \neq 0$$

- The test statistic is

$$z_0 = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} = 3.316 > 1.96 = z_{1-\alpha/2}$$

- The p value is

$$p = 2(1 - \Phi(z_0)) = 0.000913 < 0.05$$

Testing for $\mu_1 - \mu_2$ when $\sigma_1^2 = \sigma_2^2$

Method One: Derive the rejection region

When $\sigma_1^2 = \sigma_2^2 = \sigma^2$ and σ^2 is unknown,

- First, the distribution of $\bar{x} - \bar{y}$ is

$$N\left(\mu_1 - \mu_2, \left(\frac{1}{m} + \frac{1}{n}\right)\sigma^2\right).$$

- Second, we have

$$\frac{1}{\sigma^2}\sum_{i=1}^{m}(x_i - \bar{x})^2 \sim \chi^2(m-1) \quad \frac{1}{\sigma^2}\sum_{j=1}^{n}(y_j - \bar{y})^2 \sim \chi^2(n-1).$$

Then,

$$\frac{1}{\sigma^2}\left(\sum_{i=1}^{m}(x_i - \bar{x})^2 + \sum_{j=1}^{n}(y_j - \bar{x})^2\right) \sim \chi^2(m + n - 2).$$

Testing for $\mu_1 - \mu_2$ when $\sigma_1^2 = \sigma_2^2$

Method One: Derive the rejection region (Con'd)

- Let $s_w^2 = \frac{1}{m+n-2}\left(\sum_{i=1}^{m}(x_i - \bar{x})^2 + \sum_{j=1}^{n}(y_j - \bar{x})^2\right) = \frac{1}{m+n-2}\left((m-1)s_x^2 + (n-1)s_y^2\right)$.
- Then, we have $t = \frac{(\bar{x}-\bar{y})-(\mu_1-\mu_2)}{s_w\sqrt{\frac{1}{m}+\frac{1}{n}}} \sim t(m+n-2)$.
- The test statistic is

$$t = \frac{(\bar{x} - \bar{y})}{s_w\sqrt{\frac{1}{m} + \frac{1}{n}}}.$$

- The rejection region is

$$W = \{|t| > t_{1-\alpha/2}(m+n-2)\}.$$

Testing for $\mu_1 - \mu_2$ when $\sigma_1^2 = \sigma_2^2$

Method Two: Calculate the p value

When $\sigma_1^2 = \sigma_2^2 = \sigma^2$ and σ^2 is unknown,

- The p value is

$$p = P(|t| \geq |t_0|)$$

  where t is a random variable distributed as t distribution with degree of freedom $m+n-2$, that is $t \sim t(m+n-2)$, and t_0 is the specific value of the test statistic.

Testing for $\mu_1 - \mu_2$ when $\sigma_1^2 = \sigma_2^2$

Example

Consider the significance level $\alpha = 0.05$. It is calculated

- the sample means are $\bar{x} = 80.02$ and $\bar{y} = 79.98$;
- the sample standard deviations are $s_x = 0.024$ and $s_y = 0.031$;
- the pooled sample variance is

$$s_w^2 = \frac{1}{m+n-2}\left((m-1)s_x^2 + (n-1)s_y^2\right) = 0.0007253$$

  and $s_w = 0.027$.
- The test statistic is

$$t_0 = \frac{\bar{x} - \bar{y}}{s_w\sqrt{1/m + 1/n}} = 3.47 > 2.09 = t_{1-\alpha/2}$$

Example (Con'd)

- The p value is

$$p = P(|t| \geq |t_0|) = 0.00255 < 0.05 = \alpha$$

Then, we reject the null hypothesis. Thus, there is a significant difference between two methods.

Testing for $\mu_1 - \mu_2$ when $\sigma_1^2 = \sigma_2^2$

Example

A plant want to improve the abrasive resistance of a casting and nickel alloy is used to replace copper alloy as the casting material. There are two samples of castings as follows:

- Nickel alloy: 76.43, 76.21, 73.58, 69.69, 65.29, 70.83, 82.75, 72.34;
- Copper Alloy: 73.66, 64.27, 69.34, 71.37, 69.77, 68.12, 67.27, 68.07, 62.61;

Empirically, the abrasive resistance is distributed as a normal distribution and the variance remains. Is the abrasive resistance improved at the significance level $\alpha = 0.05$?

Testing for $\mu_1 - \mu_2$ when $\sigma_1^2 = \sigma_2^2$

Example (Con'd)

Suppose that x_is and y_is are respectively the samples of the abrasive resistance of nickel alloy castings and copper alloy castings, and $x_1, \cdots, x_8 \overset{\text{i.i.d.}}{\sim} N(\mu_1, \sigma^2)$ and $y_1, \cdots, y_8 \overset{\text{i.i.d.}}{\sim} N(\mu_2, \sigma^2)$. The hypotheses are

$$H_0 : \mu_1 = \mu_2 \text{ vs } H_1 : \mu_1 > \mu_2$$

The sample means are

$$\bar{x} = 73.39 \text{ and } \bar{y} = 68.276$$

and the sample standard deviations are

$$s_x = 5.234 \text{ and } s_y = 3.376.$$

Testing for $\mu_1 - \mu_2$ when $\sigma_1^2 = \sigma_2^2$

Example (Con'd)

The pooled sample standard deviation is

$$s_w = \sqrt{\frac{(m-1)s_x^2 + (n-1)s_y^2}{m+n-2}} = 4.3432$$

The test statistic is

$$t_0 = \frac{\bar{x} - \bar{y}}{s_w\sqrt{1/m + 1/n}} = 2.4334 > 1.7531 = t_{1-\alpha}$$

Then, we reject the null hypothesis. Thus, we think the abrasive resistance of the nickel alloy casting has been improved.

Testing for $\mu_1 - \mu_2$ when $\sigma_1^2 = \sigma_2^2$

Example (Con'd)

The p value is

$$p = P(t > t_0) = 0.01424 < 0.05 = \alpha$$

Then, we reject the null hypothesis. Thus, we think the abrasive resistance of the nickel alloy casting has been improved.

Testing for $\mu_1 - \mu_2$ in paired sample

Example

Scientists would like to compare the performance of two kinds of grain seeds. They select 10 different plots and each plot is divided into two equal parts, where they plant two different seeds. The grain yield per unit area in each plot is shown as follows:

Plot	I	II	III	IV	V	VI	VII	VIII	IX	X
x	23	35	29	42	39	29	37	34	35	28
y	30	39	35	40	38	34	36	33	41	31

Suppose that the grain yield per unit area is distributed as a normal distribution. Does there exist an difference between the average grain yield of two seeds at significance level $\alpha = 0.05$?

Testing for $\mu_1 - \mu_2$ in paired sample

Example (Con'd)

Solution Suppose that $x \sim N(\mu_1, \sigma_1^2)$ and $y \sim N(\mu_2, \sigma_2^2)$ and two sample are independent. It is accepted that $\sigma_1^2 = \sigma_2^2$. We could use two sample t test.

- Consider the hypotheses:

$$H_0 : \mu_1 = \mu_2 \text{ vs } H_1 : \mu_1 \neq \mu_2.$$

- The test statistic is

$$t_0 = \frac{\bar{x} - \bar{y}}{s_w/\sqrt{n/2}} = \frac{33.1 - 35.7}{4.87/\sqrt{5}} = -1.1937$$

  where $|t_0| < t_{1-\alpha/2} = 2.1009$. The sample falls in the acceptance region. We cannot reject the null hypothesis.

Testing for $\mu_1 - \mu_2$ in paired sample

Example (Con'd)

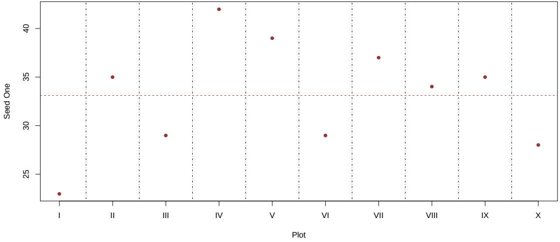
- The p value is

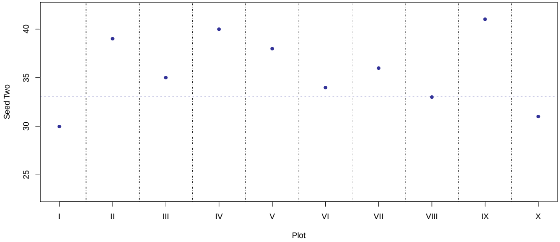
$$p = 2 \times P(|t| > |t_0|) = 0.2481$$

  which is larger than $\alpha = 0.05$. Then, we cannot reject the null hypothesis.

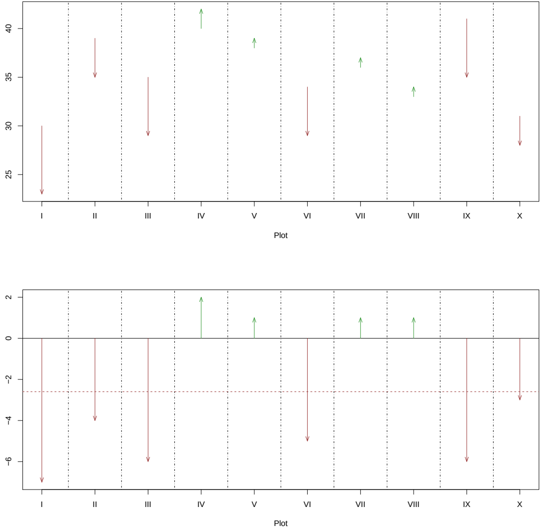
Thus, we do not think that the average grain yield of two seeds are different at significance level $\alpha = 0.05$.

Testing for $\mu_1 - \mu_2$ in paired sample





Testing for $\mu_1 - \mu_2$ in paired sample



Testing for $\mu_1 - \mu_2$ in paired sample

Introduction

- Let $d_i = x_i - y_i$;
- Under the assumption of normality, $d = x - y \sim N(\mu, \sigma_d^2)$, where $\mu = \mu_1 - \mu_2$ and $\sigma_d^2 = \sigma_1^2 + \sigma_2^2$.
- The hypotheses are

$$H_0 : \mu = 0 \text{ vs } H_1 : \mu_0 \neq 0.$$

- The test statistic is

$$t_0 = \frac{\bar{d}}{s_d/\sqrt{n}}$$

where $\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i$ and $s_d = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(d_i - \bar{d})^2}$.

Testing for $\mu_1 - \mu_2$ in paired sample

Introduction

- At the significance level $\alpha = 0.05$, the rejection region is

$$W_{\mathrm{I}} = \{|t_0| \leq t_{1-\alpha/2}(n-1)\}.$$

- The p value is

$$p_{\mathrm{I}} = P(|t| \geq |t_0|)$$

  where t is a random variable distributed as a t distribution $t(n-1)$.
- It is said to be the **paired t test**.

Testing for $\mu_1 - \mu_2$ in paired sample

Example (Con'd)

The data are calculated as follows:

Plot	I	II	III	IV	V	VI	VII	VIII	IX	X
x	23	35	29	42	39	29	37	34	35	28
y	30	39	35	40	38	34	36	33	41	31
$d = x - y$	-7	-4	-6	2	1	-5	1	1	-6	-3

The hypotheses are

$$H_0 : \mu = 0 \text{ vs } H_1 : \mu \neq 0$$

It is calculated that

$$n = 10, \quad \bar{d} = -2.6, \quad s_d = 3.5024$$

Testing for $\mu_1 - \mu_2$ in paired sample

Example (Con'd)

The test statistic is

$$t_0 = \frac{\bar{d}}{s_d/\sqrt{n}} = \frac{-2.6}{3.5024/\sqrt{10}} = -2.35$$

which is larger than $t_{1-\alpha/2}(n-1) = 2.26$. The sample falls in the rejection region.
The p value is

$$p = P(|t| \geq |t_0|) = 2P(|t| \geq 2.35) = 0.0435$$

which is smaller than $\alpha = 0.05$.
Then, we can reject the null hypothesis. Thus, we think there exists a significance difference between the average grain yield of two seeds at the significance level $\alpha = 0.05$.

Testing for σ_1^2/σ_2^2

Example

Two machine tools are used to manufacture a type of component. The diameter of a component is distributed as a normal distribution and the variance indicates the accuracy of a machine tool. We would like to compare the accuracy of two machine tools and we measure the diameter of some components from two machine tools. The data are shown as follows:

A (x)	16.2	16.4	15.8	15. 5	16.7	15.6	15.8	
B (y)	15.9	16.0	16.4	16.1	16.5	15.8	15.7	15.0

Testing for σ_1^2/σ_2^2

Introduction

- Suppose that $\boldsymbol{x} = (x_1, x_2, \cdots, x_m)$ is a sample of $N(\mu_1, \sigma_1^2)$ and $\boldsymbol{y} = (y_1, y_2, \cdots, y_n)$ is a sample of $N(\mu_2, \sigma_2^2)$.
- Consider the hypotheses are

$$\begin{array}{ll} \mathrm{I} & H_0: \sigma_1^2 \leq \sigma_2^2 \text{ vs } H_1: \sigma_1^2 > \sigma_2^2, \\ \mathrm{II} & H_0: \sigma_1^2 \geq \sigma_2^2 \text{ vs } H_1: \sigma_1^2 < \sigma_2^2, \\ \mathrm{III} & H_0: \sigma_1^2 = \sigma_2^2 \text{ vs } H_1: \sigma_1^2 \neq \sigma_2^2. \end{array}$$

  where μ_1 and μ_2 are unknown.
- The test statistic is

$$F_0 = s_x^2/s_y^2 \overset{\sigma_1^2=\sigma_2^2}{\sim} F(m-1, n-1)$$

  where s_x^2 and s_y^2 are two unbiased estimates.

Introduction (Con'd)

- The rejection region is

$$\begin{aligned} W_{\mathrm{I}} &= \{F_0 \geq F_{1-\alpha}(m-1, n-1)\} \\ W_{\mathrm{II}} &= \{F_0 \leq F_{\alpha}(m-1, n-1)\} \\ W_{\mathrm{III}} &= \{F_0 \leq F_{\alpha/2}(m-1, n-1) \\ &\quad\ \text{or } F_0 \geq F_{1-\alpha/2}(m-1, n-1)\} \end{aligned}$$

- The p value is

$$\begin{aligned} p_{\mathrm{I}} &= P(F \geq F_0) \\ p_{\mathrm{II}} &= P(F \leq F_0) \\ p_{\mathrm{III}} &= 2\min\{P(F \geq F_0), P(F \leq F_0)\} \end{aligned}$$

Testing for σ_1^2/σ_2^2

Example (Con'd)

- The hypotheses are

$$H_0 : \sigma_1^2 = \sigma_2^2 \text{ vs } H_1 : \sigma_1^2 \neq \sigma_2^2$$

- It is calculated that

$$m = 7, \quad n = 8, \quad s_x^2 = 0.1967, \quad s_y^2 = 0.2164.$$

- The test statistic is

$$F_0 = \frac{s_x^2}{s_y^2} = 0.9087$$

Testing for σ_1^2/σ_2^2

Example (Con'd)

- At the significance level $\alpha = 0.05$, the rejection region is

$$\begin{aligned} W &= \{F_0 \le F_{\frac{\alpha}{2}}(m-1, n-1) \text{ or } F_0 \ge F_{1-\frac{\alpha}{2}}(m-1, n-1)\} \\ &= \{F_0 \le 0.1756 \text{ or } F_0 \ge 5.1186\} \end{aligned}$$

- The p value is

$$\begin{aligned} p &= 2\min\{P(F \ge 0.9087), P(F \le 0.9087)\} \\ &= 2 \times \min\{0.4616, 0.5384\} = 0.9232 \end{aligned}$$

Then, we cannot reject the null hypothesis. Thus, we think there is no difference between the accuracy of two machine tools.

Testing for $Exp(1/\theta)$

Example

We would like to test whether the average lifetime of the component is not smaller than 6K hours. Suppose that the lifetime is distributed as an Exponential distribution $Exp(1/\theta)$. Five components are tested and the failure time is as follows:

$$395, 4094, 119, 11572, 6133.$$

The hypotheses are

$$H_0 : \theta \geq 6000 \text{ vs } H_1 : \theta < 6000.$$

It is also a hypothesis testing problem. It is calculated that

$$n = 5 \quad \text{and} \quad \bar{x} = 4462.6.$$

Testing for $Exp(1/\theta)$

Introduction

Suppose that $x_1, x_2, \cdots, x_n$ is a sample from an Exponential distribution $Exp(1/\theta)$, where θ is the population mean. Consider the hypothesis testing problem as

$$H_0 : \theta \geq \theta_0 \text{ vs } H_1 : \theta < \theta_0$$

The sufficient statistic for θ is $\bar{x}$. When $\theta = \theta_0$,

$$n\bar{x} = \sum_{i=1}^{n} x_i \sim Ga(n, 1/\theta_0).$$

From the property of Gamma distribution, we know

$$\chi_0^2 = \frac{2n\bar{x}}{\theta_0} \sim \chi^2(2n).$$

Testing for $Exp(1/\theta)$

Introduction

- The test statistic is $\chi^2 = \frac{2n\bar{x}}{\theta_0}$.
- The rejection region is

$$W = \{\chi^2 \leq \chi^2_\alpha(2n)\}.$$

- The p value is

$$p = P(\chi^2 \leq \chi^2_0)$$

where χ^2 is a random variable with a Chi-squared distribution $\chi^2(2n)$.

Testing for $Exp(1/\theta)$

Example (Con'd)

- The test statistic is

$$\chi_0^2 = \frac{2n\bar{x}}{\theta_0} = \frac{10 \times 4462.6}{6000} = 7.4377$$

- At the significance level $\alpha = 0.05$, the rejection region is

$$W = \{\chi_0^2 \leq \chi_\alpha^2(2n)\} = \{\chi_0^2 \leq 3.94\}$$

- The p value is

$$p = P(\chi^2 \leq \chi_0^2) = 0.3164 > \alpha = 0.05$$

Then, we cannot reject the null hypothesis. Thus, the average lifetime of the component is not smaller than 6K hours.