

Introduction to Statistics Note

2024 Spring Semester

21 CST H3Art

Chapter 6: Introduction to Inference

6.1 Estimating with Confidence

Statistical inference provides methods for drawing conclusions about a population from sample data.

Confidence Interval

A **level C confidence interval** for a parameter has two parts:

- An **interval** calculated from the data, which has the form:

$$\text{estimate} \pm \text{margin of error}$$

- A confidence level C , where C is the probability that the interval will capture the true parameter value in repeated samples. In other words, the confidence level is the success rate for the method.

Choose an SRS of size n from a population having unknown mean μ and known standard deviation σ . A level C **confidence interval** for μ is:

$$\bar{x} \pm z \times \frac{\sigma}{\sqrt{n}}$$

The critical value z is found from the **standard Normal distribution**.

The confidence level C determines the value of z , the margin of error also depends on z .

$$m = z \times \frac{\sigma}{\sqrt{n}}$$

The **margin of error** gets smaller when:

- z gets **smaller** (the same as a lower confidence level C).
- σ is **smaller**. It is easier to pin down μ when σ is smaller.
- n gets **larger**. Since n is under the square root sign, we must take four times as many observations to cut the margin of error in half.

The confidence interval for a population mean will have a specified margin of error m when the sample size is:

$$m = z \times \frac{\sigma}{\sqrt{n}} \leftrightarrow n = \left(\frac{z \times \sigma}{m} \right)^2$$

Some Cautions:

- The data should be an **SRS (简单随机抽样)** from the population.
- The **confidence interval** and **sample size formulas** are **not** correct for **other sampling methods**.
- Inference **cannot** rescue **badly produced data**.
- Confidence intervals are not resistant to outliers.
- If n is **small (<15)** and the population is not Normal, the true confidence level will be **different** from C .

- The standard deviation σ of the population **must be known**.
- The margin of error in a confidence interval **covers only random sampling errors!**

6.2 Tests of Significance

A **test of significance** (显著性检验) is a formal procedure for comparing observed data with a claim (also called a **hypothesis** (假设)) whose truth we want to assess.

We express the results of a significance test in terms of a probability, called the **P-value**, that measures how well the data and the claim agree.

A **significance test** starts with a careful statement of the claims we want to compare.

- The claim tested by a statistical test is called the **null hypothesis** (H_0). The test is designed to assess the strength of the evidence against the null hypothesis.
- The claim about the population for which we're trying to find evidence is the **alternative hypothesis** (H_a).
 - The alternative is **one-sided** if it states either that a parameter is **larger** than the null hypothesis value, or **smaller** than the null hypothesis value.
 - It is **two-sided** if it states that the parameter is **different** from the null value.

The **null hypothesis** H_0 states the claim that we seek to disprove. The probability that measures the strength of the evidence **against** a null hypothesis is called a **P-value**.

- **Small P-values** are evidence against H_0 because they say that the observed result is unlikely to occur when H_0 is true.
 - P-value small \rightarrow reject $H_0 \rightarrow$ conclude H_a (in context)
- **Large P-values** fail to give convincing evidence against H_0 because they say that the observed result is likely to occur by chance when H_0 is true.
 - P-value large \rightarrow fail to reject $H_0 \rightarrow$ cannot conclude H_a (in context)

We can compare the **P-value** with a fixed value that we regard as decisive, called the **significance level**, we write it as α :

- P-value $< \alpha \rightarrow$ reject $H_0 \rightarrow$ conclude H_a (in context)
- P-value $\geq \alpha \rightarrow$ fail to reject $H_0 \rightarrow$ cannot conclude H_a (in context)

z test for a population mean:

Draw an SRS of size n from a Normal population that has unknown mean μ and known standard deviation σ . To test the null hypothesis that μ has a specified value,

$$H_0 : \mu = \mu_0$$

calculate the one-sample z statistic

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

In terms of a variable Z having the standard Normal distribution, the P-value for a test of H_0 against

$$H_a : \mu > \mu_0 \text{ is } P(Z \geq z)$$

$$H_a : \mu < \mu_0 \text{ is } P(Z \leq z)$$

$$H_a : \mu \neq \mu_0 \text{ is } 2P(Z \geq |z|)$$

6.4 Power and Inference as a Decision

If we **reject** H_0 when H_0 is **true**, we have committed a **Type I error**.

If we **fail to reject** H_0 when H_0 is **false**, we have committed a **Type II error**.

If you insist on a **smaller significance level** (such as 1% rather than 5%), you have to take a **larger sample**.

If you insist on **higher power** (such as 99% rather than 90%), you will need a **larger sample**.

The Common Practice of Testing Hypotheses:

- State H_0 and H_a as in a test of significance.
- Think of the problem as a decision problem, so the probabilities of **Type I** and **Type II** errors are relevant.
- Consider only tests in which the probability of a **Type I error** is **no greater** than a specified α .
- Among these tests, select a test that makes the probability of a **Type II error** as **small** as possible.