Statistics I – Chapter 9 Hypothesis Testing for One Population (Part 2)

Ling-Chieh Kung

Department of Information Management
National Taiwan University

December 19, 2012

Road map

- **Testing the population mean with known population variance.**
- Four methods of hypothesis testing.
- Testing the population mean with unknown population variance.

Testing the population mean

- There are many situations to test the **population mean**.
  - Is the average monthly salary of fresh college graduates above \$22,000 (22K)?
  - Is the average thickness of a plastic bottle 2.4 mm?
  - Is the average age of consumers of a restaurant below 40?
  - Is the average amount of time spent on information system projects above six months?
- We will use hypothesis testing to test hypotheses about the population mean.
- In this section, we assume that the **population variance** is **known** or given.

Parts for hypothesis testing

- In conducting a test, write the following three parts:
  - **Hypothesis**: H_0 and H_a.
  - **Test and calculation**: The test to apply and relevant arithmetic and calculations.
  - **Decision and implication**: Reject or do not reject H_0? What does that mean?
- This is the HTAB procedure introduced by the textbook.
- While the test and calculation part requires arithmetic or software, it is the easiest and least important part.
  - Writing the correct hypothesis is the most important.
  - Writing a good concluding statement is also critical.

Testing the population mean

- When the population variance σ^2 is know, what test should be used to test the population mean?
- If the population is **normal** or if the sample size is **large** ($n \geq 30$), we have

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim \text{ND}(0, 1).$$

- The test based on this sampling distribution, the z distribution, is called **the z test**.

Example 1

- A retail chain has been operated for many years.
- The average amount of money spent by a consumer is \$60.
- A new marketing policy has been proposed: Once a consumer spends \$70, she/he can get one credit. With ten credits, she/he can get one toy for free.
- After the new policy has been adopted for several months, the manager asks: Has the average amount of money spent by a consumer increased? Let $\alpha = 0.01$.
  - Let μ be the average expenditure (in \$) per consumer after the policy is adopted. Is $\mu > 60$?
  - The population standard deviation is \$16.

Example 1: hypothesis

- The hypothesis is

$$H_0: \mu = 60$$
$$H_a: \mu > 60.$$

  - $\mu = 60$ is our **default position**.
  - We want to know whether the population mean **has increased**.

- Or it is equivalent to write

$$H_0: \mu \leq 60$$
$$H_a: \mu > 60.$$

Example 1: test and calculations

- The manager collects a sample with 100 purchasing records of consumers. The sample mean is $\bar{x} = 65$.
- Because the population variance is known and the sample size is large, we may use the z test.
- For the rejection region, we calculate the critical value x^*:

$$0.01 = \Pr(\overline{X} \geq x^*) = \Pr\left(Z \geq \frac{x^* - 60}{16/\sqrt{100}}\right)$$

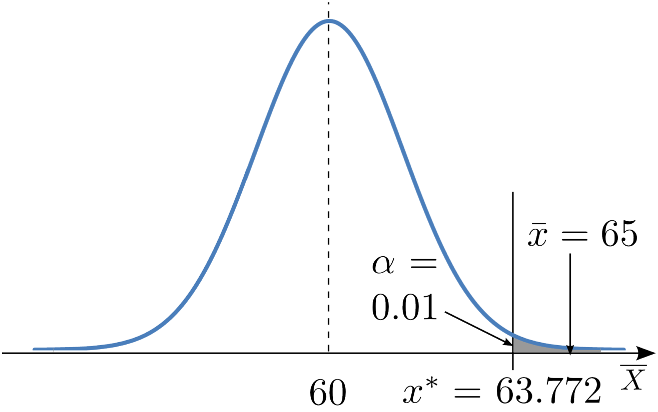
  Then

$$z_{0.01} = 2.326 \quad \Rightarrow \quad \frac{x^* - 60}{1.6} = 2.326 \quad \Rightarrow \quad x^* = 63.722.$$

- The rejection region is $(63.722, \infty)$.

Example 1: decision and implications

- Because $\bar{x} = 65$ falls in the rejection region $(63.722, \infty)$, we reject the null hypothesis.



Example 1: decision and implications

- The concluding statement:
  - Because the sample mean lies in the rejection region, we reject H_0.
  - With a 1% significance level, there is a strong evidence showing that the average expenditure per consumer is larger than \$60.
  - The new marketing policy (\$70 for one credit and ten credits for one toy) is successful: Each consumer is willing to pay more (in expectation) under the new policy.

Road map

- Testing the population mean with known population variance.
- **Four methods of hypothesis testing**.
- Testing the population mean with unknown population variance.

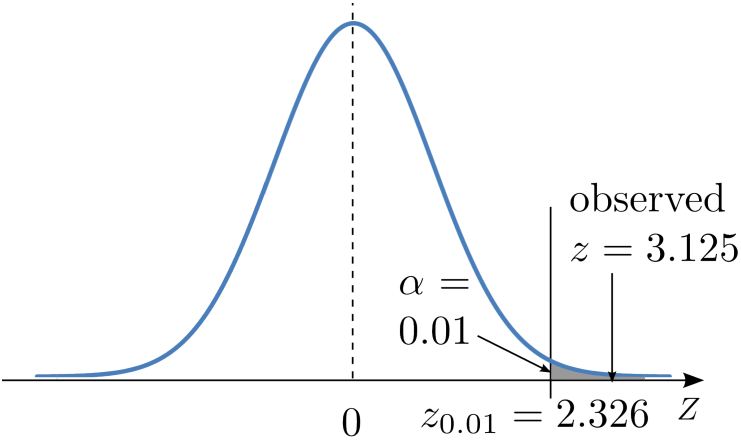
Four methods for calculations

- For the test and calculation step, there are four methods:
  - The classical method.
  - The critical-value method.
  - The p-value method.
  - The confidence interval method.
- All these four methods are **equivalent**.

The classical method

- For a statistic (e.g., the sample mean), compute its z-score, which is called the "observed z value".
- If the **observed z value** is more extreme than **the critical z value**, reject H_0.
- In the previous example:
  - The observed z value is $\frac{65-60}{1.6} = 3.125$.
  - The critical z value is $z_{0.01} = 2.326$.
  - As $3.125 > 2.326$, we reject H_0.

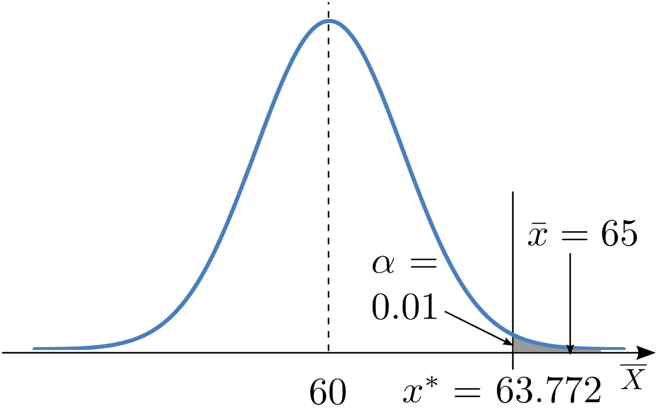
The classical method



The critical-value method

- Ignore the statistic for a moment.
- Based on the critical z value, calculate the corresponding critical value(s) in the original scale.
- If the observed **statistic** is more extreme than **the critical value**, reject H_0.
- In the previous example:
  - The observed statistic is $\bar{x} = 65$.
  - The critical value is $x^* = 60 + 1.6 z_{0.01} = 63.722$.
  - As $65 > 63.722$, we reject H_0.

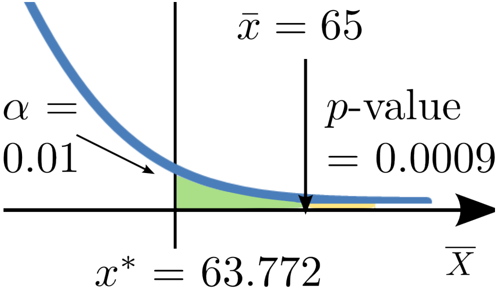
The critical-value method



The p-value method

- Based on the observed statistic, calculate the p-value.
  - The probability to observe a value that is more extreme than the observed value.
- If **the p-value** is smaller than **the significance level α**, reject H_0.
- In the previous example:
  - The p-value is $\Pr(\overline{X} > 65) = \Pr(Z > 3.125) = 0.0009$.
  - The significance level $\alpha = 0.01$.
  - As $0.0009 < 0.01$, we reject H_0.

The p-value method



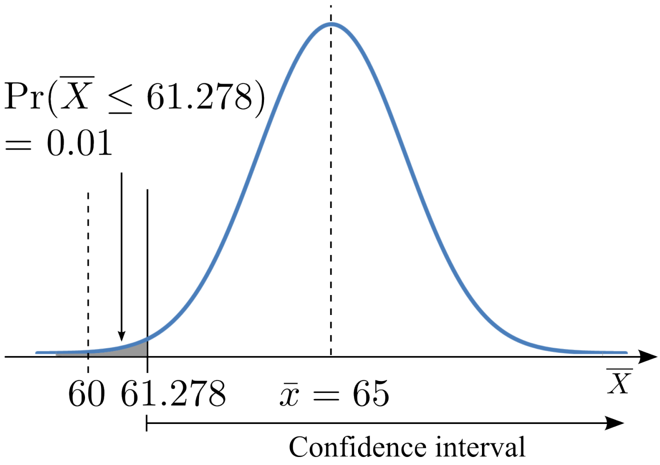
The confidence interval method

- Based on the observed statistic, calculate the confidence interval for estimating the parameter with $1-\alpha$ as the confidence level.
- If the (probably one-tailed) confidence interval **does not cover** the hypothesized parameter, we reject H_0.
- In the previous example:
  - The one-tailed confidence interval $[65-d, \infty)$ satisfies

$$0.99 = \Pr(65 - d < \overline{X}) \Rightarrow d = 3.722.$$

  - The confidence interval is $[61.278, \infty)$.
  - As $60 \notin [61.278, \infty)$, we reject H_0.

The confidence interval method



Four methods for calculations

- The four methods are **equivalent**:
  - If one says rejection, all the other three say rejection.
  - If one says no rejection, all the other three say no rejection.
- Which one to use?
  - Mostly we use **the critical-value method** and **the p-value method**.
  - When we want a rejection criterion in the original scale, use the critical-value method.
  - When we want to avoid specifying the significance level at the beginning, use the p-value method.
  - The classical method is seldom used.
  - The confidence interval method is not recommended.

Road map

- Testing the population mean with known population variance.
- Four methods of hypothesis testing.
- **Testing the population mean with unknown population variance**.

When the variance is unknown

- When the population variance σ^2 is **unknown**, the quantity $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ is unknown.
- When we use the sample variance S^2 as a substitute, we have

$$\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1),$$

  which means the quantity $\frac{\overline{X}-\mu}{S/\sqrt{n}}$ follows **the t distribution** with degree of freedom $n-1$.
- We will use **the t test** to test the population mean if the population is **normal**.
- If the **sample size is large**, we may still use the z distribution with s substituting σ.

Example 2

- We are interested in whether the students in NTU prefer the restaurants in NTU.
- One benchmark is NTUST. In a census conducted in NTUST, students are asked to rate their restaurants in a five-point scale.
- The average score is 4.6.
- We asked 60 NTU students to rate the restaurants in NTU. The average score is 4.27 and the standard deviation is 1.22.
- Do NTU students rate their restaurants differently from NTUST students?
- Suppose the scores of all NTU students are normal.

Example 2: hypothesis

- The hypothesis is

$$H_0\colon \mu = 4.6$$
$$H_a\colon \mu \neq 4.6.$$

  - μ is the average score (out of a five-point scale) of NTU restaurants rated by all NTU students.
  - Why a two-tailed test?

Example 2: test and calculations

- Because the population variance is unknown and the population is normal, we may use **the t test**.
- Let $T_n \sim t(n)$, we calculate the p-value:

$$
\begin{aligned}
& \Pr(\overline{X} < 4.27) \\
= \ & \Pr\left(T_{59} < \frac{4.27 - 4.6}{1.22/\sqrt{60}}\right) \\
= \ & \Pr(T_{59} < -2.095) \\
= \ & 0.0202.
\end{aligned}
$$

Example 2: test and calculations

- The rejection decision for various α is:

α	0.01	0.05	0.1
Comparison	$0.0202 > 0.005$	$0.0202 < 0.025$	$0.0202 < 0.05$
Decision	Do not reject	Reject	Reject

- Why $\frac{\alpha}{2}$?

Example 2: decision and implications

- Suppose the significance level is $\alpha = 0.01$.
- The concluding statement:
  - For this two-tailed test, as the p-value is larger than $\frac{\alpha}{2}$, we do not reject H_0.
  - With a 1% significance level, there is no strong evidence showing that NTU students rate their restaurants differently from NTUST students.
  - NTU do not need to change their restaurants.
- The choice of α affects the decision and implications!

Example 2 with the z test

- We may also use the z test because the sample size is **large**.
- The p-value in the z test is

$$\Pr(\overline{X} < 4.27) = \Pr\left(Z < \frac{4.27 - 4.6}{1.22/\sqrt{60}}\right)$$
$$= \Pr(Z < -2.095) = 0.01808.$$

- The p-value becomes smaller in the z test than in the t test.
- It is **easier to reject** H_0 by using the z test.
  - It is assumed that S is **close enough** to σ when n is large.
  - If one wants to be conservative, the z test should be adopted only if n is **much larger** than 30.

Example 3

- Suppose an MBA program seldom admits applicants without a work experience longer than two years.
- To test whether this is true, twenty admitted applicants are randomly selected. Prior to entering the program, they have an average work experience of 2.5 years.
- The sample standard deviation is 1.1 years.
- The population is believed to be normal.
- With a 5% significance level, is the average work experience higher than two?

Example 3: hypothesis

- Suppose the one asking the question is a potential applicant with one year of work experience. He is **pessimistic** and will apply for the program **only if** the average work experience is proven to be **less** than two years.
- The hypothesis is

$$H_0: \mu = 2$$
$$H_a: \mu < 2.$$

  - μ is the average work experience (in years) of all admitted applicants prior to entering the program.
  - To **encourage** him, we need to give him a strong evidence showing that the chance is high.

Example 3: hypothesis

- Suppose he is **optimistic** and will not apply for the program **only if** the average work experience is proven to be **greater** than two.
- The hypothesis becomes

$$H_0: \mu = 2$$
$$H_a: \mu > 2.$$

  - To **discourage** him, we need to give him a strong evidence showing that the chance is slim.
- For the remaining part, let's assume that he is optimistic and H_a is $\mu > 2$.

Example 3: test and calculations

- Because the population variance is unknown and the population is normal, we may use the t test.
- Let $T_n \sim t(n)$, we calculate the p-value:

$$\begin{aligned}\Pr(\overline{X} > 2.5) &= \Pr\left(T_{15} > \frac{2.5 - 2}{1.1/\sqrt{16}}\right) \\ &= \Pr(T_{15} > 1.818) = 0.0445.\end{aligned}$$

  - Though the significance level α is specified, we may still use the p-value method.

Example 3: test and calculations

- Conclusions:
  - For this one-tailed test, as the p-value 0.0445 is smaller than $\alpha = 0.05$, we reject H_0.
  - With a 5% significance level, there is a strong evidence showing that the average work experience is longer than two years.
  - The result is strong enough to discourage the potential applicant, who has only one year of work experience.
  - The potential applicant **should not** apply.

Example 3: a pessimistic applicant

- Suppose the applicant is pessimistic and the hypothesis is

$$H_0 : \mu = 2$$
$$H_a : \mu < 2.$$

  - The p-value is

$$\Pr(\overline{X} < 2.5) = \Pr\left(T_{15} < \frac{2.5 - 2}{1.1/\sqrt{16}}\right)$$
$$= \Pr(T_{15} < 1.818) = 0.9555.$$

  - We do not reject H_0 and cannot conclude that $\mu < 2$.
  - He **should not** apply. The choice of hypotheses does not matter.
- Is it possible that the choice of hypotheses **matters**? When?

Example 3: a pessimistic applicant

- If we do not reject H_0 when H_a is $\mu > 2$, different hypotheses result in different conclusions.

p-value		$H_a : \mu > 2$	$H_a : \mu < 2$
$\alpha = 0.05$	Reject H_0?	Reject H_0	Do not reject H_0
	Apply?	Do not apply	Do not apply
$\alpha = 0.01$	Reject H_0?	**Do no reject H_0**	Do not reject H_0
	Apply?	**Apply**	Do not apply

- Be careful in setting up the hypothesis!

Remark: finite population correction

- When we use the z test, if the population is finite ($n > 0.05N$):
  - If the population variance σ^2 is known, the standard error is

$$\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}.$$

  - If the population variance σ^2 is unknown and substituted by the sample variance s^2, the standard error is $\frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$.
  - $\sqrt{\frac{N-n}{N-1}}$ is the **finite population corrector**.
- All other steps remain the same.

Summary

- The selection of tests:

σ^2	Sample size	Population distribution	
		Normal	Nonnormal
Known	$n \geq 30$	z	z
	$n < 30$	z	Nonparametric
Unknown	$n \geq 30$	t or z	z
	$n < 30$	t	Nonparametric

  - If z test, do the finite population correction if $n > 0.05N$.
  - If t test, there is no need of doing this.