Statistics I – Chapter 9 Hypothesis Testing for One Population (Part 3)

Ling-Chieh Kung

Department of Information Management
National Taiwan University

December 26, 2012

Road map

- **Testing the population proportion**.
- Testing the population variance.
- Preview for the next semester.

Testing the population proportion

- In many situations, we need to test the **population proportion**.
  - The defective rate or yield rate of a production system.
  - The proportion of people supporting a candidate.
  - The proportion of people supporting a policy.
  - The market share of a company.
  - The proportion of people viewing a product web page that will really buy the product (conversion rate).
- How to test the population proportion?

The hypotheses

- The population proportion is denoted as p.
- A two-tailed test for the population proportion is

$$H_0 : p = p_0$$
$$H_a : p \neq p_0,$$

  where p_0 is the **hypothesized proportion**.
- In a one-tailed test, the alternative hypothesis may be either

$$H_a : p > p_0$$

  or

$$H_a : p < p_0.$$

Sample proportion

- In testing the population mean μ, we base on the sample mean $\overline{X}$ and its distribution.
  - Reject H_0 if the p-value $\Pr(\overline{X} < \bar{x}) < \alpha$ (for a left-tailed test).
  - $\overline{X}$ is a statistic and $\bar{x}$ is an observed value of it.
- In testing the population proportion p, we base on the **sample proportion** $\widehat{P}$ and its distribution.
  - Reject H_0 if the p-value $\Pr(\widehat{P} < \hat{p}) < \alpha$ (for a left-tailed test).
  - $\widehat{P}$ is a statistic and $\hat{p}$ is an observed value of it.
- What is the sampling distribution of $\widehat{P}$?

Distribution of sample proportion

- $\widehat{P}$ is a random variable:
  - $\mathbb{E}[\widehat{P}] = p$.
  - $\mathrm{Var}(\widehat{P}) = \frac{p(1-p)}{n}$.
- When the sample size is large ($n \geq 30$), $\widehat{P}$ follows the **normal** distribution approximately.
  - In practice, it is safer to assume normality if $np > 5$ and $n(1-p) > 5$. As p is unknown, we check whether $n\hat{p}$ and $n(1-\hat{p})$ are greater than five.
- In testing the population proportion, we use **the z test**.

Rejection rule: The p-value method

- Let's first discuss the p-value method.
- For a left-tailed test, we need to calculate **the p-value**

$$\Pr(\widehat{P} < \hat{p})$$

  based on the observed $\hat{p}$ and then **compare it with α**.
  - If p-value $< \alpha$, reject H_0.
  - If p-value $> \alpha$, do not reject H_0.

Example 1

- In a factory, it seems to us that the defective rate of our product is too high. Ideally it should be below 1% but some workers believe that it is above 1%.
- If the defective rate is above 1%, we should fix the machine. Otherwise, we do not do anything.
- Let p be the defective rate, the hypothesis is

$$H_0 : p = 0.01$$
$$H_a : p > 0.01.$$

  - When to adopt $H_a : p < 0.01$?

Example 1

- In several random production runs, we found that out of 1000 produced items, 14 of them are defective.
  - The observed sample proportion $\hat{p} = 0.014$.
- Suppose the significance level is set of $\alpha = 0.05$, what is our conclusion?

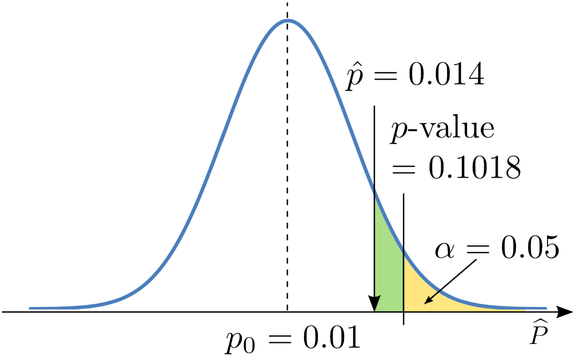
Example 1: the p-value method

- The sample proportion $\widehat{P}$ is **normal**. Moreover:
  - Its **expectation** is 0.01.
  - Its **standard error** is $\sqrt{\frac{(0.01)(0.99)}{1000}} = 0.00315$.
- The p-value is

$$\Pr\left(\widehat{P} > \hat{p}\right) = \Pr\left(Z > \frac{0.014 - 0.01}{0.00315}\right) = \Pr(Z > 1.271) = 0.1018.$$

Example 1: the p-value method

- Because the p-value is larger than α, we do not reject the null hypothesis.



Example 1: the critical value method

- The sample proportion $\widehat{P}$ is normal with
  - Its expectation is 0.01.
  - Its standard error is $\sqrt{\frac{(0.01)(0.99)}{1000}} = 0.00315$.
- The **critical value** p^* satisfies

$$\Pr\left(\widehat{P} > p^*\right) = \alpha = \Pr\left(Z > \frac{p^* - 0.01}{0.00315}\right) = 0.05.$$

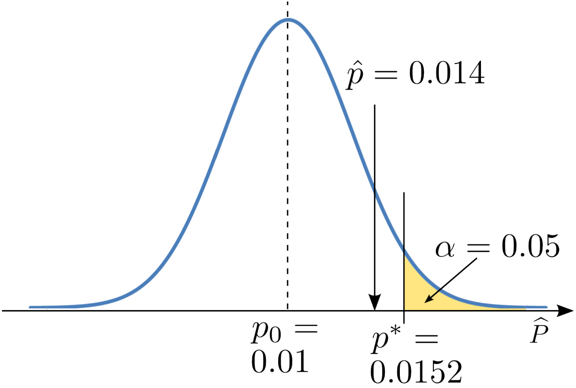
- As the critical z value is $z_{0.05} = 1.645$, we have

$$\frac{p^* - 0.01}{0.00315} = 1.645,$$

which implies $p^* = 0.0152$.

Example 1: the critical value method

- Because the observed sample proportion $\hat{p}$ is smaller than the critical value p^*, we do not reject the null hypothesis.



Example 1: decision and implications

- The concluding statement:
  - Because the p-value is larger than the significance level, we do not reject H_0.
  - With a 5% significance level, there is no strong evidence showing that the defective rate is higher than 1%.
  - We will not try to fix the machine.

Testing the population proportion

- Wait!
- The sample proportion $\widehat{P}$ is normal with
  - Its expectation is 0.01.
  - Its standard error is $\sqrt{\frac{(0.01)(0.99)}{1000}} = 0.00315$.
- One thing is strange...
- The population proportion is p. It should be $\mathbb{E}[\widehat{P}] = p$ and $\text{Var}(\widehat{P}) = \frac{p(1-p)}{n}$.
  - Why do we **use** $p_0 = 0.01$ **to substitute** p?
  - Why not $\hat{p} = 0.014$?

Using the hypothesized value

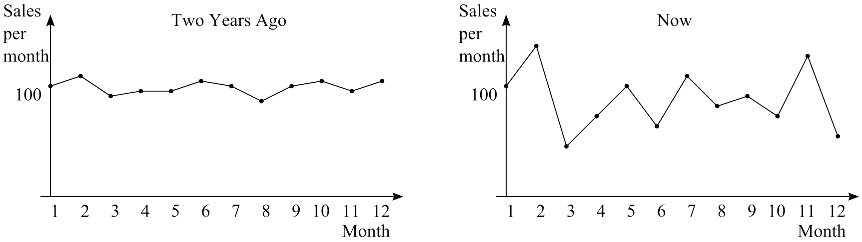
- The population proportion is p. It should be $\mathbb{E}[\widehat{P}] = p$ and $\text{Var}(\widehat{P}) = \frac{p(1-p)}{n}$.
- As p is unknown, we need a substitute. p_0 or $\hat{p}$?
- In doing hypothesis testing, it is always the case that we **assume H_0 is true**.
  - $\alpha = \Pr(\text{rejecting } H_0 | H_0 \text{ is true})$.
- If H_0 is true, it is natural that $p = p_0$ and $p \neq \hat{p}$.
- Summary:
  - For estimating p, use $\hat{p}$ as a substitute.
  - For testing p, use p_0 as a substitute.

Road map

- Testing the population proportion.
- **Testing the population variance**.
- Preview for the next semester.

Testing the population variance

- In many cases, we need to test the population variance.
  - The demand of a product seems to remain identical in average for many years. But is the variability also identical?



Testing the population variance

- Some other examples:
  - The average weight of a product is under control. But is the variance small enough?
  - If we believe the daily demand of a product D_t satisfies

    $$D_t = \mu + \epsilon_t,$$

    where μ is an estimation and ϵ_t is a random fluctuation. Is $\mathrm{Var}(\epsilon_t)$ small?
  - For tests that compare multiple populations, we may need to know whether their variances are identical. This will be discussed in Chapter 10 in the next semester.

The hypotheses

- The population variance is denoted as σ^2.
- A two-tailed test for the population proportion is

$$\begin{aligned} H_0 &: \sigma^2 = \sigma_0^2 \\ H_a &: \sigma^2 \neq \sigma_0^2, \end{aligned}$$

  where σ_0^2 is the **hypothesized variance**.
- In a one-tailed test, the alternative hypothesis may be either

$$H_a : \sigma^2 > \sigma_0^2$$

  or

$$H_a : \sigma^2 < \sigma_0^2$$

Sample variance

- To test the population variance σ^2, we base on the **sample variance** S^2 and the sampling distribution

$$\chi^2_{n-1} = \frac{(n-1)S^2}{\sigma^2} \sim \text{Chi}(n-1),$$

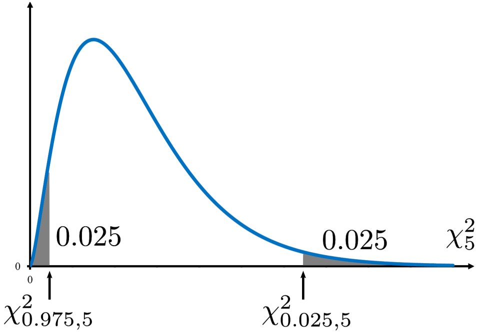
  where n is the sample size.
- The population must be **normal**!
- The test for testing the population variance is thus called **a chi-square test**.
- To utilize the chi-square distribution, let's recall the definition of **the chi-square critical value**.

Critical chi-square value

- The critical chi-square value $\chi^2_{w,n-1}$ satisfies

$$\Pr\left(\chi^2_{n-1} > \chi^2_{w,n-1}\right) = w,$$

where w is the right-tail probability.



Rejection rule

- Consider a right-tailed test ($H_a : \sigma^2 > \sigma_0^2$). With the p-value method, we will reject H_0 if

$$\Pr\left(\chi^2_{n-1} > \frac{(n-1)s^2}{\sigma_0^2}\right) < \alpha.$$

  - $\frac{(n-1)S^2}{\sigma^2} \sim \text{Chi}(n-1)$.
  - $\frac{(n-1)s^2}{\sigma_0^2}$ is the **observed chi-square value**.
  - The observed sample variance s^2 is combined with the **hypothesized population variance** σ_0^2. This is because we assume H_0 is true.

Rejection rule

- For a left-tailed test, we reject H_0 if

$$\Pr\left(\chi^2_{n-1} < \frac{(n-1)s^2}{\sigma^2}\right) < \alpha.$$

- For a two-tailed test, we reject H_0 if either

$$\Pr\left(\chi^2_{n-1} < \frac{(n-1)s^2}{\sigma^2}\right) < \frac{\alpha}{2}$$

or

$$\Pr\left(\chi^2_{n-1} > \frac{(n-1)s^2}{\sigma^2}\right) < \frac{\alpha}{2}.$$

Example 2

- Suppose we are testing

$$H_0\colon \sigma^2 = 10$$
$$H_a\colon \sigma^2 \neq 10.$$

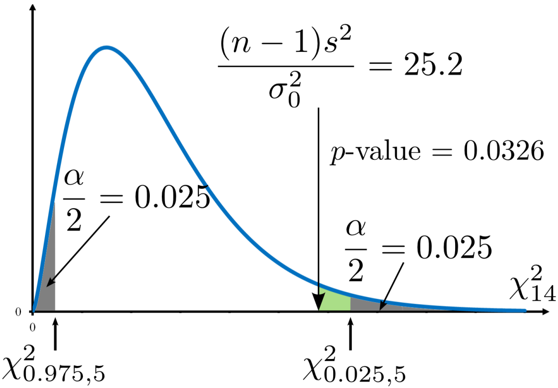
  The sample size is 15 and the sample variance is 18. Let $\alpha = 0.05$. The population is normal.

- The observed chi-square value is

$$\frac{(15-1)\times 18}{10} = 25.2.$$

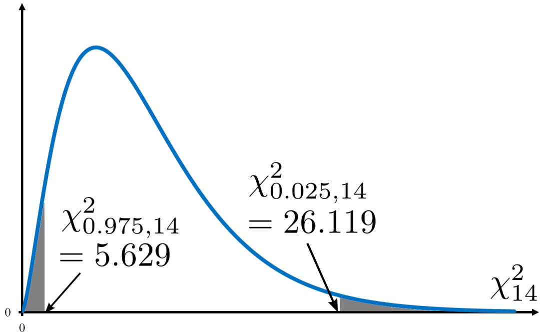
Example 2: the p-value method

- The p-value is $\Pr(\chi^2_{14} > 25.2) = 0.0326$.
- As the p-value is greater than $\frac{\alpha}{2} = 0.025$, we do not reject H_0.



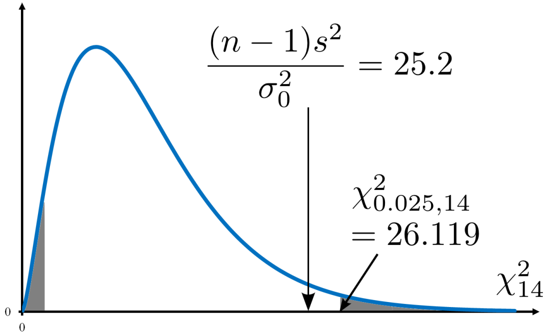
Example 2: the critical value method

- Given $\alpha = 0.05$, the critical χ^2 values for the two-tailed test are $\chi^2_{0.975,14} = 5.629$ and $\chi^2_{0.025,14} = 26.119$.



Example 2: the critical value method

- As the observed χ^2 value $25.2 < 26.119$, we do not reject H_0.
- There is no strong evidence showing that the population variance is not 10.



Remarks

- Even if $s^2 = 18$ is almost twice higher than $\sigma_0^2 = 10$, we cannot reject H_0: The evidence is not strong enough.
  - If we only allow to make a Type I error with probability 5%, we are not confident enough to claim that $\sigma^2 \neq 10$.
  - If $\alpha = 0.1$, we will reject H_0.
  - For a right-tailed test ($H_a : \sigma^2 > 10$), we will reject H_0.
- For using the chi-square test to test the population variance, the population must be normal.
  - A large sample size does not help!

Road map

- Testing the population proportion.
- Testing the population variance.
- **Preview for the next semester**.

Review for this semester

- Descriptive Statistics.
  - Chapter 2: with graphs
  - Chapter 3: with statistics.
- Probability.
  - Chapter 4: basic probability.
  - Chapter 5: discrete distributions.
  - Chapter 6: continuous distributions.
  - Moment generating functions.
  - Chapter 7: sampling distributions.

Review for this semester

- Inferential Statistics:
  - Chapter 8: Estimation for a single population.
  - Chapter 9: Hypothesis testing for a single population.
- Parameters studied: mean, proportion, and variance.

Review for this semester

Ch.	Topic	Category
1	Introduction	Introduction
2	Graphical visualization	Descriptive Statistics
3	Numerical summarization	Descriptive Statistics
4	Probability	Probability
5	Discrete distributions	Probability
6	Continuous distributions	Probability
7	Sampling distributions	Probability
8	Estimation for one population	Inferential Statistics
9	Testing for one population	Inferential Statistics

Preview for the next semester

- Chapter 10: Estimation and testing for two populations.
  - Are there relatively more girls in the business school than in the engineering school?
  - Are employees in Taiwan paid less than employees in the US?
  - Do voters prefer candidate 1 to candidate 2?
  - Do income levels become more diverse nowadays?
- Chapter 11: Analysis of variance (testing for more than two populations).

Preview for the next semester

- Chapter 12: Simple regression.
  - Determine how one variable affects another one.
  - Is there really a cause-and-effect relationship?
- Chapters 13 and 14: Multiple regression.
  - Determine how multiple variables jointly affect another one.
  - Among many possible reasons of a result, which ones are indeed influential?
- Chapter 15: Forecasting.
  - From past data to predict future data.
  - Regression is one forecasting method: how time affects things.

Preview for the next semester

- Chapter 16: Chi-square analysis.
  - Fitness test: What is the distribution of one population?
  - Independence test: Are two variables independent?
- Chapter 17: Nonparametric Statistics.
  - Randomness test: Is one experiment really random?
  - When the population is not normal and the sample size is small.

Preview for the next semester

Ch.	Topic	Category
10	Inferences about two populations	Inferential Statistics
11	Analysis of variance	Inferential Statistics
12	Simple regression	Inferential Statistics
13	Multiple regression	Inferential Statistics
14	Advanced multiple regression	Inferential Statistics
15	Forecasting	Others
16	Chi-square analysis	Inferential Statistics
17	Nonparametric Statistics	Inferential Statistics
18	Quality control	Others

Two semesters as a whole

Ch.	Category	Ch.	Category
1	Introduction	10	Inferential Statistics
2	Descriptive Statistics	11	Inferential Statistics
3	Descriptive Statistics	12	Inferential Statistics
4	Probability	13	Inferential Statistics
5	Probability	14	Inferential Statistics
6	Probability	15	Others
7	Probability	16	Inferential Statistics
8	Inferential Statistics	17	Inferential Statistics
9	Inferential Statistics	18	Others