Lecture 10
Testing Difference in Population Mean

by Hiro Kasahara

Vancouver School of Economics
University of British Columbia

Examples

We are often interested in testing if the mean of two populations is different or not.

- A/B testing
- Homework and Final Grades

A/B testing

A/B testing

Randomly assign 200 visitors into two versions of web designs.

	Click	No Click	Total visits
Design A	52	48	100
Design B	72	28	100

- 52 out of 100 visitors clicked for design A: $\hat{p}_x = 0.52$.
- 72 out of 100 visitors clicked for design B: $\hat{p}_y = 0.72$.

A/B testing

Data from Design A: $\{X_1, X_2, ..., X_{100}\}$, where $X_i \in \{0, 1\}$

Data from Design B: $\{Y_1, Y_2, ..., Y_{100}\}$, where $Y_j \in \{0, 1\}$

$$\textit{Population}: \quad \Pr(X_i = 1) = p_x \quad \text{and} \quad \Pr(Y_j = 1) = p_y.$$

$$\textit{Sample}: \quad \hat{p}_x = \frac{1}{100}\sum_{i=1}^{n} X_i = 0.52$$

$$\hat{p}_y = \frac{1}{100}\sum_{j=1}^{n} Y_j = 0.72$$

How to test $H_0 : p_y - p_x \leq 0$?

Testing $H_0 : p_y - p_x \geq 0$

1. Derive the null distribution of $\frac{(\hat{p}_y - \hat{p}_x) - 0}{\sqrt{\frac{p_x(1-p_x)}{n_x} + \frac{p_y(1-p_y)}{n_y}}}$ when $H_0 : p_y = p_x = p_0$.

2. Compute the rejection region: the unlikely values of $\frac{(\hat{p}_y - \hat{p}_x) - 0}{\sqrt{\frac{p_0(1-p_0)}{n_x} + \frac{p_0(1-p_0)}{n_y}}}$ when $H_0 : p_y - p_x \leq 0$ holds.

3. Compute the statistic $\frac{(\hat{p}_y - \hat{p}_x) - 0}{\sqrt{\frac{\hat{p}_0(1-\hat{p}_0)}{n_x} + \frac{\hat{p}_0(1-\hat{p}_0)}{n_y}}}$ from the data and reject $H_0 : p_y - p_x \leq 0$ if it falls on the rejection region.

Distribution of $\hat{p}_y - \hat{p}_x$ when $p_y = p_x = p_0$

Because

$$E(\hat{p}_y - \hat{p}_x) = p_y - p_x = 0,$$
$$Var(\hat{p}_y - \hat{p}_x) = \frac{p_0(1-p_0)}{n_x} + \frac{p_0(1-p_0)}{n_y},$$

we have

$$\frac{(\hat{p}_y - \hat{p}_x) - 0}{\sqrt{\frac{p_0(1-p_0)}{n_x} + \frac{p_0(1-p_0)}{n_y}}} \xrightarrow{d} N(0,1)$$

Therefore, if $H_0 : p_y = p_x = p_0$ holds, then

$$P\left(\frac{(\hat{p}_y - \hat{p}_x) - 0}{\sqrt{\frac{p_0(1-p_0)}{n_x} + \frac{p_0(1-p_0)}{n_y}}} \geq 1.64\right) = 0.05.$$

Distribution of $\hat{p}_y - \hat{p}_x$ when $p_y = p_x = p_0$

We replace p_0 with its consistent estimator:

$$\hat{p}_0 = \frac{1}{n_x + n_y}\left(\sum_{i=1}^{n_x} X_i + \sum_{j=1}^{n_y} Y_i\right).$$

Because $\hat{p}_0$ converges in probability to p_0, if $H_0 : p_y = p_x = p_0$ holds, then

$$P\left(\frac{(\hat{p}_y - \hat{p}_x) - 0}{\sqrt{\frac{\hat{p}_0(1-\hat{p}_0)}{n_x} + \frac{\hat{p}_0(1-\hat{p}_0)}{n_y}}} \geq 1.64\right) \approx 0.05.$$

when n is large.

Test Statistic, Critical Value, and Rejection Region

- Test statistic is
$$\frac{(\hat{p}_y - \hat{p}_x) - 0}{\sqrt{\frac{\hat{p}_0(1-\hat{p}_0)}{n_x} + \frac{\hat{p}_0(1-\hat{p}_0)}{n_y}}}$$
- Critical value and rejection region for testing $H_0 : p_y - p_x \leq 0$ at the 5 percent level is
$$\text{Critical value} : 1.64$$
$$\text{Rejection region} : [1.64, \infty]$$
- Decision rule:
$$\text{Reject } H_0 : p_y - p_x \leq 0 \text{ if } \quad \frac{(\hat{p}_y - \hat{p}_x) - 0}{\sqrt{\frac{\hat{p}_0(1-\hat{p}_0)}{n_x} + \frac{\hat{p}_0(1-\hat{p}_0)}{n_y}}} \geq 1.64$$

A/B test example

In this example, $\hat{p}_y = 0.72$, $\hat{p}_x = 0.50$, $\hat{p}_0 = 0.61$, and $n_y = n_x = 100$.

$$\frac{(\hat{p}_y - \hat{p}_x) - 0}{\sqrt{\frac{\hat{p}_0(1-\hat{p}_0)}{n_x} + \frac{\hat{p}_0(1-\hat{p}_0)}{n_y}}} = \frac{(0.72 - 0.5) - 0}{\sqrt{\frac{0.61(1-0.61)}{100} + \frac{0.61(1-0.61)}{100}}} = 3.189.$$

Because 3.198 is larger than 1.64, we reject $H_0 : p_y - p_x \leq 0$.

$\Rightarrow$ Evidence that Design B's click rate is higher than Design A's in population.

Confidence Interval and Hypothesis Testing

- 90 % confidence interval $\Leftrightarrow$ Testing at 5 % level
- 90 % confidence confidence interval is given by

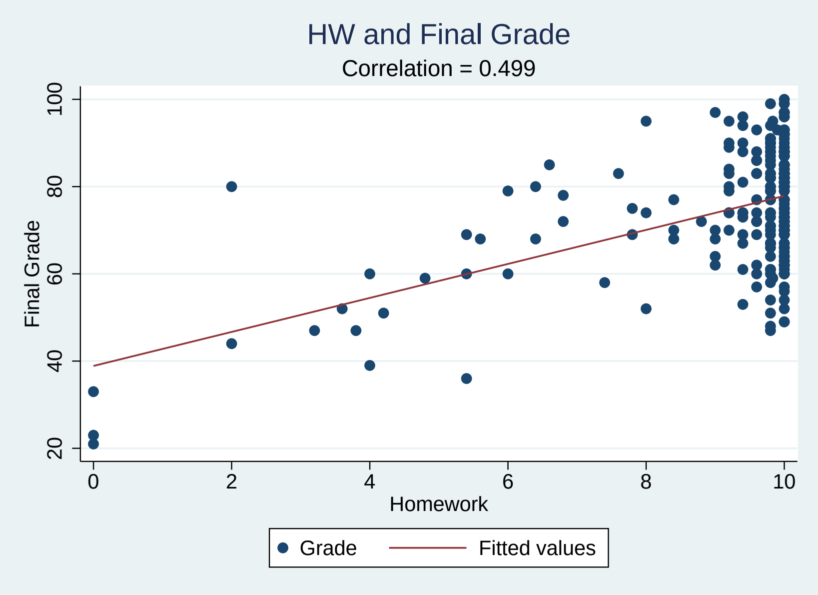
$$(0.72 - 0.5) \pm 1.64\sqrt{\frac{0.72(1-0.72)}{100} + \frac{0.5(1-0.5)}{100}}$$
$$= [0.1098, 0.3302] \geq 0$$

$\Rightarrow$ Evidence that $p_y - p_x$ is positive.

- Difference in how to compute the standard deviation of $\hat{p}_y - \hat{p}_x$ between confidence interval and testing but, in practice, they are very close.

Homework and Final Grades

Scatter Plot of HW Grade and Final Grade



Summary Statistics by Stata

Define Low HW group as students with HW grade less than 6 out of 10.

```
. gen Low_HW = 0

. replace Low_HW = 1 if hw<6

. sum grade if Low_HW==0

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       grade |        224      76.625    11.96154         47        100


. sum grade if Low_HW==1

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       grade |         16     49.3125    16.51956         21         80
```

Test for $H_0 : \mu_y - \mu_x \geq 0$

- μ_y = population mean of final grade among students who receive HW grade more than 6 out of 10.
- μ_x = population mean of final grade among students who receive HW grade less than 6 out of 10.
- Sample:

$$\bar{Y} = 76.63, \quad \bar{X} = 49.31, \quad s_y = 11.96, \quad s_x = 16.51.$$

$$n_y = 224, \quad n_x = 16.$$

How to test $H_0 : \mu_y - \mu_x \geq 0$?

Distribution of $\hat{Y} - \hat{X}$ when $\mu_y = \mu_x$

Because

$$E(\bar{Y} - \bar{X}) = \mu_y - \mu_x \quad \text{and} \quad Var(\bar{Y} - \bar{X}) = \frac{\sigma_y^2}{n_y} + \frac{\sigma_x^2}{n_x},$$

we have

$$\frac{(\bar{Y} - \bar{X}) - 0}{\sqrt{\frac{\sigma_y^2}{n_y} + \frac{\sigma_x^2}{n_x}}} \xrightarrow{d} N(0,1)$$

Because s_x^2 and s_y^2 converge in probability to σ_x^2 and σ_y^2, we replace s_x^2 and s_y^2 with σ_x^2 and σ_y^2.

Test for $H_0 : \mu_y - \mu_x \geq 0$

Applying the Central Limit Theorem,

$$\Pr\left(\frac{(\bar{Y} - \bar{X}) - 0}{\sqrt{\frac{s_y^2}{n_y} + \frac{s_x^2}{n_x}}} \geq 1.64\right) \approx 0.05$$

- Test statistic is $\frac{(\bar{Y} - \bar{X}) - 0}{\sqrt{\frac{s_y^2}{n_y} + \frac{s_x^2}{n_x}}}$.
- Rejection region at the 5 % level is $[1.64, \infty]$.
- Decision rule:

$$\text{Reject } H_0 : \mu_y - \mu_x \geq 0 \text{ if } \quad \frac{(\bar{Y} - \bar{X}) - 0}{\sqrt{\frac{s_y^2}{n_y} + \frac{s_x^2}{n_x}}} \geq 1.64$$

95 percent Confidence Interval for $\mu_y - \mu_x$

$$\bar{Y} = 76.63, \quad \bar{X} = 49.31, \quad s_y = 11.96, \quad s_x = 16.51.$$
$$n_y = 224, \quad n_x = 16.$$

The realized value of the test statistic is

$$\frac{(76.63 - 49.31) - 0}{\sqrt{\frac{(11.96)^2}{224} + \frac{(16.51)^2}{16}}} = 6.498$$

which is larger than 1.64.

$\Rightarrow$ Evidence that students who did well in HW do better in final grades than those who did not do well in HW.