107-1 Statistics

LAB8: TESTING HYPOTHESES ABOUT PROPORTIONS

助教：廖皓宇、吳家禎、賴冠宇

2018/11/16

12.1 HT Module 0: An Overview of Hypothesis Testing

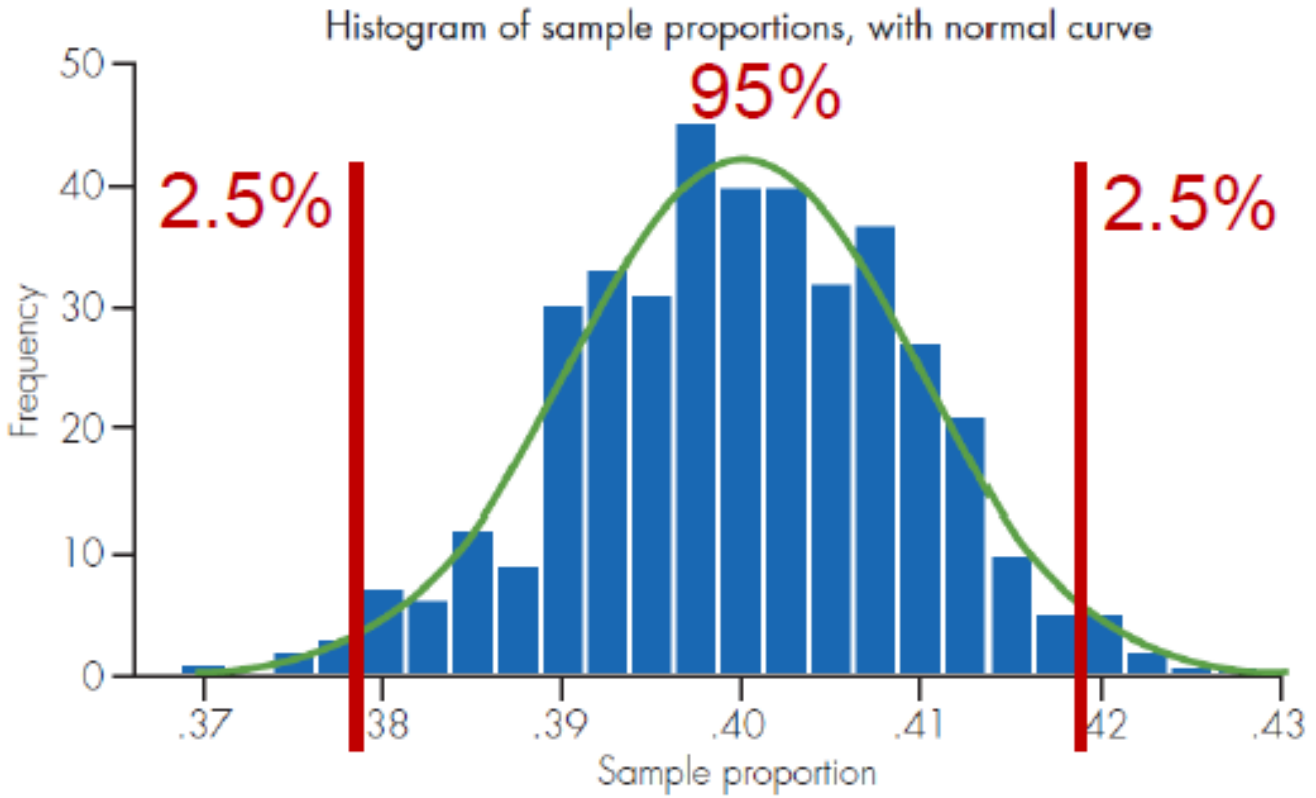
Hypothesis testing method:

uses data from a sample ***to judge*** whether or not ***a statement about a population*** may be true.

Steps in Any Hypothesis Test

1. Determine the null and alternative hypotheses.
2. Verify necessary data conditions, and if met, summarize the data into an appropriate test statistic.
3. Assuming the null hypothesis is true, find the *p*-value.
4. Decide whether or not the result is statistically significant based on the *p*-value.
5. Report the conclusion in the context of the situation.

(Type I error and level of significance)



若p=0.4是某人的認知，樣本支持某人的認知嗎？

Q1: One population proportion

- Signi*ficance level*: $\alpha = 0.01$

Step 1.

- $\boldsymbol{H_0: p = 0.4}$
- $\boldsymbol{H_1: p \neq 0.4}$

Step 2.

- $\hat{p} = 0.388, n = 4526$
- $z = \frac{\hat{p} - p}{\sqrt{\frac{p \times (1-p)}{n}}} = \frac{0.388 - 0.4}{\sqrt{\frac{0.4 \times 0.6}{4526}}} = -1.680919$

The ***z*-statistic** for the significance test is

$$z = \frac{\text{sample estimate} - \text{null value}}{\text{null standard deviation}} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Step 3.

- $p\ value = 2 \times 0.04638929 = 0.09277858$

Step 4.

- $\boldsymbol{Can\ not\ reject\ H_0}$.

Step 5.

- $Based\ on\ the\ hypotheses\ test\ result, the\ UCBA\ admission\ rate\ is\ 0.4$.

Q2: Two population proportions

- Signi*ficance level*: $\alpha = 0.01$

Step 1.

- $\boldsymbol{H_0: p_1 - p_2 = 0}$
- $\boldsymbol{H_1: p_1 - p_2 \neq 0}$

Step 2.

- $Male$: $\widehat{p_1} = 0.445, n_1 = 2691$ $\quad$ $Female$: $\widehat{p_2} = 0.304, n_2 = 1835$
- $\hat{p} = \frac{n_1\widehat{p_1}+n_2\widehat{p_2}}{n_1+n_2} = \frac{0.445\times2691+0.304\times1835}{2691+1835} = 0.388$
- $z = \frac{(\widehat{p_1}-\widehat{p_2})-0}{\sqrt{\hat{p}\times(1-\hat{p})\times(\frac{1}{n_1}+\frac{1}{n_2})}} = \frac{0.445-0.304-0}{\sqrt{0.388\times0.612\times(\frac{1}{2691}+\frac{1}{1835})}} = 9.602$

$$\hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}$$

$$z = \frac{\text{sample statistic} - \text{null value}}{\text{null standard error}} = \frac{\hat{p}_1 - \hat{p}_2 - 0}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

Step 3.

- $p\ value \approx 2 \times 0 \approx 0$

Step 4.

- $\boldsymbol{Reject\ H_0}$.

Step 5.

- $Based\ on\ the\ hypotheses\ test\ result,$
  $the\ UCBA\ admission\ rate\ between\ male\ and\ female\ is\ different.$

1116實習：類別資料的假設檢定

- 使用提供的資料(UCBAdmissions.Rdata)

Q1: One population proportion

- 單一母體比例的假設檢定與*p*值計算
  - UCB宣稱研究生入學申請錄取率為40%，檢定之

Q2: Two population proportions

- 比較兩個母體比例的假設檢定與*p*值計算
  - UCB宣稱研究生入學申請錄取率，在不同性別間無差異，檢定之

作業8 類別資料的假設檢定

- 練習題6題(Ch. 12)
  - 12.14; 12.28; 12.36; 12.54; 12.70; 12.84
- R程式練習題(繳交程式碼與執行結果)
  - 根據實習課提供的資料(UCBAdmissions.Rdata)，請分別檢視各科系的研究生入學申請錄取率，在男女間的差異，進一步評估與討論該大學在研究生入學申請是否有性別歧視(sex discrimination)的疑慮。

以 $\mathbf{Signi}\boldsymbol{ficance\ level}\mathbf{:}\ \boldsymbol{\alpha} = \mathbf{0.01}$檢定之。