有關比例的檢定

Tests on Proportions

1 基本概念

- 比例檢定

	單一母體 one population	兩個母體 two populations
proportion	H_0: $p = 0.5$ Z-統計量	H_0: $p_1 = p_2$ Z-統計量

- 理論基礎

▷ 令 $\hat{p}$ 為由樣本所算出的比例，p 為母體中真正 (但未知) 的比例，或研究者所認定之母體中真正的比例 (稱為 hypothesized value，即上例中的 0.5，有些書亦寫作 p_0)

▷ 此檢定的檢定統計量 (test statistic) 為

$$Z = \frac{\hat{p} - p}{\sqrt{\dfrac{p(1-p)}{n}}} \qquad \left(\text{理解方式} \ \Rightarrow \ \frac{\text{觀察值與認定值之差}}{\text{正規化因子}} \right)$$

▷ 在大樣本 (large sample) 下，即 n 夠大時 (較精確地說，是 $np > 5$, $n(1-p) > 5$)，此檢定統計量 (當 H_0 為正確) 為常態分配，寫作

$$Z_{\text{random}} \sim N(0,1)$$

依此性質可進行檢定，決定是否拒絕 H_0

(註: 要知道統計量所遵循的機率分配，才能夠做檢定)

- 檢定的方法

▷ 傳統檢定法 (一般教科書會先教的方法，考試較會考的方法，需要查表)

1. 由顯著水準 α 決定臨界值 Z_{critical} (critical value)
2. 比較檢定統計量的觀測值 Z_{observed} 與臨界值 Z_{critical} 的大小下結論

▷ p-value 檢定法 (實務上做研究時的主要方法，一般需用軟體才能算出 p-value)

1. 由檢定統計量的觀測值 Z_{observed} 計算出 p-value
2. 若 p-value $< \alpha \Rightarrow$ 拒絕 H_0 (反之則不拒絕)

- 討論: 單尾與雙尾 (one-tailed and two-tailed)

Q1. 在 H_0 和 H_1 的差異

Q2. 在臨界值上的差異

Q3. 在 p-value 上的差異 (記: 雙尾 p-value $= 2\times$ 單尾 p-value)

- 延伸至兩個母體的檢定

  ▷ 此時待檢定的問題變為: 比較兩母體中的比例是否相等
  此時虛無假設為 H_0: $p_1 = p_2$ (對立假設 H_1: $p_1 \neq p_2$)

  ▷ 仍需注意此處 p_1, p_2 是指母體中的比例，樣本中所得到的比例則為 $\hat{p}_1, \hat{p}_2$

  ▷ 檢定統計量 (test statistic) 為

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \qquad \left(\text{其中 } \hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}\right)$$

  ▷ 此檢定統計量在大樣本且虛無假設 H_0 正確的前提下，其隨機變數遵循常態分配

$$Z_{\text{random}} \sim N(0,1)$$

  (記得: 要知道分配才能做檢定)

  ▷ 其他流程均相同 (務必記得 p-value 法的原則: p-value $< \alpha \Rightarrow$ 拒絕 H_0)

2 單一母體比例檢定之例題

- 例題 1: 消費性貸款比例的檢定

  銀行經理發現近年來承做消費性貸款的案件大幅成長，而懷疑在所有貸款案件中消費性貸款的比例已超過 0.6，而欲進行假設檢定。其虛無假設 (H_0) 與對立假設 (H_1) 設立為:

$$\begin{aligned} \mathrm{H}_0: &\quad p = 0.6 \quad (\text{or } p \leq 0.6) \\ \mathrm{H}_1: &\quad p > 0.6 \end{aligned}$$

  銀行經理設定顯著水準 $\alpha = 0.01$。

  銀行經理決定從所有貸款案件中隨機抽取 100 件，因 $np = 60 > 5$ 且 $n(1-p) = 40 > 5$，滿足大樣本條件，故檢定統計量選擇為

$$Z_{\text{random}} = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{\hat{p} - 0.6}{\sqrt{\frac{0.6(1-0.6)}{100}}} \sim N(0,1)$$

  此次假設檢定的拒絕域 (rejection region) 為 $Z_{\text{observed}} > 2.3267$ ($=Z_{\text{critical}}$，Z_α)。

  然後銀行經理隨機抽取 100 件，發現有 72 件是屬於消費性貸款，故 $\hat{p} = 0.72$。經計算

$$Z_{\text{observed}} = \frac{0.72 - 0.6}{\sqrt{\frac{0.6(1-0.6)}{100}}} = 2.4495$$

因 2.4495>2.3267，檢定統計量之值落入拒絕域內，所以，在 0.01 的顯著水準下，拒絕虛無假設 H_0；也就是說，在 0.01 的顯著水準下，此 100 件貸款有充分證據顯示整體消費性貸款的比例 (指母體中的比例) 已超過 0.6。

▷ 討論: 如何使用 p-value 法進行分析?

- 例題 2: 樂透彩簽注 36 號比例的檢定

在 49 選 6 的大樂透遊戲中，若由電腦隨機選號，每一簽注單會出現 36 號的機率是 $\frac{6}{49}$。然而，彩券公司認為由民眾自選號碼，會對 36 號有特定好惡，亦即每一自選號碼簽注單出現 36 號的機率不是 $\frac{6}{49}$。因此，虛無假設與對立假設為:

$$\begin{aligned}\mathrm{H}_0: &\quad p = \tfrac{6}{49}\\ \mathrm{H}_1: &\quad p \neq \tfrac{6}{49}\end{aligned}$$

彩券公司設定顯著水準 α 為 0.05。

彩券公司決定從所有自選號碼簽注單中隨機抽取 2500 張，因 $np = 306.12 > 5$ 且 $n(1-p) = 2193.88 > 5$，滿足大樣本條件，故檢定統計量選擇為:

$$Z_{\text{random}} = \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{\hat{p}-\frac{6}{49}}{\sqrt{\frac{\frac{6}{49}(1-\frac{6}{49})}{2500}}} \sim N(0,1)$$

此次假設檢定的拒絕域為 $|Z_{\text{observed}}| > 1.96$ $(=Z_{\text{critical}}$，$Z_{\frac{\alpha}{2}})$。

然後彩券公司從自選號碼簽注資料庫中隨機抽取 2500 張簽注單，發現有 321 張的號碼含有 36 號。故 $\hat{p} = 0.1284$，經計算

$$Z_{\text{observed}} = \frac{0.1284-\frac{6}{49}}{\sqrt{\frac{\frac{6}{49}(1-\frac{6}{49})}{2500}}} = 0.9077$$

因 $|0.9077| \leq 1.96$，檢定統計量之值未落入拒絕域內，所以，在 0.05 的顯著水準下，不拒絕虛無假設 H_0；也就是說，在 0.05 的顯著水準下，此 2500 張簽注單沒有充分證據顯示在民眾 (指母體) 自選號碼中，對 36 號有所好惡。

▷ 討論: 如何使用 p-value 法進行分析?

3 兩個母體比例檢定之例題

- 例題 1: 擴大兩岸交流在北、高兩市支持比例差異之檢定

行政院陸委會主委懷疑臺北市民支持擴大兩岸交流的比例 p_1 與高雄市民支持擴大兩岸交流的比例 p_2 不同，為證實其看法，因此欲進行假設檢定。而其虛無假設與對立假設設立為:

$$\begin{aligned}\mathrm{H}_0: &\quad p_1 = p_2\\ \mathrm{H}_1: &\quad p_1 \neq p_2\end{aligned}$$

陸委會主委設定顯著水準 α 為 0.05。

主委了解臺灣各縣市人民多數是贊同擴大兩岸交流的，但各縣市人民支持擴大兩岸交流的比例，都未達 0.8。主委乃決定從臺北市及高雄市分別隨機選取 200 位及 250 位市民，因 $0.5 < p < 0.8$ 且 $0.2 < 1-p < 0.5$，可知 $n_1p_1 = 200p_1 > 200 \times 0.5 = 100 > 5$ 且 $n_1(1-p_1) = 200(1-p_1) > 200 \times 0.2 = 40 > 5$ 且 $n_2p_2 = 250p_2 > 250 \times 0.5 = 125 > 5$ 且 $n_2(1-p_2) = 250(1-p_2) > 250 \times 0.2 = 50 > 5$，符合大樣本條件。故檢定統計量選擇為:

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{200}+\frac{1}{250}\right)}} \sim N(0,1)$$

此次假設檢定的拒絕域為 $|Z| > 1.96$。

一切就緒後，主委逐從臺北市及高雄市分別隨機選取 200 位及 250 位市民，發現 200 位臺北市民中有 138 位支持擴大兩岸交流，250 位高雄市民中有 152 位支持擴大兩岸交流；亦即臺北市民支持的樣本比例 $\hat{p}_1$ 之值為 0.69，高雄市民支持的樣本比例 $\hat{p}_2$ 之值為 0.608。再計算北、高兩市混和樣本比例 $\hat{p}$ 之值為 0.6444。代入檢定統計量內，得出

$$Z = \frac{0.69 - 0.608}{\sqrt{0.6444(1-0.6444)\left(\frac{1}{200}+\frac{1}{250}\right)}} = 1.8057$$

故 $|Z| \leq 1.96$。

又因屬雙尾檢定，可查標準常態分配機率表求取 p-value 為：

$$p\text{-value} = 2 \times \mathrm{P}(Z > 1.8057) = 0.071$$

因檢定統計量之值未落入拒絕域內，或因 $p\text{-value} > \alpha = 0.05$，所以在 0.05 的顯著水準下，不拒絕虛無假設 H_0；也就是說，在 0.05 的顯著水準下，此 200 位臺北市民及 250 位高雄市民的意向沒有充分證據顯示臺北市市民支持擴大兩岸交流的比例 p_1 與高雄市民支持擴大兩岸交流的比例 p_2 有所不同。