

# Chapter 1: Homework

711378912 蔡宜誼

03/06/2025

設計一個類似的問題，計算 Wald test 及 Score test 所造成的信賴區間，並執行二個檢定（大樣本及小樣本），報告其結果

## 作業舉例範例：工程師學歷碩士(含)以上比例

### 事實

根據產業統計，過去科技業的工程師擁有碩士學歷的比例約為 50%。

| 學歷  | 學士(含)以下 | 碩士(含)以上 | Total |
|-----|---------|---------|-------|
| 百分比 | 50      | 50      | 100   |

### 猜想

- 隨著科技業對高階技術人才需求增加，近年來碩士學歷可能更受重視。
- 如果選擇某科技公司內的工程師來調查，他們擁有碩士學歷的比例可能高於50%。
- 利用統計檢定來檢驗這個猜想
  - $H_0 : \pi_0 = 0.50$
  - $H_a : \pi_0 > 0.50$

### 實驗

- 調查併記錄公司工程師學歷比例

### 觀察結果：

- 參與調查人數為  $n$
- 擁有碩士學歷的人數為  $x$
- 碩士比例為  $x/n$

```
n <- 40; x <- 26 # 工程師總人數40人，26人為碩士(含)學歷以上
alpha <- 0.05
pihat <- x/n; pi0 <- 0.44
```

## 區間估計 (Confidence Interval)

- Wald Test

$$\hat{\pi} \pm z_{\alpha/2} \sqrt{\frac{\hat{\pi}(1 - \hat{\pi})}{n}}$$

```
lb <- pihat + qnorm(alpha/2)*sqrt(pihat*(1-pihat)/n)
ub <- pihat + qnorm(1-alpha/2)*sqrt(pihat*(1-pihat)/n)
cat("Wald 95% CI is [", lb, ub, "]\n")
```

```
## Wald 95% CI is [ 0.5021883 0.7978117 ]
```

- Score Test

$$\hat{\pi} \left( \frac{n}{n + z_{\alpha/2}^2} \right) + \frac{1}{2} \left( \frac{z_{\alpha/2}^2}{n + z_{\alpha/2}^2} \right) \pm \sqrt{\frac{1}{n + z_{\alpha/2}^2} \left[ \left( \frac{n\hat{\pi}(1 - \hat{\pi})}{n + z_{\alpha/2}^2} \right) + \left( \frac{z_{\alpha/2}^2}{4(n + z_{\alpha/2}^2)} \right) \right]}$$

```
z2 <- (qnorm(alpha/2))^2
mu <- pihat*n/(n+z2) + .5*(z2/(n+z2))
sig <- sqrt((n*pihat*(1-pihat)/(n+z2) + z2/4/(n+z2))/(n+z2))
lb <- mu - sig
ub <- mu + sig
cat("Score 95% CI is [", lb, ub, "]\n")
```

```
## Score 95% CI is [ 0.5645095 0.709204 ]
```

## 假設檢定

### Large sample approximation

- Check  $np \geq 5$  and  $n(1 - p) \geq 5$
- $H_0$  vs  $H_a$
- $\alpha = 0.05$
- Test statistics
  - Wald Test

$$\frac{\hat{\pi} - \pi_0}{\sqrt{\hat{\pi}(1 - \hat{\pi})/n}} = t_w$$

- Score Test

$$\frac{\hat{\pi} - \pi_0}{\sqrt{\pi_0(1 - \pi_0)/n}} = t_s$$

```
(pihat-pi0)/sqrt(pihat*(1-pihat)/n)
```

```
## [1] 2.784573
```

```
(pihat-pi0)/sqrt(pi0*(1-pi0)/n)
```

```
## [1] 2.675648
```

- Critical value 1.645 (one-tailed)
- Decision:
  - Walt Test: reject  $H_0$  if  $t_w > 1.645$
  - Score Test: regject  $H_0$  if  $t_s > 1.645$

### Small Sample Exact Test

- $p$ -value is

```
1 - pbinom(x-1, n, pi0)
```

```
## [1] 0.006029371
```

- mid  $p$ -value

```
1 - pbinom(x, n, pi0) + dbinom(x, n, pi0)/2
```

```
## [1] 0.004171611
```



## Summary

根據本研究的統計檢定結果，我們對某科技公司內工程師擁有碩士（含）以上學歷的比例是否高於 50% 進行檢驗。

### 區間估計（信賴區間）

- Wald 95% 信賴區間：[ 0.5021883, 0.7978117 ]
- Score 95% 信賴區間：[ 0.5645095, 0.709204 ]
- 由於這兩個信賴區間的下限均高於 50%（特別是 Score CI 下限為 0.5645），這顯示工程師擁有碩士學歷的比例很可能高於 50%。

### 假設檢定結果

- Wald Test:

$$\text{檢定統計量 } t_w = 2.784573 > 1.645 \cdot \text{拒絕 } H_0$$

- Score Test:

$$\text{檢定統計量 } t_s = 2.675648 > 1.645 \cdot \text{拒絕 } H_0$$

- Small Sample Exact Test:

$$p - \text{value} = 0.006029371 \text{ (小於顯著水準 } \alpha = 0.05) \cdot \text{拒絕 } H_0$$

$$\text{Midp} - \text{value} = 0.004171611 \text{ (更小)} \cdot \text{進一步支持拒絕 } H_0$$



## 結論

- 由於 Wald Test、Score Test 和 Small Sample Exact Test 皆拒絕虛無假設  $H_0$ ，並且信賴區間的下限高於 50%，我們有足夠證據證明此次調查公司工程師的碩士學歷比例顯著高於 50%。
- 表示碩士（含）以上學歷在該公司內部可能是較為普遍的學歷背景，反映出科技產業對高等教育的重視程度。
- 所以我現在在念碩士。