




11. 雙樣本檢定 Two-sample tests

鄒慶士 (Ching-Shih Tsou)
台北商業技術學院資訊與決策科學所
E-mail : cstsou@mail.ntcb.edu.tw



Two-sample tests of proportion

- The implementation of two-sample tests in R is similar to one-sample test but there are differences to be aware of.
- As before, we use the command `prop.test` to handle tests of proportion problems. We just need to learn when to use it and how.

© Vince Tsou, IDS, NTCB 100年度教育部補助技專校院建立特色典範計畫 2



Example: Two surveys

- A survey is taken two times over the course of two weeks. The pollsters wish to see if there is a difference in the results as there has been a new advertising campaign run.
- Here is the data:

	Week 1	Week 2
Favorable	45	56
Unfavorable	35	47



假設設定與prop.test函數

- The standard hypothesis test is
 $H_0 : P_1 = P_2$ (or $H_0 : P_1 - P_2 = 0$)
 against the alternative (two-sided)
 $H_1 : P_1 \neq P_2$ (or $H_1 : P_1 - P_2 \neq 0$)
- The function **prop.test** is used to being called as **prop.test(x,n)** where x is the number favorable and n is the total. Here it is no different, but since there are two x's it looks slightly different.



R 指令

- We observe that the p-value is 0.9172, so we accept the null hypothesis that $P_1 = P_2$.

```
R Console
> prop.test(c(45,56),c(45+35,56+47))

2-sample test for equality of proportions with continuity
correction

data:  c(45, 56) out of c(45 + 35, 56 + 47)
X-squared = 0.0108, df = 1, p-value = 0.9172
alternative hypothesis: two.sided
95 percent confidence interval:
 -0.1374478  0.1750692
sample estimates:
 prop 1    prop 2 
0.5625000 0.5436893 
> |
```

© Vince Tsou, IDS, NTCB

100年度教育部補助技專校院建
立特色典範計畫

5



Two-sample t-test

- one-sample t-test 的統計公式：

$$t = \frac{\bar{X} - \mu}{s / \sqrt{n}}$$

1. 使用在資料近似於常態分配時
2. σ 未知

- two-sample t-test 的統計公式：

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

假設 X_i 是常態或趨近於常態分配

© Vince Tsou, IDS, NTCB

100年度教育部補助技專校院建
立特色典範計畫

6



Equal Variances

- 當我們假設兩樣本之變異數相同，則兩組資料可以合併，估計共同的變異數。
- 參見下頁 t.test 線上說明檔。

© Vince Tsou, IDS, NTCB

100年度教育部補助技專校院建立特色典範計畫

7



Equal Variances

- t.test(x,

另一筆數值向量物件

– y = NULL,

對立假設

– alternative = c(“two.sided”, “less”, “greater”),

預設為two.sided

兩個樣本mean的差距

– mu = 0,

是否為成對樣本

– paired = FALSE,

變異數是否相等

– var.equal = FALSE,

信賴區間的信心水準

– conf.level = 0.95)

© Vince Tsou, IDS, NTCB

100年度教育部補助技專校院建立特色典範計畫

8



Example: Recovery time for new drug

- Suppose the recovery time for patients taking a new drug is measured (in days). A placebo group is also used to avoid the placebo effect. The data are as follows

with drug: 15 10 13 7 9 8 21 9 14 8

placebo: 15 14 12 8 14 7 16 10 15 12

© Vince Tsou, IDS, NTCB

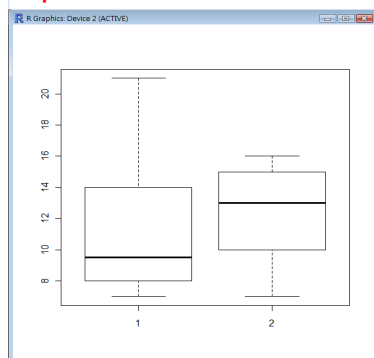
100年度教育部補助技專校院建
立特色典範計畫

9



Example: Recovery time for new drug

```
> x = c(15, 10, 13, 7, 9, 8, 21, 9, 14, 8)
> y = c(15, 14, 12, 8, 14, 7, 16, 10, 15, 12)
> boxplot(x,y)
> |
```



- 利用盒鬚圖決定
 - 變異數是否相同
(建議用var.test函數)
 - 資料是否服從常態
(呈偏態分佈)

© Vince Tsou, IDS, NTCB

100年度教育部補助技專校院建
立特色典範計畫

10



Example: Recovery time for new drug

$$H_0: \mu_1 - \mu_2 \geq 0$$

$$H_1: \mu_1 - \mu_2 < 0$$

```
> x = c(15, 10, 13, 7, 9, 8, 21, 9, 14, 8)
> y = c(15, 14, 12, 8, 14, 7, 16, 10, 15, 12)
> boxplot(x,y)
> t.test(x,y,alt="less",var.equal=TRUE)
```

Two Sample t-test

```
data: x and y
t = -0.5331, df = 18, p-value = 0.3002
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
 -Inf 2.027436
sample estimates:
mean of x mean of y
 11.4      12.3
```

- 根據此結果，我們無法拒絕虛無假設。



Unequal Variances

```
> t.test(x,y,alt="less")
```

Welch Two Sample t-test

```
data: x and y
t = -0.5331, df = 16.245, p-value = 0.3006
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
 -Inf 2.044664
sample estimates:
mean of x mean of y
 11.4      12.3
```

- 根據此結果，我們無法拒絕虛無假設。



Comparisons

Equal variances

```
> t.test(x,y,alt="less",var.equal=TRUE)
```

Two Sample t-test

data: x and y

t = -0.5331, df = 18, p-value = 0.3002

- 變異數相等與不相等的結果有些不同，但此例中結論是相同的。(無法拒絕 H_0)

Unequal variances

```
> t.test(x,y,alt="less")
```

Welch Two Sample t-test

data: x and y

t = -0.5331, df = 16.245, p-value = 0.3006

- 當變異數相等時，抽樣分配之t分配自由度較大，愈接近標準常態，兩尾面積較小，p值因此較小，虛無假設較容易被拒絕。

© Vince Tsou, IDS, NTCB

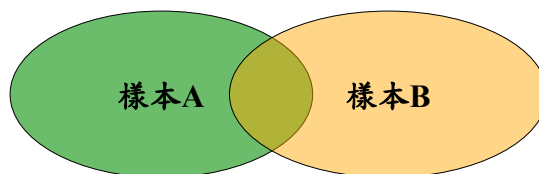
100年度教育部補助技專校院建立特色典範計畫

13



成對樣本(Matched Samples)

- 雙樣本t檢定(two sample t-test)之成對t檢定(paired t-test)，使用不同於獨立性t檢定的統計模型。
- 成對樣本，假設兩樣本中存在有共同的特點。



© Vince Tsou, IDS, NTCB

100年度教育部補助技專校院建立特色典範計畫

14



Basic model

- $Y_i = X_i + \varepsilon_i$ ，其中 ε_i 是隨機變數。

檢驗：H0: 平均數=0，H1: 平均數 \neq 0。

將 Y_i 減去 X_i ，進行單一樣本 t 檢定。

指令：`t.test(x,y,paired=TRUE)`



Example: Dilemma of two graders

- 為了公平起見，每位應徵者都接受兩位評分員評分。資料如下(閱同一份試卷，所以非獨立樣本)：

Appl. : 1 2 3 4 5 6 7 8 9 10

Grader1: 3 0 5 2 5 5 5 4 4 5

Grader2: 2 1 4 1 4 3 3 2 3 5



Example: Dilemma of two graders

```
RGui - [R Console]
檔案 編輯 視 其它 程式套件 視窗 輔助

> x=c(3,0,5,2,5,5,5,4,4,5)
> y=c(2,1,4,1,4,3,3,2,3,5)
> t.test(x,y,paired=TRUE)

Paired t-test

data: x and y
t = 3.3541, df = 9, p-value = 0.008468
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.3259550 1.6744450
sample estimates:
mean of the differences
1
```

> | 成對 t 檢定結果：
 $T = 3.541$, $df = 9$, $p\text{-value} = 0.008468 < 0.05$
 H_0 : 假設平均數差 = 0 表評分員對應徵者無偏見 拒絕 H_0
 H_1 : 假設平均數差 $\neq 0$ 表評分員對應徵者有偏見

95%的信賴區間
 樣本估計: 平均數的差異是1

© Vince Tsou, IDS, NTCB

100年度教育部補助技專校院建
立特色典範計畫

17



Example: Dilemma of two graders

```
RGui - [R Console]
檔案 編輯 視 其它 程式套件 視窗 輔助

> x=c(3,0,5,2,5,5,5,4,4,5)
> y=c(2,1,4,1,4,3,3,2,3,5)
> t.test(x,y)

Welch Two Sample t-test

data: x and y
t = 1.478, df = 16.999, p-value = 0.1577
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.4274951 2.4274951
sample estimates:
mean of x mean of y
3.8 2.8
```

> | 兩獨立樣本 t 檢定結果
 $t = 1.478$, $df = 16.999$, $p\text{-value} = 0.1577 > 0.05$
 H_0 : 假設平均數差 = 0 接受 H_0
 H_1 : 假設平均數差 $\neq 0$

95%的信賴區間
 樣本估計值: x的平均數=3.8 y的平均數=2.8

© Vince Tsou, IDS, NTCB

100年度教育部補助技專校院建
立特色典範計畫

18



Resistant two-sample tests

- 抗雜訊雙樣本檢定使用Wilcox的檢定函數(`wilcox.test()`)，用法與單一樣本中位數檢定相似。
- Example : Taxi out times
 - 比較美國和西北航空公司在Newark機場的降落滑行時間，資料集為`ewr`。

© Vince Tsou, IDS, NTCB

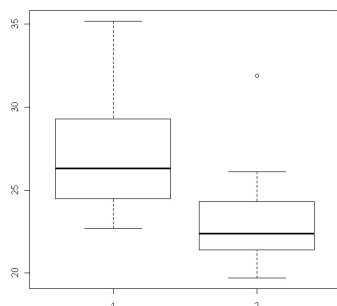
100年度教育部補助技專校院建立特色典範計畫

19



先觀察資料分佈

```
> data(ewr)
> attach(ewr)
> tmp=subset(ewr, inorout == "out", select=c("AA", "NW"))
> x=tmp[["AA"]]
> y=tmp[["NW"]]
> boxplot(x, y)
> |
```



- 呈偏態分配
- 使用Wilcox中位數檢定

© Vince Tsou, IDS, NTCB

100年度教育部補助技專校院建立特色典範計畫

20



Wilcoxon Test 檢定結果

```
> wilcox.test(x,y)

Wilcoxon rank sum test with continuity correction

data: x and y
W = 460.5, p-value = 1.736e-05
alternative hypothesis: true location shift is not equal to 0

Warning message:
In wilcox.test.default(x, y) : cannot compute exact p-value with ties
> |
```

無法很準確計算出p值

- 由wilcox.test得知，有強烈的證據可拒絕虛無假設，即接受中位數不相等之對立假設。

The End

