

Lecture 3

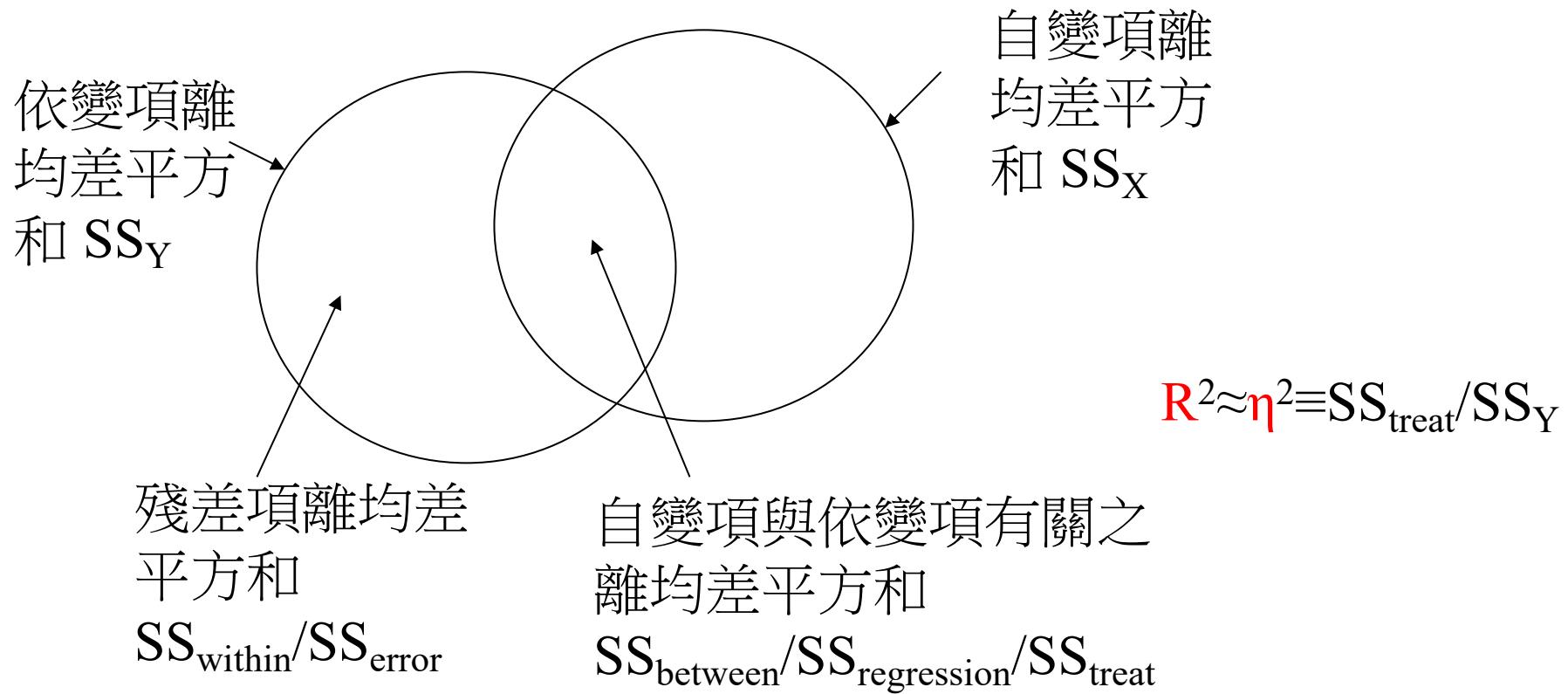
Analysis of variance, ANOVA

變異數分析_續

Instructor: 吳俊育

One-way ANOVA

- VENN DIAGRAM REPRESENTATION



ANOVA TABLE QUIZ

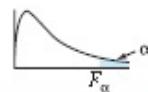
SOURCE	DF	SS	MS	F	PROB
--------	----	----	----	---	------

B:GROUP	2	50	.05
---------	---	----	-----

W:ERROR	—	—	—
---------	---	---	---

TOTAL	20	$\eta^2 =$	—
-------	----	------------	---

TABLE 8

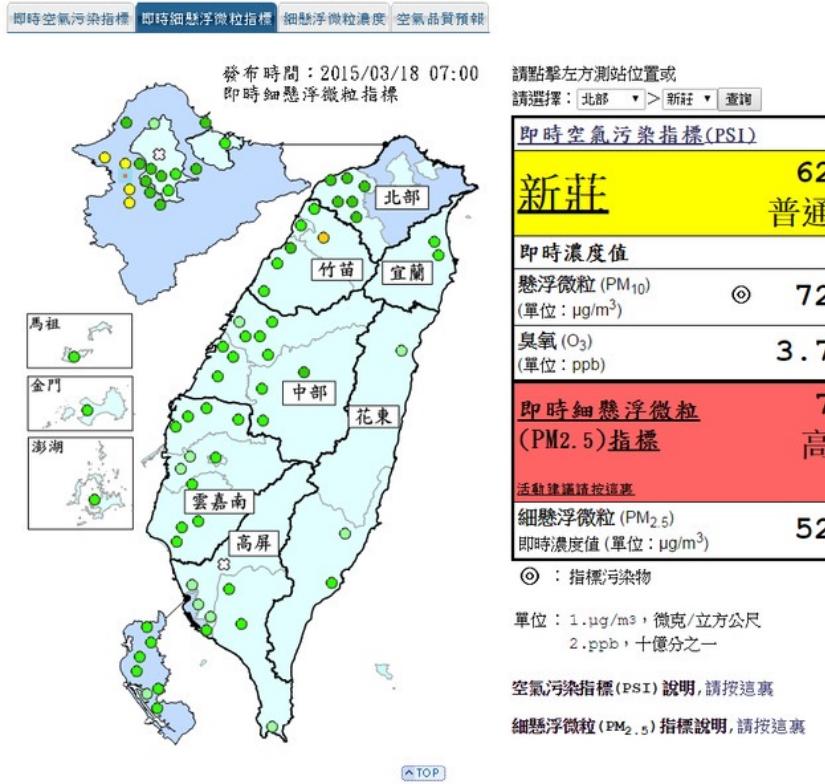
Percentage points of the F distribution (df_2 between 13 and 18)

df ₂	α	df ₁									
		1	2	3	4	5	6	7	8	9	10
18	.25	1.41	1.30	1.49	1.48	1.46	1.45	1.44	1.43	1.42	1.42
	.10	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98
	.05	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
	.025	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87

相依樣本變異數分析 (Repeated Measures ANOVA)

- 相依樣本設計 (correlated sample design)
 - 重複量數設計(repeated measures design; RM)
 - 指同一組受試者(Subject)重複接受不同實驗處理(treatment)或進行多次測量(occasion)
 - 適用於以上情境之變異數分析，稱為RM ANOVA。
 - 配對樣本設計(randomized blocks design, RB)
 - 指具有配對關係的樣本接受不同的實驗處理，每一組配對稱為一個區組 (block)
 - 如加上隨機分派，又稱為隨機區組設計 (randomized blocked design)
 - 適用於以上情境之變異數分析，稱為RB ANOVA
- Scores for the same individual/block are dependent, whereas the scores for different individuals/blocks are independent.

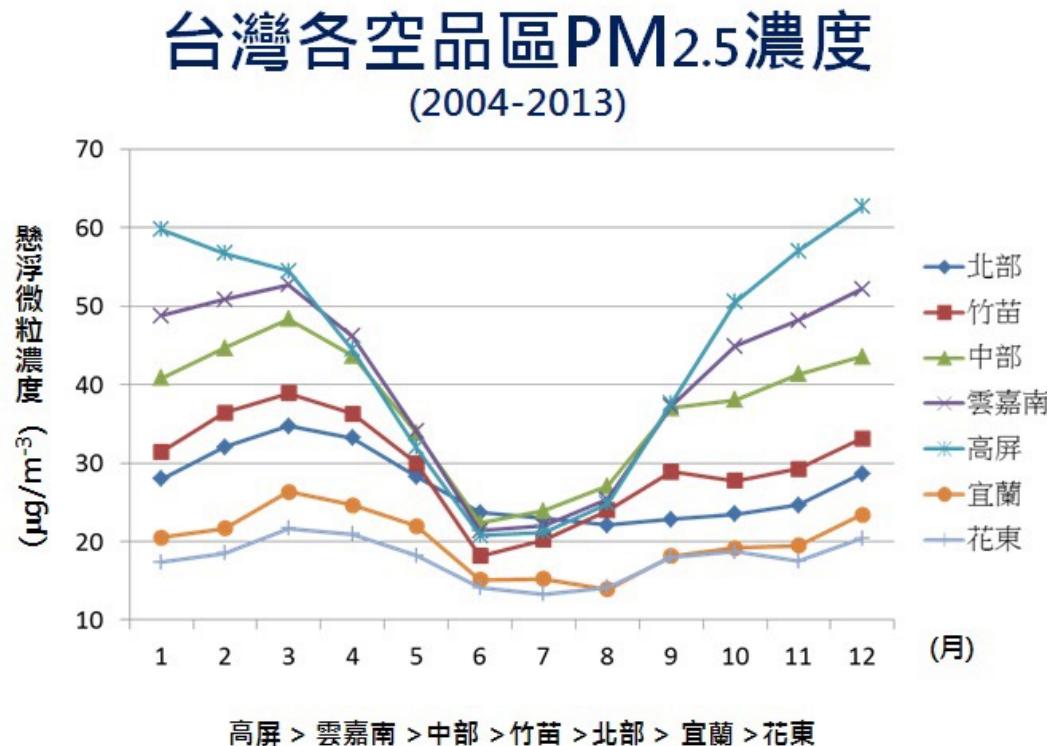
全台唯一紅色警示 新北新莊細懸浮微粒濃度高



民眾今天出門最好還是戴著口罩。

根據「環保署」即時細懸浮微粒指標監測結果，新北市新莊區細懸浮微粒濃度為7，是全台灣唯一紅色警戒。、及示還該戴著口、及示還該戴著地、縣地、新北市、新北市、城、黃門、色最好、是出眾、葉國吏/綜合報導

每月平均PM2.5濃度也有南北差距？



<https://www.thenewslens.com/feature/airpollution/19328>

單因子相依樣本設計的資料形式

- 受測者間效果(Between Subject effect: Subject/Block)：
 - 各區組(Subject/Block)平均數
 - 橫列上每一個區組的平均數
 - 反應該區組的平均水準，也就是區組同質性所造成；亦即在特定受測者(Subject)下，在自變項上所有場景的均值高低
- ☺ 受測者內因子(Within Subject effect, treatment: Occasion)：
 - 各場景(Occasion)平均數
 - 直行上自變項各水準下的平均數
 - 表示不同水準或不同水準實驗效果，也就是特定場景(Occasion)的同質性所造成；亦即在自變項上特定場景下，所有受測者得分的均值高低
- 細格(Cell)效果
 - 每一個細格只有一個觀察值，因此沒有細格內變異，沒有交互效果
 - 細格間的變異視為隨機誤差

Between Subject factor	Within Subject factor				Mean of Subjects (Between Subject mean)
	Occasion 1	Occasion 2	...	Occasion p	
Subject 1	Y_{11}	Y_{12}	...	Y_{1p}	\bar{Y}_1
Subject 2	Y_{21}	Y_{22}	...	Y_{2p}	\bar{Y}_2
:	:	:	Y_{ij}	:	:
Subject n	Y_{n1}	Y_{n2}	...	Y_{np}	\bar{Y}_n
Mean of Occasion (Within Subject mean)	$\bar{Y}_{.1}$	$\bar{Y}_{.2}$...	$\bar{Y}_{.p}$	$\bar{Y}_{..}$

單因子相依樣本設計的資料形式

Ex：4名學生接受三次統計態度量表施測，請問學生在三次施測得分是否有顯著的差異 (i.e. 人/區組在受測者內因子的各水準均值是否有統計顯著差距)

受測者	施測			
	前測1	後測2	延宕測3	
1	2	4	9	5
2	10	12	14	12
3	22	29	30	27
4	29	31	33	31
☺	15.75	19	21.5	

受測者內因子內的水準，在相依樣本資料裡，又可被稱為場景(occasion)

總觀察值 $N=12$,

水準數: $p = 3$, 受測者/區組數 $n = 4$,

受測者間:區組:各受測者/各群組

受測者內:自變項A:統計態度量表(因子)

因子內的水準: 不同施測($p=3$)

細格效果: $Y_{ij} = \bar{Y}_{..} + \bar{Y}_{i.} + \bar{Y}_{.j} + e_{ij}$

第*i*個人在第*j*個場景內所出現的隨機殘差/誤差
(error/residual)

第*i*個人在第*j*個場景的觀察值

第*i*個人在所有場景內觀察值平均值

第*j*個場景內所有人的觀察值平均值

Advantage of using RM-ANOVA

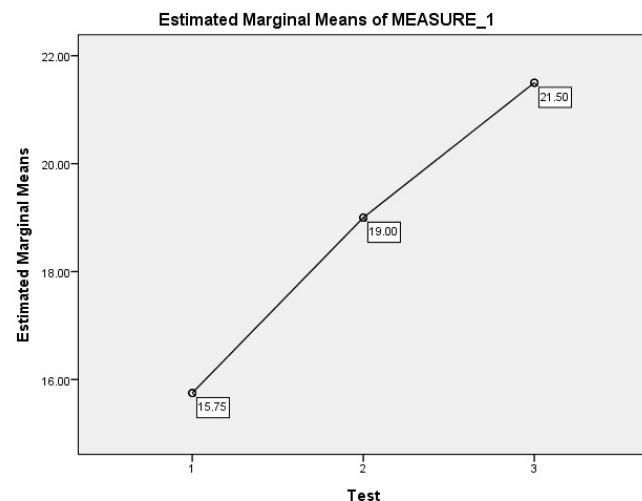
- 因為重複量測分數之間相關性高，所以必須使用 Repeated Measures 分析技巧以檢驗均值差異。
- $H_0: \mu_{\text{test1}} = \mu_{\text{test2}} = \mu_{\text{test3}}$
- RM-ANOVA：
 - Reduce overall variability by using a common subject pool for all treatments
 - Remove subject differences from the error term, leaving the error components independent from occasion to occasion or cell to cell.

e.g. $SS_r \perp SS_{B.\text{subject}}$

		Correlations		
		Test1	Test2	Test3
Test1	Pearson Correlation	1	.984*	.985*
	Sig. (2-tailed)		.016	.015
	N	4	4	4
Test2	Pearson Correlation	.984*	1	.996**
	Sig. (2-tailed)	.016		.004
	N	4	4	4
Test3	Pearson Correlation	.985*	.996**	1
	Sig. (2-tailed)	.015	.004	
	N	4	4	4

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).



RM-ANOVA, 基本假設

- 常態性(normality): 依變項樣本的母體為常態分配。
- 變異數同質性(homogeneity of variance): 每組樣本內的變異數均相等。(SPSS: Levene檢定)
- 抽自各母體的各組隨機樣本互為獨立，變異數的加減具獨立且可加性。
- 球面性假設 (sphericity)
 - 適用於相依樣本變異數分析的基本假設。係指同一組樣本在不同水準上的依變項得分，兩兩配對相減後的差的變異數必須相等 (homogeneity of difference variance)。也就是說，不同的受試者在不同水準間配對或重複測量，其變動情形應具有一致性。

Mauchly 球形檢定^b

測量:MEASURE_1

受試者內效應項	Mauchly's W	近似卡方分配	df	顯著性	Epsilon ^a		
					Greenhouse-Geisser	Huynh-Feldt	下限
測驗	.126	5.649	5	.369	.667	1.000	.333

檢定正交化變數轉換之依變數的誤差共變量矩陣的虛無假設，是識別矩陣的一部份。

a. 可用來調整顯著性平均檢定的自由度。改過的檢定會顯示在 "Within-Subjects Effects" 表檢定中。

b. Design: 裁距

受試者內設計: 測驗

單因子相依樣本設計的資料形式

Ex：五名學生參加四種科學素養測驗結果，學生在四種測驗的成績是否有顯著的差異 (i.e. 人/區組在受測者內因子的各水準均值是否有統計顯著差距)

總觀察值 $N=20$,

水準數: $p = 4$, 受測者/區組數 $n = 5$,

受測者間: 區組: 各受測者/各群組

受測者內: 自變項 A: 科學素養測驗(因子)

因子內的水準: 不同測驗($p=4$)

細格效果: $Y_{ij} = \bar{Y}_{..} + \bar{Y}_{i.} + \bar{Y}_{.j} + e_{ij}$

第*i*個人在第*j*個場景
的觀察值

第*i*個人在所有場景
內觀察值平均值

第*j*個場景內所有人的
觀察值平均值

第*i*個人在第*j*個場景
內所出現的隨機殘
差/誤差
(error/residual)

以母體表示RM-ANOVA的模型

$$\mu + \pi_{i.} + \pi_{.j} + e_{ij}$$

- μ : grand mean
(average subject + average occasion)
- $\pi_{i.}$: the deviation with i th subject from the average subject
- $\pi_{.j}$: the deviation with j th occasion from the average occasion
- e_{ij} : the experimental error with i th subject under the j th occasion.

A	B	C	D	E	F
受試者	測驗一	測驗二	測驗三	測驗四	Block mean (between subject)
1	16.6	15.8	14.7	14.4	15.375
2	16.9	15.6	15.1	13.8	15.35
3	17.1	16.1	14.8	13.9	15.475
4	17.2	16.3	15.2	14.1	15.7
5	16.8	15.7	15.3	14.3	15.525
mean of 測驗	16.92	15.9	15.02	14.1	15.485

細格效果($Y_{ij} = \bar{Y}_{..} + \bar{Y}_{i.} + \bar{Y}_{.j} + e_{ij}$)的變異量可以被分解為:

$$SS_T = SS_{B,Subject} + SS_A + SS_r$$

所有人在所有場景
的離均差平方和

所有人的離均差平
方和

隨機殘差/誤差的離
均差平方和

所有場景的離均差
平方和

單因子相依樣本ANOVA

• VENN DIAGRAM REPRESENTATION

$$SS_T = SS_{B.\text{Subject}} + SS_A + SS_r$$

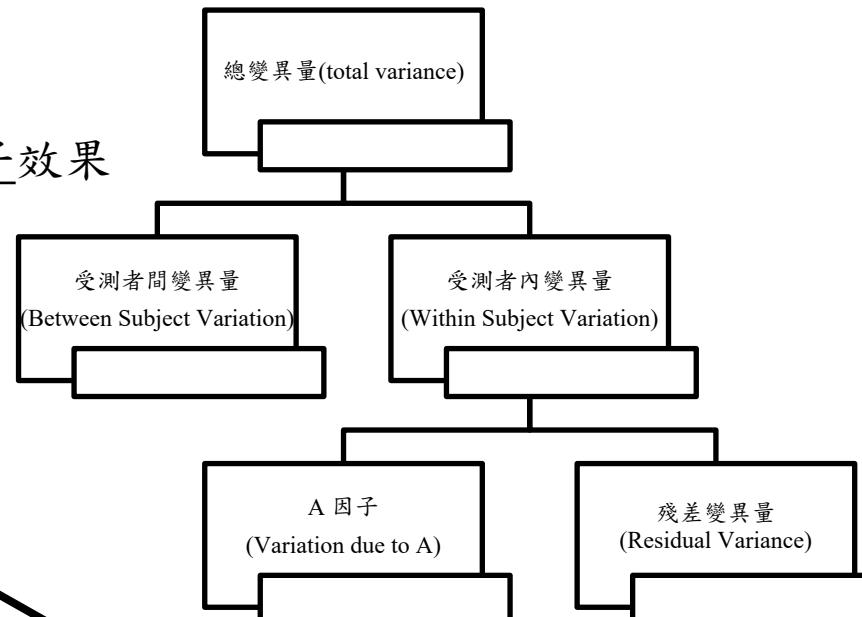
依變項離均差
平方和 SS_T

殘差項離均差平方和
 $SS_{\text{within}} / SS_{\text{residual}} / SS_{\text{error}}$

人/區組均差平方和
 $SS_{B.\text{subject}}$

人/區組的效果

受測者內因子效果
 $SS_{W.\text{subject}}$



受測者內因子(自變項)與依
變項有關之離均差平方和
 SS_{treat} / SS_A

單因子相依樣本ANOVA F考驗與摘要表

總分數
SS

$$SS_T = SS_{total} = \sum_{i=1}^n \sum_{j=1}^p (x_{ij} - \bar{x}_{total})^2 = \sum_{i=1}^n \sum_{j=1}^p x_{ij}^2 - \frac{\left(\sum_{i=1}^n \sum_{j=1}^p x_{ij} \right)^2}{N} = \sum_{i=1}^n \sum_{j=1}^p x_{ij}^2 - \frac{G^2}{N}$$

人/區組間SS

$$SS_{B.Subject} = \sum_{i=1}^n \frac{\left(\sum_{j=1}^p x_{ij} \right)^2}{p} - \frac{G^2}{N}$$

人/區組內SS

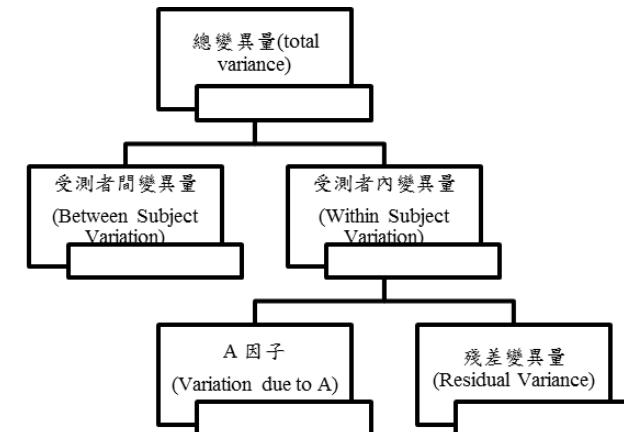
$$SS_{W.Subject} = SS_T - SS_{B.Subject}$$

場景/水準SS

$$SS_{Occasion} = \sum_{j=1}^p \frac{\left(\sum_{i=1}^n x_{ij} \right)^2}{n} - \frac{G^2}{N}$$

隨機誤差/殘差SS

$$SS_r = SS_{W.Subject} - SS_{Occasion}$$



卦画(Source)	(SS)	(df)	(MS)	F 保
體 (Between Subject)	$SS_{Subject}$	$n - 1$		
體 (Within Subject)	$SS_{W.Subject}$	$n(p - 1)$		
體 (Occasion)	$SS_{Occasion}$	$p - 1$	$MS_{Occasion} = \frac{SS_{Occasion}}{df_{Occasion}}$	$F_{Occasion} = \frac{MS_{Occasion}}{MS_r}$
體 (Residual/error)	SS_r	$(n - 1)(p - 1)$	$MS_r = \frac{SS_r}{df_r}$	
卦 (Total)	SS_T	$N - 1$		

Test if
 $H_0: \mu_{test1} = \mu_{test2} = \mu_{test3}$

Further Reading

- Winer, B. J., Brown, D. R., & Michels, K. M. (1991). *Statistical principles in experimental design*. McGraw-Hill.
- 林清山(2003)。心理與教育統計學。台北：東華。Ch14-2。ISBN 9789576365102

效應值與事後比較

AFTER ANOVA : EFFECT SIZE & POST HOC TESTS

關聯強度分析

- 統計顯著性 (statistical significance)
 - 基於機率理論，提供『自變項效果相對於隨機變化的強度檢驗』統計詮釋
 - 例如，以F考驗來決定自變項效果的統計意義
- 實務顯著性 (practical significance)
 - 反應自變項效果在環境中強度的真實意義
 - 常用 ω^2 、 η^2 、 D 量數表示
 - 也稱為臨床顯著性 (clinical significance)

效果量係數，Effect Size

- 效果量係數
 - 用來衡量自(獨)變項強度的統計量。
- D 量數
 - 最簡單的效果量
 - 兩組平均數之標準化差異程度
 - 標準化差異越大，表示自變項可解釋依變項變異程度越強(亦即，自變項之不同_____均值差異越____)
- f 量數
 - 適用於當平均數數目大於2時

$$D = \frac{\mu_1 - \mu_2}{\sigma_\varepsilon}$$

$$f = \sqrt{\frac{\sum_{j=1}^p \frac{(\mu_j - \mu)^2}{p}}{\sigma_\varepsilon^2}} = \sqrt{\frac{\sum_{j=1}^p \frac{\alpha_j^2}{p}}{\sigma_\varepsilon^2}}$$

$$\hat{f} = \sqrt{\frac{\sum_{j=1}^p \frac{(\bar{Y}_j - \bar{Y}_G)^2}{p}}{\hat{\sigma}_\varepsilon^2}} = \sqrt{\frac{\sum_{j=1}^p \frac{\hat{\alpha}_j^2}{p}}{\hat{\sigma}_\varepsilon^2}} = \sqrt{\frac{(p-1)(MS_b - MS_w)}{np / MS_w}}$$

η^2 (eta square) 量數

- η^2 為ANOVA分析中最常用之效果量(Effect Size)的指標
 - 可類比為迴歸分析中的 R^2 (為 X 對 Y 解釋強度的指標)

$$\hat{\eta}^2 = \frac{SS_B}{SS_T}$$

- 樣本數小時，為母體的偏估計數，需以下式進行調整，以得到不偏估計數 (Wherry, 1931)

$$\tilde{\eta}^2 = 1 - \frac{N-1}{N-p} (1 - \hat{\eta}^2)$$

- 淨 η^2 (partial η^2) 量數
 - 扣除了其他效果項的影響後的關聯強度量數

$$partial \ \hat{\eta}^2 = \frac{SS_B}{SS_B + SS_W}$$

ω^2 (omega square) 量數

- 類似於迴歸分析的 R^2

- 定義式

$$\omega^2 = \frac{\sigma_B^2}{\sigma_B^2 + \sigma_W^2}$$

- ω^2 量數

- 為組間變異與總變異的比值
 - 表示依變項變異量能被自變項解釋的百分比
 - 亦即自變項與依變項的關聯強度

- 樣本估計式：

$$\hat{\omega}^2 = \frac{SS_b - (p-1)MS_w}{SS_{total} + MS_w} = \frac{(p-1)(F-1)}{(p-1)(F-1) + np}$$

Rule of Thumb of Effect Sizes

	Small	Medium	Large	Very big
Cohen's d	0.2	0.5	0.8	2
r (correlation coefficient)	0.1	0.3	0.5	-
R^2 : % of variance explained	1 %	9 %	25 %	-
R^2 : (\approx Partial eta-squared)	0.02	0.13	0.26	-
Generalized eta-squared η^2	0.02	0.13	0.26	-
Partial eta-squared	0.01	0.06	0.14	-
Omega squared	0.059	0.138	>0.138	-

References for Effect Size Measures

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, 37(3), 379–384.

<https://doi.org/10.3758/BF03192707>

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Earlbaum Associates.

Levine, T. R., & Hullett, C. R. (2002). Eta squared, partial eta squared, and misreporting of effect size in communication research. *Human Communication Research*, 28(4), 612–625.

<https://doi.org/10.1111/j.1468-2958.2002.tb00828.x>

Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, 6(2), 135–147. <https://doi.org/10.1016/j.edurev.2010.12.001>

One-way ANOVA的結論

- 當 F 大於事先設定可容許第一型錯誤大小 (e.g. $\alpha=0.05$) 的危險檢定值時，我們就結論：並不是因子內所有的水準平均值都是一樣的！(組別來自於不同的母體)

$$F = \frac{MS_B}{MS_W} \geq F_\alpha(G-1, n-G) \quad \rightarrow \quad \begin{cases} H_0: \mu_1 = \mu_2 = \dots = \mu_G \\ H_1: \text{至少有兩組}\mu\text{不相等} \end{cases}$$

- 也就是說、類別變數（因子(way), e.g. 教學法）的主要效果（main effect）裡的四個組別（水準(Level), e.g. ABCD）有統計顯著的差異存在(statistically significant difference among levels)
- 所以，ANOVA是一種檢驗因子之主要效果的整體性檢定(Omnibus test)。我們想知道因子內的水準平均值是否統計上相同。

因子之主要效果

- 固定效果模式 (fixed effect model)
 - 當一個研究的自變項（因子）的水準個數（g組），包括了該變項所有可能的水準數（G組），也就是樣本的水準數等於母體的水準數 ($g = G$) 。
 - 例如比較大學四個年級學生的曠課次數，此時自變項為年級，具有四個水準，而母體亦為四個年級。
- 隨機效果模式 (random effect model)
 - 研究所取用的自變項，只包含特定的一些水準，而並非包括所有可能的類別，即樣本的水準數小於母體的水準數 ($g < G$) 。
 - 例如教育學者比較不同地區的學校教學方法的成效有所不同，因此隨機選取幾個地區的一些學校共四所（自變項），該研究所關心的四個水準，可以說是隨機自教學方法的母體中，隨機取用得來的。

因子之主要效果

**固定效果模式（ fixed effect model ）與 隨機效果模式
（ random effect model ）**

- If a factor is fixed we would be interested in testing the equality of the population means for those levels. If a factor were random, we would be interested in testing whether or not the variance of the effects of the factor levels is zero for the range of levels of interest. (Ryan, 2007)

One-way ANOVA的結論

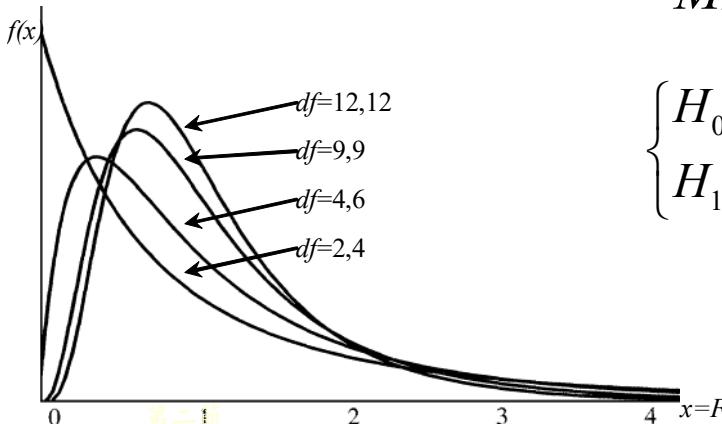
- 兩個變異數的比值稱為F統計量

$$F = \frac{\hat{\sigma}_b^2}{\hat{\sigma}_w^2} = \frac{MS_B}{MS_W} = \frac{SS_B/df_B}{SS_W/df_W}$$

- F統計量的機率分配為F分配
 - F 值越大，表示研究者關心的組平均數的分散情形較隨機誤差的變異來得大
 - 若大於臨界值，研究者即可獲得拒絕H0的結論

$$F = \frac{MS_B}{MS_W} \geq F_{\alpha}(G-1, n-G)$$

$\begin{cases} H_0: \mu_1 = \mu_2 = \dots = \mu_G \\ H_1: \mu_g \text{ 不全相等} \end{cases}$
Reject!



當 F 大於事先設定可容許第一型錯誤大小
(e.g. $\alpha=0.05$)的危險檢定值時 → 並不是所有的因子內水準的平均值都是一樣的！

One-way ANOVA的結論

$$\begin{cases} H_0: \mu_1 = \mu_2 = \dots = \mu_G \\ H_1: \text{至少有兩組}\mu\text{不相等} \end{cases}$$

- 回到ANOVA檢定的最開始，我們是相信虛無假設的，也就是說：「在沒有足夠的統計證據可以拒絕虛無假設前，我們相信因子內不同水準的平均數是來自同一個母體的」

Ex. A、B、C、D四種教學法做數學成就表現測試

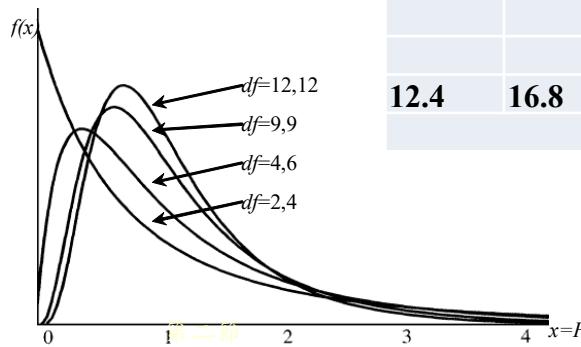
教學法			
A	B	C	D
12	16	12	16
16	20	15	16
10	14	18	15
14	18	15	18
10	16	16	12
		15	14
			15
12.4	16.8	15.17	15.17
		14.91	



ID	組別	分數
1	1	12.00
2	2	16.00
3	3	10.00
4	4	14.00
5	5	10.00
6	6	16.00
7	7	20.00
8	8	14.00
9	9	18.00
10	10	16.00
11	11	12.00
12	12	15.00
13	13	18.00
14	14	15.00
15	15	16.00
16	16	15.00
17	17	16.00
18	18	16.00
19	19	15.00
20	20	18.00
21	21	12.00
22	22	14.00
23	23	15.00



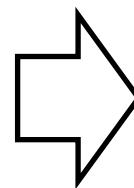
$$F = \frac{MS_B}{MS_W} \geq F_{\alpha}(G-1, n-G)$$



One-way ANOVA的結論

- 一旦主要效果被證明是具有顯著差異，我們就可以針對不同的水準來比較兩兩配對的均值比較，以進行下一步的討論→這些細部的討論，就是所謂的事後比較（Post-Hoc comparison）（e.g. A v.s. B, B v.s. C...）

$$\begin{cases} H_0 : \mu_1 = \mu_2 = \dots = \mu_G \\ H_1 : \text{至少有兩組 } \mu \text{ 不相等} \end{cases}$$



$$\text{test1} = \begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{cases}$$

$$\text{test2} = \begin{cases} H_0 : \mu_2 = \mu_3 \\ H_1 : \mu_2 \neq \mu_3 \end{cases}$$

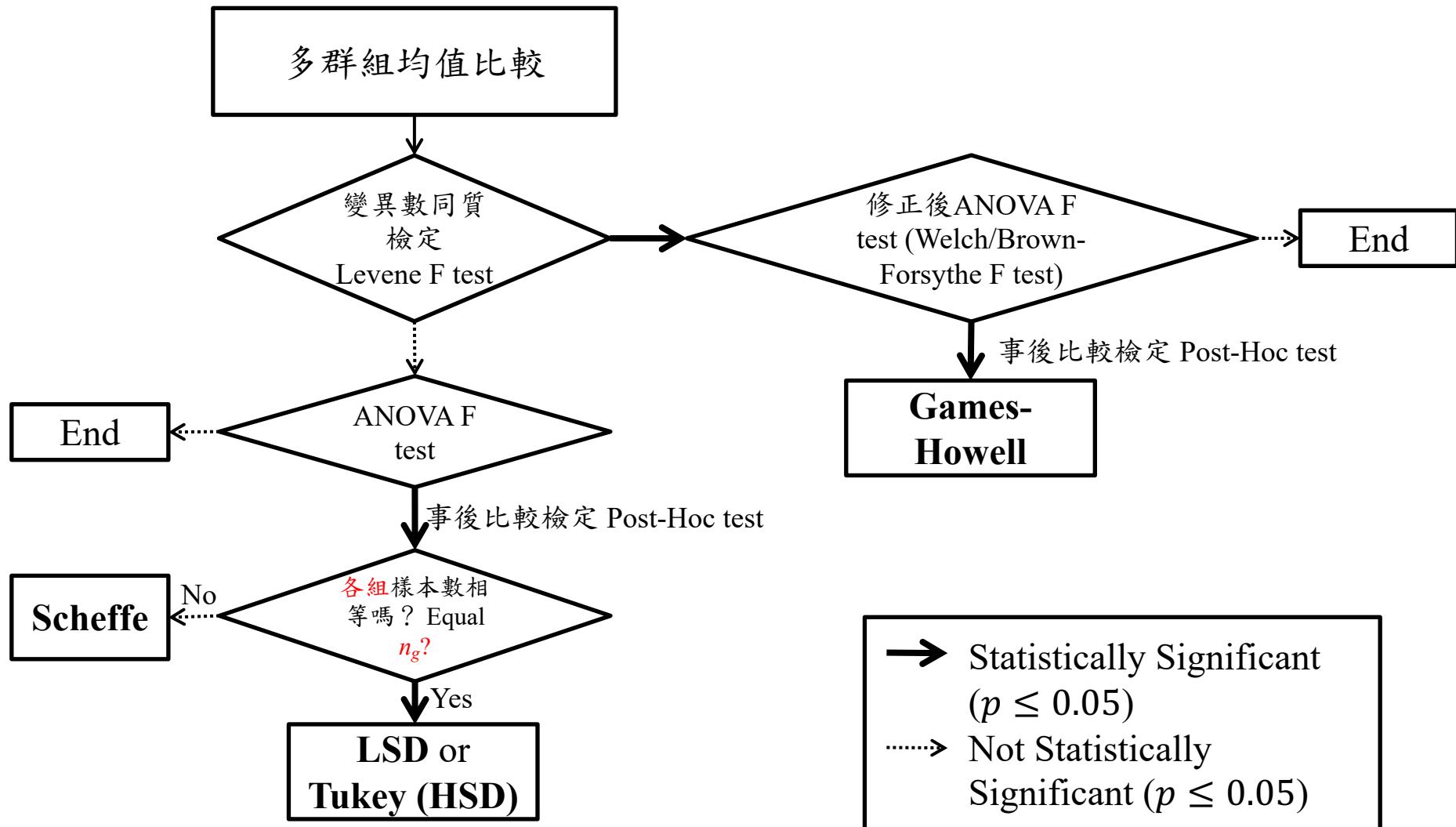
$$\text{test3} = \begin{cases} H_0 : \mu_1 = \mu_3 \\ H_1 : \mu_1 \neq \mu_3 \end{cases}$$

事後比較（Post-Hoc comparison）



Note: 實線為建議使用的事後檢定法

事後比較 (Post-Hoc comparison)



變異數同質檢定

變異數同質檢定

$$H_0 : \sigma_1^2 = \dots = \sigma_g^2$$

- 變異數同質檢定的方法：
 - Bartlett's test (1949)
 - Hartley's F-max test (1950)
 - Levene F test (1960)
- Bartlett法適用於當資料分佈是常態分佈時。但它對資料分佈型態非常敏感，資料非常態常會造成錯誤的統計顯著推論。依據 Conover etc. (1981), Levene檢定法較不易受資料分佈影響，所以 SPSS提供 Levene F 檢定法。

變異數同質檢定

- 變異數同質檢定

$$H_0 : \sigma_1^2 = \dots = \sigma_g^2$$

變異數同質性檢定

數學成就

Levene 統計量	分子自由度	分母自由度	顯著性
.573	3	19	.639

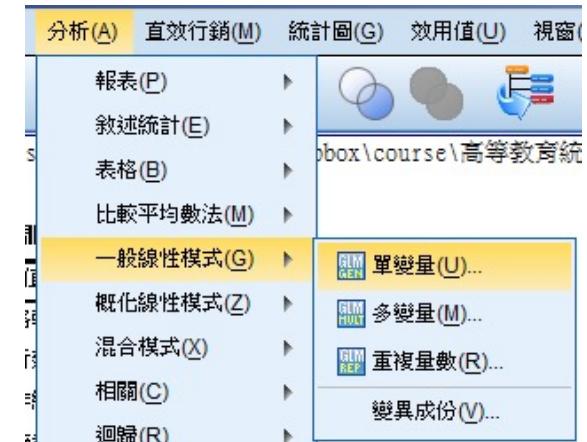
誤差變異量的 Levene 檢定等式^a

依變數：數學成就

F	df1	df2	顯著性
.573	3	19	.639

檢定各組別中依變數誤差變異量的虛無假設是相等的。

a. Design: 截距 + 組別



事後比較

- 變異數同質時（當各組樣本數相同時）
 - Tukey's HSD法:將所有的配對比較視為一體，使整個研究的第一類型錯誤維持衡定，第一類型錯誤是一種實驗誤差（experiment-wise error）
 - Fisher's LSD法又稱為Fisher擔保t檢定（Fisher's protected t-test），表示這個t檢定是以 F 考驗達到顯著之後所進行的後續考驗，同時也在 F 考驗的誤差估計下所進行

$$t = \frac{\bar{Y}_j - \bar{Y}_k}{\sqrt{s_p^2 \left(\frac{1}{n_j} + \frac{1}{n_k} \right)}} = \frac{\bar{Y}_j - \bar{Y}_k}{\sqrt{MS_{within} \left(\frac{1}{n_j} + \frac{1}{n_k} \right)}}$$

HSD法

- HSD法原理

- 在常態性、同質性假設成立下，各組人數相等的一種以族系誤差率的控制為原則的多重比較程序
- 稱為誠實顯著差異（Honestly Significant Difference）
- 所謂誠實，就是在凸顯LSD法並沒有考慮到實驗與族系面誤差的問題
- 代價是降低了統計考驗的檢定力。以HSD法所得到的顯著性，會比沒有考慮型一錯誤膨脹問題的檢定方法來的高（例如如果比較次數為三次，HSD的

值

為會是LSD法的三倍）
- Kramer則將Tukey的方法加以延伸至各組樣本數不相等的情況下，由於原理相同，故合稱為Tukey-Kramer法

$$Q = \frac{\bar{Y}_j - \bar{Y}_k}{\sqrt{\frac{MS_{within}}{2} \left(\frac{1}{n_j} + \frac{1}{n_k} \right)}}$$

Scheffe's method

- 原理

- 一種以F考驗為基礎，適用於 n_g 不相等的多重比較技術
- 此一方法對分配常態性與變異一致性兩項假定之違反頗不敏感，且所犯第一類型錯誤（type I error）的機率較小。可以說是各種方法中最嚴格、統計檢定力（ $1 - \beta$ ）最低的一種多重比較。
- Cohen (1996) 甚至認為Scheffe執行前不一定要執行F整體考驗，因為如果F考驗不顯著，Scheffe考驗亦不會顯著，但是如果F整體考驗顯著，那麼Scheffe檢定則可以協助研究者尋找出整體考驗下的各種組合效果

$$F = \frac{\frac{(\bar{Y}_j - \bar{Y}_k)^2}{p-1}}{MS_{within} \left(\frac{1}{n_j} + \frac{1}{n_k} \right)}$$

變異數同質假定違反的多重比較

- Dunnett's T3法

- 調整臨界值來達成族系與實驗面的錯誤機率，使型一機率控制在一定的水準下

$$q_j = \frac{s_j^2}{n_j} \quad \hat{v}_{jk} = \frac{(q_j + q_k)^2}{\frac{q_j^2}{n_j - 1} + \frac{q_k^2}{n_k - 1}}$$

- s_j^2 表示有 n_j 個人的第 j 組變異數，表示各平均數變異誤估計數

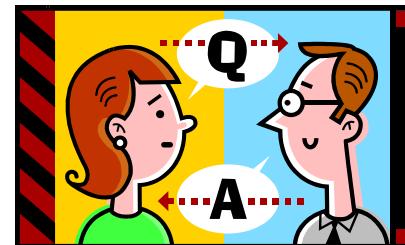
變異數同質假定違反的多重比較

- Games-Howell法
- 原理
 - 計算出調整自由度 \hat{v}_{jk} 後，直接與查自於 Studentized range distribution的 q_{cv} 臨界值相比，來決定顯著性
 - 當各組人數大於50時Games-Howell法所求出的機率估計會較T3法正確，類似於Dunnett另外提出的C法

$$\frac{|\bar{Y}_j - \bar{Y}_k|}{\sqrt{\frac{1}{2}(q_j + q_k)}} \geq q_{cv}$$

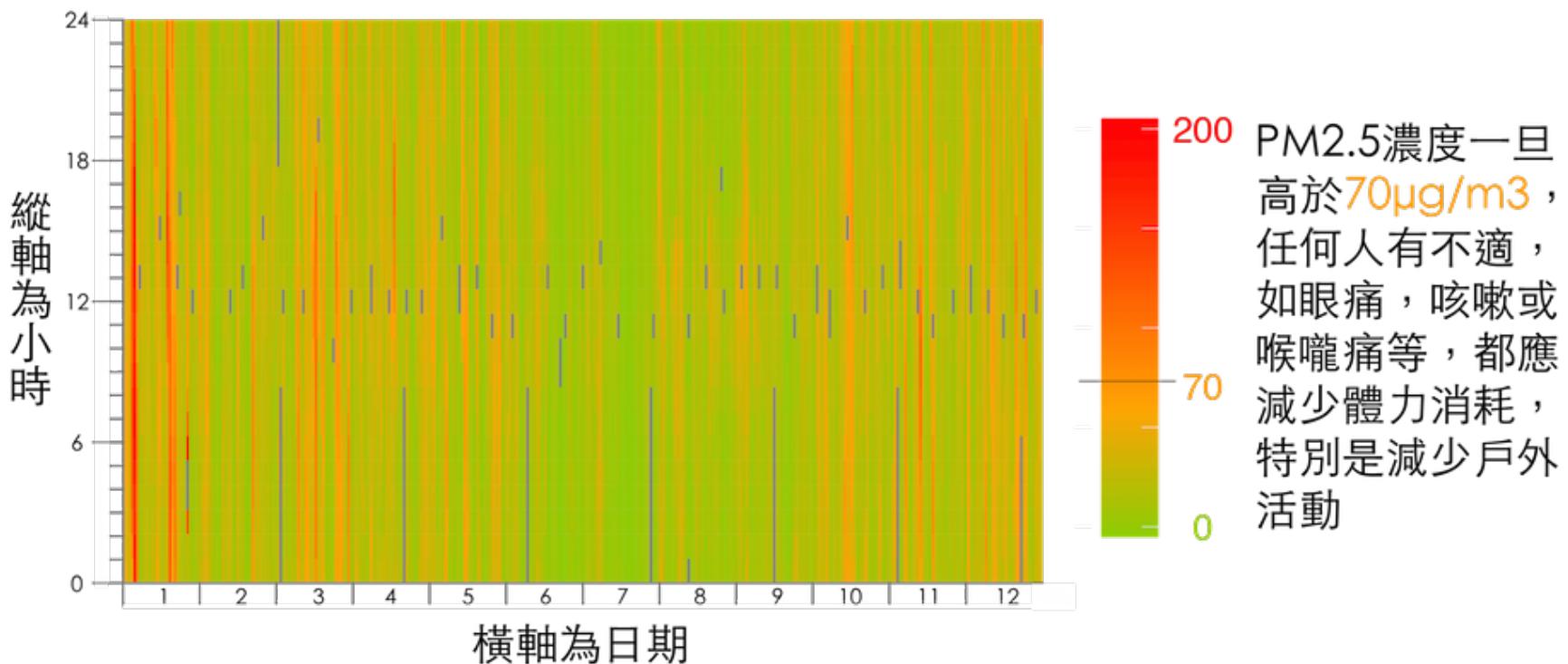
Thank you very much!!

Q & A



年度PM2.5統計

馬祖2014年PM2.5濃度表



<https://www.thenewslens.com/feature/airpollution/19328>

台灣各地年度PM2.5統計

北北基	桃竹苗	中彰投	雲嘉南	高屏	宜花東	離島
基隆	平鎮	二林	斗六	大寮	冬山	金門
淡水	桃園	大里	台西	小港	宜蘭	馬公
陽明山	大園	竹山	台南	仁武	關山	馬祖
三重	龍潭	西屯	安南	左營	台東	
土城	中壢	沙鹿	朴子	林園	花蓮	
士林	三義	忠明	嵙背	前金		
大同	頭份	南投	麥寮	前鎮		
菜寮	苗栗	埔里	善化	屏東		
中山	竹東	彰化	新港	恆春		
古亭	新竹	線西	新營	美濃		
永和	湖口	豐原	嘉義	復興		
汐止				楠梓		
松山				鳳山		
板橋				潮州		
林口				橋頭		
觀音						
新店						
新莊						
萬里						
萬華						

<https://www.thenewslens.com/feature/airpollution/19328>