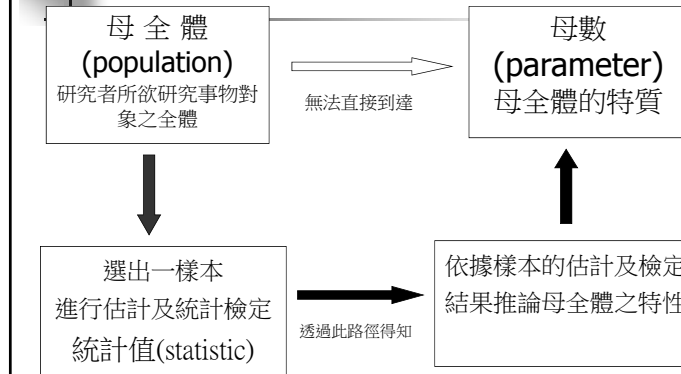


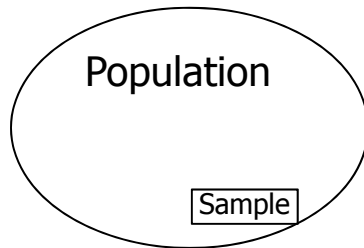
統計分析與軟體應用

何佩珊副教授：
分機：2159
psho@kmu.edu.tw
辦公室：國際學研大樓5F IR514
(口腔健康促進及照護研究室)

統計分析的原理



Study question is focused on population,
but the fact is.....



- 母數(parameter):
- 統計量(statistic): 為母數的估計值(estimate)

抽樣的原因

- 母群體數量太大
- 時效上不允許
- 觀察會造成破壞
- 根本無法接觸
- 經濟上不允許
- 普查正確性不見得高

隨機抽樣(Random Sampling)

- 有一定機率, 可經由數學方法計算出抽樣誤差大小
- 為推論統計之理論基礎

非隨機抽樣 (Non-random sampling)

- 不能經由數學方法計算出誤差量之大小
- 推論統計在非隨機抽樣下難以發揮其功能
- 僅具有參考價值

隨機抽樣的基本原理

- 每個個體被抽中的機會是隨機的 (random)
 - 隨機: 非人為能控制的
- 母全體中的每一個個體都有一個已知, 而且大於零的被抽到機會

Statistical methods for different purposes

- For describing magnitude/effect size
- For describing efficiency
- Providing statistical evidences for study hypothesis
- Describing effect sizes while controlling (adjusting) for confounders

For describing magnitude/effect size

描述測量值的大小

- Means, standard deviations,
- Number of subjects, proportions
- Odds ratios (OR), risk ratios (RR)
- Correlation or regression coefficients
- Plots
- Survival curves

Descriptive Statistics

9

For describing efficiency

描述測量值的精確度

- standard error (SE),
- 95% confidence intervals (95% CI)

Descriptive Statistics

10

Confidence Interval (CI)

- 信賴區間
- 估計未知參數，並提供一個誤差範圍來標示此估計好壞程度
 - 有多少信心- 信賴水準(Confidence Level)
 - 某候選人民調支持率為48%(誤差範圍±3%)
- 95%CI
 - 有95%的把握, 此區間會涵蓋估計值的母數

定量資料描述性統計-範例

變數名稱(單位)	Mean	SD	Median	25%	75%	95%CI
住院天數(天)	8.6	5.72	8	5	11	(6.24, 10.96)
年齡(歲)	41.24	20.1	41	25	57.5	(32.94, 49.54)
體溫(華氏、度)	98.31	0.68	98.2	98	98.8	(98.03, 98.59)
白血球(千)	7.84	3.21	7	5	11	(6.51, 9.17)

定性資料描述性統計- 頻率分配表

國中生父母教養風格基本分布

教養風格	人數	百分比	95%CI
權威式	451	37.43%	(34.74% , 40.20%)
溺愛式	242	20.08%	(17.92% , 22.44%)
威權式	320	26.56%	(24.14% , 29.12%)
不參與式	192	15.93%	(13.98% , 18.11%)

Providing statistical evidences for study hypothesis

用來回答yes/no問題

- Hypothesis testing, including
 - Chi-square tests, Fisher's exact test, ...
 - t-tests, ANOVA,
 - Nonparametric tests
 - Wilcoxon signed rank test
 - Wilcoxon rank sum test/Kruskal-Wallis test

Inferential Statistics

14

Describing effect sizes while controlling (adjusting) for confounders

探討多個變項對結果的影響

- Models
 - Multiple linear regression (concepts of adjustment, confounding and interaction)
 - Multi-way ANOVA, ANCOVA
 - Multiple (multivariable) logistic regression for categorical outcome
 - Cox regression for time-to-event outcome

Descriptive and Inferential Statistics

15

Inferential Statistics

如何決定資料分析方法?

decisions based on

- What are the **study objectives**?
- What do we want to prove in our studies?

17

Examples of Study hypothesis

- Investigate the outcome (Y) in different setting of effects (X)
- Whether changes in factors (X) will result in different response level (Y)
- Other study questions....

18

判定統計分析方法

- 依**研究目標**設計分析方法
- 變數型態
 - 連續性(continuous)變數
 - 類別性(categorical)變數
- 研究角色
 - Outcome (response, dependent, Y)
 - Design (independent, X)

變數型態及轉換

等比尺度
> 等距尺度
> 序位尺度
> 類別尺度

判定統計分析的基本原理

- 觀察不同的X對Y的影響
- 不同X的設定(design Variables)導致Y的變化(outcome Variables)

變項的研究角色

- | | |
|---|--|
| <ul style="list-style-type: none"> ■ Design Variables <ul style="list-style-type: none"> ■ Predictors <ul style="list-style-type: none"> ■ X ■ 實驗或研究上的控制變因 ■ 多半已知且容易被測量的變因 <ul style="list-style-type: none"> ■ Independent <ul style="list-style-type: none"> ■ 自變項 | <ul style="list-style-type: none"> ■ Outcome Variables <ul style="list-style-type: none"> ■ Response <ul style="list-style-type: none"> ■ Y ■ 實驗或研究上的觀察值 ■ 多半實驗後才知或以後被估計的對象 <ul style="list-style-type: none"> ■ Dependent <ul style="list-style-type: none"> ■ 依變項 |
|---|--|

		Design (X)	
		Continuous	Categorical
Outcome (Y)	Continuous	Pearson Correlation coeff. Linear regression <i>Spearman Correlation coeff.</i>	t-test / ANOVA Linear models <i>Wilcoxon Rank-Sum test</i> <i>Kruskal-Wallis test</i>
	Categorical	Logistic Categorical analysis	χ^2 Fisher's Exact Categorical analysis Logistic analysis

統計檢定的概念

機率理論為統計學的基礎

- 在利用統計方法作推論時, 須利用**機率分配(Probability Distribution)**之理論作為依據

機率分配(probability distribution)

- 任何**隨機變數**皆有其自己的機率分布, 可以用**機率分配**來表示。
- 機率分配:
 - 將變項中可能發生的項目或數值放在**X軸**
 - **Y軸**則為**0到1**→該項目或數值發生的**機率大小**。

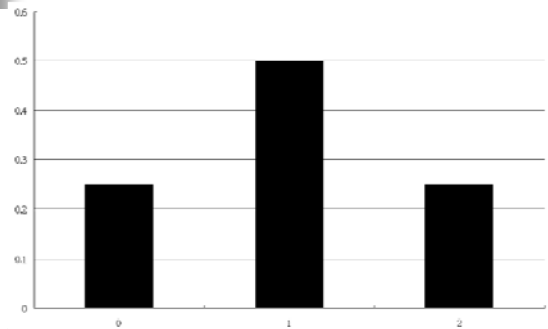
隨機變數(Random Variable) (X)

- 任何統計上實驗的變數皆是隨機變數
- 各隨機變數有各自出現的機率
- 即每一隨機變數皆配備一
統計分布(statistical distribution)

隨機變數(續)

- 若以**X**代表一個隨機變數
- **X**之每一個可能值(possible value)皆有其發生**機率(probability)**
- **Ex1.** 投擲一公正硬幣兩次
 - **X** → 硬幣出現正面次數之隨機變數
 - **X** → 0, 1, 2

投擲2個硬幣得到正面次數的 機率分布.



機率分配與隨機變數

- 由於所有的隨機變數會概括全部樣本空間的所有可能事件
- 事件彼此間皆為互斥
- 因此機率分配必須滿足下列兩點：
 - $0 \leq \text{Pr}(X) \leq 1$
 - $\sum \text{Pr}(X) = 1$ ，對所有可能的X值。

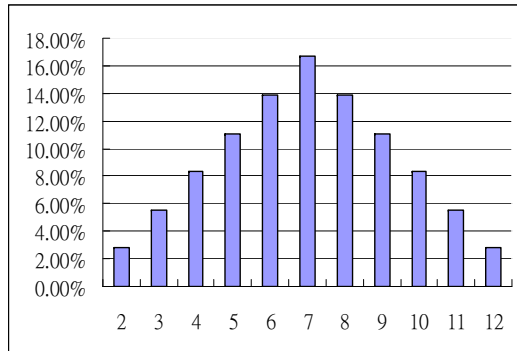
隨機變數機率分佈之例題

- Ex2. 擲兩個公平骰子
 - 隨機變數(X)為兩個骰子的總和
 - 依機率分配之原理(law of probability)，此隨機變數之所有發生可能為：

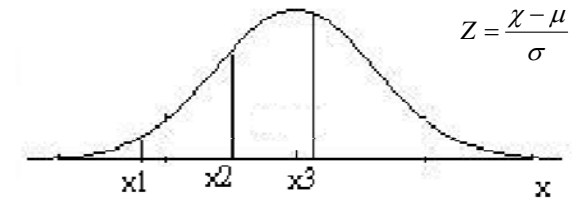
隨機變數機率分佈之例題(續)

X	Possible combination	Frequency	Percentage (%)
2	1,1	1	1
3	1,2,2,1	2	6
4	1,3,3,1,2,2	3	8
5	1,4,4,1,2,3,3,2	4	11
6	1,5,5,1,2,4,4,2,3,3	5	14
7	1,6,6,1,2,5,5,2,3,4,4,3	6	16
8	2,6,6,2,3,5,5,3,4,4	5	14
9	3,6,6,3,4,5,5,4	4	11
10	4,6,6,4,5,5	3	8
11	5,6,6,5	2	6
12	6,6	1	3
Total		36	100

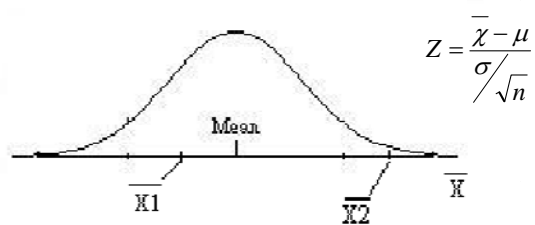
投擲2個骰子點數總和隨機變數的機率分布.



連續隨機變數之機率分布



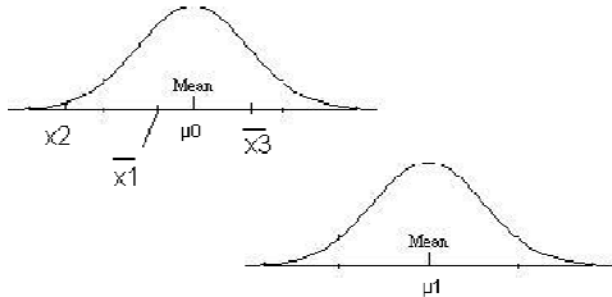
樣本平均值抽樣分布



統計檢定的概念-範例

- 想評估DM P't血壓收縮壓是否與正常人相同
 - 已知正常人收縮壓的平均值母數(μ_0)為90mmHg
 - 有3組DM P't樣本, 其平均收縮壓分別為 $\bar{x}_1, \bar{x}_2, \bar{x}_3$
 - DM P't收縮壓的平均值母數為 μ_1
- 檢定假說
- 統計結果判讀

統計檢定的概念



統計檢定的概念(續)

- 針對未知母群體母數提出假設, 藉由樣本資料評估此假設成立的可能性
- 虛無假說(null hypothesis) H_0
 - 無意義的假設, 無因果關係
- 對立假說(alternative hypothesis) H_1
 - 替代假設, 希望應該是對的假設

統計檢定的概念(續)

- 欲比較男女性平均身高是否相等?
- 欲評估男性平均體重是否高於女性?
- 欲了解性別是否會影響教育程度?

統計檢定結果的解讀-P值

- 實際上, 若兩變項間是沒有關係的, 則統計上得到相同結論的機率
- 當此機率小於0.05時, 即會以兩變項是有關係的作為結論

統計檢定結果的解讀(續)

- P-值 < 0.05
 - 虛無假設是錯的
 - 達顯著性差異,顯著性相關
- P-值 \geq 0.05
 - 無法證明虛無假設是錯的

		Design (X)	
		Continuous	Categorical
Outcome (Y)	Continuous	Pearson Correlation coeff. Linear regression Spearman Correlation coeff.	t-test / ANOVA Linear models Wilcoxon Rank-Sum test Kruskal-Wallis test
	Categorical	Logistic Categorical analysis	χ^2 Fisher's Exact Categorical analysis Logistic analysis

One sample test

- EX. 比較一組20歲男性大學生樣本的SBP是否與全國同年齡男性的SBP相同
- 評估此組樣本是否來自預設的母群體
- 檢定假說
 - H0: 此組大學生樣本來自母群體的SBP平均值與全國同年齡男性相同
 - H1: 此組大學生樣本來自母群體的SBP平均值與全國同年齡男性不同

43

One sample t-test - Example

	N	Mean	SD	SE	P
薪資差值	474	17403.48	10814.62	496.73	<0.001

- 此樣本中共有474名樣本, 平均薪資差值為17403.48元, 與薪資差值平均為20000元的母群體相比較, 此樣本來自的母群體薪資差值平均顯著不同(P值<0.001)

44

How to interpret significance

統計顯著性

- 以統計檢定後p-value 來決定，而不管差距大小。
- 每一個p-value對應一個問題。

46

每一個p-value對應一個問題

實驗組的 cholesterol	比較對象	P-value
181.38±7.01	?	<.0001
181.38±7.01	?	<.0001
181.38±7.01	?	0.0740

47

每一個p-value對應一個問題

實驗組的 cholesterol	比較對象	P-value
181.38±7.01	0	<.0001
181.38±7.01	對照組 282.40±2.88	<.0001
181.38±7.01	對照組 195.80±13.22	0.0740

48

臨床顯著性

- 由經驗或文獻資料來決定統計估計值之大小，是否會使得臨床上的作用有顯著性的不同。

49

情況一：

- 當估計值很小，但因樣本數很大使得估計值顯著.....
- 若只講統計顯著性不講估計值大小，讀者會誤以為影響重大.....
- 例：使用音樂減壓介入後，實驗組與對照組心跳數相差一下，且達統計顯著性。
- True Story : nothing excited !

50

情況二：

- 當估計值很大，但因樣本數很小使得不具統計顯著性.....
- 只講不具顯著性，讀者會誤以為沒事。
- 例：住在變電所周圍的居民得腦癌的機率是其他地區的10倍，但不達統計顯著！
- True Story : 基本上10倍非常驚人！建議應進一步求證，用較大的樣本數來證明。

51

良心的統計報告必須

- 統計估計值(臨床顯著性)與
- 檢定結果(統計顯著性)
- 都呈現否則容易造成誤導！

52

沒有達顯著性怎麼辦？

- 表示沒有足夠證據證明
早先設定之研究假設具
顯著性

53

其原因可能：

- 樣本數不夠
 - 嘗試增加樣本, 增加前可先估算
- 亂度太大(standard deviation)
 - 利用降低亂度的分析方法:
分層(stratification), 配對(match), 迴歸
(regression), etc...

54

其原因可能：

- 沒用對適當的統計方法或是變數型式
 - 是否有選出適合資料特性之統計方法
 - 迴歸中控制變項的增減
 - 數字或類別, 比較憂鬱指數或是達憂鬱症之比例
 - etc...

55

其原因可能：

- 真的是沒有差異
 - 未達臨床顯著性
- 真的需要顯著性嗎？

56

2-sample t-test (Equal and Unequal Variance)

- EX. 比較男女之SBP是否不同
- Y: SBP (continuous)
- X: Sex (2 categories)
- 2-sample t-test 男女生SBP平均值的差異是否具有統計上意義
- 檢定假說
 - H0: 不同性別的SBP平均值是相同的
 - H1: 不同性別的SBP平均值是不同的

57

F test (Levene test)

- p-value of F test ≥ 0.05 , then use equal variance 2 sample t-test
- p-value of F test < 0.05 , then use unequal variance 2 sample t-test

58

2-sample t-test - Example

性別	N	Mean	SD	P value
男	258	21140.39	12557.06	<0.001
女	216	12939.95	5661.085	

此樣本中共收集男性樣本其薪資差值平均值(SD)為21140.39(12557.06)元, 女性樣本有216名, 薪資差值平均值(SD)為12939.95(5661.08)元, 經統計檢定達統計上顯著意義(p值<0.001), 故認為男女性的平均薪資差值是不相同的

59

同樣是2 samples

- 獨立樣本
 - 兩個樣本資料自於互相獨立的不同個體
 - 除了分兩組的變項外, 個體間互相獨立沒有關係
- Paired t-test
 - 兩個相依樣本的比較
 - 同一組樣本, 重複收集同一種測量值2次
 - 樣本有配對處理

60

Paired t-test

- Ex. 比較服藥前後SBP是否不同
- Diff=服藥前SBP-服藥後SBP
- 評估Diff的平均值是否等於“0”
- 檢定假說
 - H0: 服藥前後SBP值是相同的
 - H1: 服藥前後SBP值是不同的

61

資料型態如下表所示:

編號	X_{i1} 測量一	X_{i2} 測量二	$d_i = X_{i1} - X_{i2}$
1	X_{11}	X_{12}	d_1
2	X_{21}	X_{22}	d_2
⋮	⋮	⋮	⋮
i	X_{i1}	X_{i2}	d_i
⋮	⋮	⋮	⋮
n	X_{n1}	X_{n2}	d_n

在此情形之下,要做檢定應針對兩次測量之差值 d_i 作檢定,因此以 d_i 取代 X_i ,配對t檢定事實上就是一種單樣本t檢定。

62

Paired t-test - Example

不同時間脈搏之比較

	N	Mean	SD	P value
時間1 - 時間3	12	-0.1	0.159545	0.052663

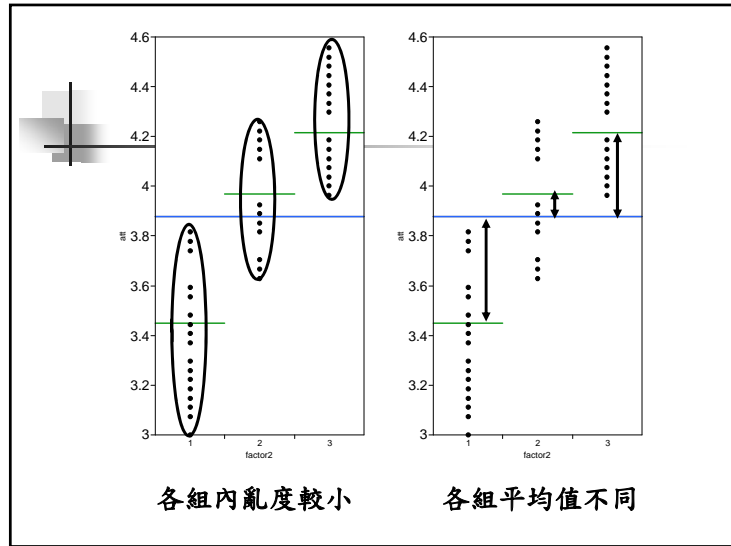
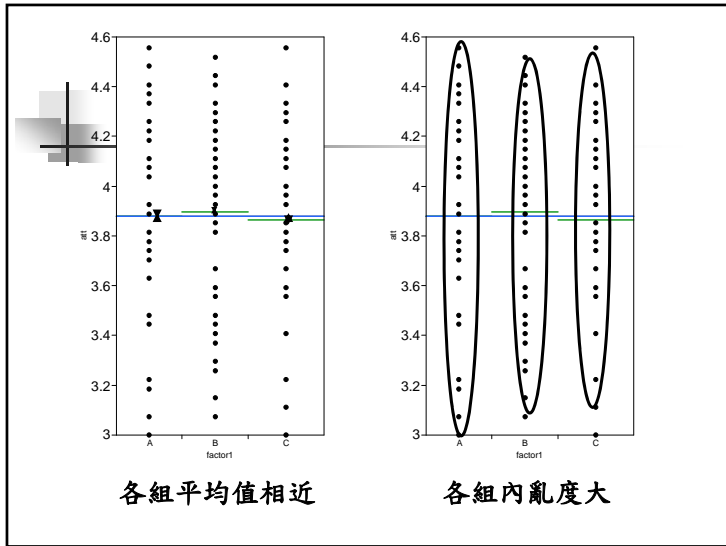
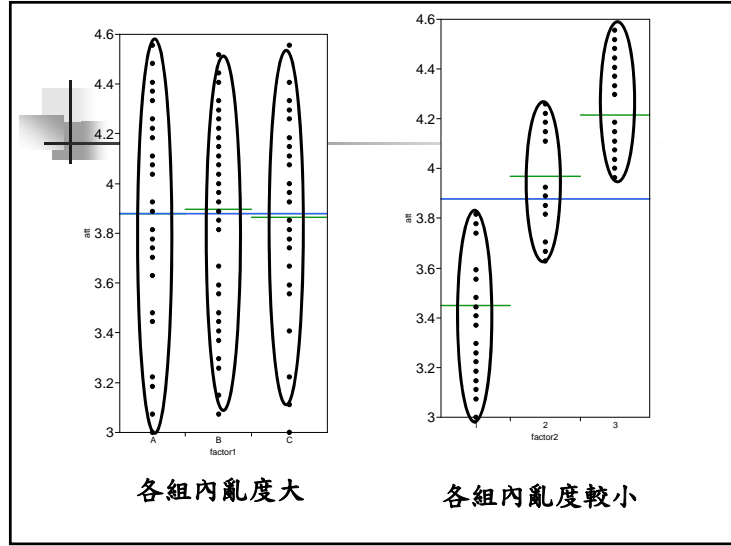
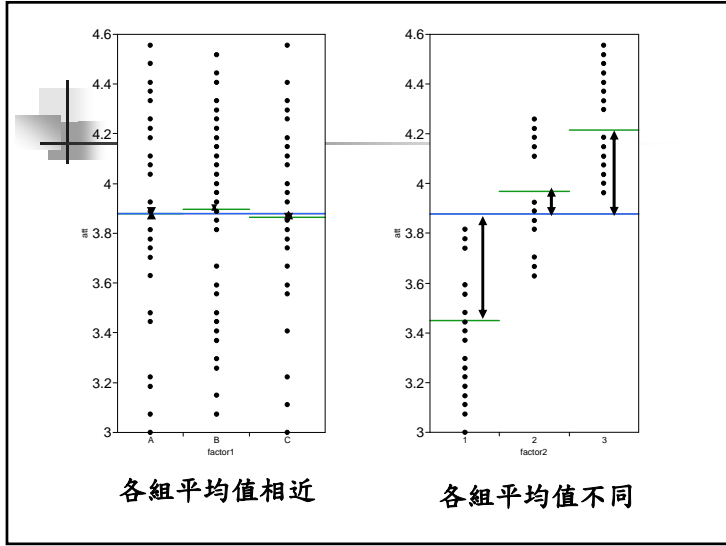
此樣本中時間1脈搏與時間3脈搏心跳數差值的平均值(SD)為-0.1(0.16),以統計檢定其前後測是否相同,結果未達統計顯著意義(P值=0.052),因此由本研究結果認為第1次(pulse1)與第3次(pulse3)測量的心跳數並無差別

63

ANOVA

- A, B, C三地之SBP是否不同
- Y: SBP (continuous)
- X: 地點 (3 categories)
- ANOVA ABC三地SBP平均值的差異是否具統計上意義
- 檢定假說
 - H0: 三地的SBP平均值皆相同
 - H1: 至少有兩地的SBP平均值是不同的

64



- Within group variation
 - Within sum of squares
 - Within SS= $\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$
- Between group variation
 - Between sum of squares
 - Between SS= $\sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{y}_i - \bar{\bar{y}})^2$

69

So, if means are different

- Then, Between SS should be larger than Within SS
 - where
 - Within SS= $\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$
 - Between SS= $\sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{y}_i - \bar{\bar{y}})^2$

70

ANOVA table

Source	Sum of Squares (SS)	df	Mean Squares (MS=SS/df)	F	p-value
Between	BSS	k-1	BMS=BSS/(k-1)	F=BMS/WMS	Prob(Fk-1, n-k>F)
Within	WSS	n-k	WMS=WSS/(n-k)		
Corrected total	Total SS=BSS+WSS				

$$totalSS = \sum_{i=1}^k \sum_{j=1}^{n_i} \left(y_{ij}^2 - \frac{\bar{y}_i^2}{n} \right)$$

p-value < 0.05,
達顯著性差異

71

ANOVA - Example

不同職別薪資差值之比較

職別種類	N	Mean	SD	P value	post-comparison
職員	363	13742.49	5973.77	<0.001	管理人員>保全人員
保全人員	27	15861.11	2415.31		管理人員>職員
管理人員	84	33719.94	13424.34		

72

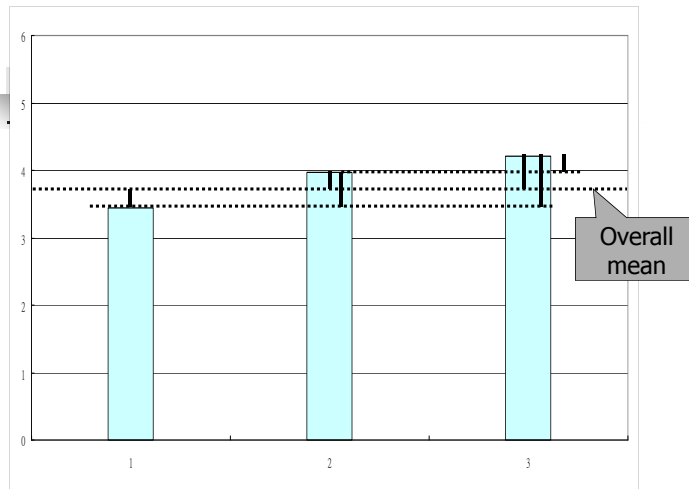
ANOVA - Example

- 比較不同職別薪資差值, 共收集363名職員樣本, 其薪資差值平均值(SD)為13742.49(5973.77)元, 保全人員有27名樣本, 薪資差值平均值(SD)為15861.11 (2415.31)元, 管理人員有84名樣本, 其薪資差值平均值(SD)為33719.94(13424.34)元; 以ANOVA進行統計檢定, 結果認為不同職別的薪資差值是不相同的($P < 0.001$), 且經過事後檢定發現管理人員的薪資差值顯著高於保全人員及職員。

73

Post-hoc comparison

When ANOVA significant...



75

ANOVA表格範例一：

Group	n	Y1			Y2			...
		mean	±	sd	mean	±	Sd	
Group1	22	√	±	√	√	±	√	
Group2	22	√	±	√	√	±	√	
Control	22	√	±	√	√	±	√	
p-value from ANOVA		<0.0001			0.1111			
Significant pairs from Tukey's pairwise comparison		Group 2 > Group 1 Group 2 > control						

76

ANOVA表格範例二:

Group	n	mean	±	sd	p-value from ANOVA	Significant pairs from Tukey's pairwise comparison
Group1	22		±		<0.0001	Group 2 > Group 1 Group 2 > control
Group2	22		±			
Control	22		±			
Sex						
Men	32		±		0.5555	
Women	23		±			
...						

77

Design (X)

		Design (X)	
		Continuous	Categorical
Outcome (Y)	Continuous	Pearson Correlation coeff. Linear regression Spearman Correlation coeff.	t-test / ANOVA Linear models Wilcoxon Rank-Sum test Kruskal-Wallis test
	Categorical	Logistic Categorical analysis	χ^2 Fisher's Exact Categorical analysis Logistic analysis

Chi-square test

- 比較兩類別變項是否有相關
- 嚼食檳榔與罹患口腔癌的關係
 - Y: 是否有口腔癌 (categorical)
 - X: 是否嚼檳榔 (categorical)

Chi-square test -Example

- 比較兩種預防頭部傷害患者早期癲癇發作的藥劑A與B，在產生有害反應上是否有差異時，資料整理如下表

有害反應藥劑	是	否	總計
A	10	37	47
B	7	38	45
總計	17	75	92

卡方檢定結果表格

使用抗生素是否會影響住院時間				
抗生素	住院時間			P value
	<=5天	5-10天	10天以上	
是	1(14.29)	3(42.86)	3(42.86)	0.3698
否	8(44.44)	5(27.78)	5(27.78)	

卡方檢定結果描述

- 本研究共有25名住院病人; 有使用抗生素共有7位, 其中住院天數<=5天的共1位(14.29%), 住院天數5-10天的共3位(42.86%), 住院天數10天以上的共3位(42.86%); 共有18位住院病人無使用抗生素, 其中8位病人的住院天數<=5天(44.44%), 住院天數5-10天的共5位(27.78%), 住院天數10天以上的共5位(27.78%), 經卡方檢定得知未達統計上顯著意義($X^2=1.9901$, $df=2$, $P=0.3698$), 因此本研究認為住院天數分組不受有否使用抗生素的影響

使用 χ^2 test 之注意事項

- use this test only if the following two Conditions are satisfied:
 - No more than 20% of the cells should have expected values < 5.
 - No cell should be expected value < 1.

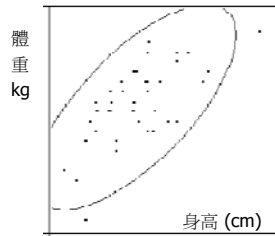
Design (X)

		Continuous	Categorical
--	--	------------	-------------

Outcome (y)	Design (X)	
	Continuous	Categorical
Continuous	Pearson Correlation coeff. Linear regression Spearman Correlation coeff.	t-test / ANOVA Linear models Wilcoxon Rank-Sum test Kruskal-Wallis test
Categorical	Logistic Categorical analysis	χ^2 Fisher's Exact Categorical analysis Logistic analysis

相關係數 (Correlation coefficient)

- 檢定兩個連續變項呈現線性相關之程度
- 沒有單位，為一個從-1~1的數值



$$r=0.7092$$

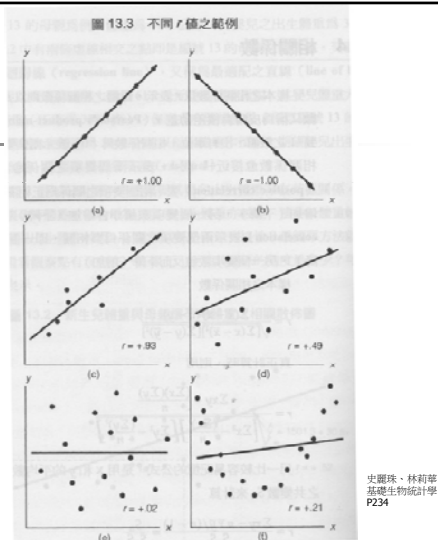
$$P<0.001$$

身高與體重成正相關

相關係數(續)

- 又稱之為皮爾森相關係數(Pearson correlation coefficient)
- Spearman correlation coefficient
- 正相關: 一個數值增加(下降), 另一個數值亦會增加(下降)
- 負相關: 一個數值增加(下降), 另一個數值亦會下降(增加)

相關係數-續



相關係數之限制

- 相關係數是要測量兩變項與一直線接近程度，而非線性關係之正確性
- 極端值(outlier)對相關係數影響很大
- 當樣本數很小時(<50對)，相關係數不穩定
- 兩變項間相關很強，不代表兩者間有因果關係

Models

Linear regression
and
Logistic regression

		Design (X)	
		Continuous	Categorical
Outcome (Y)	Continuous	Pearson Correlation coeff. <i>Linear regression</i> <i>Spearman Correlation coeff.</i>	t-test / ANOVA Linear models <i>Wilcoxon Rank-Sum test</i> <i>Kruskal-Wallis test</i>
	Categorical	<i>Logistic regression</i> Categorical analysis	χ^2 Fisher's Exact Categorical analysis <i>Logistic analysis</i>

The purpose of regression

- 由已知的自變項預測依變項的狀況
- 考慮(調整)其他自變數後, 評估某一自變項對依變項的影響
- 比較不同自變數對依變數影響的重要性

Multiple Linear Regression:

Y (continuous)

$$y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \dots + \beta_k X_k$$

- regression coefficients or parameter estimates

Dependent Variable Y

Independent Variables X_1, X_2, \dots, X_k

Parameter estimates 的取得是
 X_1, X_2, \dots, X_k 彼此互相調整後而得!

Example of Linear regression

Table 24. Factors related to oral health knowledge of mothers by the multiple regression analysis.

Variable	Coefficient	95% CI	P-value	Coefficient ¹	95% CI	P-value
Nationality						
Native						
Mainland Chinese	-2.65	(-4.06, -1.24)	0.0003	-1.74	(-3.19, -0.29)	0.0198
Southeastern Asian	-6.39	(-7.49, -5.29)	<.0001	-5.35	(-6.55, -4.15)	<.0001
Age of mother						
<25 years old						
25-30 years old	3.82	(2.31, 5.33)	<.0001			
>30 years old	5.70	(4.25, 7.15)	<.0001			

Coefficient: Simple regression coefficient Coefficient¹: Multiple regression coefficient

- Outcome: 口腔健康知識分數(Continuous variable)

Chi-square VS Logistic regression

- *Chi-square* 口腔癌是否與嚼檳榔有關
- *Logistic regression* 嚼食檳榔者是不嚼檳榔者多少倍的機率會得口腔癌

Multiple logistic Regression:

Y binary, nominal or ordinal

$$\text{logit}(y) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

- regression coefficients will be used to compute odds ratios (OR)
- Y變項為 Yes/No, Disease/Non-disease, Recover/Not-recover,.....

Example of Logistic regression

Variable	Alive/Death (Crude)			Alive/Death (Adjusted#)		
	OR	95%CI	p-value	OR	95%CI	p-value
Age_gp <=45yrs	1.91	(1.39 , 2.63)	<0.001	1.76	(1.19 , 2.61)	0.005
45-60yrs	1.58	(1.19 , 2.10)	<0.001	1.36	(0.96 , 1.93)	0.079
>60yrs	1.00			1.00		

adjusted by Histological grade, Lymph node, Pathologic Tumor Size

- Outcome: 是否存活超過5年(Binary variable)

Introduction of SPSS

- ## What is SPSS?
- Statistical Package for the Social Science
 - SPSS Inc.
 - Excellent in
 - 簡易操作
 - 對硬體需求低
- One of many many spread sheet based statistical softwares!

time	cholesterol	group
1	1019	1.00
2	684	1.00
3	810	1.00
4	950	1.00
5	810	1.00
6	978	1.00
7	1002	1.00
8	1110	1.00
9	884	2.00
10	606	2.00
11	638	2.00
12	708	2.00
13	780	2.00
14	600	2.00
15	1320	2.00
16	750	2.00
17	594	2.00
18	750	2.00

變項: 每一種測量值稱之

筆數

One observation key in as one row in the dataset. Excel, JMP, SPSS or

IBM SPSS19.0

date	men	wk
01/01/1989	11357.97	1
02/01/1989	10605.95	1
03/01/1989	10990.67	4
04/01/1989	6583.75	3
05/01/1989	8007.09	2
06/01/1989	9639.00	2
07/01/1989	9398.37	3
08/01/1989	10395.53	3
09/01/1989	11663.13	3
10/01/1989	12805.22	3
11/01/1989	13006.25	0