



複回歸分析、中介與調節效果檢驗、路徑分析： SPSS實作技巧

成功大學統計學研究所

陳晉群



-
- ◆ SPSS版本：SPSS 17.0
 - ◆ 資料名稱：Student Performance Data Set
 - ◆ 資料來源：<https://archive.ics.uci.edu/ml/datasets/Student+Performance>
 - ◆ 資料筆數：395
 - ◆ 變數數量：33
 - ◆ 本資料蒐集一校學生的各項資訊以及第一次、第二次期中考和期末考成績
 - ◆ `student_math.sav` 下載連結

<https://drive.google.com/drive/folders/1q5iW4viK1jwASwxQNigYBGsJwQk6VzGd?usp=sharing>



複迴歸分析

◆ 探討多個自變數(x)與依變數(y)之間的關係

◆ 依變數必須為連續型(數值)

◆ 複迴歸方程式

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon$$

◆ 其中 β_0 為截距項， β_1, \dots, β_k 為迴歸係數， ε 為殘差項

◆ 估計式

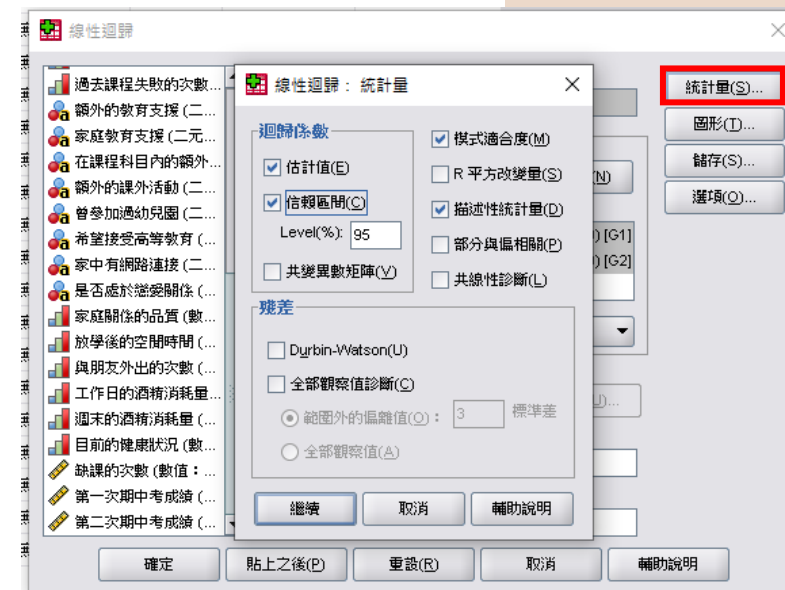
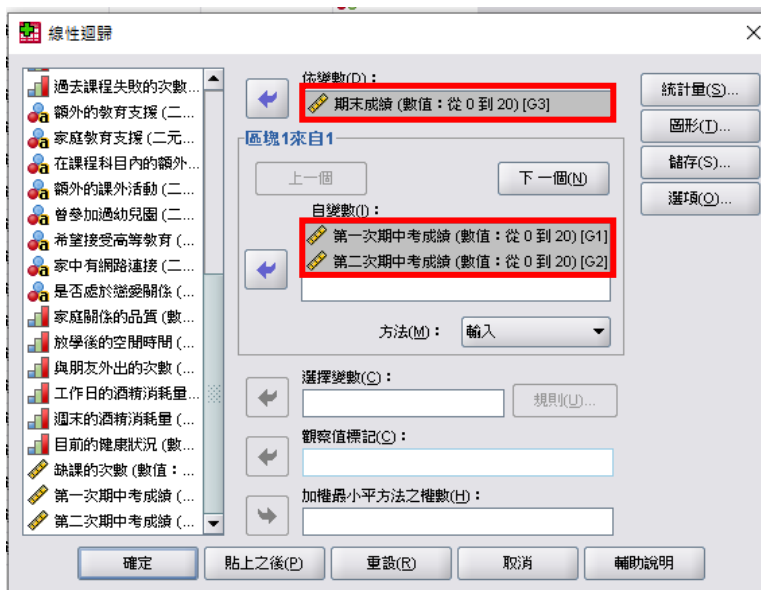
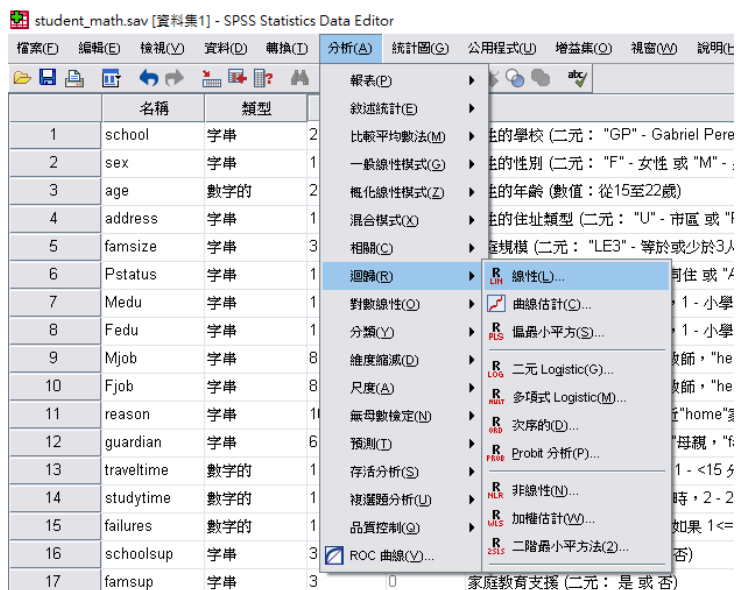
$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \cdots + \hat{\beta}_k x_k$$

◆ 殘差誤差需滿足常態、獨立、同質變異數



複迴歸分析：SPSS操作

- ◆ 研究問題：研究兩次期中考的成績對於期末成績的影響



複迴歸分析：SPSS操作

◆ 研究問題：研究兩次期中考的成績對於期末成績的影響

敘述統計

	平均數	標準離差	個數
期末成績 (數值：從 0 到 20)	10.42	4.581	395
第一次期中考成績 (數值：從 0 到 20)	10.91	3.319	395
第二次期中考成績 (數值：從 0 到 20)	10.71	3.762	395

期中跟期末呈高度正相關

相關

	期末成績 (數值：從 0 到 20)	第一次期中考成績 (數值：從 0 到 20)	第二次期中考成績 (數值：從 0 到 20)
Pearson 相關			
期末成績 (數值：從 0 到 20)	1.000	.801	.905
第一次期中考成績 (數值：從 0 到 20)	.801	1.000	.852
第二次期中考成績 (數值：從 0 到 20)	.905	.852	1.000
顯著性(單尾)			
期末成績 (數值：從 0 到 20)		.000	.000
第一次期中考成績 (數值：從 0 到 20)	.000		.000
第二次期中考成績 (數值：從 0 到 20)	.000	.000	
個數			
期末成績 (數值：從 0 到 20)	395	395	395
第一次期中考成績 (數值：從 0 到 20)	395	395	395
第二次期中考成績 (數值：從 0 到 20)	395	395	395

模式摘要^b

模式	R	R 平方	調整後的 R 平方	估計的標準誤	Durbin-Watson 檢定
1	.907 ^a	.822	.821	1.937	1.849

a. 預測變數 (常數), 第二次期中考成績 (數值：從 0 到 20), 第一次期中考成績 (數值：從 0 到 20)

b. 依變數: 期末成績 (數值：從 0 到 20)

Anova^b

模式	平方和	df	平均平方和	F	顯著性
1 迴歸	6799.215	2	3399.608	906.134	.000 ^a
殘差	1470.694	392	3.752		
總數	8269.909	394			

模型據預測及解釋能力

a. 預測變數 (常數), 第二次期中考成績 (數值：從 0 到 20), 第一次期中考成績 (數值：從 0 到 20)

b. 依變數: 期末成績 (數值：從 0 到 20)

係數^a

模式		未標準化係數		標準化係數	t	顯著性	B 的 95.0% 信賴區間		共線性統計量	
		B 之估計值	標準誤差	Beta 分配			下界	上界	允差	VIF
1	(常數)	-1.830	.335		-5.458	.000	-2.489	-1.171		
	第一次期中考成績 (數值：從 0 到 20)	.153	.056	.111	2.728	.007	.043	.264	.274	3.651
	第二次期中考成績 (數值：從 0 到 20)	.987	.050	.810	19.909	.000	.889	1.084	.274	3.651

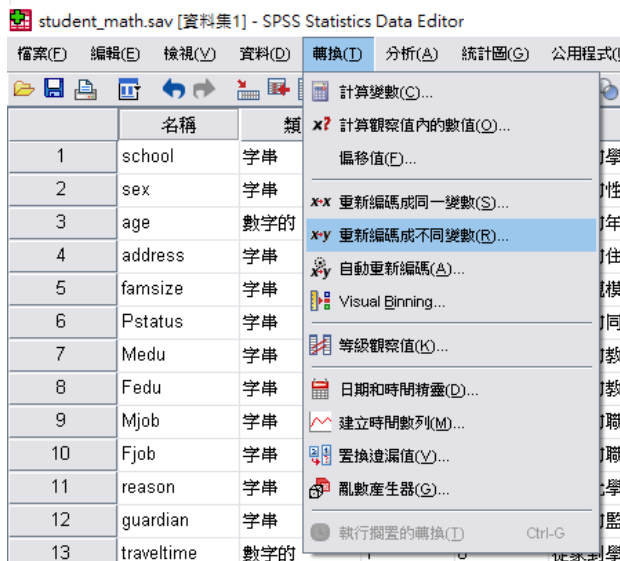
a. 依變數: 期末成績 (數值：從 0 到 20)

係數為正：期中成績對期末成績有顯著且正向的影響

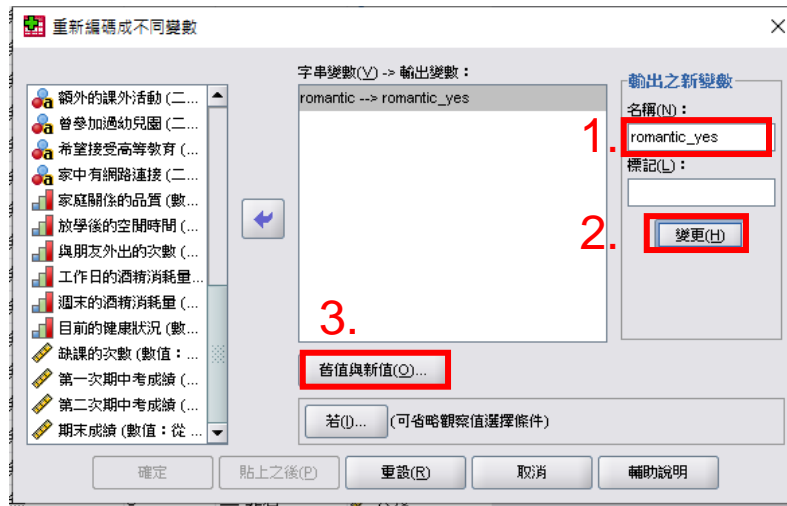
複迴歸分析：SPSS操作(虛擬變數)

◆ 研究問題：研究兩次期中考成績以及是否談戀愛對於期末成績的影響

①

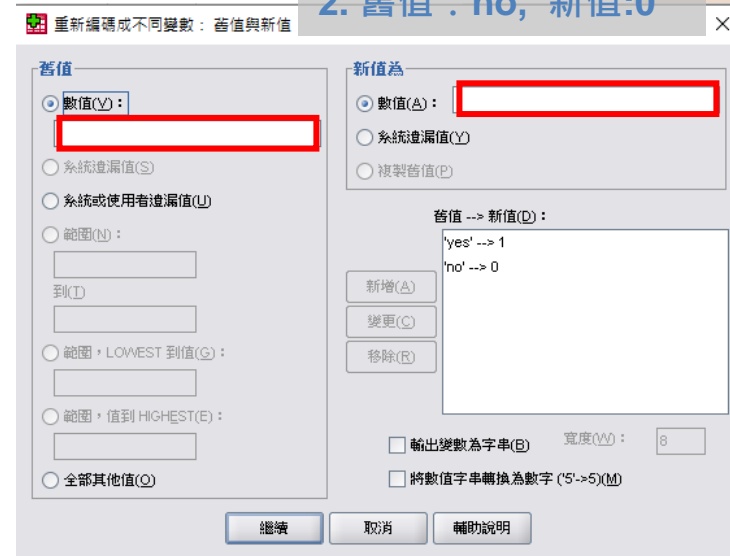


②



③

分別輸入
1. 舊值：yes, 新值:1
2. 舊值：no, 新值:0



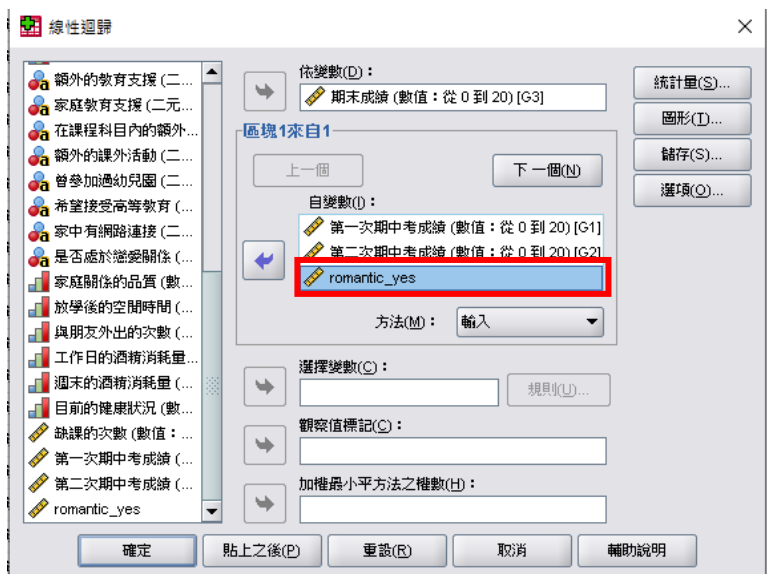
④

33	G3	數字的	2	0
34	romantic_yes	數字的	8	2

資料型態需確認是「尺度」

複迴歸分析：SPSS操作(虛擬變數)

- ◆ 研究問題：研究兩次期中考成績以及是否談戀愛對於期末成績的影響



模式摘要^b

模式	R	R 平方	調過後的 R 平方	估計的標準誤	Durbin-Watson 檢定
1	.907 ^a	.823	.822	1.932	1.863

a. 預測變數：(常數), romantic_yes, 第一次期中考成績 (數值：從 0 到 20), 第二次期中考成績 (數值：從 0 到 20)

b. 依變數：期末成績 (數值：從 0 到 20)

Anova^b

模式		平方和	df	平均平方和	F	顯著性
1	迴歸	6809.768	3	2269.923	607.845	.000 ^a
	殘差	1460.141	391	3.734		
	總數	8269.909	394			

a. 預測變數：(常數), romantic_yes, 第一次期中考成績 (數值：從 0 到 20), 第二次期中考成績 (數值：從 0 到 20)

b. 依變數：期末成績 (數值：從 0 到 20)

係數^a

模式		未標準化係數		標準化係數	t	顯著性	B 的 95.0% 信賴區間		共線性統計量	
		B 之估計值	標準誤差	Beta 分配			下界	上界	允差	VIF
1	(常數)	-1.690	.345		-4.903	.000	-2.368	-1.012		
	第一次期中考成績 (數值：從 0 到 20)	.164	.056	.119	2.905	.004	.053	.275	.270	3.697
	第二次期中考成績 (數值：從 0 到 20)	.974	.050	.800	19.462	.000	.876	1.072	.267	3.739
	romantic_yes	-.351	.209	-.036	-1.681	.094	-.761	.059	.975	1.025

a. 依變數：期末成績 (數值：從 0 到 20)

複迴歸分析：SPSS操作(虛擬變數)

◆ 虛擬變數 【SPSS - 虛擬編碼】 - 永析統計及論文諮詢顧問

*未命名標題1 [資料集0] - IBM SPSS Statistics Data

檔案(F) 編輯(E) 檢視(V) 資料(D) 轉換(T) 分

	教育程度	var	var
1	1.00		
2	2.00		
3	3.00		
4	3.00		
5	2.00		
6	2.00		
7	2.00		
8	3.00		
9	3.00		
10	3.00		
11	1.00		
12	1.00		
13	1.00		
14	1.00		
15	1.00		
16			

*未命名標題1 [資料集0] - IBM SPSS Statistics Data Editor

檔案(F) 編輯(E) 檢視(V) 資料(D) 轉換(T) 分析(A) 直效行銷 統計圖(G)

10:

	教育程度	var
1	1.00	
2	2.00	
3	3.00	
4	3.00	
5	2.00	
6	2.00	
7	2.00	
8	3.00	
9	3.00	

- 計算變數(C)...
- 程式設計轉換...
- 計算觀察值內的數值(O)...
- 偏移值(F)...
- 重新編碼成同一變數(S)...
- 重新編碼成不同變數(R)...
- 自動重新編碼(A)...
- 建立虛擬變數**
- Visual Binning...
- 最適 Binning

建立虛擬變數

變數(V):

建立虛擬變數:

教育程度

主要效果虛擬變數

建立主要效果虛擬項

根名稱 (每個選定變數一個):

教育程度dummy

巨集名稱:

雙向互動

為所有雙向互動建立虛擬項目

根名稱(R):

巨集名稱:

三向互動

為所有三向互動建立虛擬項目

根名稱(R):

巨集名稱:

此對話框需要 Python Essentials

確定 貼上(P) 重設(R) 取消 說明

虛擬變數標籤

使用數值標籤

使用數值

數值次序

遞增(O)

遞減

巨集

省略巨集定義的第一個虛擬種類

附註: 依慣例巨集名稱以!開頭。

測量層次使用

不要為小數變數值建立虛擬項目

建立所有變數的虛擬項

複迴歸分析：SPSS操作(虛擬變數)

◆ 虛擬變數 【SPSS - 虛擬編碼】 - 永析統計及論文諮詢顧問

檔案(F) 編輯(E) 檢視(V) 資料(D) 轉換(T) 插入(I) 格式(O) 分析(A) 直效行銷 統計圖(G) 公用程式(U)



```
SPSSINC CREATE DUMMIES VARIABLE=教育程度  
ROOTNAME1=教育程度dummy  
/OPTIONS ORDER=A USEVALUELABELS=YES
```

建立虛擬變數

Variable Creation

	Label
教育程度 dummy_1	教育程度=1.00
教育程度 dummy_2	教育程度=2.00
教育程度 dummy_3	教育程度=3.00

	教育程度	教育程度 dummy_1	教育程度 dummy_2	教育程度 dummy_3	var
1	1.00	1.00	.00	.00	
2	2.00	.00	1.00	.00	
3	3.00	.00	.00	1.00	
4	3.00	.00	.00	1.00	
5	2.00	.00	1.00	.00	
6	2.00	.00	1.00	.00	
7	2.00	.00	1.00	.00	
8	3.00	.00	.00	1.00	
9	3.00	.00	.00	1.00	
10	3.00	.00	.00	1.00	
11	1.00	1.00	.00	.00	
12	1.00	1.00	.00	.00	
13	1.00	1.00	.00	.00	
14	1.00	1.00	.00	.00	
15	1.00	1.00	.00	.00	
16					

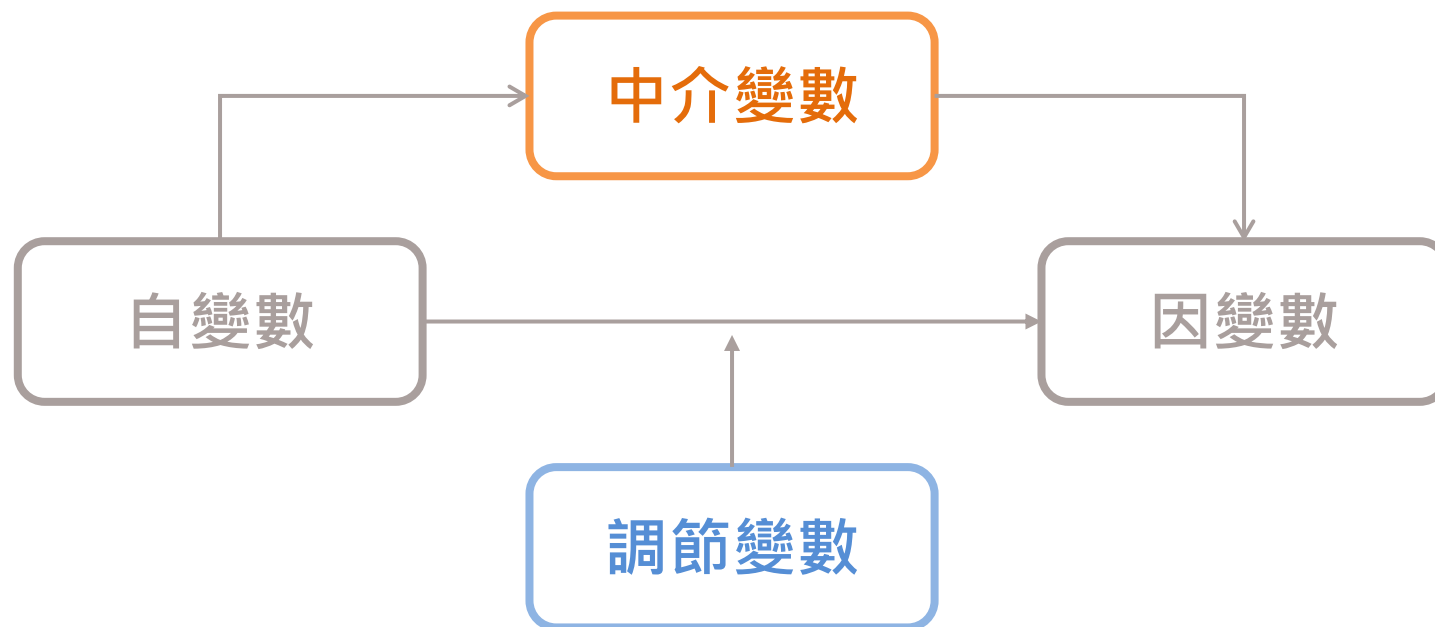
中介與調節變數檢驗

◆ 中介效應

- 自變數透過中介變數影響因變數
- 「透過什麼過程影響」
- 部分中介、完全中介

◆ 調節效應

- 調節變數影響自變數和因變數之間的關係
- 「在什麼情況下影響」



中介變數檢驗

◆ 研究問題：通勤時間是否影響第二次期中考成績，並假設學習時間為可能的中介變量



◆ 分析流程

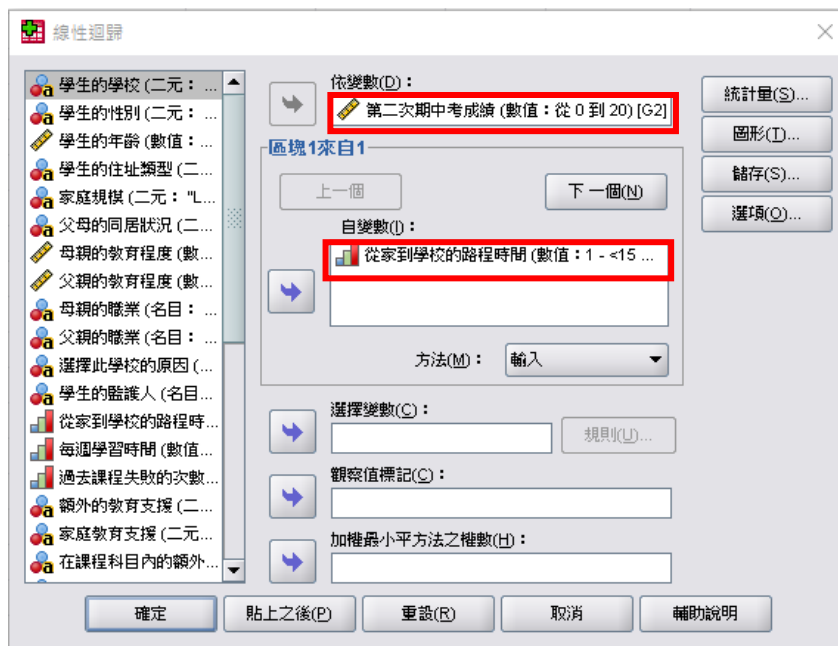
1. 確認通勤時間對於**學習時間**有影響 (A)
2. 確認**學習時間**對於第二次期中考成績有影響 (B)
3. 確認通勤時間對於第二次期中考成績有影響 (C)
4. 當同時考慮通勤時間和學習時間(C')，若通勤時間對於成績的影響變得**不顯著或削弱**，則稱學習時間為中介變量



中介變數檢驗：SPSS操作

◆ 分析流程

1. 確認通勤時間對於第二次期中考成績有影響 (C)



係數^a

模式	未標準化係數		標準化係數	t	顯著性
	B 之估計值	標準誤差	Beta 分配		
1 (常數)	11.910	.432		27.571	.000
從家到學校的路程時間 (數值：1 - <15 分鐘，2 - 15 到 30 分鐘，3 - 30 分鐘到 1 小時，或 4 - >1 小時)	-.826	.269	-.153	-3.073	.002

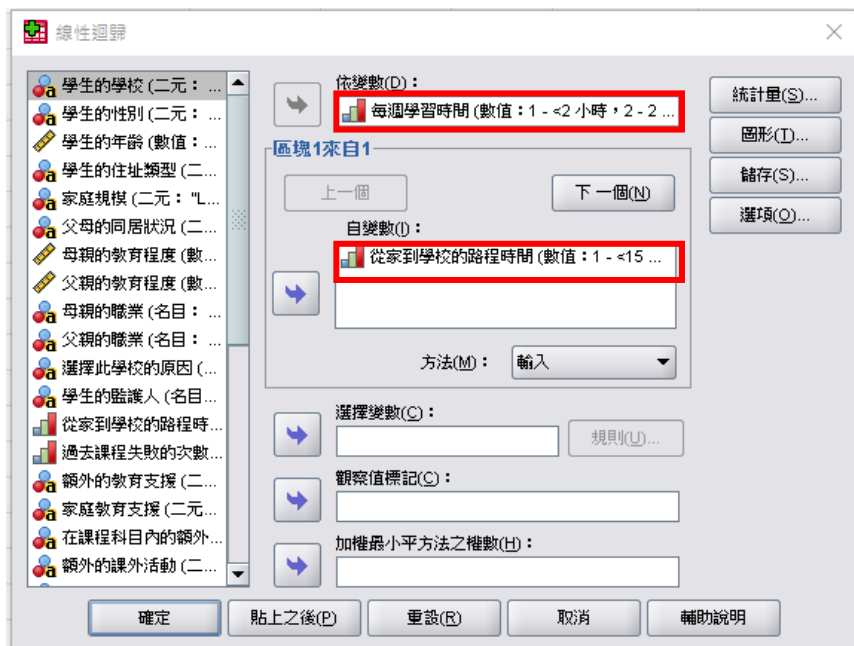
a. 依變數: 第二次期中考成績 (數值：從 0 到 20)

顯著！

中介變數檢驗：SPSS操作

◆ 分析流程

2. 確認通勤時間對於學習時間有影響 (A)



係數^a

模式	未標準化係數		標準化係數	t	顯著性
	B 之估計值	標準誤差	Beta 分配		
1 (常數)	2.211	.097		22.789	.000
從家到學校的路程時間 (數值：1 - <15 分鐘, 2 - 15 到 30 分鐘, 3 - 30 分鐘到 1 小時, 或 4 - >1 小時)	-.121	.060	-.101	-2.011	.045

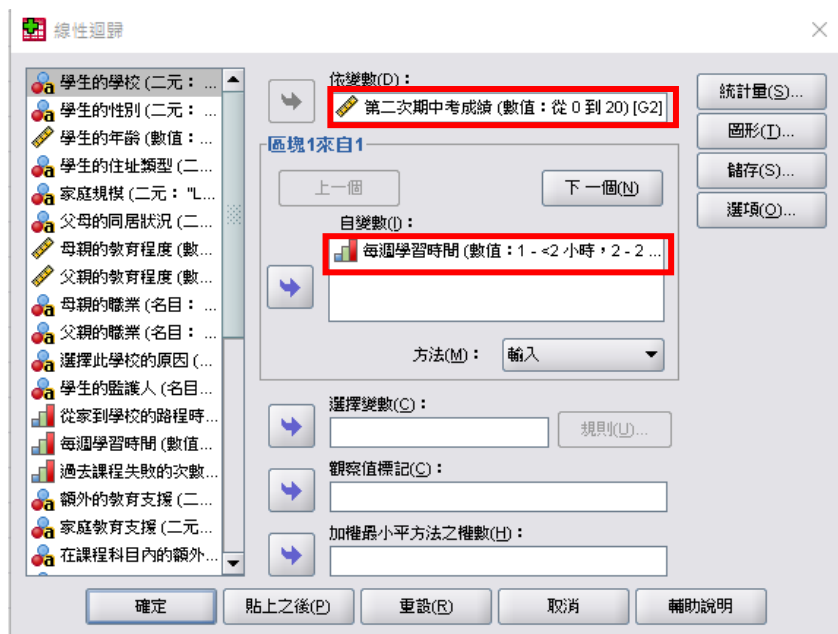
a. 依變數: 每週學習時間 (數值：1 - <2 小時, 2 - 2 到 5 小時, 3 - 5 到 10 小時, 或 4 - >10 小時)

顯著！

中介變數檢驗：SPSS操作

◆ 分析流程

3. 確認學習時間對於第二次期中考成績有影響 (B)



係數^a

模型	未標準化係數		標準化係數	t	顯著性
	B 之估計值	標準誤差	Beta 分配		
1 (常數)	9.474	.493		19.215	.000
每週學習時間 (數值：1 - <2 小時, 2 - 2 到 5 小時, 3 - 5 到 10 小時, 或 4 - >10 小時)	.609	.224	.136	2.719	.007

a. 依變數: 第二次期中考成績 (數值：從 0 到 20)

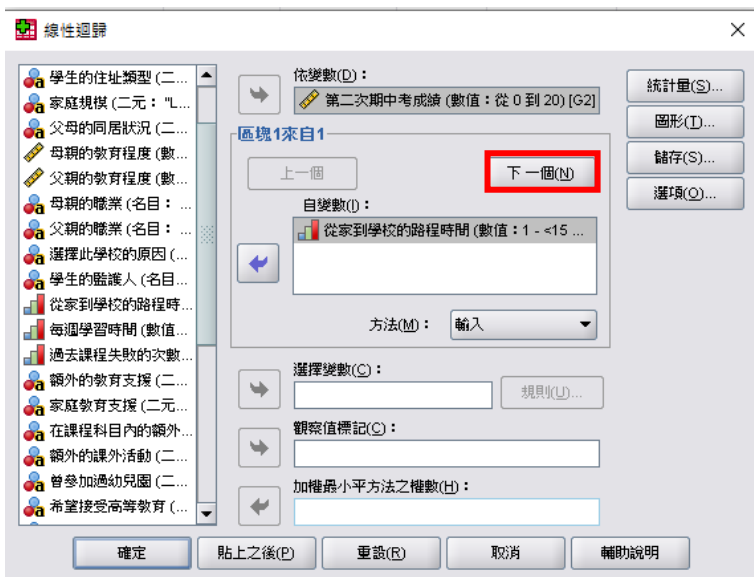
顯著!

中介變數檢驗：SPSS操作

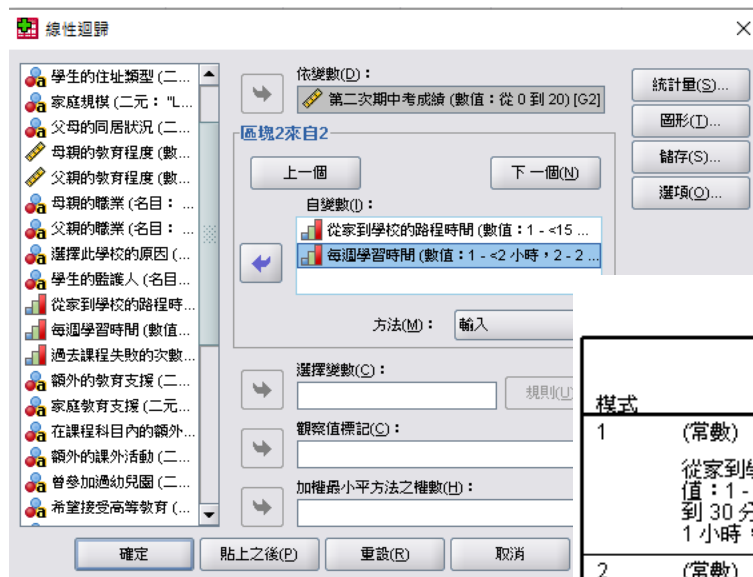
◆ 分析流程

4. 當同時考慮通勤時間和學習時間(C')，若通勤時間對於成績的影響變得**不顯著或削弱**，則稱學習時間為中介變量

①



②



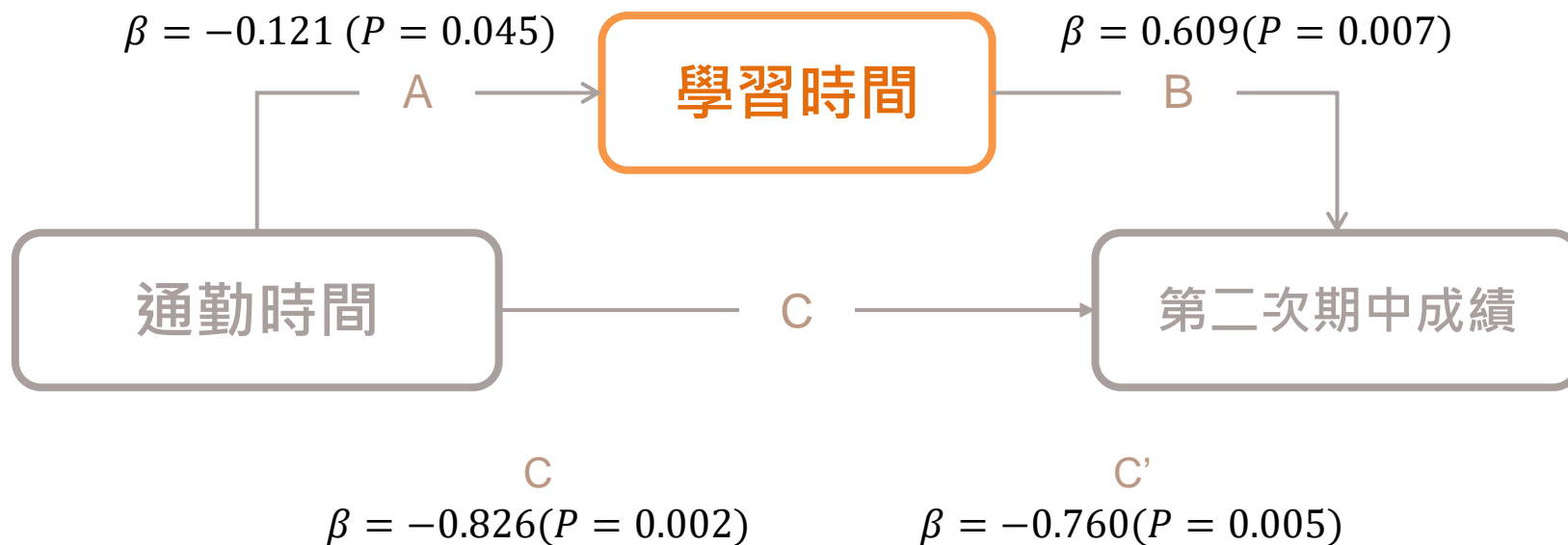
仍顯著→學習時間為部分中介
(若不顯著則為完全中介)

模式	未標準化係數		標準化係數		t	顯著性
	B 之估計值	標準誤差	Beta 分配			
1 (常數)	11.910	.432			27.571	.000
從家到學校的路程時間 (數值：1 - <15 分鐘, 2 - 15 到 30 分鐘, 3 - 30 分鐘到 1 小時, 或 4 - >1 小時)	-826	.269	-.153		-3.073	.002
2 (常數)	10.705	.654			16.367	.000
從家到學校的路程時間 (數值：1 - <15 分鐘, 2 - 15 到 30 分鐘, 3 - 30 分鐘到 1 小時, 或 4 - >1 小時)	-760	.269	-.141		-2.830	.005
每週學習時間 (數值：1 - <2 小時, 2 - 2 到 5 小時, 3 - 5 到 10 小時, 或 4 - >10 小時)	.545	.223	.122		2.443	.015

a. 依變數：第二次期中考成績 (數值：從 0 到 20)

中介變數檢驗

- ◆ 研究問題：通勤時間是否影響第二次期中考成績，並假設學習時間為可能的中介變量
- ⇒ 學習時間為部分中介變量，代表通勤時間對成績的影響並非完全透過學習時間這個中介變量來發揮

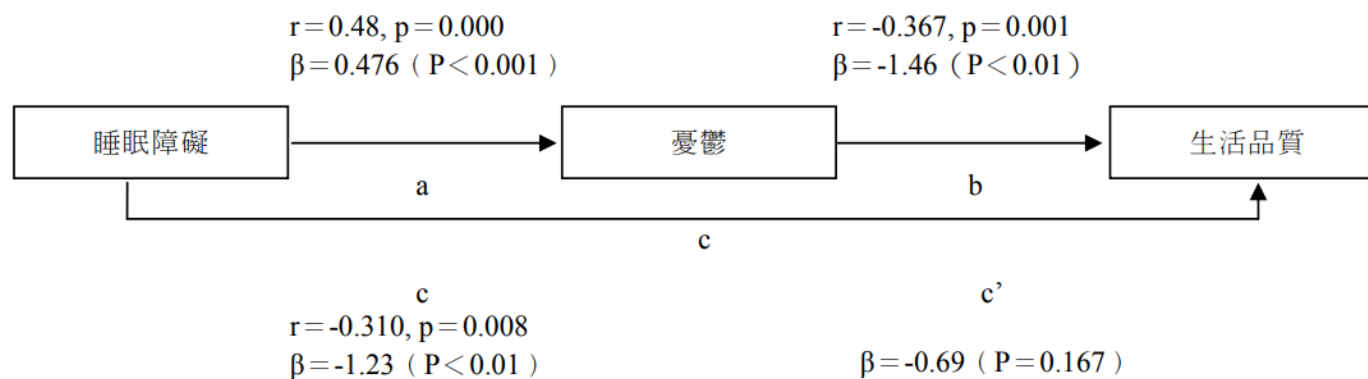


* β 為通勤時間之估計係數



接受肝動脈栓塞術前肝癌病患之睡眠品質、憂鬱與生活品質間之關係

憂鬱在睡眠障礙影響病患生活品質的過程中為**完全中介變量** ($P = 0.167$)，亦即當肝癌病患睡眠障礙產生時，會導致病患出現憂鬱的情緒，而憂鬱進一步使病患之生活品質變差，睡眠障礙的產生並不直接對生活品質造成影響。

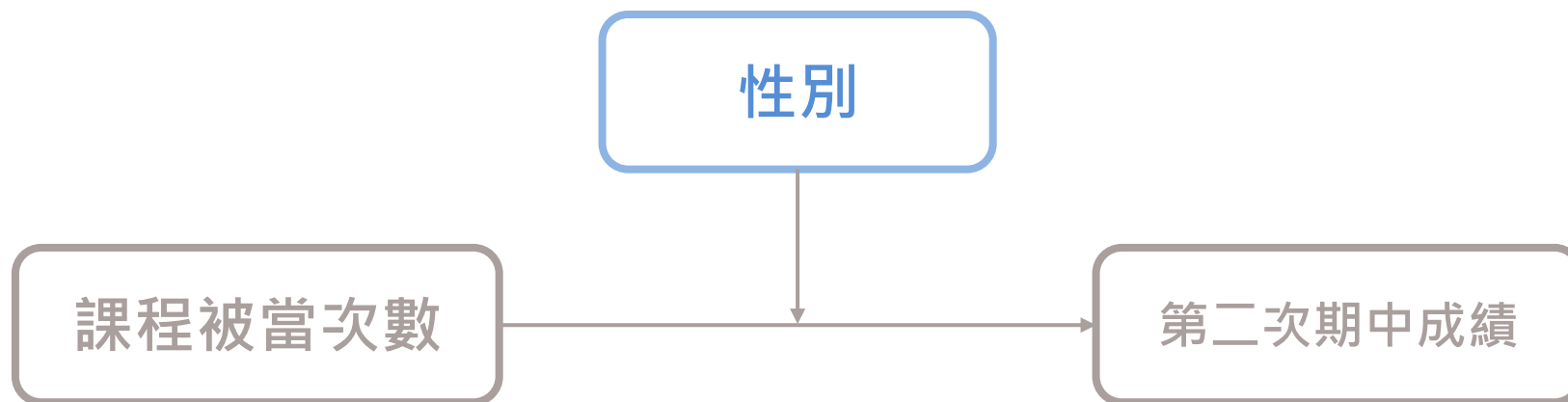


利用多階段迴歸分析測試憂鬱在睡眠障礙與生活品質間之中介效果

	自變項	依變項	β	標準誤	Standardized β	P
Step 1	睡眠障礙	憂鬱	0.476	0.103	0.48	0.000
Step 2	睡眠障礙	生活品質	-1.226	0.446	-0.310	0.008
Step 3	憂鬱	生活品質	-1.461	0.440	-0.367	0.001
Step 4	睡眠障礙	生活品質	-0.69	0.494	-0.175	0.167
	憂鬱		-1.128	0.498	-0.283	0.027

調節變數檢驗

◆ 研究問題：性別對學生課程被當次數與第二次期中考成績關係的調節作用



◆ 分析流程

1. 配適期中成績與被當次數之迴歸模型
2. 配適期中成績與被當次數及性別之迴歸模型
3. 配適期中成績與被當次數、性別及 (被當次數×性別) 之迴歸模型

* (被當次數×性別) 通常被稱為調節變量或交互作用項



調節變數檢驗：SPSS操作

- ◆ 研究問題：性別對學生課程被當次數與第二次期中考成績關係的調節作用
- ◆ 操作流程

1. 建立性別變數(sex)的虛擬變數

The image shows two SPSS dialog boxes for creating a dummy variable. The first dialog, titled "重新編碼成不同變數" (Recode into Different Variables), shows the variable "sex" being recoded into "sex_male". The second dialog, titled "重新編碼成不同變數：舊值與新值" (Recode into Different Variables: Old and New Values), shows the mapping of old values to new values.

分別輸入
1. 舊值：M, 新值:1
2. 舊值：F, 新值:0

Dialog 1: 重新編碼成不同變數
名稱: sex_male
字串變數(Y) -> 輸出變數:
sex --> sex_male
輸出之新變數:
名稱(N): sex_male
標記(L):
變更(H)
舊值與新值(O)...
若(I)... (可省略觀察值選擇條件)
確定 貼上之後(P) 重設(R) 取消 輔助說明

Dialog 2: 重新編碼成不同變數：舊值與新值
舊值:
 數值(V):
系統遺漏值(S)
系統或使用者遺漏值(U)
範圍(N):
到(I):
範圍, LOWEST 到值(G):
範圍, 值到 HIGHEST(E):
 全部其他值(O)
新值為:
 數值(A):
 系統遺漏值(Y)
 複製舊值(P)
舊值 --> 新值(D):
'F' --> 0
'M' --> 1
新增(A) 變更(C) 移除(R)
 輸出變數為字串(B) 寬度(W): 8
 將數值字串轉換為數字 ('5' -> 5)(M)
繼續 取消 輔助說明

調節變數檢驗：SPSS操作

- ◆ 研究問題：性別對學生課程被當次數與第二次期中考成績關係的調節作用
- ◆ 操作流程

1. 建立性別變數(sex)的虛擬變數

The image shows two SPSS dialog boxes for creating a dummy variable. The first dialog, titled "重新編碼成不同變數", shows the variable "sex" being recoded into "sex_male". The second dialog, titled "重新編碼成不同變數：舊值與新值", shows the mapping of old values to new values.

分別輸入

1. 舊值：M, 新值:1
2. 舊值：F, 新值:0

Dialog 1: 重新編碼成不同變數

- 名稱(N): sex_male
- 字串變數(Y) -> 輸出變數: sex --> sex_male
- 輸出之新變數: 名稱(N): sex_male
- 若(I)... (可省略觀察值選擇條件)

Dialog 2: 重新編碼成不同變數：舊值與新值

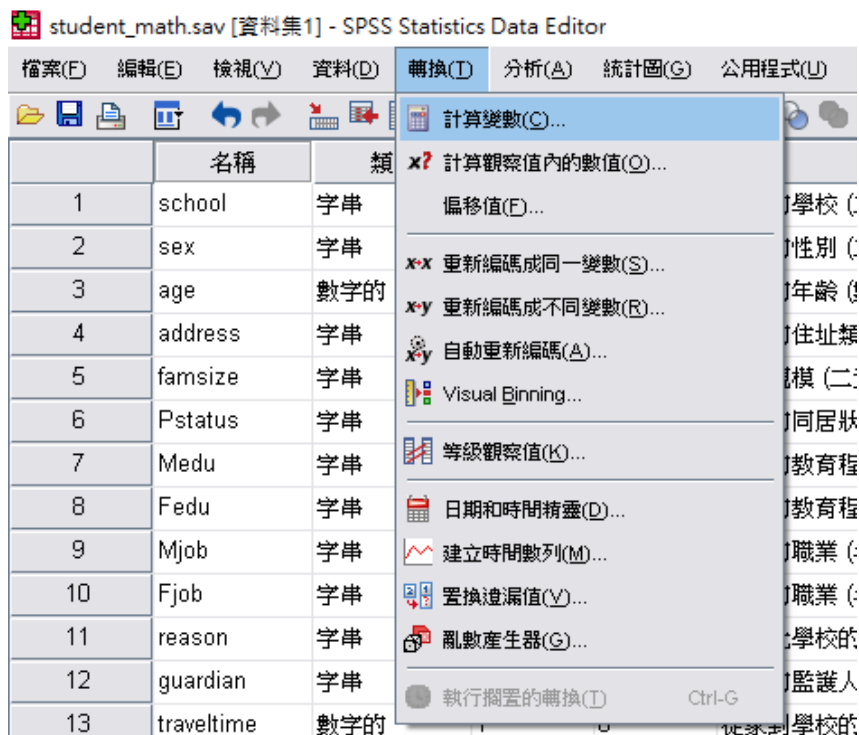
- 舊值: 數值(Y): []
- 新值為: 數值(A): []
- 舊值 --> 新值(D): 'F' --> 0, 'M' --> 1
- 輸出變數為字串(B) []
- 將數值字串轉換為數字 ('5' -> 5)(M) []

調節變數檢驗：SPSS操作

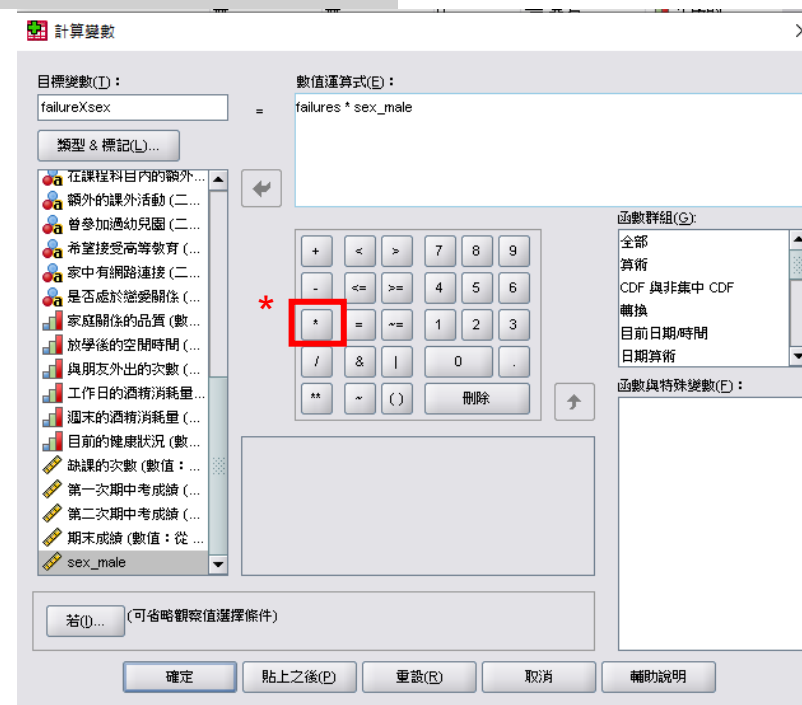
◆ 研究問題：性別對學生課程被當次數與第二次期中考成績關係的調節作用

◆ 操作流程

2. 建立調節變量(被當次數×性別)



目標變數：failureXsex



調節變數檢驗：SPSS操作

◆ 研究問題：性別對學生課程被當次數與第二次期中考成績關係的調節作用

◆ 操作流程

3. 建立階層迴歸模型

The image displays three sequential screenshots of the SPSS '線性迴歸' (Linear Regression) dialog box, illustrating the process of building a hierarchical regression model. Each window shows a list of variables on the left and a list of variables added to the model on the right. The '依變數(D):' (Dependent Variable) is consistently '第二次期中考成績 (數值: 從 0 到 20) [G2]'. The '自變數(I):' (Independent Variable) is '過去課程失敗的次數 (數值: n 如果 1<=...)' in the first window, 'sex_male' in the second, and 'failureXsex' in the third. The '下一個(N)' (Next) button is highlighted in red in each window, indicating the progression through the steps. The '方法(M):' (Method) is set to '輸入' (Stepwise) in all three windows.

調節變數檢驗

◆ 研究問題：性別對學生課程被當次數與第二次期中考成績關係的調節作用

➤ 估計迴歸模型：

$$\hat{y} = 10.783 - 1.30x_1 + 1.135x_2 - 0.986x_1x_2$$

➤ 男性($x_2 = 1$)

$$\begin{aligned} \hat{y} &= 10.783 - 1.30x_1 + 1.135 \times 1 - 0.986x_1 \times 1 \\ &= (10.783 + 1.135) - (1.30 + 0.986)x_1 \\ &= 11.915 - 2.286x_1 \end{aligned}$$

➤ 女性($x_2 = 0$)

$$\begin{aligned} \hat{y} &= 10.783 - 1.30x_1 + 1.135 \times 0 - 0.986x_1 \times 0 \\ &= 10.783 - 1.30x_1 \end{aligned}$$

係數^a

模式		未標準化係數		標準化係數	t	顯著性	B 的 95.0% 信賴區間	
		B 之估計值	標準誤差	Beta 分配			下界	上界
1	(常數)	11.316	.194		58.269	.000	10.934	11.697
	過去課程失敗的次數 (數值：n 如果 1 ≤ n < 3，否則為 4)	-1.800	.238	-.356	-7.550	.000	-2.269	-1.331
2	(常數)	10.942	.253		43.219	.000	10.444	11.440
	過去課程失敗的次數 (數值：n 如果 1 ≤ n < 3，否則為 4)	-1.824	.237	-.361	-7.684	.000	-2.291	-1.357
	sex_male	.806	.353	.107	2.282	.023	.112	1.500
3	(常數)	10.783	.263		40.941	.000	10.265	11.301
	x_1 過去課程失敗的次數 (數值：n 如果 1 ≤ n < 3，否則為 4)	-1.300	.345	-.257	-3.766	.000	-1.979	-.621
	x_2 sex_male	1.135	.386	.151	2.944	.003	.377	1.893
	failureXsex	-.986	.474	-.150	-2.081	.038	-1.917	-.055

a. 依變數：第二次期中考成績 (數值：從 0 到 20)



調節變數檢驗

◆ 研究問題：性別對學生課程被當次數與第二次期中考成績關係的調節作用

➤ 估計迴歸模型：

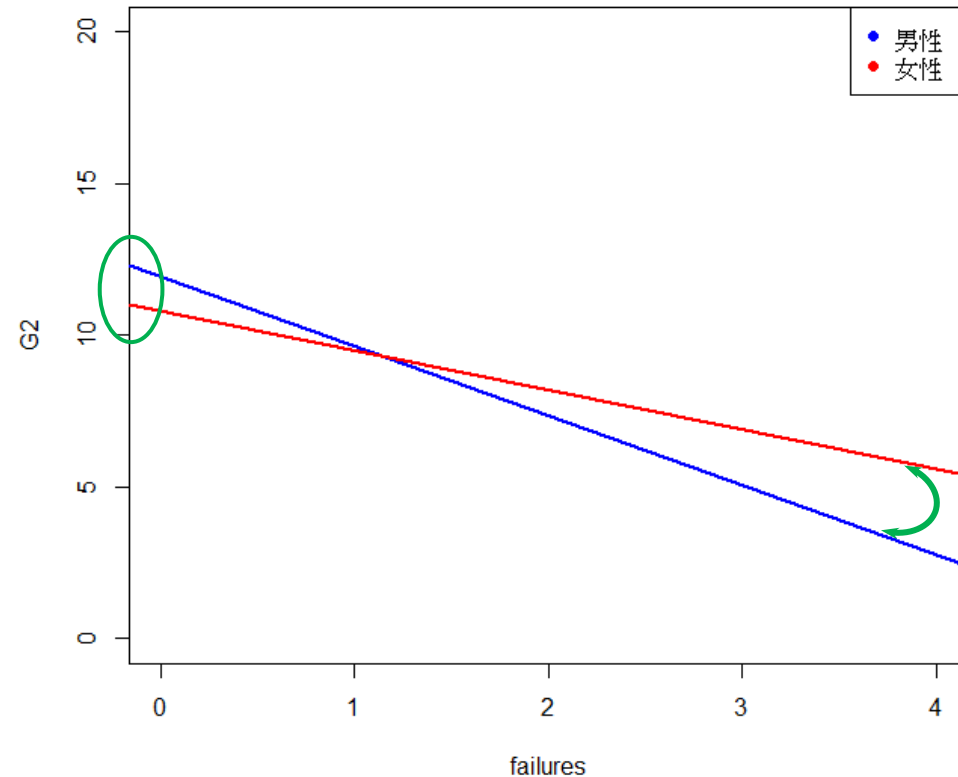
$$\hat{y} = 10.783 - 1.30x_1 + 1.135x_2 - 0.986x_1x_2$$

➤ 男性($x_2=1$)

$$\begin{aligned}\hat{y} &= 10.783 - 1.30x_1 + 1.135 \times 1 - 0.986x_1 \times 1 \\ &= (10.783 + 1.135) - (1.30 + 0.986)x_1 \\ &= 11.915 - 2.286x_1\end{aligned}$$

➤ 女性($x_2=0$)

$$\begin{aligned}\hat{y} &= 10.783 - 1.30x_1 + 1.135 \times 0 - 0.986x_1 \times 0 \\ &= 10.783 - 1.30x_1\end{aligned}$$

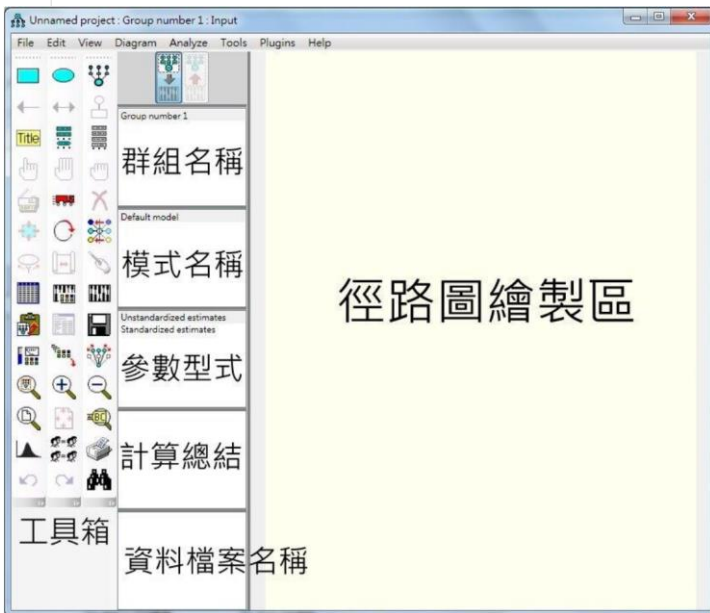


路徑分析

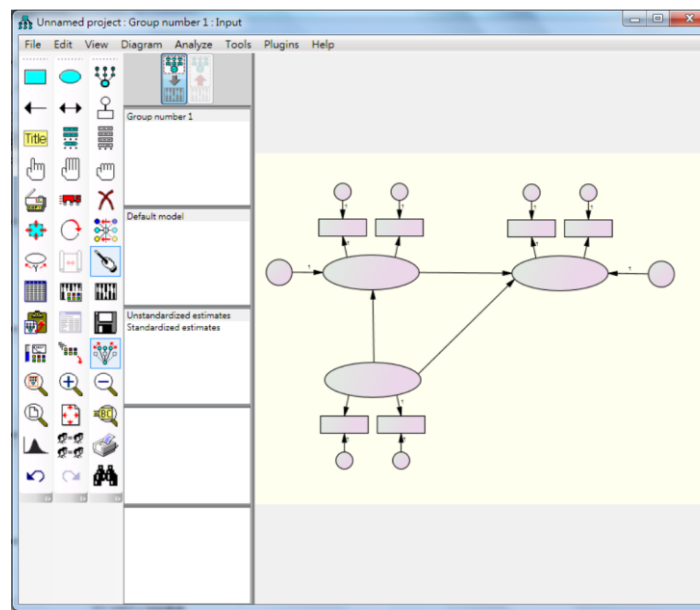
基本路徑分析教學：

<https://www.youtube.com/watch?v=M0tsd2snBFM>

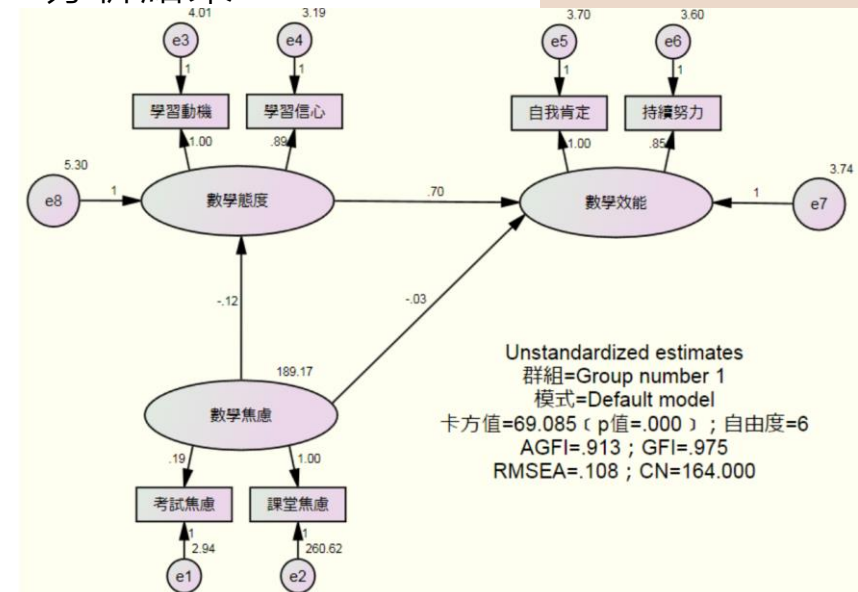
AMOS軟體介面



繪製徑路圖



分析結果



圖表 30. 完成參數估計後的徑路圖



THANKS

聯絡信箱：r28111553@gs.ncku.edu.tw

