

Propensity Score 傾向分數的使用

醫學研究部生統小組：陳韻仔 博士

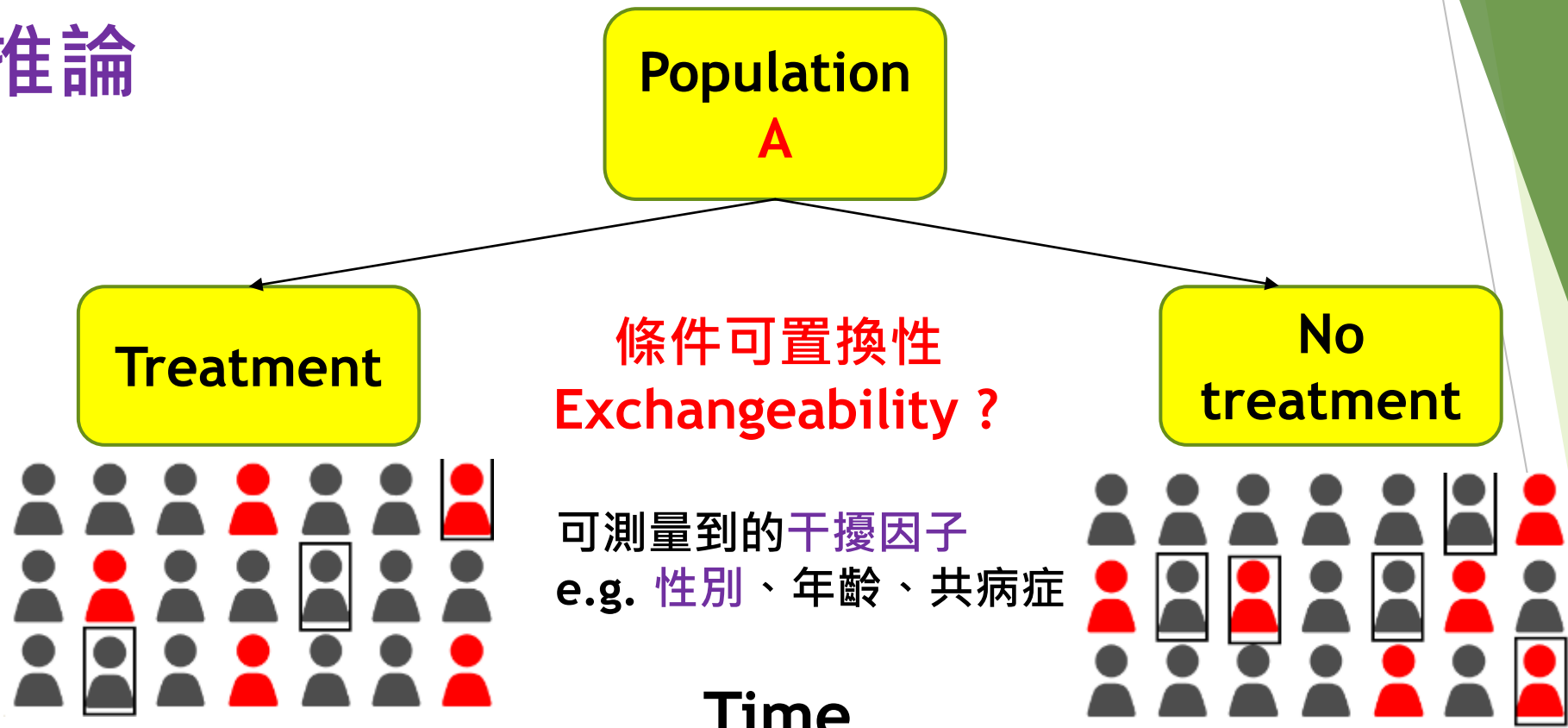
授課日期：111年12月7日

什麼是 傾向分數 (Propensity-score)?

- ▶ 傾向分數分析的目的：**降低混淆（干擾）效果**
- ▶ 治療組（實驗組）與非治療組（對照組）：
 - ▶ 干擾變項（X）的分佈相似
- ▶ 可視為代表所有預測變項的綜合分數（summary score）
 - ▶ 有相同分數的2個個案，儘管他們可能實際上是控制組，他們有相同的預測機率會成為治療組

因果推論

觀察性研究



如何進行因果推論?

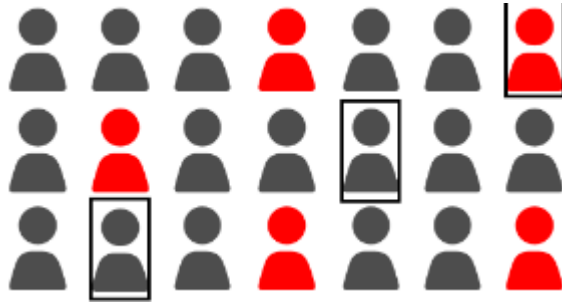
多變項迴歸校正 (multivariable regression adjustment)

因果推論

觀察性研究

Population
A

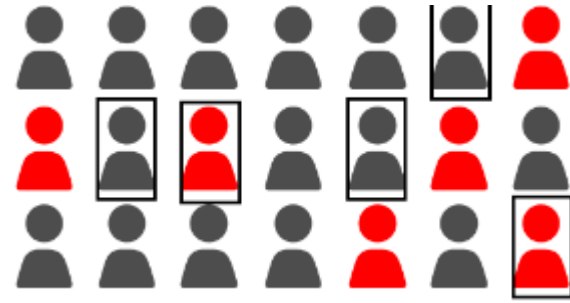
Treatment



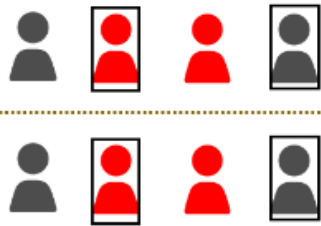
條件可置換性
Exchangeability ?

可測量到的混淆因子
e.g. 性別

No
treatment



多變項迴歸校正 (multivariable regression adjustment) → 事件數或樣本數比較小 ?



Propensity Score (PS)

計算傾向值 (使用Logistic Regression)

SPSS 操作步驟

分析 → 迴歸(R) → 二元Logistic

*無標題2 [資料集1] - IBM SPSS Statistics 資料編輯器

檔案(E) 編輯(E) 檢視(V) 資料(D) 轉換(T) 延伸(X) 視窗(W) 說明(H)

報告(P)
敘述統計(E)
貝氏統計資料(B)
表格(B)
比較平均數法(M)
一般線性模型(G)
概化線性模型(Z)
混合模型(X)
相關(C)
迴歸(R)
對數線性(O)
神經網路(W)
分類(E)
維度縮減(D)
比例(A)
無母數檢定(N)
預測(I)
存活(S)
複選題(U)
遺漏值分析(Y)...
多重插補(I)
複式樣本(L)
模擬(U)...
品質控制(Q)

自動線性建模(A)...
線性(L)...
曲線估計(C)...
局部最小平方(S)...
二元 Logistic...
多項式 Logistic(M)...
序數(D)...
Probit...
非線性(N)...
加權估計(W)...
二階最小平方(Q)...
最適尺度 (CATREG)(Q)...

	id	mdr	gp	ef_final	dm	htn	mi	af	b
1	1	222982		27.20	1	1	0	1	
2	2	283582		26.90	0	1	1	1	
3	3	292321		31.50	1	0	1	0	
4	4	319334		30.20	1	1	0	1	
5	5	323000		34.40	0	0	0	1	
6	6	324075					1	0	
7	7	327264					1	0	
8	8	340088					1	0	
9	9	360953					0	1	
10	10	363086					0	0	
11	11	368017					0	0	
12	12	380953					0	0	
13	13	381077					0	1	
14	14	381782					1	0	
15	15	471583					0	0	
16	16	504882					0	0	
17	17	548244					1	0	
18	18	604966					1	1	
19	19	620795					0	0	
20	20	773174		33.50	0	0	1	0	



Logistic 迴歸

應變數(D): gp_new [gp_new]

區塊(B)1 / 1

共變數(C): age, sex, ef_final, dm, htn, mi, af, bb_g, mra_g

方法(M): 進入

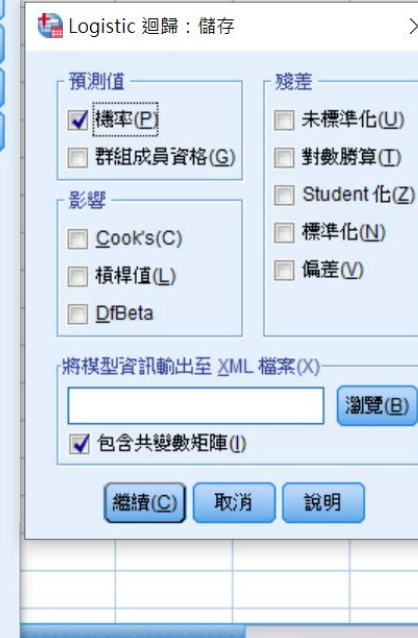
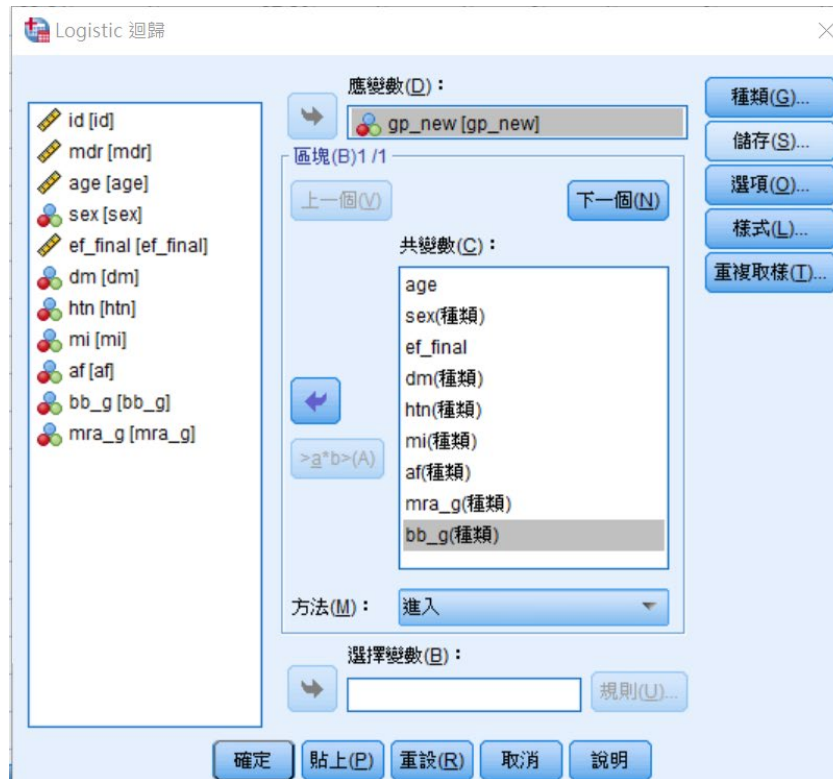
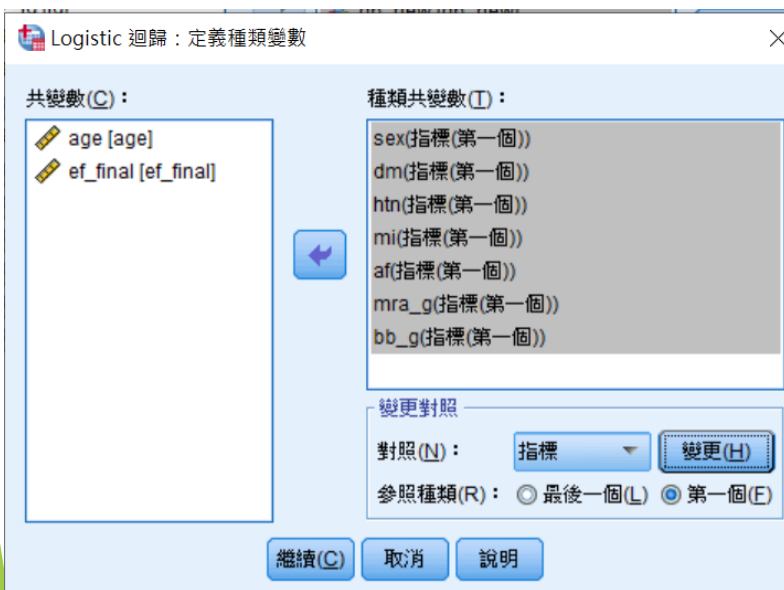
選擇變數(B):

確定 貼上(P) 重設(R) 取消 說明

計算傾向值 (使用Logistic Regression)

SPSS 操作步驟

分析 → 迴歸(R) → 二元Logistic →
種類 → 定義種類 (類別) 變數 → 儲存 (預測值：機率 P)



PRE_1
.24882
.27003
.19716
.29439
.18979
.40181
.30750
.22864
.35028
.20505
.18593
.37815
.16281
.16200
.50430
.41441
.22270
.28449
.16712
.23341

計算傾向值 (使用Logistic Regression)

SAS 操作步驟

```
libname data 'F:\中榮醫研部-生統小組\全院教育課程規劃-2022oct\111年度第四季全院生統教育課程\20221207 傾向分數的使用\PS matching'; /*建立SAS資料集*/
```

```
/*STEP 1: Estimating the Propensity Score (PS)*/
```

```
proc sort data=data.arni_f; by descending gp_new; run;
```

```
Title j=center height=12 pt font=Arial Bold Italic "PS-logistic regression model fitting";
```

```
Proc logistic data=data.arni_f;  
class sex dm htn mi af bb_g mra_g /* (ref=first)*/;  
model gp_new= age sex ef_final dm htn mi af bb_g mra_g / lackfit;  
/*requests Hosmer and Lemeshow goodness-of-fit test*/  
output out=out_ps prob=ps xbeta=logit_ps; /*create new data set: out_ps*/  
run;  
/*new variable: ps:propensity score logit_ps: logit of propensity score*/
```

計算傾向值 (使用Logistic Regression)

R 操作步驟

```
#Propensity score model
```

```
psmodel <- glm (gp_new ~ age + sex + dm + htn ,  
  family = binomial (link = "logit"), data =data_ipwt )  
summary (psmodel)
```

```
#Value of propensity score for each subject  
ps_w <- predict (psmodel, type = "response")  
ps_case <- cbind (ps_w, data_ipwt )
```

```
#合併時千萬不可以有missing data  
#把算好的PS存成excel.csv檔
```

```
setwd("D:/助理研究員/中榮醫研部-生統小組/全院教育課程規劃-2022oct/111年第4季/20221207 傾向  
分數的使用/PS matching")  
write.table(ps_case, "ps_case.csv", quote=F, row.names=F)
```


傾向分數 (Propensity Score) 的適用情形

- ▶ 以下兩種情況時，可使用傾向性評分匹配，藉由**邏輯迴歸模型**來決定各實驗組與對照組的評分
 - ▶ 在研究中，實驗組與對照組可直接比較的**個體數量很少**，若直接將兩者進行比對，容易產生非常偏倚的結果
 - ▶ 在研究中，當衡量**個體特徵的變數很多**時，若想要從對照組中找出一組各項變數都與實驗組相同或相近的子集便會變得非常困難

Logistic regression

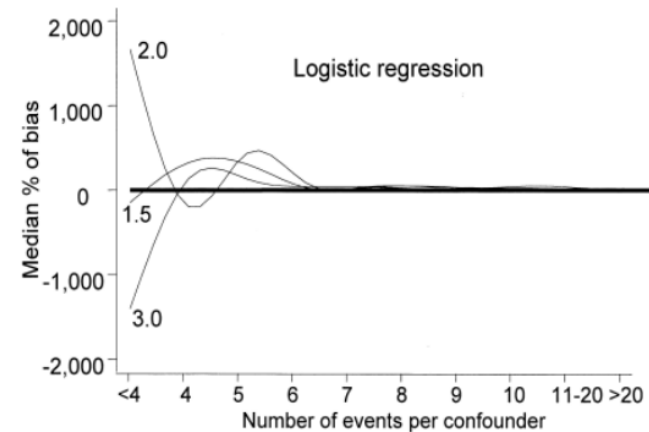
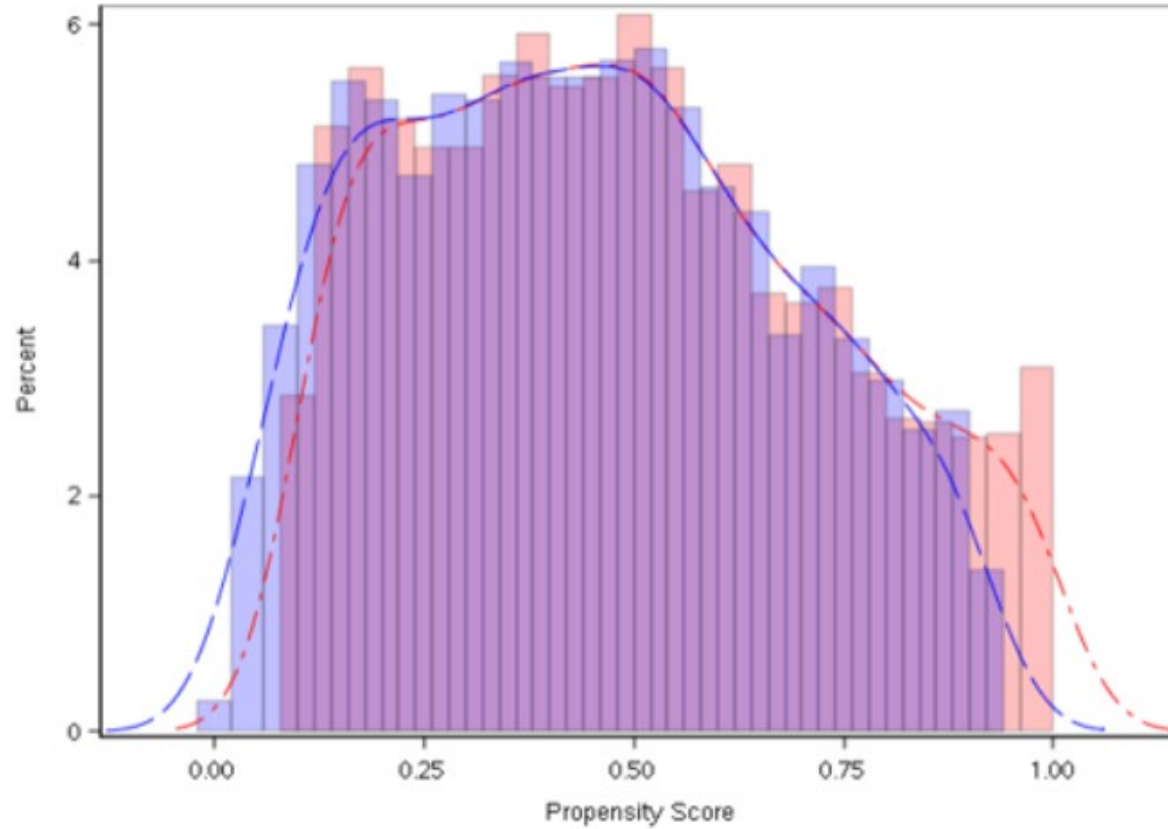


FIGURE 1. Median percentage of bias with the logistic regression, by strength of the exposure and number of events per confounder. In the logistic regression, the bias declines as the number of events per confounder increases. Values greater than zero indicate an overestimation of the effect of the exposure on the outcome. Negative values indicate an underestimation of the effect of the exposure on the outcome.

Propensity Score 的限制

理想世界 (隨機分派實驗)

PS 密度分佈圖：重疊性高的分布

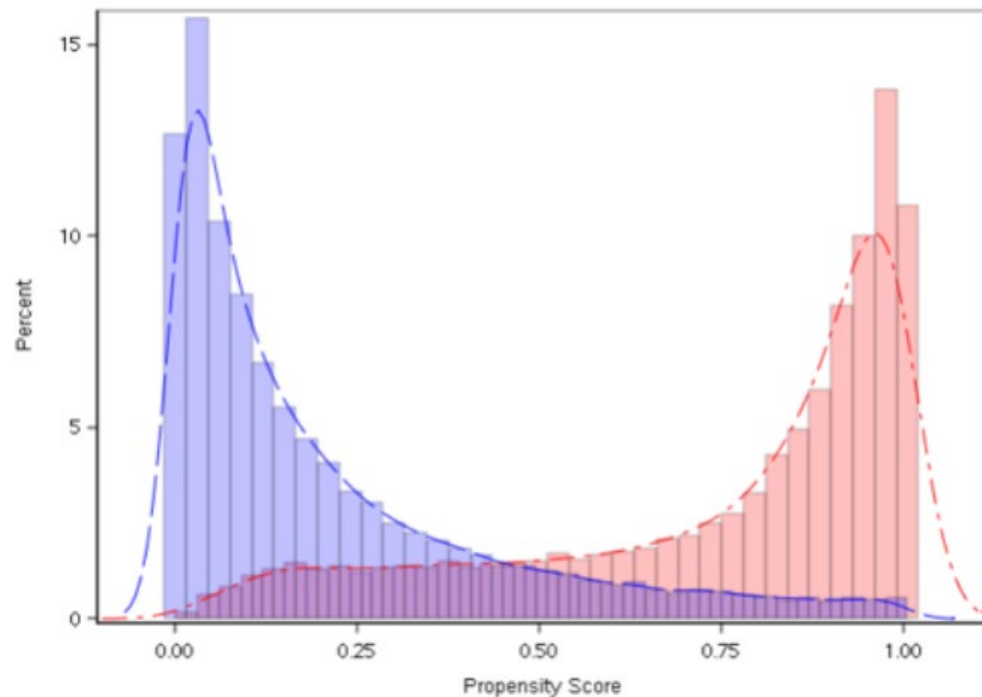


二群的重疊性很高 (good overlap)，表示**基線特徵的差異很小**，此時以 PS 進行干擾控制會得到很好的效果

Propensity Score 的限制

真實世界

PS 密度分佈圖：重疊性低的分布



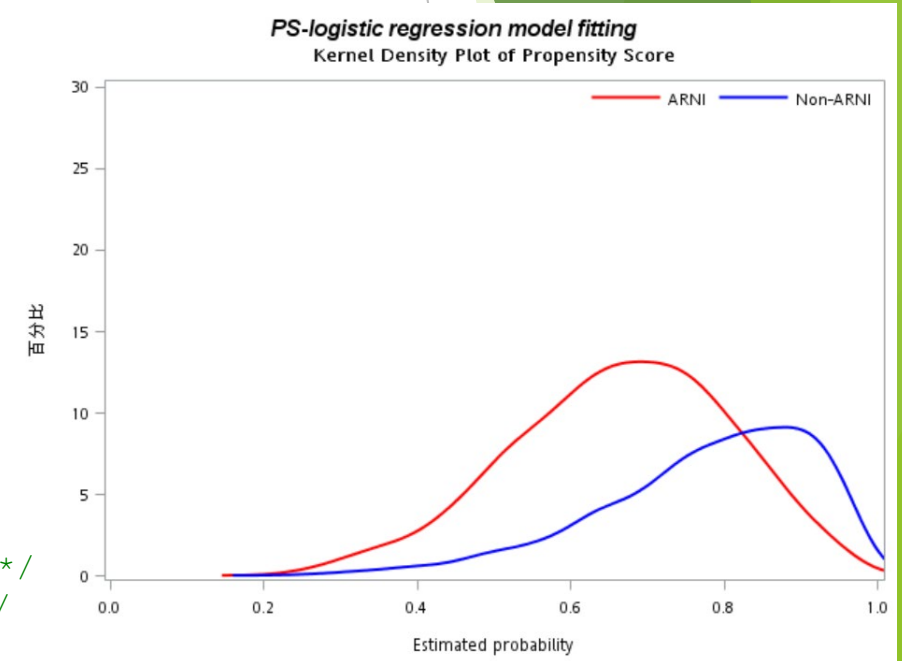
常常二組治療的PS 密度分布重疊性很低 (poor overlap)，此時 PS 在極大與極小值附近比例的太高，表示兩組基線特徵存在過大的差異

雖然使用配對或 IPTW 的方式仍然可以有效控制干擾，但是通常會偏離目標族群，直接影響了兩組的平衡及大大的降低準確度

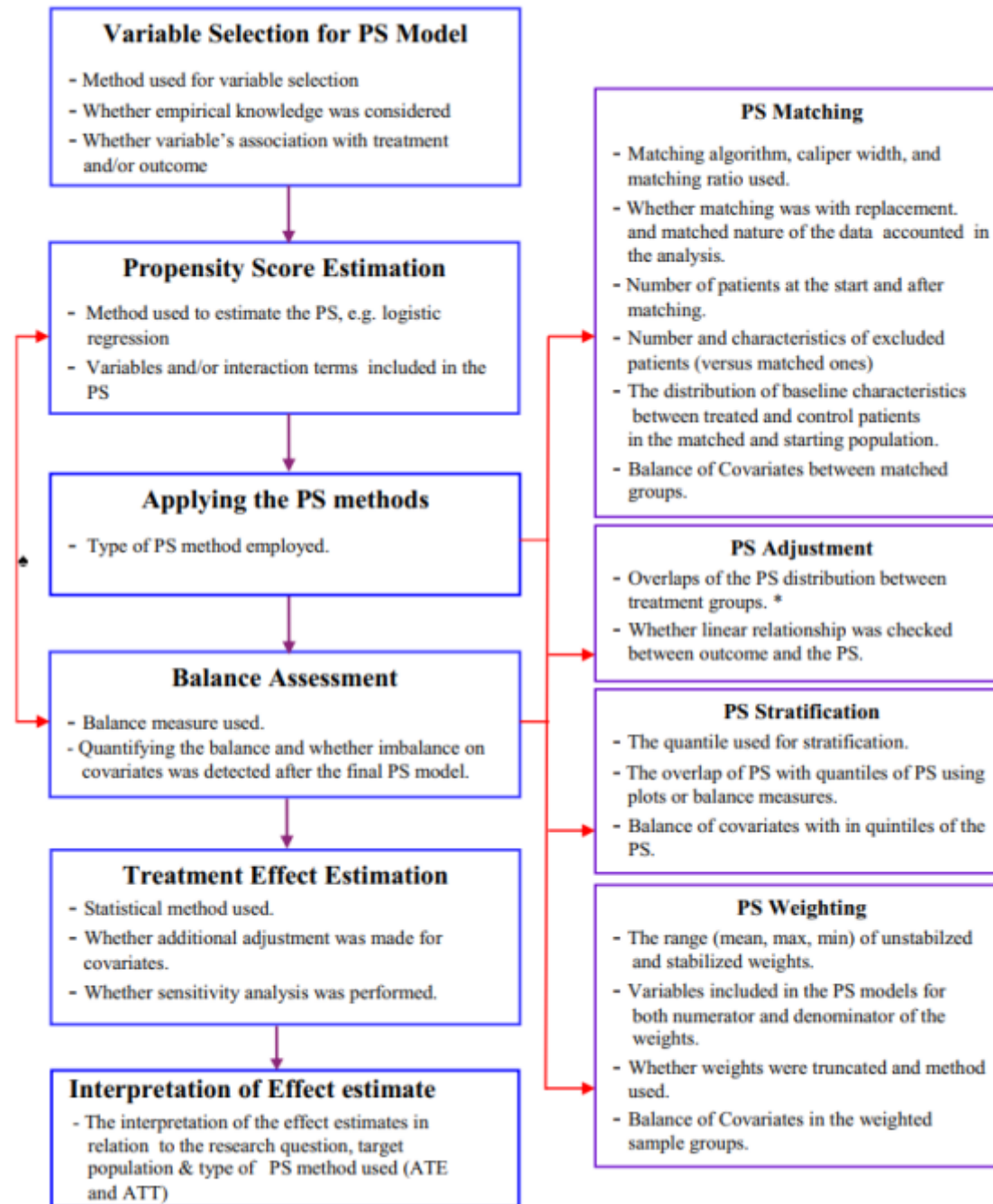
計算傾向值 (使用Logistic Regression)

SAS 操作步驟

```
/* (2) Creating Kernel Density Plot of Propensity Score */  
title2 "Kernel Density Plot of Propensity Score";  
data gp;  
    set out_ps(keep=gp_new ps); /*將原本PS值的變項, 依照組別給予不同的欄位名稱*/  
    if gp_new="1" then ARNI=ps;  
    else Non_ARNI=ps;  
run;  
  
proc sgplot data=gp;  
    density ARNI /scale=percent /*若以個數呈現: Count*/  
        type=kernel  
        legendlabel='ARNI' /*組別的名稱*/  
        LINEATTRS=(COLOR=red); /*線條顏色*/  
    density Non_ARNI /scale=percent  
        type=kernel  
        legendlabel='Non-ARNI'  
        LINEATTRS=(color=blue); /*利用不同組別(不同變項)各畫一條但在同一張圖上*/  
    axis label='Estimated probability' max=1 min=0 /*x軸名稱與最大最小值等的設定*/  
    yaxis max=30;  
    keylegend / noborder location=inside position=topright; /*設定組別標籤的格式*/  
run;
```



使用傾向分數



使用傾向分數的趨勢

Table 3. The frequency of the different PS methods and balance assessment

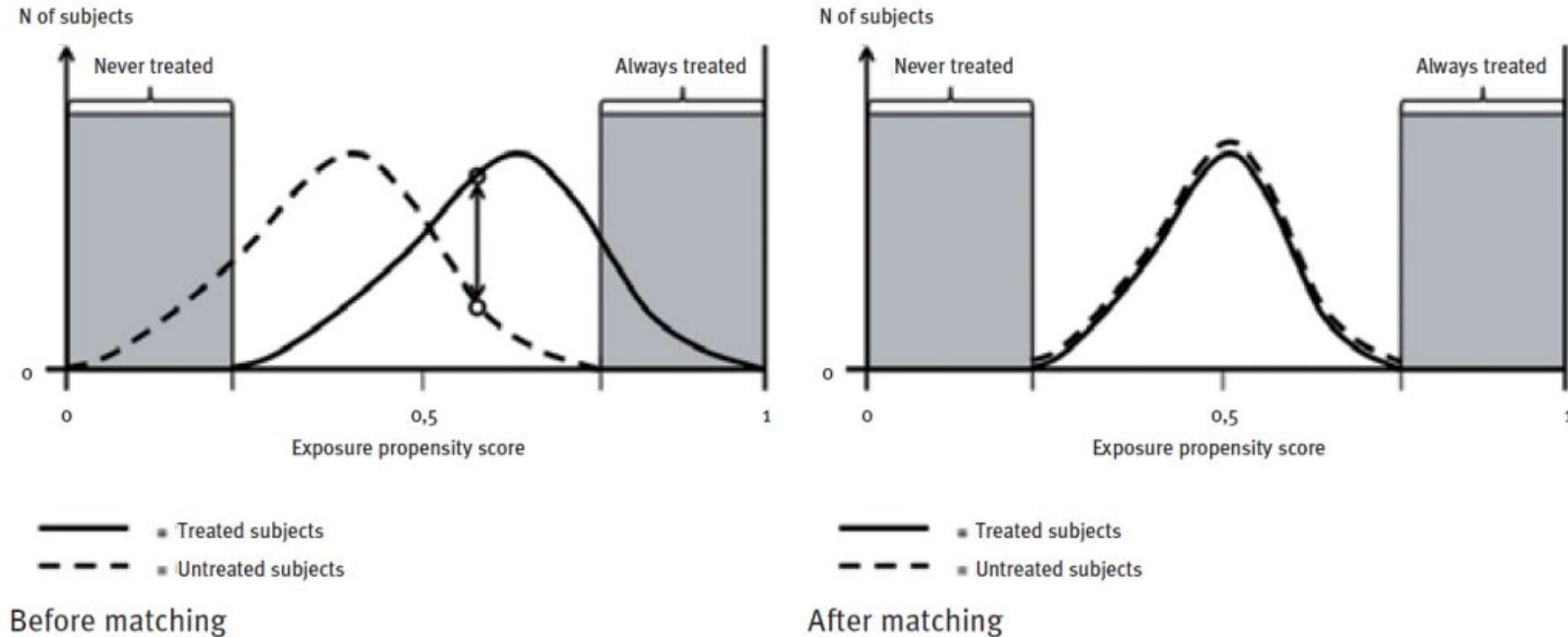
Method	Number of articles (n)	Balance checked
PS matching ^a	204 (68.9)	144 (70.6)
1:1 matching	118	92
1:2 matching	3	2
1:3 matching	4	3
1:4 matching	5	3
Covariate adjustment using PS	62 (20.9)	25 (40.3)
Stratification using PS ^a	41 (13.9)	17 (41.5)
Quintiles of PS	21	10
Deciles of PS	8	1
Quartiles of PS	3	2
Tertiles of PS	5	3
IPTW	21 (7.1)	15 (71.4)
Mixed ^b	26 (8.8)	18 (69.2)

Ali MS, et al. J Clin Epidemiol, 2015.

傾向分數的四種應用方式

- ▶ **配對 (matching)**
- ▶ 加權 (weighting)
- ▶ 分層 (stratification)
- ▶ 迴歸調整 (covariate adjustment in regression model)

Propensity-score Matching



Propensity-score Matching

▶ 優點

- ▶ 相較於分層分析與統計控制，PSM更能有效使兩組的共變項更均勻
- ▶ 很直覺的提供像隨機試驗那樣的報告（i. e. 透明化呈現兩組的基本特性在表格中）

▶ 缺點

- ▶ 損失許多人數的控制組，無法將結果推論到整個群體
- ▶ 由於人數下降，因此統計檢定力（Power）下降（i. e. 會得到較不顯著之結果）

傾向分數配對 (PS matching) 的步驟

- ▶ 計算傾向值 (使用Logistic regression)
- ▶ 進行得分匹配，主要可歸類於以下三種方法
 - ▶ 最鄰近匹配 (Nearest neighbor matching 或 **Greedy matching**)：以傾向得分為依據，在對照組中尋找一個或多個與傾向值相同或相近的樣本作為配對對象，在本次的SPSS範例中便採用此種方法
 - ▶ 半徑匹配(Radius matching)：設定一個常數R (為一個區間，一般為小於傾向得分標準差的四分之一)，並將實驗組得分值與對照組得分值的差異在R內的進行配對
 - ▶ 核匹配 (Kernel Matching)：傾向性評分匹配與核匹配結合後可以通過加權評分增加個別較重要變數的權重，而權重數可由核函數計算得出

傾向分數配對 (PS matching) 的步驟

▶ SPSS操作步驟

- ▶ 在SPSS 19版以後，便可以用外掛方式在SPSS中使用PSM功能
- ▶ 在SPSS 21版以後，可以在功能表「資料」下使用「傾向分數對照」

【例題】

某研究想要瞭解使用「喝酒」(1：喝酒、0；不喝酒)

與「高血壓」之間的關係，

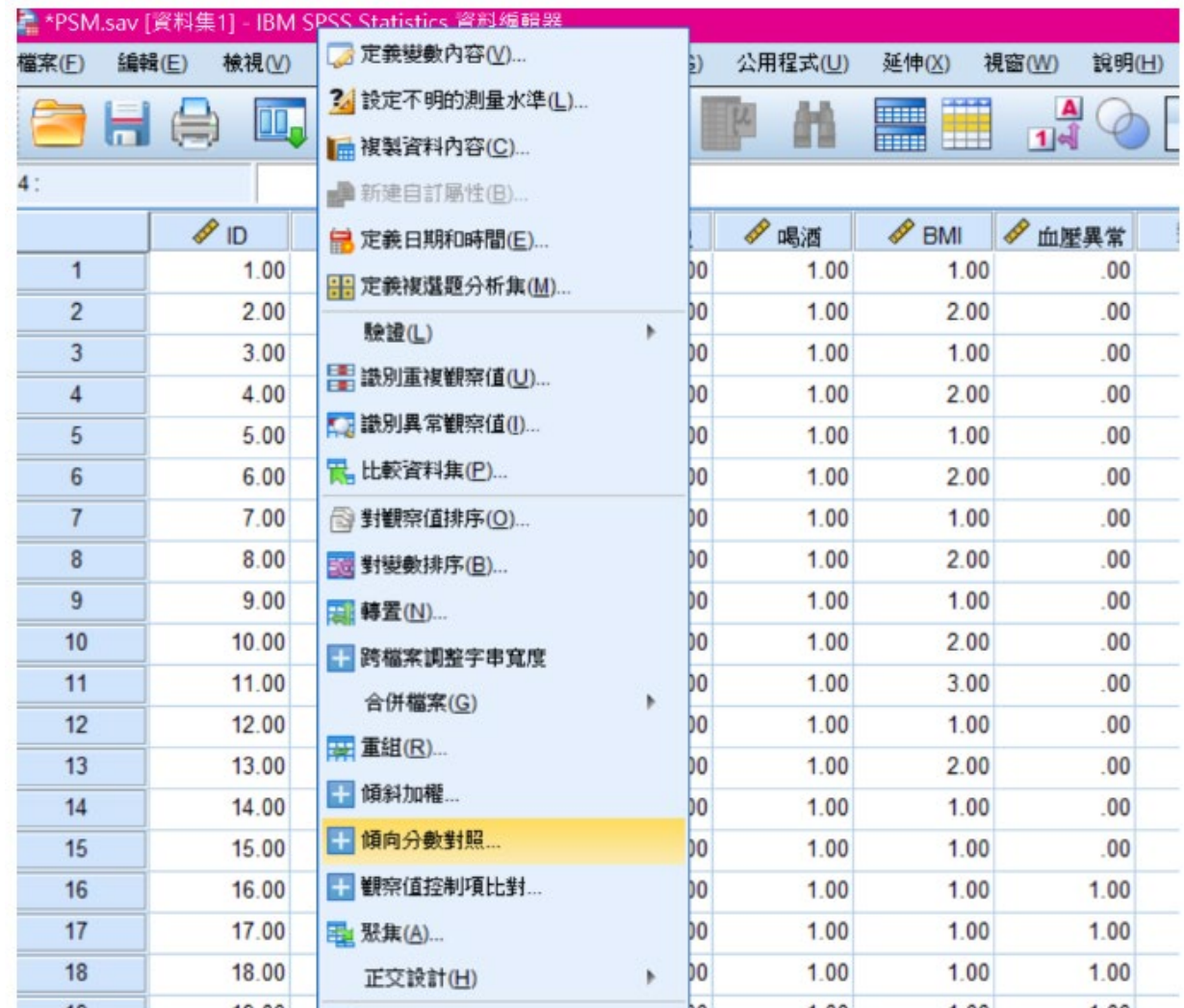
打算使用某項調查資料抽取實驗組與對照組樣本，

進一步觀察「喝酒」與「高血壓」之間的關聯性

傾向分數配對 (PS matching) 的步驟

▶ SPSS操作步驟

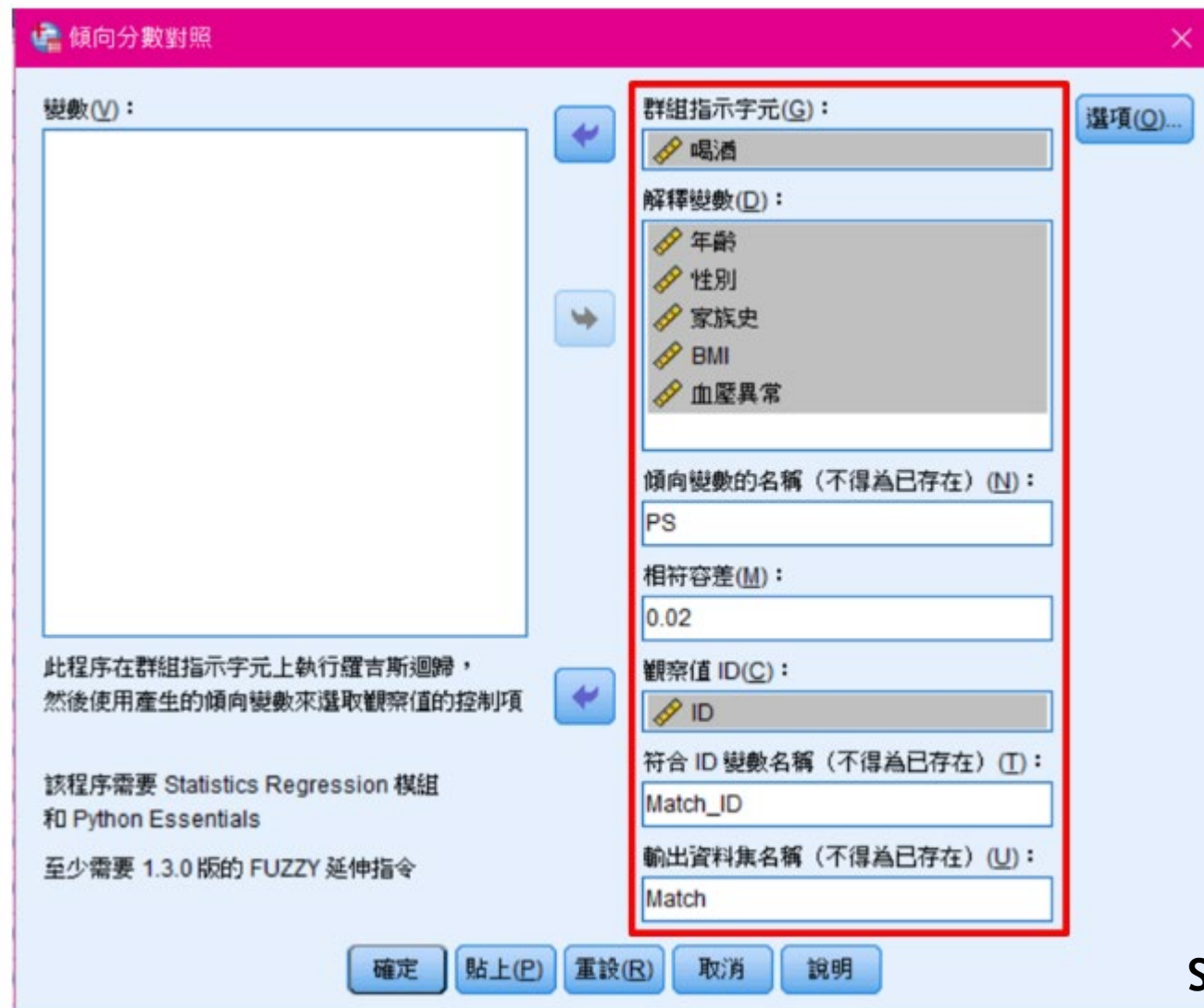
▶ 資料→傾向分數對照



傾向分數配對 (PS matching) 的步驟

▶ SPSS操作步驟

- ▶ 資料→傾向分數對照
- ▶ 變數根據需求進行配置



傾向分數配對 (PS matching) 的步驟

▶ SPSS操作步驟

- ▶ 資料→傾向分數對照
- ▶ 將變數根據需求進行配置
- ▶ 點選選項



傾向分數配對 (PS matching) 的步驟

▶ SPSS操作步驟

- ▶ 資料→傾向分數對照
- ▶ 將變數根據需求進行配置
- ▶ 點選選項
- ▶ 報表結果

		B	S.E.	Wald	自由度	顯著性	Exp(B)
步驟 1 ^a	年齡	-.241	.142	2.874	1	.090	.786
	性別	-1.675	.355	22.212	1	.000	.187
	家族史	1.809	.370	23.931	1	.000	6.106
	BMI	-.342	.232	2.167	1	.141	.710
	血壓異常	.585	.410	2.036	1	.154	1.795
	常數	2.089	.583	12.824	1	.000	8.080

a. 步驟 1 上輸入的變數：[%1:, 1:

* 上表以「喝酒」(1為喝酒；0為不喝酒)作為應變數，其他需要調整的變數作為自變數建構迴歸模型，藉此能得出每一個觀察對象的PS值。

傾向分數配對 (PS matching) 的步驟

▶ SPSS操作步驟

- ▶ 資料→傾向分數對照
- ▶ 將變數根據需求進行配置
- ▶ 點選選項
- ▶ 報表結果

Match Type	Count
Exact Matches	34
Fuzzy Matches	23
Unmatched Including Missing Keys	102
Unmatched with Valid Keys	102
Sampling	without replacement
Log file	none
Maximize Matching Performance	yes

- 上表顯示「精準匹配」有34對
- 「模糊匹配」有23對，共匹配成功57對

傾向分數配對 (PS matching) 的步驟

▶ SPSS操作步驟

- ▶ 資料→傾向分數對照
- ▶ 將變數根據需求進行配置
- ▶ 點選選項
- ▶ 報表結果

Case Control Match Tolerances

Match Variables	Value	Fuzzy Match Tries	Incremental Rejection Percentage
Exact (All Variables)	.	1968.000	98.272
PS	.020	1934.000	98.811

Tries is the number of match comparisons before drawing.
Rejection percentage shows the match rejection rate. Rejections are attributed to the first variable in the BY list that causes rejection.

* 上表顯示了匹配過程，在「精準匹配」下匹配了1968次，約有1.728%匹配成功；接著再進行「模糊匹配」（即以當初設定的「相符容差」值0.02進行匹配），共匹配1934次，約有1.189%匹配成功。

傾向分數配對 (PS matching) 的步驟

▶ SPSS操作步驟

- ▶ 資料→傾向分數對照
- ▶ 將變數根據需求進行配置
- ▶ 點選選項
- ▶ 報表結果
- ▶ 匹配後的檔案

	ID	年齡	性別	家族史	喝酒	BMI	血壓異常	PS	Eligible_cases	match_id
1	1.00	1.00	1.00	.00	1.00	1.00	.00	.45801	6.00	164.00
2	2.00	2.00	1.00	1.00	1.00	2.00	.00	.74225	.00	.
3	3.00	3.00	1.00	.00	1.00	1.00	.00	.34274	2.00	192.00
4	4.00	4.00	1.00	1.00	1.00	2.00	.00	.63991	2.00	231.00
5	5.00	1.00	1.00	1.00	1.00	1.00	.00	.83766	.00	.
6	6.00	2.00	1.00	1.00	1.00	2.00	.00	.74225	1.00	179.00
7	7.00	3.00	1.00	.00	1.00	1.00	.00	.34274	3.00	189.00
8	8.00	4.00	1.00	1.00	1.00	2.00	.00	.63991	1.00	230.00
9	9.00	1.00	1.00	1.00	1.00	1.00	.00	.83766	.00	.
10	10.00	1.00	1.00	.00	1.00	2.00	.00	.37515	1.00	213.00

*輸出的資料庫會多了幾個之前設定的新變數，

「PS」為使用上述迴歸模型計算出的傾向性評分；

「Eligible_cases」表示對照組有幾個符合條件的觀察值對象；

「Match_id」表示匹配成功的ID。

傾向分數配對 (PS matching) 的步驟

▶ SPSS操作步驟

- ▶ 資料→傾向分數對照
- ▶ 將變數根據需求進行配置
- ▶ 點選選項
- ▶ 報表結果
- ▶ 匹配後的檔案
- ▶ 檔案整理：資料→選取觀察值

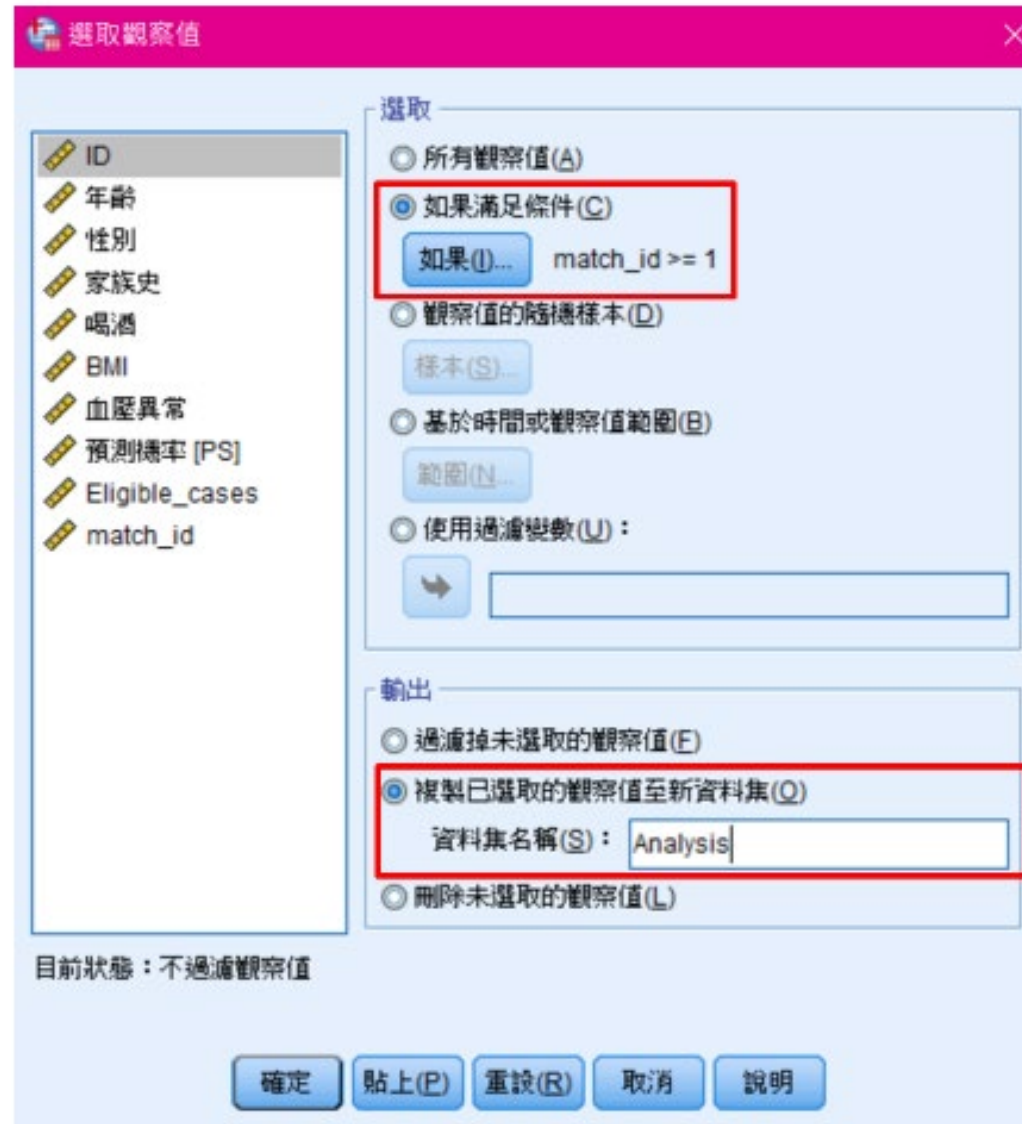
The screenshot displays the SPSS Statistics interface with the 'Select Cases' menu option highlighted in red. The menu is open, showing various options for data manipulation. The background shows a data editor window with columns for 'ID', '喝酒', 'BMI', '血壓異常', and 'PS'. The 'Select Cases' option is located at the bottom of the menu.

ID	喝酒	BMI	血壓異常	PS
1	1.00	1.00	.00	
2	1.00	2.00	.00	
3	1.00	1.00	.00	
4	1.00	2.00	.00	
5	1.00	1.00	.00	
6	1.00	2.00	.00	
7	1.00	1.00	.00	
8	1.00	2.00	.00	
9	1.00	1.00	.00	
10	1.00	2.00	.00	
11	1.00	3.00	.00	
12	1.00	1.00	.00	
13	1.00	2.00	.00	
14	1.00	1.00	.00	
15	1.00	1.00	.00	
16	1.00	1.00	1.00	
17	1.00	1.00	1.00	
18	1.00	1.00	1.00	

傾向分數配對 (PS matching) 的步驟

▶ SPSS操作步驟

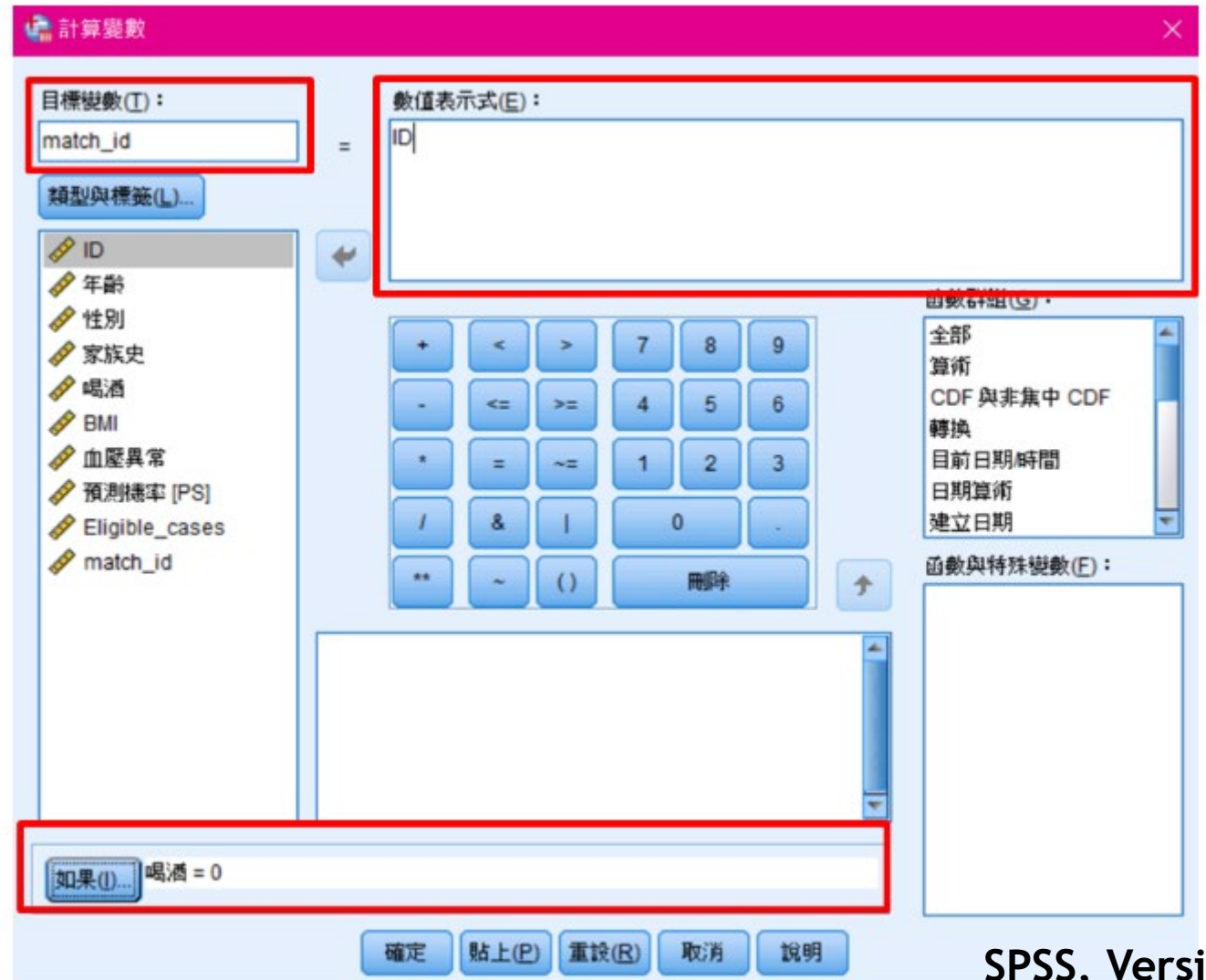
- ▶ 資料→傾向分數對照
- ▶ 將變數根據需求進行配置
- ▶ 點選選項
- ▶ 報表結果
- ▶ 匹配後的檔案
- ▶ 檔案整理：資料→選取觀察值



傾向分數配對 (PS matching) 的步驟

▶ SPSS操作步驟

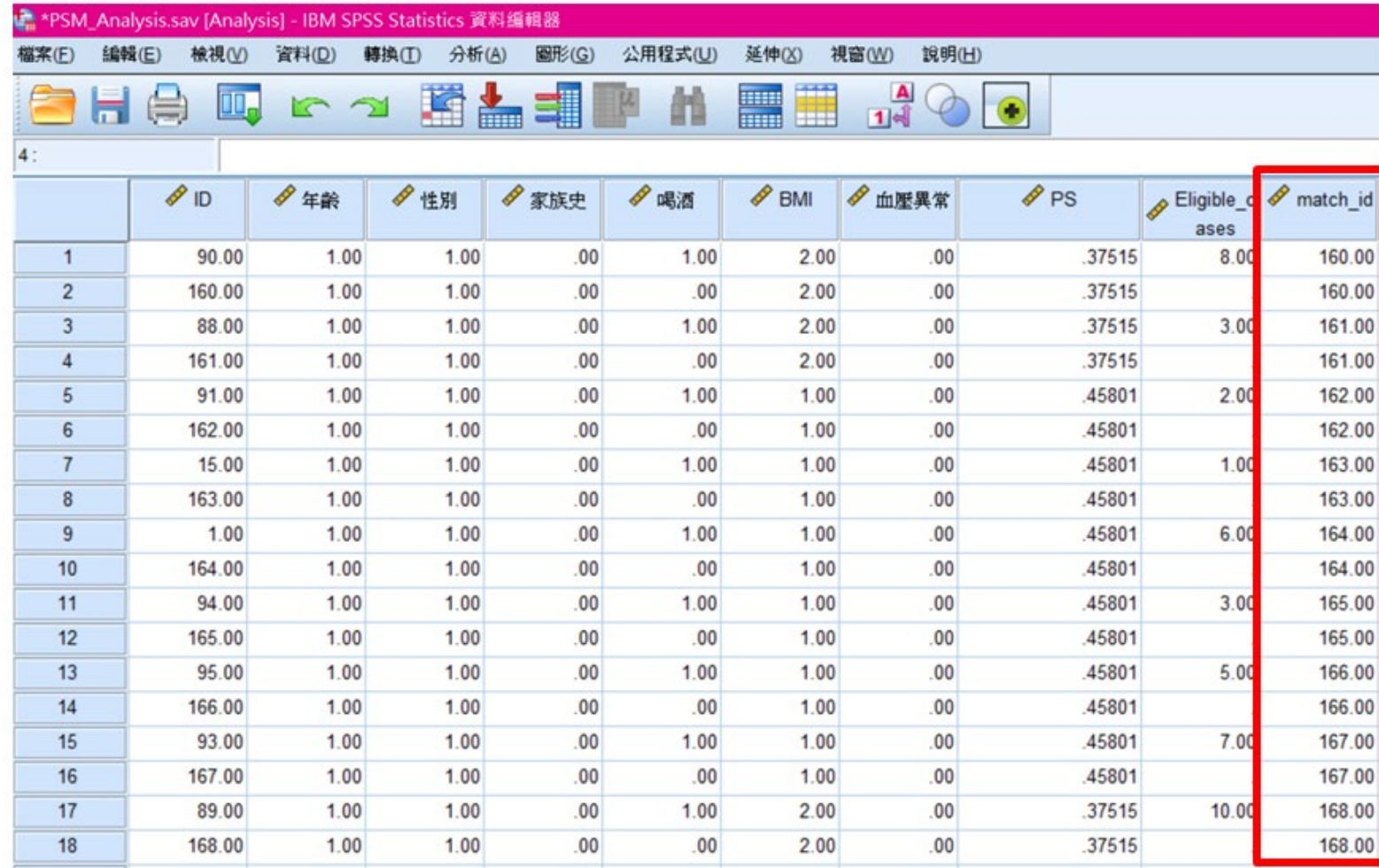
- ▶ 資料→傾向分數對照
- ▶ 將變數根據需求進行配置
- ▶ 點選選項
- ▶ 報表結果
- ▶ 匹配後的檔案
- ▶ 檔案整理：資料→選取觀察值
- ▶ 設置匹配成功的ID標示：轉換→計算變數



傾向分數配對 (PS matching) 的步驟

SPSS操作步驟

- ▶ 資料→傾向分數對照
 - ▶ 將變數根據需求進行配置
 - ▶ 點選選項
 - ▶ 報表結果
 - ▶ 匹配後的檔案
 - ▶ 檔案整理：資料→選取觀察值
 - ▶ 設置匹配成功的ID標示：轉換→計算變數
 - ▶ 完成傾向性評分匹配
- 在「match_id」欄按滑鼠右鍵，選擇「遞增排序」
 - * 如此一來「match_id」便能看出兩兩配對成功的ID樣本，完成傾向性評分匹配。



*PSM_Analysis.sav [Analysis] - IBM SPSS Statistics 資料編輯器

	ID	年齡	性別	家族史	喝酒	BMI	血壓異常	PS	Eligible cases	match_id
1	90.00	1.00	1.00	.00	1.00	2.00	.00	.37515	8.00	160.00
2	160.00	1.00	1.00	.00	.00	2.00	.00	.37515		160.00
3	88.00	1.00	1.00	.00	1.00	2.00	.00	.37515	3.00	161.00
4	161.00	1.00	1.00	.00	.00	2.00	.00	.37515		161.00
5	91.00	1.00	1.00	.00	1.00	1.00	.00	.45801	2.00	162.00
6	162.00	1.00	1.00	.00	.00	1.00	.00	.45801		162.00
7	15.00	1.00	1.00	.00	1.00	1.00	.00	.45801	1.00	163.00
8	163.00	1.00	1.00	.00	.00	1.00	.00	.45801		163.00
9	1.00	1.00	1.00	.00	1.00	1.00	.00	.45801	6.00	164.00
10	164.00	1.00	1.00	.00	.00	1.00	.00	.45801		164.00
11	94.00	1.00	1.00	.00	1.00	1.00	.00	.45801	3.00	165.00
12	165.00	1.00	1.00	.00	.00	1.00	.00	.45801		165.00
13	95.00	1.00	1.00	.00	1.00	1.00	.00	.45801	5.00	166.00
14	166.00	1.00	1.00	.00	.00	1.00	.00	.45801		166.00
15	93.00	1.00	1.00	.00	1.00	1.00	.00	.45801	7.00	167.00
16	167.00	1.00	1.00	.00	.00	1.00	.00	.45801		167.00
17	89.00	1.00	1.00	.00	1.00	2.00	.00	.37515	10.00	168.00
18	168.00	1.00	1.00	.00	.00	2.00	.00	.37515		168.00

使用SAS檢視PS的分佈

變數: ps (估計機率)

gp_new	N	平均值	標準差	標準誤差	最小值	最大值
0	1101	0.6658	0.1413	0.00426	0.2373	0.9552
1	1101	0.6660	0.1410	0.00425	0.2456	0.9551
Diff (1-2)		-0.00015	0.1411	0.00602		

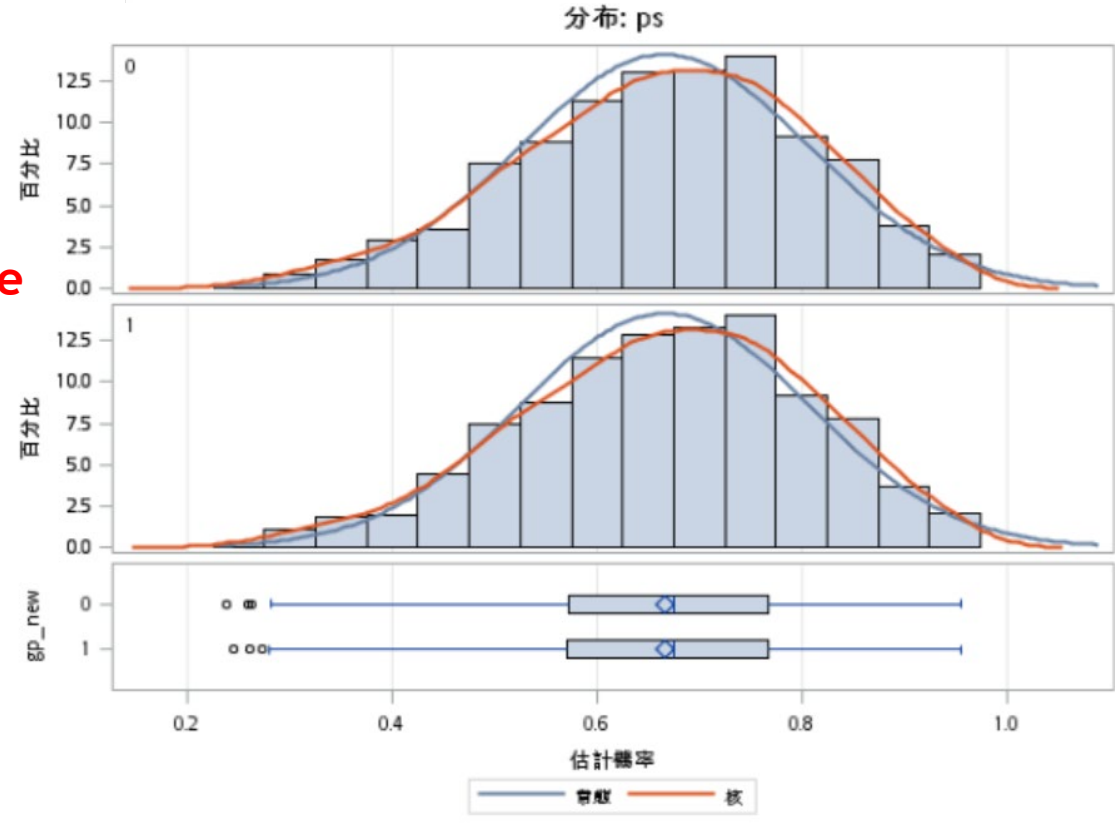
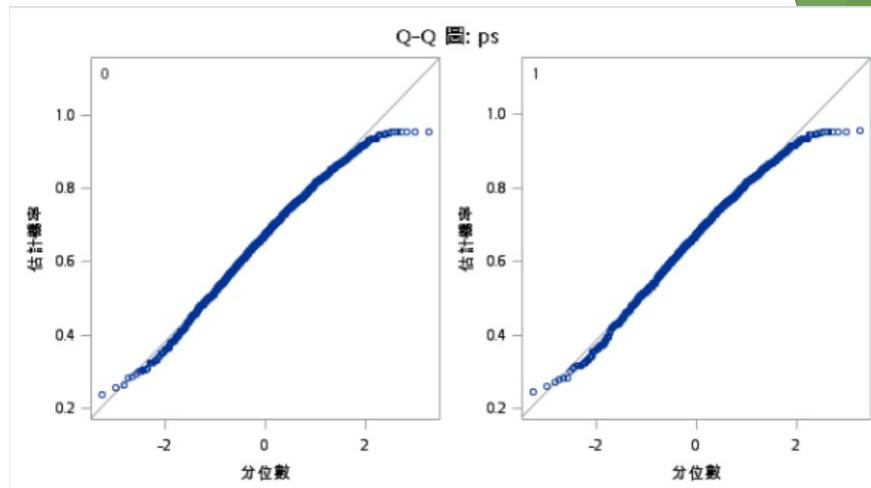
gp_new	方法	平均值	95% CL 平均值	標準差	95% CL 標準差
0		0.6658	0.6575 0.6742	0.1413	0.1356 0.1474
1		0.6660	0.6576 0.6743	0.1410	0.1353 0.1471
Diff (1-2)	集區	-0.00015	-0.0119 0.0116	0.1411	0.1371 0.1454
Diff (1-2)	Satterthwaite	-0.00015	-0.0119 0.0116		

Mean Difference

方法	變異數	自由度	t 值	Pr > t
集區	均等	2200	-0.02	0.9805
Satterthwaite	不均等	2200	-0.02	0.9805

變異數相等性

方法	分子自由度	分母自由度	F 值	Pr > F
Folded F	1100	1100	1.00	0.9419



傾向分數的四種應用方式

- ▶ 配對 (matching)
- ▶ 加權 (**weighting**)
- ▶ 分層 (stratification)
- ▶ 迴歸調整 (covariate adjustment in regression model)

傾向分數加權 (IPTW) 的方法

PS weighting

- Inverse probability of the treatment (IPTW)
 - Average treatment effect
 - $1/PS$ in treated individuals
 - $1/(1-PS)$ in untreated individuals

Jackson JW, et al. Curr Epidemiol Rep, 2017
Austin PC. Multivar Behav Res, 2011

傾向分數加權 (IPTW) 的方法

treatment=1 (治療組) → 重新加權為 1-PS
treatment=0 (對照組) → 維持原本的 PS



圖 C 為各基線特徵的絕對標準化均值差 (Absolute Standardized Mean Difference, ASMD)：值愈小表示有無治療的二組之間的差異愈小，一般來說 ASMD 小於 0.1 表示有良好的平衡 (圖 C 虛線)

傾向分數加權 (IPTW) 的方法

▶ 優點

- ▶ 保留所有個案，減少因配對被排除的個案損失，結果外推性高
- ▶ 可以延伸到處理設限資料 (censoring) 與時間相依混淆因子 (time-dependent confounding) 的處理

▶ 缺點

- ▶ 加權後圖表的呈現不直觀，不好理解
- ▶ 極端權重的影響：IPTW 的加權值很容易受到極大或極小的 PS 所影響，此時會導致評估效果出現偏差且變異較大

去除極端權重，或使用 Stabilized weight，即可中和極端權重對結果的影響

計算傾向值 (使用Logistic Regression)

SAS 操作步驟

```
/*IPTW*/  
proc logistic data=data.arni_f;  
  class sex dm htnmi af bb_g mra_g /* (ref=first)*/;  
  model gp_new= age sexef_final dm htnmi af bb_g mra_g /  
lackfit; /*requests Hosmer and Lemeshow goodness-of-fit test*/  
  output out=data.out_ps_r2 prob=ps xbeta=logit_ps; /*create  
new data set: out_ps*/  
run; /*new variable: ps:propensity score  
logit_ps: logit of propensity score*/  
  
data data.out_ps_r2; set data.out_ps_r2;  
weight2 = .;  
if gp_new =1 then weight2 = 1/ps2;  
else if gp_new=0 then weight2 = 1/(1-ps2); run;
```

計算傾向值 (使用Logistic Regression)

R 操作步驟

```
#Propensity score model
```

```
psmodel <- glm (gp_new ~ age + sex + dm + htn ,  
  family = binomial (link = "logit"), data =data_ipwt )  
summary (psmodel)
```

```
#Value of propensity score for each subject  
ps_w <- predict (psmodel, type = "response")  
ps_case <- cbind (ps_w, data_ipwt )  
#合併時千萬不可以有missing data
```

```
#creat weights  
weight_r <- ifelse ( ps_case$gp_new == 1, 1/(ps_w), 1/(1-ps_w))  
iptw_case <- cbind (ps_case, weight_r )
```

```
setwd("D:/助理研究員/中榮醫研部-生統小組/全院教育課程規劃-2022oct/111年第4季/20221207 傾向分數的使用  
/PS matching")  
write.table(iptw_case, "iptw_case.csv", quote=F, row.names=F)
```

傾向分數的四種應用方式

- ▶ 配對 (matching)
- ▶ 加權 (weighting)
- ▶ 分層 (stratification)
- ▶ 迴歸調整 (covariate adjustment in regression model)

PS stratification

Table 1. Effect of stratification

Propensity Score			Propensity Score		
Quintile	Subjects	Person-Years	Quintile Limits	Mean	Min, Max
<i>All Quintiles</i>					
Temazepam	93,011	280,712	---	0.31	0.02, 0.97
Zopiclone	54,592	126,002	---	0.47	0.03, 0.97
<i>Quintile 1</i>					
Temazepam	26,957	107,774	0, 0.17	0.12	0.02, 0.17
Zopiclone	2,563	10,409	0, 0.17	0.13	0.03, 0.17
<i>Quintile 2</i>					
Temazepam	22,402	79,934	0.17, 0.30	0.23	0.17, 0.30
Zopiclone	7,119	25,645	0.17, 0.30	0.24	0.17, 0.30
<i>Quintile 3</i>					
Temazepam	17,833	47,472	0.30, 0.42	0.36	0.30, 0.42
Zopiclone	11,687	31,988	0.30, 0.42	0.37	0.30, 0.42
<i>Quintile 4</i>					
Temazepam	14,784	29,458	0.42, 0.55	0.48	0.42, 0.55
Zopiclone	14,737	30,325	0.42, 0.55	0.49	0.42, 0.55
<i>Quintile 5</i>					
Temazepam	11,035	16,073	0.55, 1.00	0.64	0.55, 0.97
Zopiclone	18,486	27,635	0.55, 1.00	0.66	0.55, 0.97

Arbogast PG, Seeger JD. Summary variables in observational research: propensity scores and disease risk scores.

PS stratification

- Advantages

- Transparency in that balance on covariates achieved through use of the propensity score can be shown explicitly when using stratification.
- Many readers of the research result will either be familiar with the technique of stratification or find it easy to understand so they can follow what was done and be able to better interpret the results of the analysis.

- Disadvantages

- In order to be transparent many tables may be required, making for a potentially unwieldy presentation.
- Residual confounding within strata may cause bias.

Seeger J., Rassen J. Propensity scores in pharmacoepidemiology. ICPE MONTREAL, 2013.

Austin PC. Multivar Behav Res, 2011

What PS cannot do

- Only unbiased as the predictors included in their calculation.
- When improperly modeled, PS cannot provide unbiased estimates of treatment effects.
- Do not inform the research about the effect of any individual variable that was used to create the score.
- Change from sample to sample and will vary with any change in the variables used to calculate them.

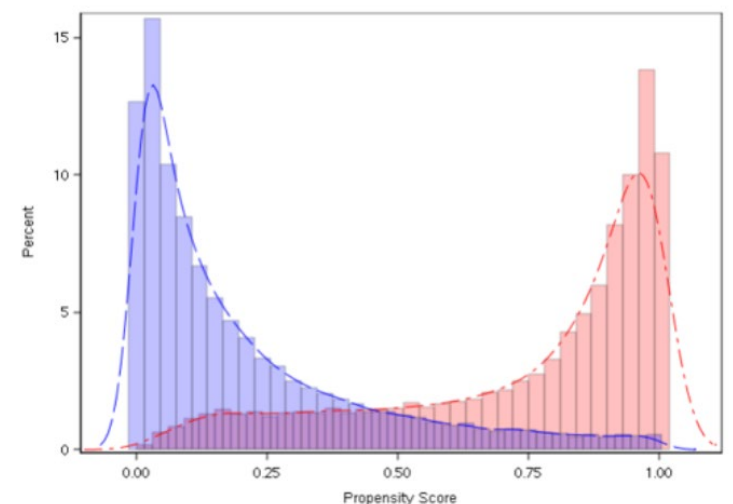
雖然使用配對或 IPTW 的方式仍然可以有效控制干擾，
但是**通常會偏離目標族群**，直接影響了兩組的平衡
及大大的**降低準確度** → **建模不正確**

千萬不要把配入 PS 的變數，再拿去做分層分析

Beal SJ, Kupzyk KA. J Early Adolescence, 2014

真實世界

PS 密度分佈圖：重疊性低的分布



RCT vs. PS matching

- Differences

- Balance is by construction, not by design.
- Balance is only among the measured covariates.
- No balance among unmeasured covariates is implied.

- Similarities

- Can be treated analytically like an RCT.
- Equivalence at baseline like an RCT.

Seeger J., Rassen J. Propensity scores in pharmacoepidemiology. ICPE MONTREAL, 2013.
Previous lecture powerpoint by Dr. Yeh.

Thank you for listening

The background features abstract, overlapping geometric shapes in various shades of green, ranging from light lime to dark forest green. These shapes are primarily located on the right side of the frame, creating a modern, layered effect against the white background.