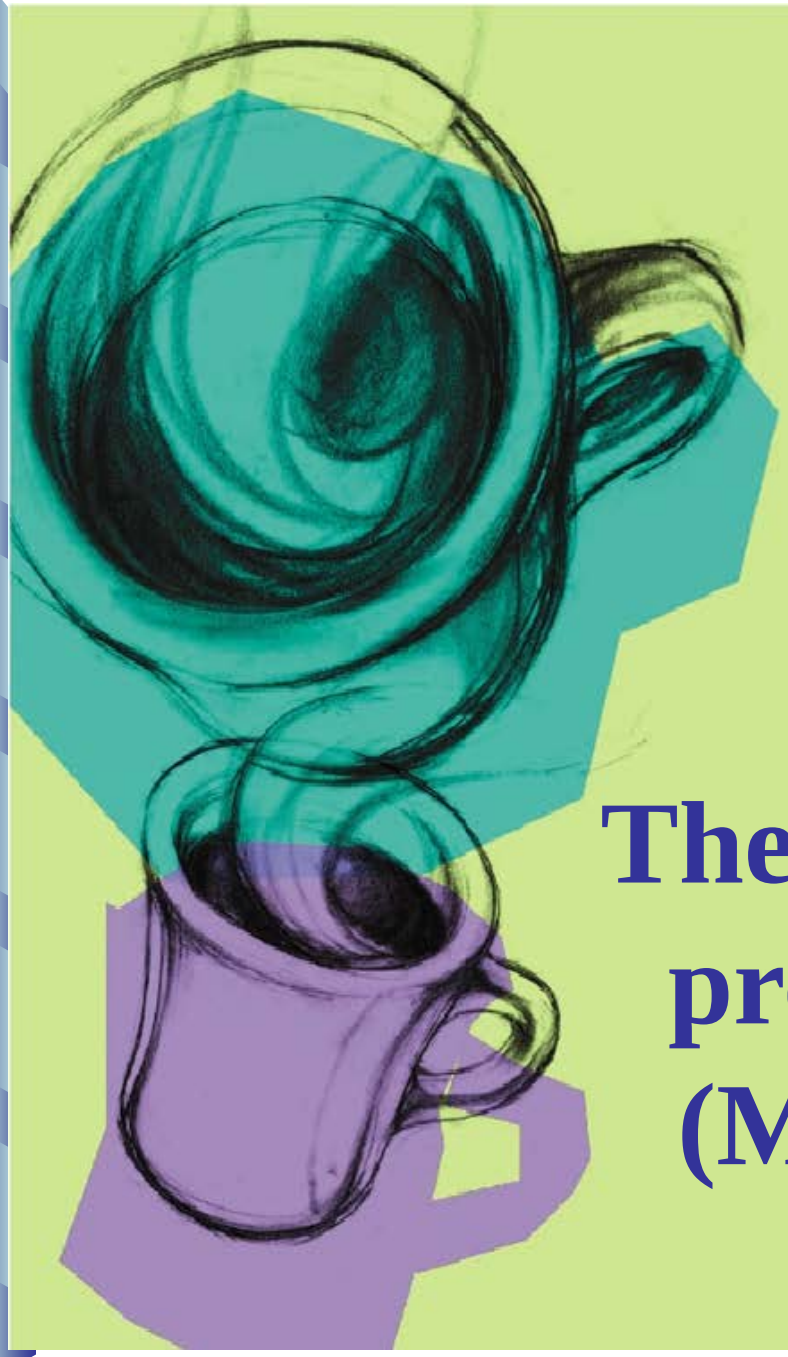


CHAPTER

The GLM Univariate procedure in SPSS (Multiple Factors)

Randomized Block Designs

Introduce a design that has its basic focus on a single factor or multiple factors, but uses an additional factor (called a blocking factor) to account for the effects of dissimilar groups of experimental units on the value of the response variable.

Randomized Block Designs

Suppose we are interested in a single factor with k treatments (levels). Sometimes there is so much variation in the values of the response variable within each treatment that the use of a completely randomized design will fail to detect differences among the treatment means when such differences exist.

This is because it is often not possible to decide whether the variation among the sample means for the different treatments is due to differences among the treatment means or whether it is due to variation within the treatments (i.e., variation in the values of the response variable within each treatment).

Randomized Block Designs

If a large portion of the variation within the treatments is due to one extraneous variable, then it is often appropriate to use a randomized block design instead of a completely randomized design. In a randomized block design, the extraneous source of variation is isolated and removed so that it is easier to detect differences among the treatment means when such differences exist.

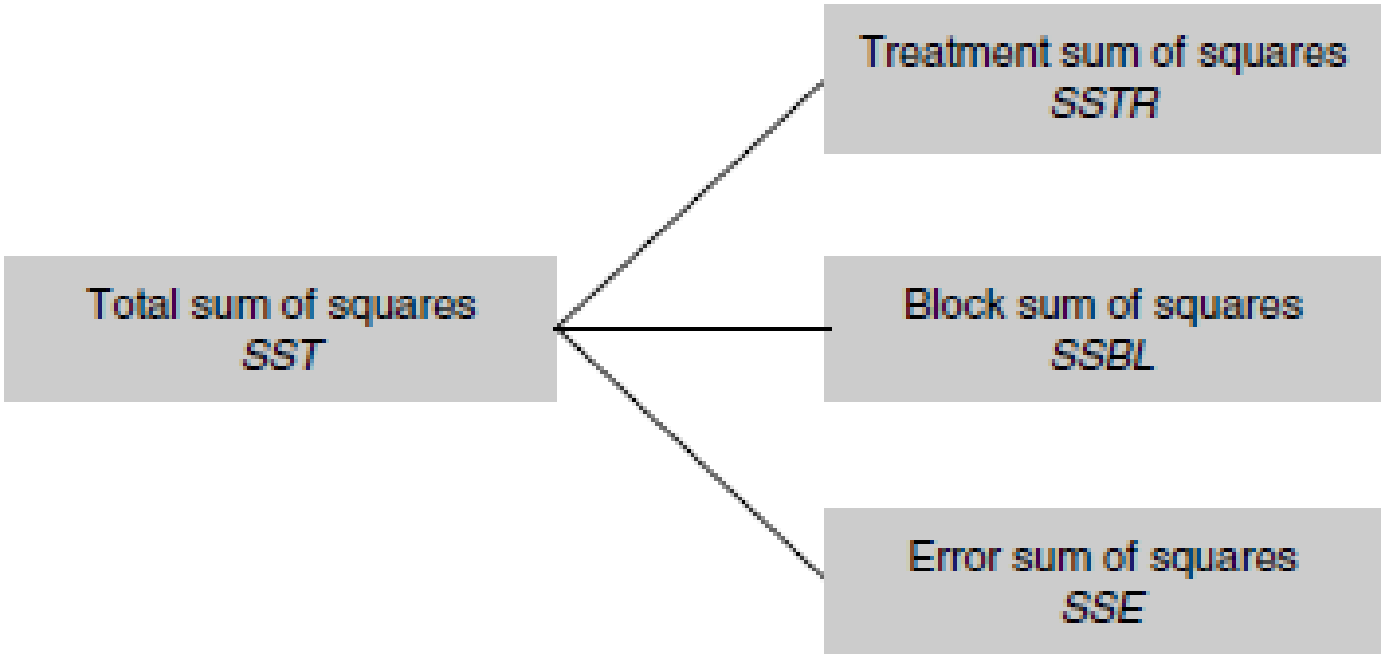
Although a randomized block design is not always appropriate or feasible, it is often a viable alternative to a completely randomized design in the presence of a single extraneous source of variability. Consider the following example.

Randomized Block ANOVA: The Logic

In the randomized block design, we can represent the response variable by a model that relates a value of that variable to the effects associated with the treatment factor, T, and the blocking factor, B. We write our response as

Response = Overall Mean + Treatment Factor Effect + Blocking Factor Effect + Error.

Partitioning of the total sum of squares for a randomized block design

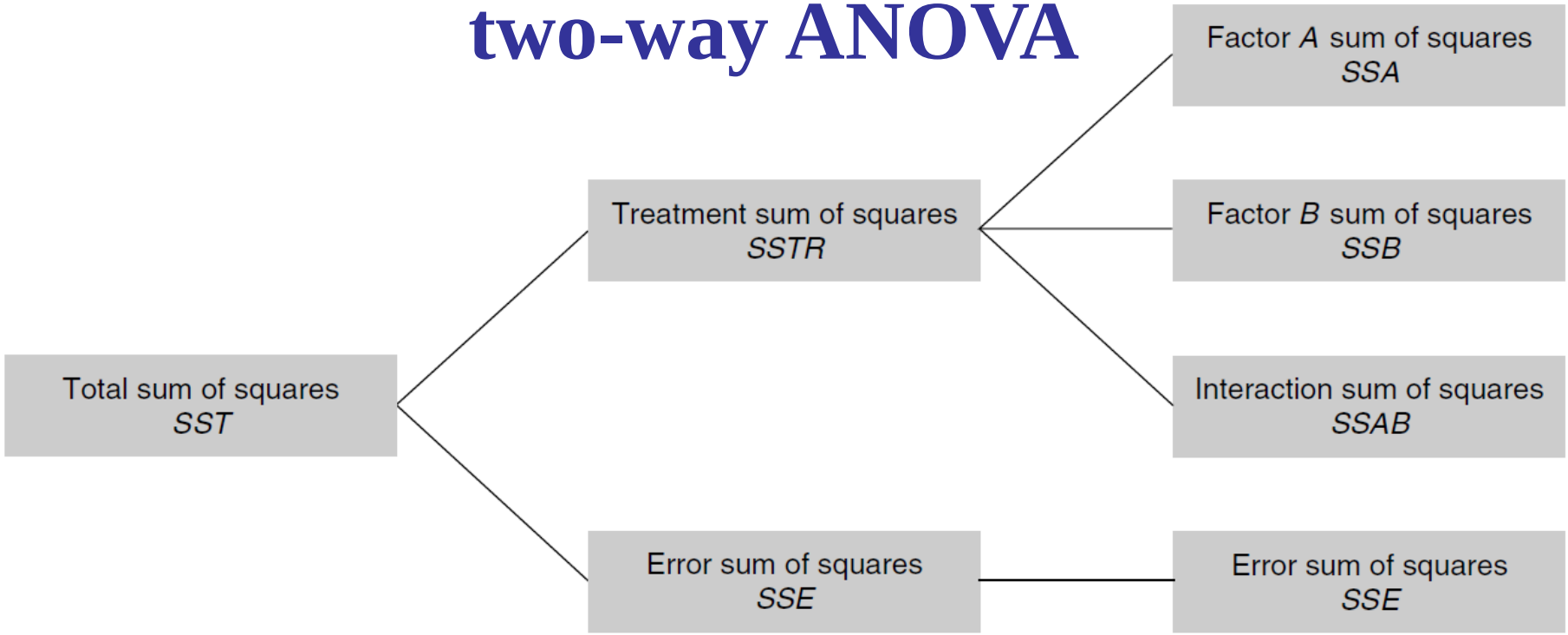


Randomized Block ANOVA Identity for Sums of Squares

The total sum of squares equals the treatment sum of squares plus the block sum of squares plus the error sum of squares; that is

$$SST = SSTR + SSBL + SSE.$$

Partitioning of the total sum of squares in two-way ANOVA



For the two-way model, we have the following effects: the main effect due to Factor *A*, the main effect due to Factor *B*, and the interaction effect between Factors *A* and *B*. We write our response as

$$\text{Response} = \text{Overall Mean} + A \text{ Main Effect} + B \text{ Main Effect} + AB \text{ Interaction Effect} + \text{Error}.$$

ANOVA table format for a randomized block design

Source	df	SS	$MS = SS/df$	F-statistic
Treatment	$k-1$	$SSTR$	$MSTR = \frac{SSTR}{k-1}$	$F_{TR} = \frac{MSTR}{MSE}$
Block	$b-1$	$SSBL$	$MSBL = \frac{SSBL}{b-1}$	$F_{BL} = \frac{MSBL}{MSE}$
Error	$(k-1)(b-1)$	SSE	$MSE = \frac{SSE}{(k-1)(b-1)}$	
Total	$kb-1$	SST		

Equations for randomized block ANOVA

Sums of Squares In a Randomized Block ANOVA

For a randomized block ANOVA, we have the following defining formulas for the three sums of squares.

Sum of Squares	Defining Formula
Total, *SST*	$\sum(x-\overline{x})^2$
Treatment, *SSTR*	$b\sum(\overline{T}_i-\overline{x})^2$
Block, *SSBL*	$k\sum(\overline{B}_j-\overline{x})^2$
Error, *SSE*	$\sum(x-\overline{T}_i-\overline{B}_j+\overline{x})^2$

Here k = number of treatments
b = number of blocks
$\overline{T}_i$ = mean of the b observations for treatment i, for $i = 1, \ldots, k$
$\overline{B}_j$ = mean of the k observations for block j, for $j = 1, \ldots, b$
$\overline{x}$ = mean of all the kb observations
x = individual observation among the kb observations

The Elements of randomized block ANOVA

Assumptions for Randomized Block ANOVA

1. *Independence/dependence of observations:*
   - Fixed effect blocking factor: The observations of the response variable are independent of one another across all treatment-block combinations.
   - Random effect blocking factor: The observations of the response variable in different blocks are independent of one another. The observations within a block may be correlated. The correlation (if any) between two observations in a block is the same for all pairs of observations in a block. This correlation (if any) is the same from one block to another.
2. ***Normal distributions:*** For each treatment-block combination, the response variable is normally distributed.
3. ***Equal standard deviations:*** The standard deviations of the response variable are the same for all treatment-block combinations.

Key Fact C.8

Elements of a Randomized Block ANOVA

Elements of a Randomized Block ANOVA

Assumptions

The three assumptions for randomized block ANOVA (Key Fact C.8)

Notation:

k = number of levels of the treatment factor T

b = number of blocks (levels of the blocking factor B)

n = total number of observations = kb

Elements of a Randomized Block ANOVA

Test for Treatment Factor T effect. The null and alternative hypotheses are:

H_0: There is no effect due to the treatment factor T.

H_a: There is an effect due to the treatment factor T.

The test statistic is

$$F_{TR} = \frac{MSTR}{MSE},$$

which has an F-distribution with $\text{df} = (k-1, (k-1)(b-1))$ if the null hypothesis is true.

Elements of a Randomized Block ANOVA

Test for Blocking Factor B effect (Optional). The null and alternative hypotheses are:

H_0: There is no effect due to the blocking factor B.

H_a: There is an effect due to the blocking factor B.

The test statistic is

$$F_{BL} = \frac{MSBL}{MSE},$$

which has an F-distribution with df $= (b-1, (k-1)(b-1))$ if the null hypothesis is true. This test is optional, since the treatment factor is the factor of intrinsic interest.

Randomized Block Designs Example

Analgesic (止痛藥) *Effectiveness* : Suppose that we want to compare the times it takes three analgesics-Brand A, Brand B, and Brand C - to relieve a headache.

Let μ_A, μ_B, and μ_C denote, respectively, the mean times it takes Brand A, Brand B, and Brand C to relieve a headache. Then the hypotheses to be tested are:

H_0: All three mean relief times are equal

H_1: Not all three means are equal.

With 30 people as subjects, explain how to carry out this hypothesis test using a (a). completely randomized design. (b). randomized block design.

Which experimental design is probably better?

Analgesic (止痛藥) *Effectiveness Example*

Times, in minutes, until headache relief

Age group	Analgesic brand		
	Brand *A*	Brand *B*	Brand *C*
16–20	26	15	25
21–25	17	16	19
26–30	24	20	19
31–35	18	20	22
36–40	22	19	27
41–45	27	19	27
46–50	28	27	28
51–55	23	27	25
56–60	33	22	32
61–65	34	23	33

Analysis of covariance (ANCOVA)

Analysis of covariance (ANCOVA) is a general linear model which blends ANOVA and regression. ANCOVA evaluates whether the means of a dependent variable (DV) are equal across levels of a categorical independent variable (IV) often called a treatment, while statistically controlling for the effects of other continuous variables that are not of primary interest, known as covariates (CV) or nuisance variables.

Example on ANCOVA

Mathematically, ANCOVA decomposes the variance in the DV into variance explained by the CV(s), variance explained by the categorical IV, and residual variance. Intuitively, ANCOVA can be thought of as 'adjusting' the DV by the group means of the CV(s).

Purpose of including covariates in ANOVA

1. To reduce within-group error variance: If we can explain some of this 'unexplained' variance (SSR) in terms of covariates, then we reduce the error variance, allowing us to more accurately assess the effect of the experimental manipulation

2. Elimination of Confounds: In any experiment, there may be unmeasured variables that confound the results (i.e. a variable that varies systematically with the experimental manipulation). Once a possible confounding variable has been identified, it can be measured and entered into the analysis as a covariate.

Using GLM Univariate to Perform ANCOVA

Example: Proponents of a government works program want to see if it helps people into better jobs, controlling for their salary before entering the program. A sample of potential program participants were followed, some of whom were randomly selected for enrollment in the program, while others were not.

This information is collected in workprog.sav. Use the GLM Univariate procedure to perform an analysis of covariance (ANCOVA) on the incomes after the program. An extra assumption of ANCOVA is that there is no significant interaction between the covariate and factor, so begin by fitting a model with an interaction term.

Example on ANCOVA

workprog.sav - SPSS 資料編輯程式

檔案(F) 編輯(E) 檢視(V) 資料(D) 轉換(T) 分析(A) 統計圖(G) 公用程式(U) 視窗(W) 輔助說明(H)

1 : age 16

	age	marital	incbef	incaft	ed	gender	reside	prog	var
1	16	Unmarried	8.00	12.00	Did not comple	Male	1	0	
2	17	Unmarried	8.00	10.00	High school de	Female	1	0	
3	17	Unmarried	8.00	11.00	Did not comple	Female	1	0	
4	19	Married	9.00	18.00	Did not comple	Female	3	1	
5	18	Married	7.00	12.00	Did not comple	Female	3	0	
6	17	Married	8.00	15.00	Did not comple	Female	2	1	
7	17	Married	8.00	13.00	High school de	Female	3	0	
8	21	Unmarried	9.00	22.00	Did not comple	Female	2	1	
9	18	Unmarried	7.00	18.00	Did not comple	Male	1	1	
10	16	Unmarried	7.00	9.00	Did not comple	Male	1	0	
11	17	Married	6.00	8.00	Did not comple	Female	5	0	
12	18	Unmarried	10.00	20.00	High school de	Female	1	1	
13	17	Unmarried	6.00	14.00	Did not comple	Female	1	1	
14	19	Unmarried	8.00	16.00	Did not comple	Female	1	1	
15	20	Unmarried	12.00	25.00	Some college	Female	1	0	
16	18	Unmarried	9.00	20.00	High school de	Female	1	1	
17	19	Unmarried	10.00	18.00	High school de	Male	1	1	
18	19	Unmarried	9.00	21.00	Did not comple	Female	1	1	
19	20	Married	11.00	20.00	Some college	Female	3	1	
20	17	Married	8.00	17.00	Did not comple	Female	2	1	
21	20	Married	7.00	9.00	Did not comple	Female	2	0	
22	19	Married	8.00	10.00	Did not comple	Male	3	0	
23	19	Married	11.00	26.00	Some college	Male	2	1	

Example on ANCOVA

workprog.sav - SPSS 資料編輯程式

檔案(F) 編輯(E) 檢視(V) 資料(D) 轉換(T) 分析(A) 統計圖(G) 公用程式(U) 視窗(W) 輔助說明(H)

	名稱	類型	寬度	小數	標記	數值	遺漏	欄	對齊	測量
1	age	數字的	4	0	Age in years	無	無	8	右	尺度
2	marital	數字的	4	0	Marital status	{0, Unmarried}...	無	8	右	尺度
3	incbef	數字的	8	2	Income before the program	無	無	8	右	尺度
4	incaft	數字的	8	2	Income after the program	無	無	6	右	尺度
5	ed	數字的	4	0	Level of education	{1, Did not complet	無	8	右	尺度
6	gender	字串	2	0	Gender	{f, Female}...	無	8	左	名義的
7	reside	數字的	4	0	Number of people in househ	無	無	8	右	尺度
8	prog	數字的	4	0	Program status	無 ...	無	8	右	尺度
9										
10										
11										
12										
13										
14										
15										
16										
17										

Example on ANCOVA

To run a GLM Univariate analysis, from the menus choose:

Analyze

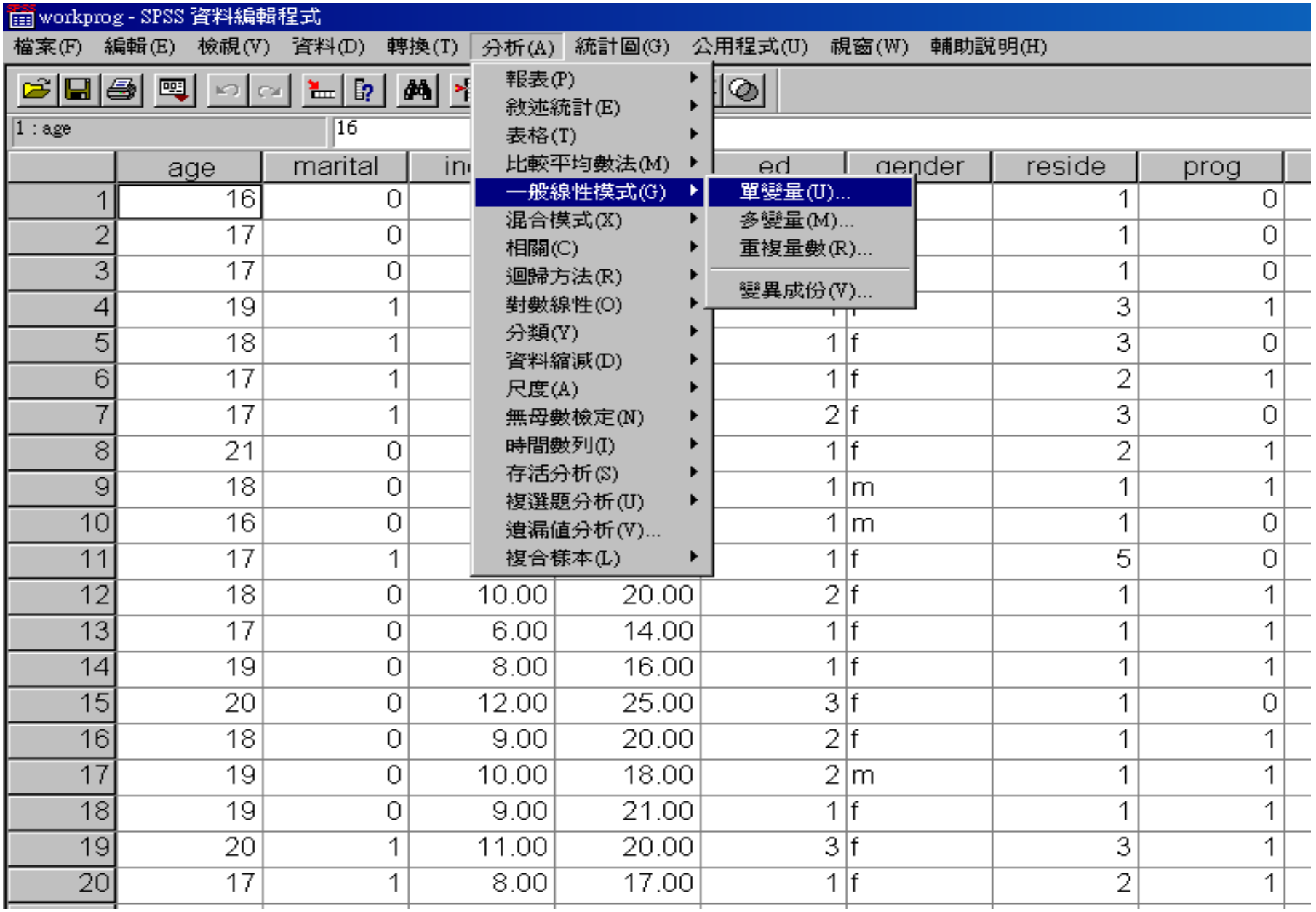
General Linear Model

Univariate...

Click Reset to restore the default settings.

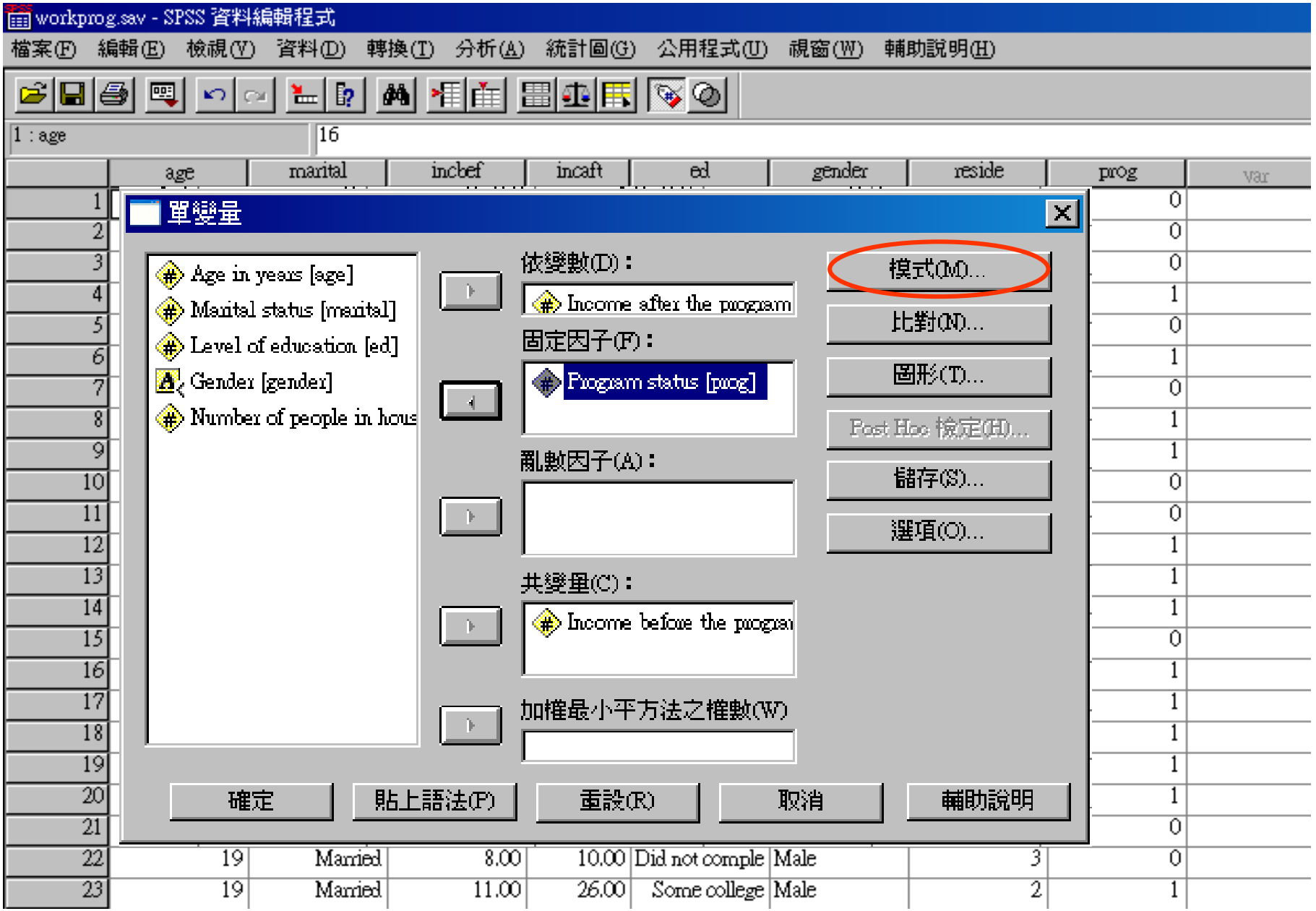
- Select Income after the program as the dependent variable.
- Select Program status as the fixed factor.
- Select Income before the program as the covariate.

Example on ANCOVA

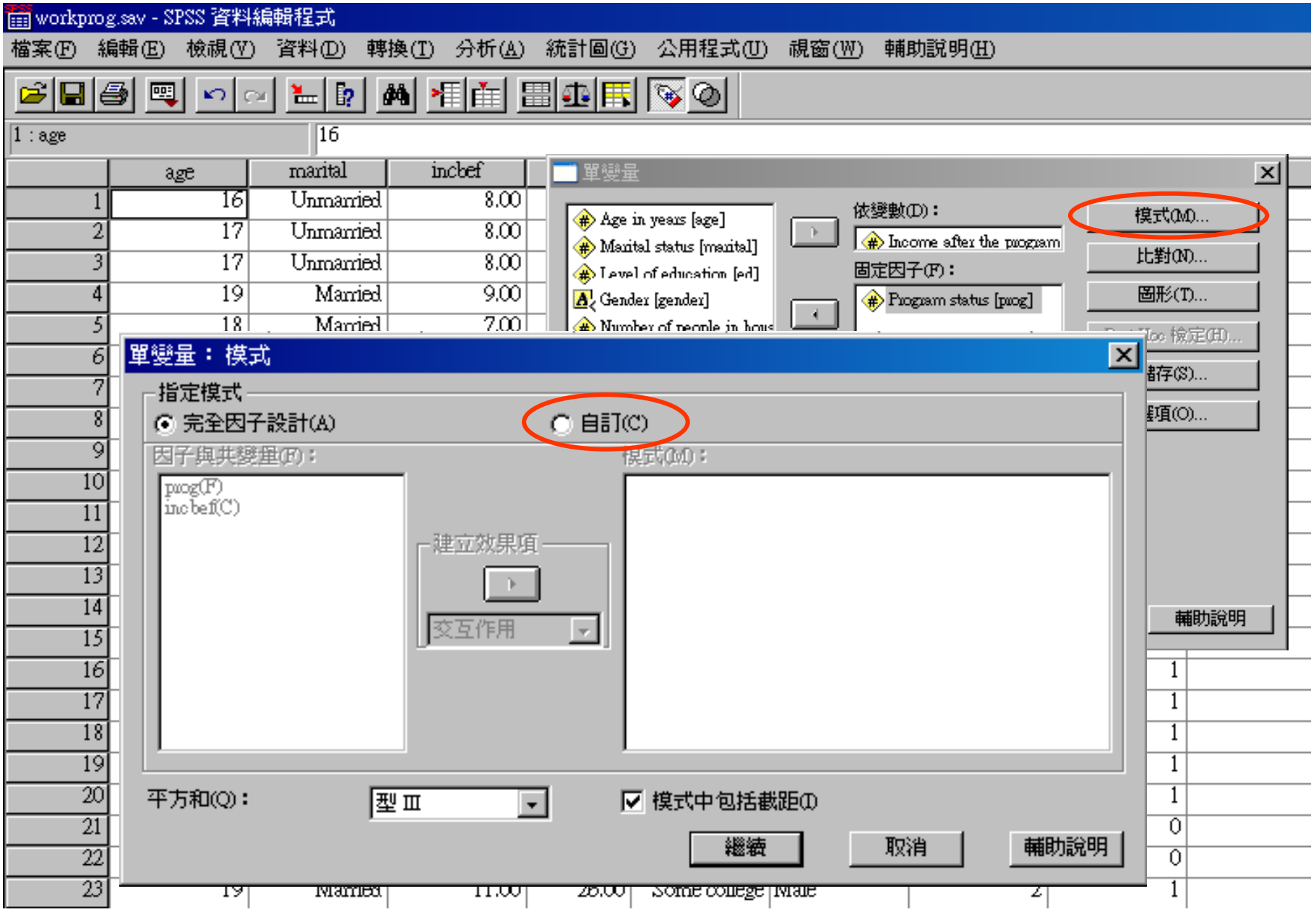


	age	marital			ed	gender	reside	prog
1	16	0					1	0
2	17	0					1	0
3	17	0					1	0
4	19	1					3	1
5	18	1			1	f	3	0
6	17	1			1	f	2	1
7	17	1			2	f	3	0
8	21	0			1	f	2	1
9	18	0			1	m	1	1
10	16	0			1	m	1	0
11	17	1			1	f	5	0
12	18	0	10.00	20.00	2	f	1	1
13	17	0	6.00	14.00	1	f	1	1
14	19	0	8.00	16.00	1	f	1	1
15	20	0	12.00	25.00	3	f	1	0
16	18	0	9.00	20.00	2	f	1	1
17	19	0	10.00	18.00	2	m	1	1
18	19	0	9.00	21.00	1	f	1	1
19	20	1	11.00	20.00	3	f	3	1
20	17	1	8.00	17.00	1	f	2	1

Example on ANCOVA



Checking Homogeneity of the Covariate Coefficients

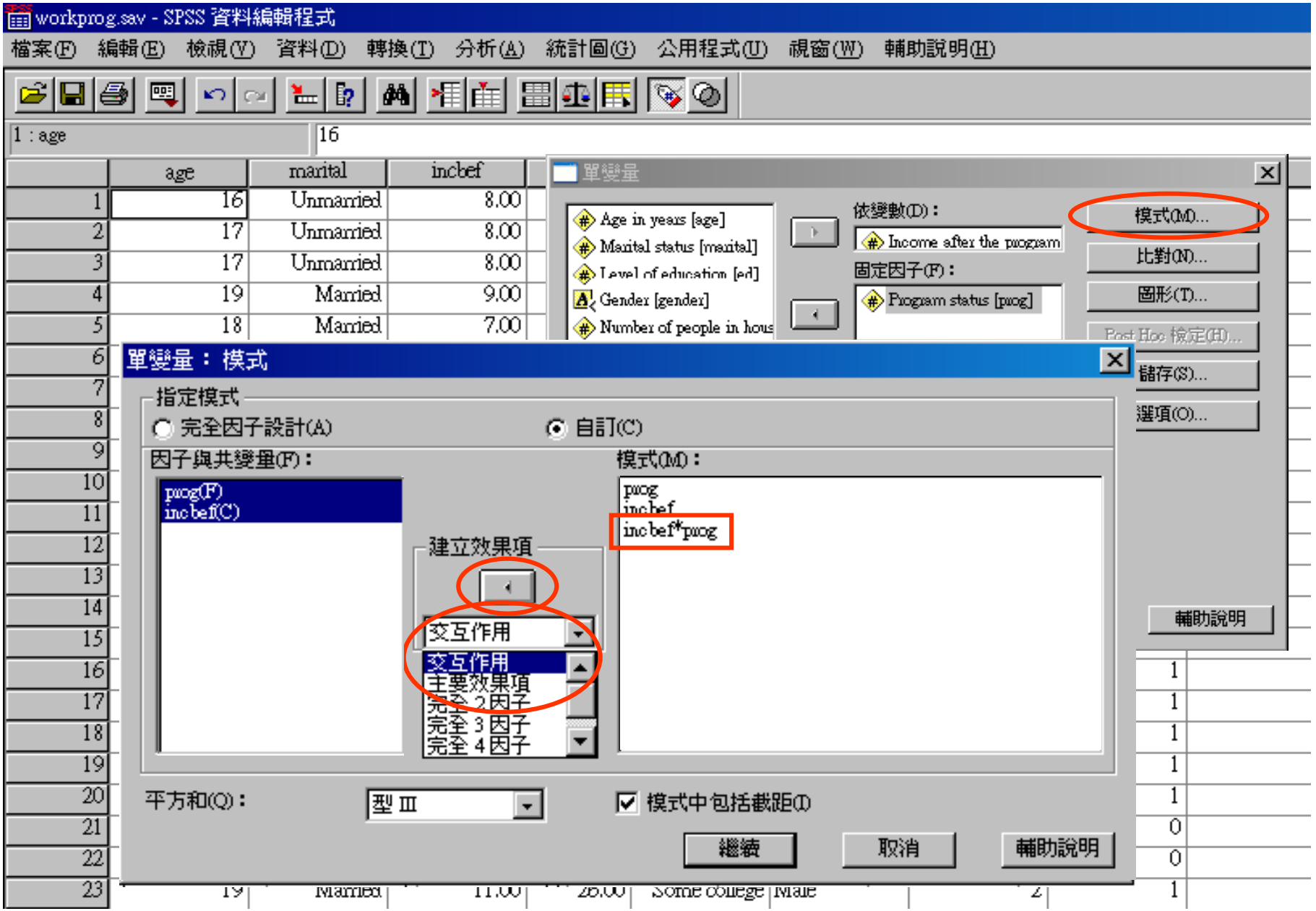


Checking Homogeneity of the Covariate Coefficients

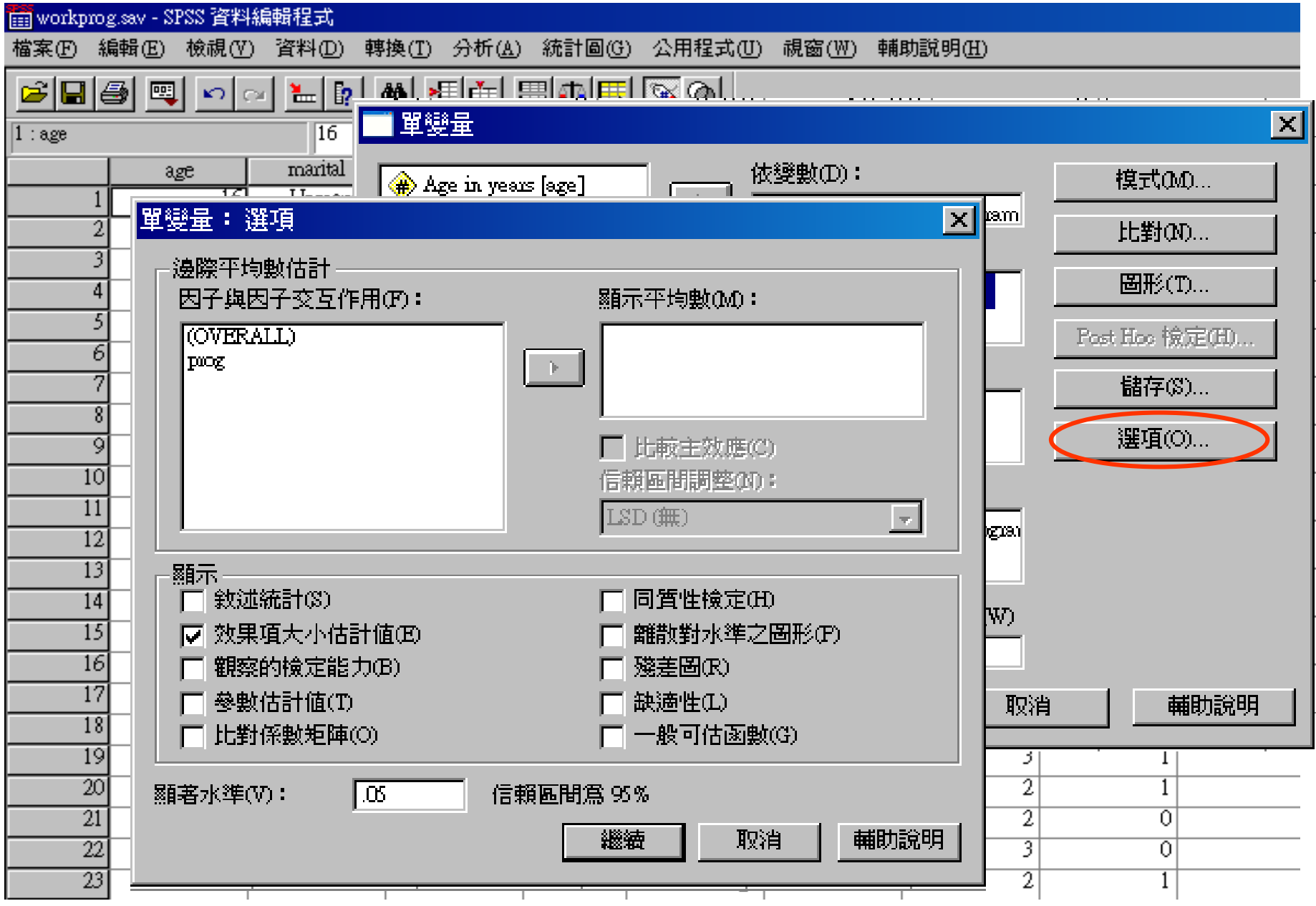
Click Model.

- Select Custom as the model type.
- Select prog(F) and incbef(C) in the Factors and Covariates list.
- Select Main effects from the Build Term(s) drop-down list and select the main effects to the model.
- Select prog(F) and incbef(C) in the Factors and Covariates list.
- Select Interaction from the Build Term(s) drop-down list and select the interaction term to the model.

Checking Homogeneity of the Covariate Coefficients



Example on ANCOVA



Example on ANCOVA

```
UNIANOVA
  incaft  BY prog  WITH incbef
  /METHOD = SSTYPE(3)
  /INTERCEPT = INCLUDE
  /PRINT = ETASQ
  /CRITERIA = ALPHA(.05)
  /DESIGN = prog incbef incbef*prog .
```

Example on ANCOVA

受試者間因子

		個數
Program status	0	517
	1	483

受試者間效應項的檢定

依變數: Income after the program

來源	型 III 平方和	自由度	平均平方和	F 檢定	顯著性	淨相關 Eta 平方
校正後的模式	12295.033[a]	3	4098.344	429.755	.000	.564
截距	131.271	1	131.271	13.765	.000	.014
prog	106.795	1	106.795	11.199	.001	.011
incbef	7152.586	1	7152.586	750.025	.000	.430
prog * incbef	4.292	1	4.292	.450	.502	.000
誤差	9498.318	996	9.536			
總和	297121.000	1000				
校正後的總數	21793.351	999				

a. R 平方 = .564 (調過後的 R 平方 = .563)

Example on ANCOVA

受試者間效應項的檢定

依變數: Income after the program

來源	型 III 平方和	自由度	平均平方和	F 檢定	顯著性	淨相關 Eta 平方
校正後的模式	12295.033[a]	3	4098.344	429.755	.000	.564
截距	131.271	1	131.271	13.765	.000	.014
prog	106.795	1	106.795	11.199	.001	.011
incbef	7152.586	1	7152.586	750.025	.000	.430
prog * incbef	4.292	1	4.292	.450	.502	.000

Its partial eta squared term is near 0, showing it accounts for a negligible amount of variation compared to the error term.

These results mean that you can assume homogeneity of the coefficient for the covariate across the levels of the factor.

Running the Analysis

Recall the GLM Univariate dialog box.

Click Model.

Select Full factorial as the model type

Click Continue.

Click Options in the GLM Univariate dialog box.

Select Descriptive statistics, Homogeneity tests, Spread vs. level plot, and Parameter estimates in the Display group.

Click Continue.

Click OK in the GLM Univariate dialog box.

Running the Analysis

These selections produce an analysis of covariance to assess the effect of program participation, controlling for salary before entering the program.

Summary

- By specifying an interaction between the covariate and factor, you are able to test the homogeneity of the covariate parameter estimates across levels of the factor. Since the interaction term was not significant, indicating the covariate parameter estimates are homogenous, you proceeded with an analysis of covariance and found that participation in the program increases salary by $4,357, on average.

- If the interaction term was significant, you could use the model with the interaction term, with the understanding that assessing the effect of program participation is complicated by the presence of the interaction. That is, when this interaction term is significant, the difference between participation and nonparticipation changes for different values of Income before the program.

Conclusions

The GLM Univariate procedure is useful for modeling the linear relationship between a dependent scale variable and one or more categorical and scale predictors.

- If you have only one factor, you can alternatively use the One-Way ANOVA procedure.

- If you only have covariates, use the Linear Regression procedure for more model-building, residual-checking, and output options.

End of The Section