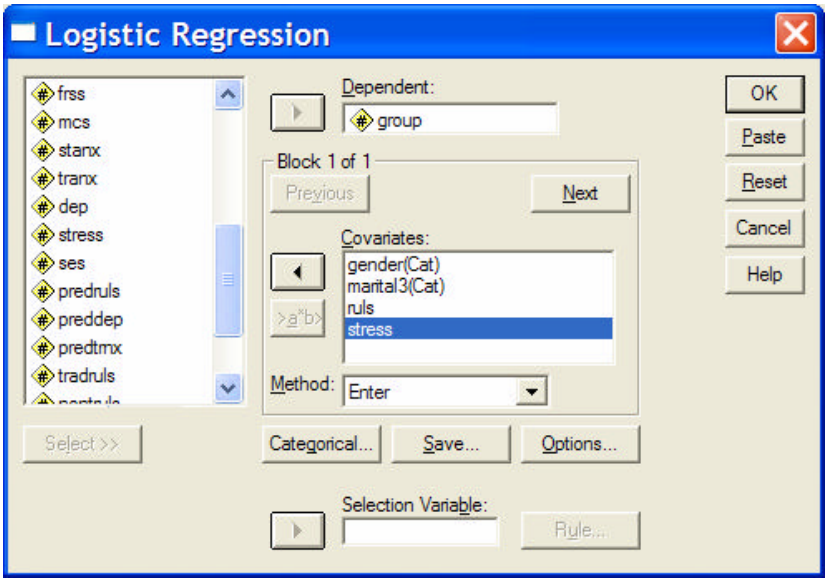
Binary Logistic Regression

Main Effects Model

Logistic regression will accept quantitative, binary or categorical predictors and will code the latter two in various ways. Here's a simple model including a selection of variable types -- the criterion variable is traditional vs. non-traditionally aged college students and the predictors are gender, marital status, loneliness and stress.

Analyze → Regression → Binary Logistic



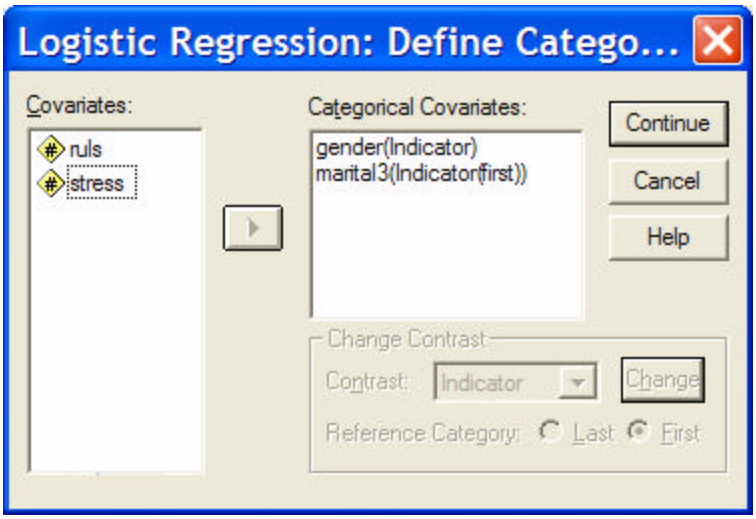
Move the criterion variable into the "Dependent:" window

Move the predictors into the "Covariate" window

Notice that you can select a subset of cases from here which gives the same result as using Data → Select Cases.

Click "Categorical" to specify the categorical variables.

Specify the entry method -- here Enter means to add all variables to the model simultaneously. Both blockwise entry (adding all predictors in that block) and blockwise selection (selecting "best" predictors from that block and entering them one at a time) are possible



Move each categorical predictor into the "Categorical Covariates:" window.

The default coding is dummy codes ("Indicator coding") with the highest coded group as the comparison group ("reference group").

I've changed the comparison group for marital3 to be the lowest coded group (1)

Highlight the categorical variable you want to specify, then choose whether the comparison group should be the "First" or the "Last" then click "Change".

Notice the specification of each predictor is given in the "Categorical Covariates:" window.

The SPSS output specifies the coding, etc. in the first part of the output.

Dependent Variable Encoding

Original Value	Internal Value
traditional	0
nontraditional	1

The coding for the criterion variable is given first -- The largest coded group is identified as the "target"

Categorical Variables Codings

		Frequency	Parameter coding (1)	Parameter coding (2)
MARITAL3	single	242	.000	.000
	married	121	1.000	.000
	other	42	.000	1.000
GENDER	male	180	1.000	
	female	225	.000	

The coding for all categorical predictors is specified.

Marital3

- Single will be the comparison group
- 1st parameter will compare married to single
- 2nd parameter will compare other to single

Gender

- female will be the comparison group
- the parameter will compare male to female

Classification Table[a,b]

			Predicted		
			GROUP		Percentage Correct
	Observed		traditional	nontraditional	
Step 0	GROUP	traditional	204	0	100.0
		nontraditional	201	0	.0
	Overall Percentage				50.4

a. Constant is included in the model.

b. The cut value is .500

"Block 0"

The process is inherently stepwise -- for forming and testing nested hierarchical models.

The first step is to compute and enter just the constant -- even if you've specified only a single "block" of variables, as in this case.

The classification table tells the # and % or cases correctly classified by the model.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	-.015	.099	.022	1	.881	.985

Regression weights and a test of the H0: b = 0 for the variables in the equation (only the constant for Block 0) is provided.

Variables not in the Equation

			Score	df	Sig.
Step 0	Variables	GENDER(1)	5.138	1	.023
		MARITAL3	263.571	2	.000
		MARITAL3(1)	163.823	1	.000
		MARITAL3(2)	47.559	1	.000
		RULS	27.890	1	.000
		STRESS	8.912	1	.003
	Overall Statistics		277.780	5	.000

The contribution of each predictor were it added alone into the equation on the next step is "foretold".

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	360.959	5	.000
	Block	360.959	5	.000
	Model	360.959	5	.000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	200.468	.590	.786

Classification Table[a]

			Predicted		
			GROUP		Percentage Correct
	Observed		traditional	nontraditional	
Step 1	GROUP	traditional	198	6	97.1
		nontraditional	30	171	85.1
	Overall Percentage				91.1

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	GENDER(1)	-.081	.370	.048	1	.826	.922
	MARITAL3			61.774	2	.000	
	MARITAL3(1)	5.974	.760	61.774	1	.000	393.029
	MARITAL3(2)	23.064	5793.703	.000	1	.997	1.0E+10
	RULS	.091	.017	29.943	1	.000	1.095
	STRESS	-.072	.028	6.448	1	.011	.930
	Constant	-4.480	.706	40.255	1	.000	.011

a. Variable(s) entered on step 1: GENDER, MARITAL3, RULS, STRESS.

"Block 1"

Step -- tests the contribution of the specific variable(s) entered on this step

Block -- tests the contribution of all the variables entered with this block

Model -- tests the fit of the whole model

These are all the same for a model with a single set of predictors that are entered simultaneously.

2 R^2 values are presented to estimate the fit of the model to the data -- both are transformations of the -2log likelihood values.

The reclassification table shows the accuracy of the model.

If you are going to do classification predictions, asymmetries and/or accuracy can sometimes be improved by adjusting the cutoff value from the default of .5 (in the options window).

Interpreting the model:

Gender does not contribute to the model

- the negative B indicates that the target group (nontraditional) tends to have more of those coded "0" (females) than of these coded "1" (males) - but not significantly, after controlling for the other predictors

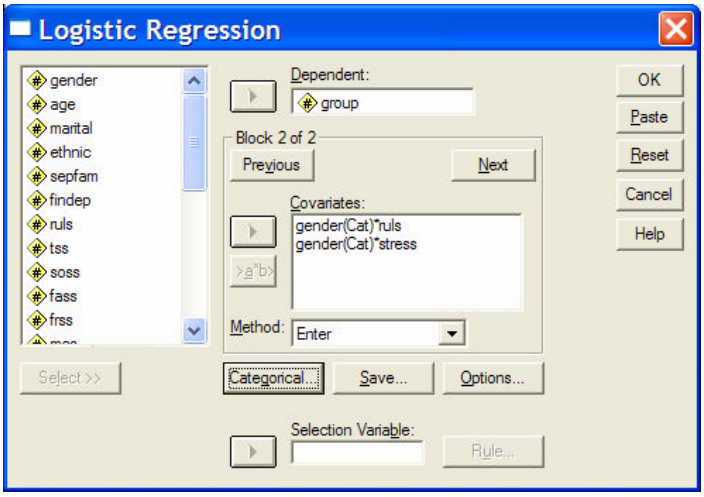
Marital does contribute to the model

- there is a test of the overall contribution of multi-category variables, as well as a test of the contribution of each parameter
- the positive B for married vs. single indicates more married (1) in the nontraditional (1) group, after controlling…
- other vs. single does not contribute to the model, after controlling for the other predictors

Loneliness does contribute to the model, with higher average loneliness for the nontraditional (1) group, after controlling..

Stress does contribute to the model, with lower average stress for the nontraditional (1) group, after controlling…

A second step added the interactions of gender with loneliness and stress.



Click “Next” to specify what will be entered on the second Block

To inter an interaction highlight both terms of the interaction (by holding down “Ctrl” as you click each predictor) and then click the “>a*b>” button

The interaction will be calculated as the product of the terms, or the term parameters in the case of categorical variables (following the specifications given earlier)

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	14.158	2	.000
	Block	14.158	2	.000
	Model	375.117	7	.000

Again “Step” and “Block” are the same because of the blockwise entry of the two interaction terms.

The “Step” and “Block” X^2 tests tell us that the model was improved by the inclusion of these terms. Remember that tests the “average contribution” of the included terms and so you one or more of multiple included terms can be significant without a significant improvement to the model.

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	186.123	.634	.819

As would be expected, the fit of the model improved, according to both of the R^2 calculations.

Classification Table[a]

			Predicted		
			GROUP		Percentage Correct
	Observed		traditional	nontraditional	
Step 1	GROUP	traditional	201	3	98.5
		nontraditional	12	189	94.0
	Overall Percentage				96.3

a. The cut value is .500

The % correct reclassification also improved, especially for the non-traditional group.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	GENDER(1)	.259	1.325	.038	1	.845	1.296
	MARITAL3			61.736	2	.000	
	MARITAL3(1)	5.983	.762	61.736	1	.000	396.781
	MARITAL3(2)	23.072	5800.267	.000	1	.997	1.0E+10
	RULS	.093	.023	16.509	1	.000	1.098
	STRESS	-.064	.038	2.803	1	.094	.938
	GENDER(1) by RULS	-.005	.033	.022	1	.883	.995
	GENDER(1) by STRESS	-2.179	.057	4.099	1	.000	1.882
	Constant	-4.641	.933	24.742	1	.000	.010

a. Variable(s) entered on step 1: GENDER * RULS , GENDER * STRESS .

Interpreting the model:

Gender does not contribute to the model including the interactions either (though always check -- contributions can change with the adding or deleting of predictors -- as the “colinearity mix” changes!)

Marital does contribute to the model -- with the same pattern as in the main effects model

Loneliness does contribute to the model, with higher average loneliness for the nontraditional (1) group, after controlling..

Stress does not contribute to the model -- though it contributed to the main effect model

The **Gender * Loneliness interaction** does not contribute to the model

The **Gender * Stress interaction** does contribute to the model

- the negative weight for this interaction means the slope of the relationship between group and stress is less positive for the males (1) and more positive for the females (0)