60080079 Introduction to Statistical Methods
Semester 2 2023-2024
Handout 10

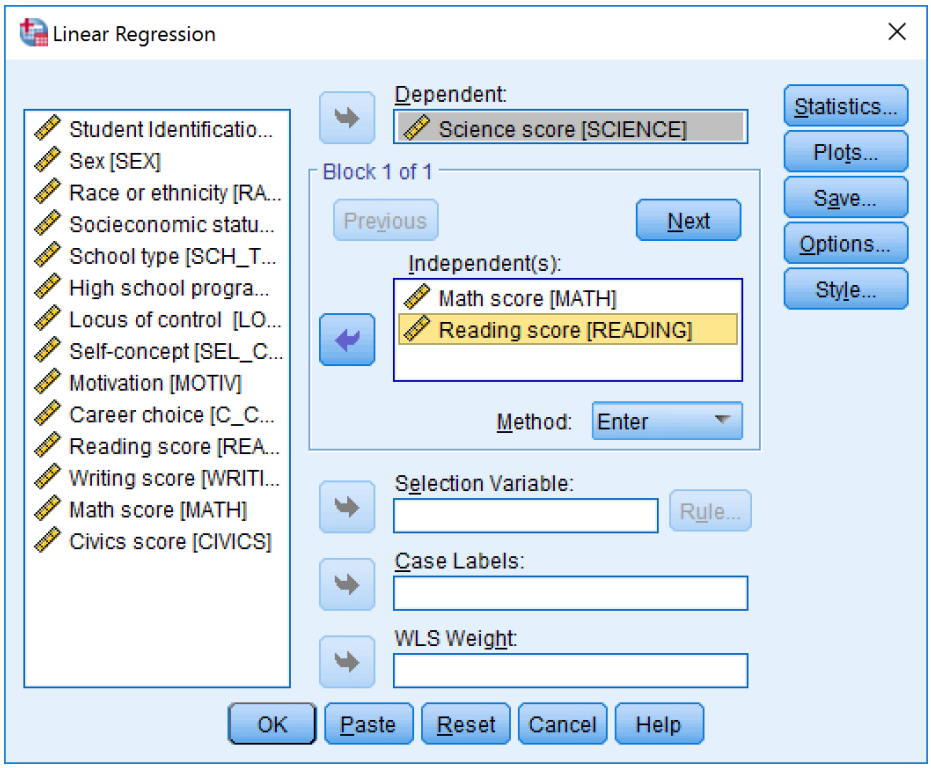
A Brief Introduction to Regression Inferences in SPSS

I. Multiple Regression

Open the **hsb.sav** file in SPSS. As with simple linear regression, multiple regression analysis can be carried out in SPSS under:

Analyze → Regression → Linear.

The main difference is that in the **Independent(s)** box, we can have more than one predictor. For example, assume we are interested in predicting Science score using Math and Reading scores.



Below are some results from the requested basic multiple regression analysis.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.733[a]	.537	.536	6.61231

a. Predictors: (Constant), Reading score, Math score

The R^2 in the **Model Summary** table shows that Math and Reading scores can account for 53.7% of the variance in Science score. (R is just the square-root of R^2.)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	30329.351	2	15164.675	346.839	.000[b]
	Residual	26102.382	597	43.723		
	Total	56431.733	599			

a. Dependent Variable: Science score

b. Predictors: (Constant), Reading score, Math score

The **ANOVA** tables shows that the computed F-statistic is 346.839, and the corresponding p-value is close to zero. This means that, taken together, Math and Reading scores can predict Science score.

Do we know from this result whether Math score alone can predict Science score? How about Reading score?

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	10.765	1.597		6.740	.000
	Math score	.345	.039	.335	8.829	.000
	Reading score	.445	.036	.463	12.213	.000

a. Dependent Variable: Science score

The **Coefficients** table gives us the estimates of the regression coefficients, and the corresponding t-statistics and p-values.

How should the coefficient estimates be interpreted?

What does it mean for the Math score to be significant (i.e., $t = 8.829$ and p-value is close to zero)? How about the Reading score?

Do we know from the **ANOVA** and **Coefficients** results whether Math score alone can predict Science score? How about Reading score?

What analysis is necessary for us to determine whether Math score alone can predict Science score? How about Reading score?

Note: As with simple linear regression, we can request additional analyses when conducting multiple regression analysis, specifically, using the **Statistics** and **Save** options.

Exercise: Suppose we are interested in:

1) A 90% confidence interval of the regression coefficients;
2) Predicting the Science score of an individual whose Math and Reading scores are 41.90 and 46.90, respectively; and
3) Constructing a 90% confidence interval of the mean prediction for the same Math and Reading scores as in 2).

Verify that you will get the following results:

Model		Unstandardized Coefficients		90.0% Confidence Interval for B	
		B	Std. Error	Lower Bound	Upper Bound
1	(Constant)	10.765	1.597	8.134	13.397
	Math score	.345	.039	.281	.410
	Reading score	.445	.036	.385	.505

a. Dependent Variable: Science score

MATH	READING	SCIENCE	PRE_1	LMCI_1	UMCI_1
40.20	33.60	39.00	39.59681	38.67601	40.51761
41.90	46.90	36.30	46.10262	45.44138	46.76386
41.90	41.60	44.40	43.74395	43.07127	44.41664

What is the predicted score for an individual whose Math and Reading scores are 42.4 and 48.5, respectively? Note that none of the 600 students in the HSB data has this particular combination of Math and Reading scores. (Answer: 46.98729)