BIOSTATS 640 – Normal Theory Multiple Linear Regression
Frequently Asked Questions (FAQ)
Incremental Sum of Squares and Partial F Tests

This FAQ is about **incremental sum of squares ("change in sum of squares"), analysis of variance,** and **partial F tests.** It explains a "shorthand" notation for variance explained in model building.

Illustration

An educator examined the relationship between number of hours devoted to reading each week (Y) and the independent variables social class (X_1), number of years school completed (X_2), and reading speed measured by pages read per hour (X_3). The analysis of variance table obtained from a stepwise regression analysis on data for a sample of 19 women over the age of 60 is shown.

The following **"shorthand"** summarizes incremental regression sum of squares explained as each of three predictors are added to the model, step by step:

Step 1 - **(X_3)** says: X_3 is fit

Step 2 - **($X_2|X_3$)** says: X_2 is added to the model that already contains X_3.

Step 3 - **($X_1|X_2, X_3$)** says: X_1 is added to the model that already contains X_3 and X_2 .

Source		DF	Sum of Squares	
Regression	(X_3)	1	1058.628	= Regression SSQ (X_3)
	($X_2 \mid X_3$)	1	183.743	= **"Extra regression SSQ" = Δ regression SSQ →** = Regression SSQ (X_2, X_3) - Regression SSQ (X_3)
	($X_1 \mid X_2,X_3$)	1	37.982	= **"Extra regression SSQ" = Δ regression SSQ →** = Regression SSQ (X_1,X_2, X_3) - Regression SSQ (X_2,X_3)
Residual		15	363.300	= Residual SSQ (X_1, X_2, X_3)
Total, corrected		18	1643.653	= $\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2$

Tip - It is possible to extract the 3 analysis of variance tables, one for each step, because the total corrected sum of squares is constant. Note how it is possible to obtain the residual sum of square values in two ways!

Step 1 Model

$$Y = \beta_0 + \beta_1 X_3 + \text{error}$$

Source	DF	Sum of Squares
Regression on X_3	1	**1058.628**
Residual	17	**585.025 = (1643.653 – 1058.628) = (363.300 + 37.982 + 183.743)**
Total, corrected	18	$1643.653 = \sum(y_i - \bar{y})^2$ **fixed**

Step 2 Model

$$Y = \beta_0 + \beta_1 X_3 + \beta_2 X_2 + \text{error}$$

Source	DF	Sum of Squares
Regression on X_3, X_2	2	**1242.371 = (1058.628 + 183.743)**
Residual	16	**401.282 = (1643.653 – 1242.371) = (363.300 + 37.982)**
Total, corrected	18	$1643.653 = \sum(y_i - \bar{y})^2$ **fixed**

Step 3 Model

$$Y = \beta_0 + \beta_1 X_3 + \beta_2 X_2 + \beta_3 X_1 + \text{error}$$

Source	DF	Sum of Squares
Regression on X_3, X_2, X_1	3	**1280.353 = (1058.628 + 183.743 + 37.982)**
Residual	15	**363.300 = (1643.653 – 1280.353) = (363.300)**
Total, corrected	18	$1643.653 = \sum(y_i - \bar{y})^2$ **fixed**