MET MA 603: SAS Programming and Applications

Proc Reg

The Regression Procedure

The **Regression Procedure** fits linear regression models to a dataset. This course will only cover simple linear regression, which have a single explanatory variable.

The **Model Statement** specifies the dependent and independent variables. In the example below, Weight is the dependent variable, and Height is the independent variable. In other words, Height is being used to predict Weight.


```
proc reg data = Height_Weight_Age ;
model   Weight = Height ;
quit ;
```

Note: Proc Reg is an “interactive procedure”.

The Regression Procedure (cont.)

The **Analysis of Variance output** shows information about the fit of the model. Usually, a p value of less than 0.05 is considered to be a good fit. The R-squared value indicates how much of the variance in the dependant variable can be explained by the independent variable. It ranges from 0 – 1.

The **Parameter Estimates** show the coefficients of the fitted regression equation. In the example used, the regression equation would be:

$$Weight = -143.02692 + 3.89903 * Height$$

The Residual Plot shows the distribution of residuals (actual – predicted value). Randomly distributed residuals indicate an unbiased model.

Practice

Use the Regression Procedure to create the simple linear regression models specified below:

Model	Dependent Variable	Independent Variable
1	Weight	Age
2	Height	Weight
3	Weight	Age

Practice

Use the Regression Procedure to create a simple linear regression models using the scores.sas7bdat dataset, such that exam2 is the dependant variable and exam1 is the independent variable. Does this model have a good fit? Why or why not?

Readings

- Textbook section 9.10