

Analysis of Covariance (ANCOVA) or Dummy Variable Regression (DVR)

Required Statements:

CLASS or CLASSES: As in ANOVA the CLASS statement defines which variables signify qualitative groupings in the data. The format of the CLASS statement is identical to ANOVA.

MODEL: Since this class of models contains a mixture of the ANOVA and REGRESSION models, all the rules for MODEL statements mentioned earlier apply. Main effects and interactions for CLASS variables can be defined as before as well as continuous variables such as pH. In ANCOVA, analysis of variance is done on responses which have been adjusted for another continuous variable, such as pH. In DVR a new effect may come into the MODEL statement which is a crossed term between a CLASS effect and a continuous effect. These models are used to compare regressions between two or more groups, like varieties. The interpretation of the terms in these models will be discussed later in the examples.

Additional Statements and Options:

MEANS and LSMEANS: Only the LSMEANS statement should be used in ANCOVA models. These will be the means for the given effects after they have been adjusted for the continuous variable or covariate. Such models are useful only when experiments have identifiable and measurable gradients or differing levels of some extraneous effect such as moisture, pretreatment levels of weed infestation, or as in the case of our example, pH levels. This provides statistical control (controlling for unanticipated variability after the experiment) as apposed to experimental control (controlling for known variability before the experiment is conducted, i.e. blocking). As usual there are assumptions required for these analyses. (Handouts on ANCOVA are available upon request).

CONTRAST: The CONTRAST statement may be used as in ANOVA. In ANCOVA this statement would be used to compare specific combinations of treatments. For DVR the contrast allows the user to compare regression coefficients and response curves for the corresponding models.

MODEL Statement Options: All options mentioned under regression are valid - - more useful ones in DVR are SOLUTIONS, INT and NOINT. The last two become important in interpreting the meaning of the parameter estimates provided by the SOLUTION option.

Example 7 - ANCOVA (RCB)

```
PROC GLM;  
  CLASS VAR FERT BLOCK;  
  MODEL YIELD = BLOCK VAR FERT VAR*FERT pH;  
  
  LSMEANS VAR FERT VAR*FERT / PDIFF STDERR;
```

NOTES: This is identical to the ANOVA model given earlier except for the addition of pH to the MODEL statement. Notice that this term does not appear in the CLASS statement, therefore it is a continuous effect and will have only 1 df. The LSMEANS statement calculates the adjusted means for the effects listed.

Example 8 - DVR with implied intercept:

```
PROC GLM;  
  CLASS VAR;  
  MODEL YIELD = VAR VAR*pH / SOLUTION;  
  
  CONTRAST 'Slopes' VAR*pH 1 -1;
```

NOTES: In this example two regression lines are fit simultaneously, one for variety VAR1 and the other for VAR2. Both models are simple linear regressions of yield on pH. SAS will automatically put in an overall intercept. Thus, the estimates for the VAR effect listed by SOLUTION will measure the differential effect of the intercepts of the two models. The effects listed under VAR*pH are the estimated slope coefficients of each variety. The CONTRAST statement tests the equality of these slopes or whether or not the regression lines are parallel.

Example 9 - DVR with explicit intercepts:

```
PROC GLM;  
  CLASS VAR;  
  MODEL YIELD = VAR VAR*pH/SOLUTION NOINT;  
  
  CONTRAST 'Intercepts'  VAR 1 -1;  
  CONTRAST 'Slopes'  VAR*pH 1 -1;  
  CONTRAST 'Lines'  VAR 1 -1,  
              VAR*pH 1 -1;
```

NOTES: The MODEL statement is the same as Example 8 except for the addition of the NOINT option. This has the effect of forcing direct estimation of the intercept parameters instead of their differential effect. The first two contrasts compare intercepts and slopes respectively while the third tests for coincident lines. Note that this is a 2 degree of freedom test that simultaneously looks at intercepts and slopes.