

collard.sas: Explanation of code

Goals of code:

- Fitting a 2-way ANOVA model
- After the ANOVA options for main effects and cell means
- Analyzing data from an RCBD

Fitting a 2-way ANOVA model: `model conc = clarify size clarify*size;`

Earlier, we used `proc glm;` to fit ANOVA models. `proc glm;` can also be used to fit more complicated models. Just name the components on the right-hand side of the model statement. The interaction is given by two terms connected with an asterisk.

Remember that anything identifying groups must be named in a class statement before being used in the model statement. You only need the factor names in the class statement; you do not need the interaction there. The interaction only goes in the model statement.

`Proc glm` provides three sets of F tests:

At the top of the output: box with model and error. This is the overall test of the entire model: It compares the equal means (intercept only) model to the specified model. Probably not useful for a two-way ANOVA and certainly not useful when you have blocks in the model.

The error line in this output is where you will find the MSE, the estimate of the pooled error variance.

Type III (partial) F tests: These are the tests that we are using in this class because I believe they answer the most interesting questions. I usually ignore the type I (sequential) F tests.

After the ANOVA:

You can do all the “after the ANOVA” analyses using the principles discussed in the one-way ANOVA example. You just name the factor of interest in the `estimate` or `lsmeans` statement. You can name multiple factors, e.g. `lsmeans clarify size;` to get both the size `lsmeans` and the `clarify lsmeans` from one statement.

If you want information or comparisons between individual treatments, put the interaction, e.g. `clarify*size`, in the `lsmeans` or `estimate` statement (example in the `proc glimmix` also works in `proc glm`).

Alternate way to fit models: `proc glimmix;`

`Proc glm` is one of the very first SAS procedures; it dates from the early 1970's. Over the years, SAS has added options to old procs and added a lot of new procs. Some of the new procedures provide useful output not in the earlier procs.

proc glimmix can be used to analyze a huge number of models, including everything that you would use proc glm for. glimmix provides improved output and some new features that are not in proc glm. Two that are useful now are improved results for differences of means (either treatment means or factor lsmeans) and a letter display for differences.

The syntax for proc glimmix is exactly the same as for proc glm, when you want to fit models with normally distributed errors. Same class statement; same model statement.

The lsmeans /diff option provides nicer results for pairwise differences between means.

The lsmeans /lines option provides a compact display of which groups are significantly different from each other. adjust= names the multiple comparisons adjustment; it can be used with either /diff or /lines.

Analyzing data from a randomized complete block design:

No code, because this is identical to the steps above. Just leave the interaction out of the model, because it can not be estimated without replication within a block.