SAS refresher

Brenna Aegerter, Vegetable Crops Advisor
UCCE San Joaquin County
bjaegerter@ucdavis.edu

Michael Cahn, Irrigation and Water Resources Advisor UCCE Monterey County
mdcahn@ucdavis.edu

Purchasing SAS
http://software.ucdavis.edu

MacOS Availability			
Version	Platform	Cost	License
6.12	MacOS 9	$32.00	Expiration: June 30,2009
Unix Availability			
Version	Platform	Cost	License
9.1.3	Linux	$32.00	Expiration: June 30,2009
	Solaris	$32.00	Expiration: June 30,2009
Windows Availability			
Version	Platform	Cost	License
9.2 Windows 64bit	Windows XP x64 Vista	$32.00	Expiration: June 30,2009
9.2 Windows 32bit	Windows XP Vista	$32.00	Expiration: June 30,2009
9.1.3 Standalone	Windows 2000 Windows ME Windows XP	$32.00	Expiration: June 30,2009

Resources for help with SAS

- UC DAVIS STATS LAB
- SAS-SPECIFIC BOOKS
  - **SAS publishing:** http://support.sas.com/publishing/resources.html
  - Books also available from booksellers (UCD bookstore, Amazon, and others)
  - UCD library
- SAS DOCUMENTATION
- ONLINE RESOURCES

Online resources for help with SAS

- SAS "workshops" from University of Idaho

http://www.uiweb.uidaho.edu/ag/statprog/sas/workshops/workshops.htm

- Choosing the Right Statistical Test (plus SAS code)

http://www.ats.ucla.edu/stat/mult_pkg/whatstat/choosestat.html

- Jorge Dubcovsky's UCD PLS205 course syllabus and notes (includes SAS code)

http://www.plantsciences.ucdavis.edu/agr205/schedule.htm

- P-value calculator (put in F, t or chi-square, etc and degrees freedom)

http://www.graphpad.com/quickcalcs/PValue1.cfm

Outline

- DATA step - 3 ways to handle
- Testing the assumptions of ANOVA
- Transformations and conversions
- PROC step
  - CLASS statement
  - MODEL statement
  - MEANS and LSMEANS statements
    - t-tests and multiple comparisons
  - CONTRAST statements
  - Combining data from multiple years, locations
  - Exporting ouput to Excel

"MEANS" VERSUS "LSMEANS"

MEANS - These are what is usually meant by mean (average) and are computed by summing all the data points and dividing by the total # of points. They are also referred to as arithmetic means and they are based on the data only.

LSMEANS - Least Squares Means can be defined as a linear combination (sum) of the estimated effects (means, etc) from a linear model. These means are based on the model used.

When missing values do occur, the two will differ (and LSMEANS would be preferable)

Combining Similar Experiments

- Before performing combined analysis across locations, years, etc., make sure the error mean squares are homogeneous
- Use Bartlett's chi-square test if you have more than two mean squares
- Use an F-test (i.e. Larger MS/Smaller MS) if you have two mean squares
- For most of us, trials repeated at "random" locations over the years really should be analyzed as mixed models with year/location as a random effect (but that's another class…)