**Final Exam: SOIL 5112
Tuesday, April 30, 2013
8:00 am**

1. **The soil N test discussed by Dr. Bushong was originally developed in \_\_\_\_\_\_\_\_ and was called the \_\_\_\_\_\_\_\_\_\_\_\_\_\_**
2. **Worldwide fertilizer prices paid by farmers has \_\_\_\_\_\_\_\_\_\_\_\_\_ in the last 10 years**

**Doubled**

**Tripled**

**Quadrupled**

1. **What is the definition of a “critical level”**



1. **What would the critical level be for the data above.**

**Using Cate Nelson \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Using a quadratic model \_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Using a linear plateau \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

1. **For the graph included below, please modify the SAS code for a linear-plateau model**

**proc nlin data = one best = 3;
parms b0= \_\_\_\_ to \_\_\_\_ by 0.01 b1=\_\_\_\_\_ to \_\_\_\_\_ by \_\_\_ njoint=\_\_\_\_\_ to \_\_\_\_\_ by \_\_\_;**

1. **Plant to plant differences in corn grain yield averaged (AJ article)**

**47 bu/ac**

**4.7 bu/ac**

**14.7 bu/ac**

**104.7 bu/ac**

**COVARIANCE**

1. **The assumptions that must be considered when using analysis of covariance are….**
2. **When analyzed as a dependent variable, covariate needs to be \_\_\_\_\_\_\_**
3. **Covariance can be viewed as “a linear regression adjustment” within analysis of variance (T or F)**
4. **What are the dangers of analyzing data using ANOVA when data is not normal? What can be done to fix this?**
5. **Fill in the SAS program below so as to properly use the covariate “prep” (pre plant soil test P)**

**Data one;**

**input rep trt yield prep;**

**cards;**

**1 1 30 42**

**1 2 35 40**

**proc glm;**

**class \_\_\_\_\_\_ \_\_\_\_\_\_\_;**

**model \_\_\_\_\_ = \_\_\_\_\_ \_\_\_\_\_ \_\_\_\_\_\_;**

**lsmeans \_\_\_\_\_\_;**

**run;**

1. **What analysis was discussed in class that could possibly be used to account for underlying spatial variability?**
2. **Spatial variability in production fields was demonstrated to occur at**

**1 ft x 1ft**

**8 rows \* 20 ft in length**

**Field to field**

1. **Name three causes of spatial variability encountered in agricultural production**
2. **Third dimension of stability analysis discussed in class whereby a surface response model would be generated using the original Env. Mean versus Treatment mean and \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.**
3. **How many years (locations, sites, etc.) of data are required to generate a meaningful regression equation for use in stability analysis?**
4. **Stability analysis conducted on the Magruder Plots showed that \_\_\_\_\_\_ applications appeared to be beneficial in \_\_\_\_\_\_\_\_\_\_ environments.**
5. **You have an experiment with 3 reps and 12 treatments. The 12 treatments consist of a full factorial arrangement, where there are 4 nitrogen rates (NR) and 3 varieties (VAR).**

**Treatment N Rate Variety**

**1. 0 TAM101**

**2. 40 TAM101**

**3. 80 TAM101**

**4. 120 TAM101**

**5. 0 KARL**

**6. 40 KARL**

**7. 80 KARL**

**8. 120 KARL**

**9. 0 DUSTER**

**10. 40 DUSTER
11. 80 DUSTER**

**12. 120 DUSTER**

1. **SAS program if you analyze this as a full factorial**
2. **SAS program if you analyze this as a rep-treatment model**
3. **If a treatment\*environment interaction is significant what does it say about how treatment must be interpreted?**
4. **What about treatment\*year?**
5. **What advantages of 4 versus 3 reps were discussed in class?**
6. **What does CGIAR stand for?**
7. **What is a “synergistic” interaction? Graph would help (label the axes)**
8. **What is an “antagonistic” interaction? Graph would help**
9. **Two trials: LMSE = 58000 SMSE = 24000, dfe (both trials) = 20**

**Compute the F statistic. \_\_\_\_\_\_ Based on your knowledge of the table values, should these trials be combined?**

**(F values on the board)**

1. **I want to know what “percent of the mean” difference you need to say there are differences in treatments? (more or less, and why)**
2. **(2 treatment means were 2500 and 3400 kg/ha). Using your answer in 26, what would this be in kg/ha? (for this data)**

1. **Fill in the blanks below on how you would use PROC CORR to establish the relationship between yield and NDVI with population, disease, height, and BYDV (barley yellow dwarf virus).**

**\_\_\_\_\_\_ \_\_\_\_\_\_;**

**var \_\_\_\_\_\_ \_\_\_\_\_\_;**

**with \_\_\_\_\_\_ \_\_\_\_\_\_ \_\_\_\_\_\_\_ \_\_\_\_\_\_\_;**

1. **In order to merge two data sets that have rep, trt, yield, and location as identifiers, fill in the blanks below as to how this would be accomplished.**

**data loc1;**

**proc \_\_\_\_\_; by \_\_\_\_ \_\_\_\_\_ \_\_\_\_\_ ;**

**data loc1;**

**proc \_\_\_\_\_; by \_\_\_\_ \_\_\_\_\_ \_\_\_\_\_ ;**

**data comb; \_\_\_\_\_\_ \_\_\_\_\_\_ \_\_\_\_\_; by \_\_\_\_ \_\_\_\_ \_\_\_\_;**

1. **1. If you want to identify that you have a character variable variety (e.g., TAM101, OK101, HUSKER1, KSU2, CSU2), followed by rep and treatment (both in numeric form) provide an example of how this will look in the input statement.**

**data one;**

**input \_\_\_\_\_ \_\_\_\_\_\_ \_\_\_\_\_\_ \_\_\_\_\_\_;**

**cards;**

1. **In order for SAS to understand that you have missing data, what must be entered within that cell?**
2. **The very first “PROC” procedure that you should run in any program is ?**
3. **Which of the following have to be true in order to use an independent variable as a covariate?**

**a. the covariate has to be independent of “trt”**

**b. treatment must be significant when the covariate is analyzed as a dependent variable**

**c. must be collected before treatments are applied**

**d. must be collected after treatments are applied**

1. **What are the assumptions of analysis of variance?**
2. **When should “LSMEANS” be used to replace the normally computed “MEANS?” (2 answers)**
3. **\_\_\_\_\_\_\_\_ ensures that you will have an estimate of experimental error**
4. **\_\_\_\_\_\_\_\_ ensures that you will have an unbiased estimate of experimental error**
5. **When you have missing data, what sums of squares should be used?**
6. **LSD’s cannot be used when the treatment structure includes \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**
7. **What is the main reason for blocking?**
8. **If there isn’t a known “gradient” within a field trial, what experimental design is recommended?**
9. **SED times \_\_\_\_\_\_\_ is generally what would be computed using what mean separation procedure?**
10. **What is the main difference between the scientific method and the experimental method?**
11. **What kind of error is incurred if a scientist “excludes” data that does not conform to his/her hypotheses?**
12. **Good researchers aren’t necessarily characterized by being smart, but by…….**
13. **What is autocorrelation?**
14. **For the example below, from the 2 linear regression equations, is there a**

**a. significant difference in the intercept components? \_\_\_\_\_\_\_\_\_\_\_**

**b. significant difference in the slope components? \_\_\_\_\_\_\_\_\_\_\_\_**

****

****

1. **For the data below (Yield by N rate study under zero-till and conventional tillage, what would the Cate-Nelson critical level (N Rate) be for the two tillage systems? (Draw the 2 cross bars for full credit.**
2. **Below is a GLM for 1971 and 1981, treatments 1-6 from Experiment 502 in Lahoma Oklahoma. Using the analysis provided, answer the following questions.**
3. **When should the Type III sums of squares be used instead of Type I?**
4. **Was there a need to use Type III sums of squares in this case for these 2 years of data?**
5. **Should treatment means have been interpreted over years or by year?**
6. **What statistic did you use to make the decision in #c?**
7. **Was there a treatment mean(s) (either year) that stood out, whereby you suspected an outlier?**
8. **What statistic tells you that was likely a wheat experiment and not a corn trial?**
9. **What is meant by REP(YR)?**
10. **What is REP(YR) used for?**
11. **Compute the SED for this experiment.**The SAS System 14:11 Thursday, April 25, 2013
The GLM Procedure

 Class Level Information

 Class Levels Values

 YR 2 1971 1981

 REP 4 1 2 3 4

 TRT 7 1 2 3 4 5 6 7

 Number of Observations Read 56

 Number of Observations Used 56

Dependent Variable: kgha

 Sum of

 Source DF Squares Mean Square F Value Pr > F

 Model 19 8205054.58 431844.98 4.15 0.0001

 Error 36 3747087.72 104085.77

 Corrected Total 55 11952142.30

 R-Square Coeff Var Root MSE kgha Mean

 0.686492 14.41917 322.6233 2237.460

 Source DF Type I SS Mean Square F Value Pr > F

 YR 1 1431808.560 1431808.560 13.76 0.0007

 REP(YR) 6 495662.840 82610.473 0.79 0.5810

 TRT 6 3436281.580 572713.597 5.50 0.0004

 YR\*TRT 6 2841301.597 473550.266 4.55 0.0016

 Source DF Type III SS Mean Square F Value Pr > F

 YR 1 1431808.560 1431808.560 13.76 0.0007

 REP(YR) 6 495662.840 82610.473 0.79 0.5810

 TRT 6 3436281.580 572713.597 5.50 0.0004

 YR\*TRT 6 2841301.597 473550.266 4.55 0.0016

 Tests of Hypotheses Using the Type III MS for REP(YR) as an Error Term

 Source DF Type III SS Mean Square F Value Pr > F

 YR 1 1431808.560 1431808.560 17.33 0.0059

 Level of ---------kgha---------

 TRT N Mean Std Dev

 1 8 1860 458

 2 8 1890 631

 3 8 2264 190

 4 8 2278 513

 5 8 2356 474

 6 8 2451 201

 7 8 2560 205

 Level of Level of ------kgha-------

 YR TRT N Mean Std Dev

 1971 1 4 2264 118

 1971 2 4 2467 151

 1971 3 4 2399 130

 1971 4 4 2387 369

 1971 5 4 2367 144

 1971 6 4 2380 192

 1971 7 4 2514 148

 1981 1 4 1455 200

 1981 2 4 1313 131

 1981 3 4 2130 140

 1981 4 4 2169 668

 1981 5 4 2345 710

 1981 6 4 2522 210

1. 7 4 2606 266
2. **What is this formula for?**

 **square root (2\*MSE/reps) or square root (2\*s2/reps)**

1. **For the 3D scatter plot below, fill in the blanks for the program used to generate this output (variables are YP0 (yield potential) on the Z, Year on the X and RI0N (response index) on the Y). This is data from Experiment 502 that we looked at in class (long-term NPK trial at Lahoma).**

**proc** \_\_\_\_\_ ;

scatter \_\_\_\_\_\_\_ \* \_\_\_\_\_\_\_ = \_\_\_\_\_\_\_/shape='pyramid';

**run**;

****

1. **If I had a fourth variable, “variety” (in addition to the 3 reported) where there were 2 different varieties evaluated, how could I look at this, on this same graph?**
2. **For this data set (visual observation), was there a relationship between RI0N and YP0 (yield)?**
3. **What does the following program do?**

proc iml;
dens={0 100 600 1200}; \*\*
p=orpol(dens);
t=nrow(p);
do i=1 to t;
pr=abs(p[,i]);
pr[rank(abs(p[,i]))]=abs(p[,i]);
do j=t to 1 by -1;
if pr[j] > 1.e-10 then scale=pr[j];
if abs(p[j,i]) < 1.e-10 then p[j,i]=0;
end;
p[,i]=p[,i]/scale;
end;
print p;
run;

****

1. **You have an experiment with 4 N Rates (0, 20, 40, 60 kg N/ha) and 2 Tillage systems (Conventional and Zero-Till). Using the coefficients for equally spaced treatments above, produce the proper SAS statement for the following contrasts. (actual statement has to work in SAS, no errors).**
2. **N rate linear**
3. **N rate quadratic**
4. **N rate linear \* tillage**
5. **N rate quad \* tillage**