**Final Exam: SOIL 5112
Friday, May 1, 2015
8:00 am**

1. What is the definition of a “critical level” as we understand it in agriculture?



1. What would the critical level be for the data above. (include your estimated lines)

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 in the Agronomy Journal article that you reviewed, averaged

47 bu/ac

4.7 bu/ac

14.7 bu/ac

104.7 bu/ac

1. Why is plant to plant variability important?

COVARIANCE

1. The assumptions that must be considered when using analysis of covariance are….
2. When analyzed as a dependent variable, the selected covariate needs to be

1. Covariance can be viewed as “a linear regression adjustment” within analysis of variance

(T or F)

1. 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 yield = \_\_\_\_\_\_\_\_\_\_\_\_;

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

run;

1. 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. How many years (locations, sites, etc.) of data are required to generate a meaningful regression equation for use in stability analysis?
3. Stability analysis conducted on the Magruder Plots showed that \_\_\_\_\_ applications appeared to be beneficial in \_\_\_\_\_\_\_\_\_\_ environments.
4. 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 (e.g., location) 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 CIMMYT stand for? Where is it located and who founded CIMMYT?

1. Draw what a “synergistic” interaction, and “antagonistic” interaction look like? (x axis should have N rates of 0, 40, 80 and 120 kg/ha and the y axis should be grain yield)
2. 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?

1. What “percent of the mean” difference do you need in your work to say a significant difference exists (more or less, and why) (dependent variable being analyzed is your choice)
2. Write a PROC CORR program to determine the relationship between yield and NDVI with population, disease, height, and BYDV (barley yellow dwarf virus).
3. In order to merge two data sets that have rep, trt, yield, and location as identifiers, show what the SAS code might look like to do so.
4. 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.
5. In order for SAS to understand that you have missing data, what must be entered within that cell?
6. The very first “PROC” procedure that you should run in any program is ?
7. What are the assumptions of analysis of variance?
8. When should “LSMEANS” be used to replace the normally computed “MEANS?” (2 answers)

1. \_\_\_\_\_\_\_\_ ensures that you will have an estimate of experimental error
2. \_\_\_\_\_\_\_\_ ensures that you will have an unbiased estimate of experimental error
3. When you have missing data, what sums of squares should be used?
4. LSD’s cannot be used when the treatment structure includes \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
5. If there isn’t a known “gradient” within a field trial, what experimental design is recommended?
6. How is SED (standard error of the difference between two equally replicated means) computed?
7. SED times \_\_\_\_\_\_\_ is generally what would be computed using what mean separation procedure?
8. \_\_\_\_\_\_\_\_\_\_ method is a broad term and that encumbers the reporting of survey statistics
9. \_\_\_\_\_\_\_\_\_\_ method is more specific and that includes formulating a hypothesis, putting together a treatment structure to test the hypothesis, collecting structured data, analyzing the data, and interpreting the results.
10. What kind of error is incurred if a scientist “excludes” data that does not conform to his/her hypotheses?
11. Good researchers aren’t necessarily characterized by being smart, but by…….
12. What is autocorrelation?
13. Below is a GLM for 2013 and 2014, treatments 1-14 from Experiment 502 in Lahoma Oklahoma.

The SAS System 10:14 Friday, April 24, 2015 48

The GLM Procedure

 Class Levels Values

 yr 2 2013 2014

 rep 4 1 2 3 4

 trt 14 1 2 3 4 5 6 7 8 9 10 11 12 13 14

Dependent Variable: mgha

 Sum of

 Source DF Squares Mean Square F Value Pr > F

 Model 33 17.84415509 0.54073197 7.42 <.0001

 Error 77 5.61150327 0.07287667

 Corrected Total 110 23.45565836

 R-Square Coeff Var Root MSE mgha Mean

 0.760761 12.33944 0.269957 2.187756

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

 yr 1 7.00818598 7.00818598 96.17 <.0001

 rep(yr) 6 0.90599232 0.15099872 2.07 0.0662

 trt 13 3.70091890 0.28468607 3.91 <.0001

 yr\*trt 13 6.22905789 0.47915830 6.57 <.0001

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

 yr 1 6.64232863 6.64232863 91.14 <.0001

 rep(yr) 6 0.71528736 0.11921456 1.64 0.1486

 trt 13 3.97101267 0.30546251 4.19 <.0001

 yr\*trt 13 6.22905789 0.47915830 6.57 <.0001

 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 6.64232863 6.64232863 55.72 0.0003



1. This figure illustrates treatment means and the SD (standard deviation) about each treatment mean for 2013 and 2014. What would be the yield difference needed to declare that an observed difference in treatment yields was significant? (For this example, you have to use residual error (accounting for main effects of yr, trt, and yr\*trt). (see 42e below)

When should the Type III sums of squares be used instead of Type I?

Was there a need to use Type III sums of squares in this case for these 2 years of data?

1. Should treatment means have been interpreted over years or by year?
2. What statistic did you use to make the decision in #a?
3. What statistic tells you that this was likely a wheat experiment and not a corn trial?
4. What is meant by REP(YR)?
5. What is REP(YR) used for?

Compute the SED for this experiment. (include the formula for SED)

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 \_\_\_\_\_\_\_ \* \_\_\_\_\_\_\_ = YP0\_\_\_\_\_\_/shape='pyramid';

**run**;

****

1. If I had a fourth variable, “variety” (in addition to the 3 reported) and where there were 2 different varieties evaluated, how could I look at this, on this same graph?

****

1. You have an experiment with 4 N Rates (0, 30, 60, 90 kg N/ha) and 2 varieties (Junco and Pavon). 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 \* variety
5. N rate quad \* variety
6. What is the formula for CV, or Coefficient of Variation?
7. Why are the CV’s from wheat experiments generally less than that found for corn trials evaluating the same treatment effects?
8. Is it possible to have a CV, %, greater than 100? (yes/no, why?)
9. For the following treatment structure, please fill in the blanks below for a completely random design (CRD) and a randomized complete block design (both with 4 replicates)

|  |  |  |
| --- | --- | --- |
| Treatment | P rate | Foliar/Preplant |
|   | kg/ha |  |
| 1 | 0 | -- |
| 2 | 20 | Foliar |
| 3 | 40 | Foliar |
| 4 | 0 | -- |
| 5 | 20 | Preplant |
| 6 | 40 | Preplant |
|  |  |  |
|  |  | **CRD** |  | **RCBD** |   |
|  |  | Source of variation | df | Source of variation | df |
|  |  | Total (4\*6)-1 | 23 | Total (4\*6)-1 | 23 |
|   |   |   |   | block | \_\_\_ |
|  |  | **treatment** | \_\_\_ | **treatment** | **\_\_\_** |
|  |  | error | \_\_\_ | error | \_\_\_ |

1. To obtain the correct coefficients for rates that are not \_\_\_\_\_\_\_\_\_\_\_ spaced, place the rates of 0, 40, 130, 220 in the program below.

proc iml;
dens={ };
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;

**BONUS:**

Dr. Borlaug’s first job in Mexico was with?

Japanese wheat cultivar that had reduced height and improved tillering \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ and that was used for initial crosses that led to SONORA 64 and Pitic62

The first crop that Dr. Borlaug worked on at the University of Minnesota under his professor Elvin Stakman was Linum usitatissimum. What is the common name ?

Once residing in Mexico, Dr. Borlaug’s true genius was seen in his moving wheat selections from where to where so as to complete wheat crosses, 2 cycles/year?

Why is genetic variability in plant breeding so important?

What is the definition of “environmental mean” when used for data analysis?

What are the two products coming from the animal industry that are global warming gasses?

**STUDENT Questions**

1. What is the growth stage in maize at which nitrogen deficiency is noticeable in Oklahoma?
2. Based on the modifications so far (OSU Hand Planter), what is its potential impact on agriculture and general life of farmers in the developing world
3. What modifications would you suggest to the current OSU hand planter.
4. From Melissa Golden’s work on wheat (rust), what is one item that the optimum growth window could be based on?
5. Question: what percent of California's rice production does the Sacramento Valley produce?
6. Provide a reason why significant efforts and money should be poured into the conversion of C3 rice to C4 rice?
7. (Dr. Zoca) What nutrient recently became a soil fertility and plant mineral nutrition problem in areas where it was not a limiting factor before?
8. What is the threshold temperature from the original INSEY (in-season-estimated-yield)?
9. As a self employed individual what three "roles" do you fill when making decisions?