**Final Exam: SOIL 5112
Friday, May 5, 2017
7:30 am**



1. Collecting NDVI data using the Green seeker sensor is now an easy task. How often should NDVI data be collected for any given trial (wheat or maize), over the entire season?

**As many times as you**

1. Would you expect CV’s to be higher in corn or wheat experiments? Why?

**Yes (corn) because corn does not produce tillers**

1. Using daily CV data coming from NDVI readings, what stage of maize growth was identified where increased variability over a very short period of time stood out?

**Tasseling stage**

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

**This is the level at which an increase in the variable X does not result in an increase in Y**

1. From Mullen et al (1999). The percent increase in atmospheric CO2 due to a worldwide decrease of 3, 2, and 1% in soil organic matter of arable land is estimated to be 20 mg kg-1,12.5 mg kg-1, and 5 mg kg-1, respectively. Decrease in soil organic matter accounted for
**6** to **25** of the 80 mg kg-1 (340-260) increase in atmospheric CO2 over the last 150 years.
2. What is the world atmospheric CO2 concentration today **405ppm**
3. Why is plant to plant variability important in both the developed and developing world?
**The resolution (plant to plant) is the same for both. And both can treat this variability, plant to plant (high tech VRT equipment, and OSU hand planter)**
4. What is the product of using “Replication” in field experiments?

**Ensures having an estimate of experimental error**

1. What is the product of utilizing “Randomization” in field experiments?

**Ensures have an unbiased estimate of experimental error**

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

Using Cate-Nelson **42**

Using a quadratic model **60**

Using a linear plateau **42**

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

Proc nlin data = one best = 3;
parms b0= \_\_0\_\_ to \_0.3\_\_\_ by 0.01 b1=\_\_0.01\_\_\_ to \_\_0.02\_\_\_ by \_0.01\_\_ njoint=\_\_30\_\_\_ to \_\_50\_\_\_ by \_1\_\_;

1. For the following equation coming from a linear plateau model,

Y = 1.2\*x, when x < 5,
Y = 15 when X >=5

What was the slope (b1) **1.2**

What was the plateau **15**

What was the joint? **5**

1. Why are “abstracts” so important (you can use either a research, teaching, or extension perspective)

**They are like a “hand shake”, an abstract tells exactly what the research is about and gives a brief description of the results.**

1. When is it appropriate to combine data coming from different years and/or locations?

**It is never appropriate to combine data**

1. What was the experimental design for Experiment 222 and Experiment 502, discussed in class?

**Randomized complet block design (RCBD)**

1. Formula for SED:

**SED = square root (2\*MSE/reps)**

1. Formula for LSD:

**LSD = 2\*SED**

1. Experimental Method should attempt to remove **sources of variation** other than the effects being evaluated within the treatment structure
2. Definition of an “independent variable”
**Variables you can control**
3. Definition of a “dependent variable”

**Variables you cannot control or what you want to measure.**

1. In the SAS statement below, fill in the blanks with the independent variables or dependent variables

(mgha = yield , Mg/ha, gn = grain nitrogen, %)
data one;
input rep trt mgha gn;
cards;
1 1 2.5 0.2
1 2 2.9 0.3
Proc **glm**;
model **Yield gn** = **rep trt**;
means **trt**;
run;

1. The ideal number of treatments for an agricultural field experiment with 3 replications is

A. 3

**B. 13**

C. 23

D. 33

1. If you have too many treatments included in your field experiment the following can happen;
2. **The number of degrees of freedom in the mean square error term will go up**
3. **Differences required to detect significance get smaller**
4. Nothing happens, SAS gets a break
5. What text book was discussed as being a source of science, chemistry, and comprehensive knowledge that students need to employ more frequently?
**Dictionary**
6. If the contrast coefficients for a linear comparison were equally spaced N rates, for which there were four rates, what would the contrast coefficients be? What about quadratic? (USE first page)

Contrast ‘n-rate-lin’ trt **-3 -1 1 3**;

Contrast ‘n-rate-quad’ trt **1 -1 -1 1**;

\*Using the orthogonal contrast coefficients on the board, provide the coefficients for a

1. Linear effect of N rate (5 rates, equally spaced)
2. Quadratic effect of N rate (5 rates, equally spaced)
3. Comparison of the bottom N rate (1) versus the highest N rate (5)
4. For the rep Nrate model below (one year, one location), analyzing yield, identify the error/problem in the SAS program

Proc glm;
class rep Nrate ;
model yield = rep Nrate **rep\*Nrate**;
means Nrate;
run;

1. Does Experiment 502 have a Factorial arrangement of treatments? Is there such a thing as a “Factorial Design?” **No, its factorial treatment structure**
2. Research station in England that started “modern day” long-term research?

**Rothamsted station**

1. Mean separation procedure recommended for use when missing data is present

**LSMean**

1. Can continuous ‘years’ be used instead of different ‘locations’ to add to the ‘site-year’ total?

**Yes because different years have different environmental conditions**

1. How was the CV (data collected over an entire season) used to decipher/determine the best time to “possibly” apply fertilizer? **When CV was high, its and indicator of poor stands and N rates should be decreased, when the CV was low, homogeneity of plant stands was good, and N rates could be higher**

35. List 3 causes of plant to plant variability

**Genetics, variation in emergence due to planting depth, soil compaction, soil texture**

1. OSU work from 46 transects around the world showed that regardless of yield level, plant-to-plant variability in corn grain yield can be expected and averaged more than **2765** kg/ha or **47** bu/ac over sites and years.
2. The Cate-Nelson procedure developed at North Carolina State University was the precursor to what statistical methodology that has been discussed in class? Do they come up with the same results?

**Linear plateau model, almost the same**

1. Definitions:

r- **Correlation coefficient**

 r2- **Coefficient of determination (Shows percentage of y explained by X)**

 R- **Coefficient of determination using multiple variables**

 SED - **standard error of the difference between equally replicated means) SED = square root (2\*MSE/reps)**

 CV - **coefficient of variation CV = (SD/mean) \*100**

Yield Potential - **Highest possible yield obtainable with ideal management conditions**

Yield Goal - **Yield per acre that you hope to grow. It’s the average yield for the past 3 to 5years + 20%**

1. Stability analysis when used in plant breeding attempts to identify those genetic-biotypes that perform well in all **environments**

COVARIANCE

1. The assumptions that must be considered when using analysis of covariance are….

**“Treatment” effect from your AOV, was not significant when the covariate was analyzed as a dependent variable**

1. When analyzed as a dependent variable, the selected covariate needs to be
**not significant**
2. 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 “pH” (pre plant soil pH)

Data one;

input rep trt yield pH;

cards;

1 1 30 42

1 2 35 40

Proc glm;

Class **rep trt**;

Model yield = **rep trt PH**;

LSMeans **trt**;

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. How many years (locations, sites, etc.) of data are required to generate a meaningful regression equation for use in stability analysis?

**As many as possible**

1. Stability analysis conducted on the Magruder Plots showed that **potassium** applications appeared to be beneficial in **water/ moisture stress** environments.

Why is genetic variability in plant breeding so important?

**The need for a wide range in genetics assures selection for different traits that may arise in the future**

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

**It’s the average of all treatments in a particular environment/year**

1. 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

Proc glm;

class rep nrate variety;

Model yield = rep nrate variety nrate\*variety

Means nrate variety nrate\*variety

Run;

1. SAS program if you analyze this as a rep-treatment model

Proc glm;

class rep trt;

model yield = rep trt;

means trt;

run;

1. If a treatment\*environment (e.g., location) interaction is significant what does it say about how treatment must be interpreted?

**It means that treatment response needs to be interpreted by environment**

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)

Synergistic Antagonistic

Y

 Y

 X X

1. Two trials: LMSE = 58000 SMSE = 24000, dfe (both trials) = 20

Compute the F statistic. \_**2.4\_\_\_\_\_**

Based on your knowledge of the table values, should these trials be combined?

**Absolutely not**

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)

**5% and above**

1. Write a PROC CORR program to determine the relationship between yield and NDVI with population, disease, height, and BYDV (barley yellow dwarf virus).

proc corr;

var population disease height bydv yield;

run;

1. In order for SAS to understand that you have missing data, what must be entered within that cell?

**Period**

1. The very first “PROC” procedure that you should run in any program is?

 **Proc print**

1. When should “LSMEANS” be used to replace the normally computed “MEANS?”

**When you have missing data**

1. When you have missing data, what sums of squares should be used?

 **Type III**

1. LSD’s cannot be used when the treatment structure includes **sequential rates/ numeric**
2. If there isn’t a known “gradient” within a field trial, what experimental design is recommended?

 **CRD**

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 **g3d**;

Scatter **Year** \* **RION** = YP0\_\_\_\_\_\_/shape='pyramid';

run;



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 (data first has to be sorted by variety)

Data one;

Proc sort; by variety;

Proc glm; by variety;

contrast ‘nrate\_linear’ trt -3 -1 1 3;

1. N rate quadratic

contrast ‘nrate quad’ trt 1 -1 -1 1;

1. N rate linear \* variety

contrast ‘nlin\_variety’ nrate\*variety -3 -1 1 3 3 1 -1 -3;

N rate quad \* variety

contrast ‘nquad\_variety’ nrate\*variety 1 -1 -1 1 -1 1 1 -1;

1. Why are the CV’s from wheat experiments generally less than that found for corn trials evaluating the same treatment effects? Higher plant populations, higher tillering
2. Is it possible to have a CV, %, greater than 100? (yes/no, why?)
**Abosolutely yes. Think about the formula. If you have a low mean and high variability, CV’s can be 1000 %. Pay attention to this as you analyze data**
3. 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, kg/ha | Foliar/Preplant |
|   |  |  |
| 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 | **3** |
|  |  | treatment | **5** | treatment | **5** |
|  |  | error | **18** | error | **15** |

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

proc iml;
dens={ **0 40 130 220** };
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:**

1. Dr. Norman Borlaug, recipient of the 1970 Nobel Peace Prize, stated the following

*…It is my belief that all who are born into the world have the moral right to the basic ingredients for a decent, human life*." How do we know he really meant it and that it wasn’t just a bunch of words?

His work proved as much

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

**Rockefeller foundation**

1. Japanese wheat cultivar that had reduced height and improved tillering **Norin 10** and that was used for initial crosses that led to SONORA 64 and Pitic62
2. 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? **Flax**
3. 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? What was the name used for this? **Chapingo near Toluca and Yaqui valley**

**Shuttle breeding**

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

**Methane and carbon dioxide**

1. What does the acronym CIMMYT stand for? (either English or Spanish versions will work).

**International Maize and Wheat Improvement Center**

**Centro Internacional de Mejoramiento de Maiz y Trigo**