

Introduction: The following data are from Cochran and Cox's famous experimental design book. This is the same Gertrude Cox after whom Cox Hall is named, who founded the departments of statistics at NCSU and Chapel Hill and who was instrumental in getting the Research Triangle Institute started. The data are old, from a 1943 experiment on corn hybrids run here at the North Carolina Agricultural Experiment Station. It is estimated that commercial companies like Monsanto test at least a hundred thousand lines a year. You might be interested in investigating how yields have increased over time. Of course in 1943 few people had access to mechanical calculation methods so most analyses would have been done with just a pencil and paper. The blocks are different locations around North Carolina and the yields are in pounds per plot so they cannot well be compared to today's yield data.

1. Input the data and compute and list the block and treatment means. Here are the data
 Each row is a block. In each row the first number is the block number, followed by 4 pairs of line and yield numbers. The "line" is our treatment and in modern trials is referred to as "entry". If you're not familiar with crop terminology you could think of these as 13 cultivars or varieties. We will treat the lines as fixed effects here but often in crops entries are viewed as random effects for the purpose of computing heritability coefficients*.

1	3	25.3	6	19.9	9	29.0	11	24.6
2	3	23.0	4	19.8	8	33.3	12	22.7
3	10	16.2	11	19.3	12	31.7	13	26.6
4	2	27.3	5	27.0	8	35.6	11	17.4
5	7	23.4	8	30.5	9	30.8	10	32.4
6	4	30.6	5	32.4	6	27.2	10	32.8
7	1	34.7	5	31.1	9	25.7	12	30.5
8	3	34.4	5	32.4	7	33.3	13	36.9
9	1	38.2	2	32.9	3	37.3	10	31.3
10	2	28.7	4	30.7	9	26.9	13	35.3
11	1	36.6	4	31.1	7	31.1	11	28.4
12	1	31.8	6	33.7	8	27.8	13	41.1
13	2	30.3	6	31.5	7	39.3	12	26.7

2. Here is some code for plotting. This lets you visualize the design and the data.

```
proc g3d; scatter line*block=Yield/shape="prism" zmin=0;
title "Corn Yields";
Title2 "from Cochran and Cox";
Title3 "1943 NC field trial";
run;
proc means data=corn; var yield; class block;
output out=out1 mean=mnyield;
proc gplot; plot mnyield*block;
symbol1 v=dot i=needle;
run;
```

In the graph window, click the file menu button (at the top) and then choose export graph in the popup menu. Include the 3D graph in your report. Also modify the code above to produce and plot LINE means (treatment means) and plot against the line numbers 1 through 13. Add the treatment mean plot to your report.

3. Compute the ANOVA table treating the blocks as fixed.
- (a) Compute the LINE LSMEANS and include them with the ANOVA table in your report.
 - (b) Are the LSMEANS the same as the ordinary means? Which ones do you think would be better for your comparisons?
 - (c) Comparing to our book or class notes, is this strictly an intrablock (within blocks) analysis or have we recovered the interblock information?

4. The blocks are randomly selected locations around North Carolina.

- (a) How many locations were used?
- (b) This is a balanced incomplete block design. Knowing that, what are b __, k __, r __, t __ and λ __ ? (numbers!)
- (c) Why do you think the experimenters used several randomly selected locations around the state? It would have been a lot easier to do locations near Raleigh. Do location considerations justify what we did in part 3 or do they suggest a preferable analysis?

5. Repeat the analysis in part 3 but this time treat the locations as random (LINES are still fixed effects for our current). Be sure to include the treatment F test and the LSMEANS in your output.

- (a) Summarize the differences in the F tests for LINES between the intrablock analysis and the analysis that recovers interblock information.
- (b) List three columns of treatment (LINE) means: Line numbers in the first column then ordinary means in the second column, LSMEANS from the intra block analysis in the third and LSMEANS from the analysis that recovered the interblock information in the fourth.

Optional (not graded) Using $X=1, 2,$ and 3 to denote the type of mean and Y to denote the mean, plot Y by X , connecting with lines the 3 points that go with each treatment. Those lines will give a nice visual on the effect of these different analyses.

* Note: I have talked about heritability on a per unit basis, per plot for example. In crops heritabilities are more often computed on an entry mean basis so that, in a simple case with r plots per family (or line) $\sigma_F^2 / (\sigma_F^2 + \sigma^2)$ becomes $\sigma_F^2 / (\sigma_F^2 + \sigma^2 / r)$.