

Week 8--IES 612-STA 4-573-STA 4-576.doc

IES 612/STA 4-573/STA 4-576 Winter 2009

Design Issues (Ch 14 text)

DESIGN OF EXPERIMENT (or **DESIGNED EXPERIMENT**) = "process of establishing a framework through which the comparison of treatments or groups can be made in terms of recorded response" (OL 14, p. 829)

"Road Map" of:

1. Definition of the "experimental unit," ie what object, element, will be investigated.
2. What data (ie variables) will be measured on each EU.
3. How many and how will the EU's be collected.
4. How will the EU's be assigned to treatments
5. Which statistical techniques will be used to analyze the data

Note: balance between the control of condition and depiction of reality must be maintained - "ecological validity" - when can you use the lab vs. when must you work in the field?

STUDY TYPES

1. **Observational** - factors not manipulated, sampling from populations where factors (trts) already present and want to compare populations with respect to some response (e.g. samples, polls, surveys, epi. studies). Results can only be deemed "relational"!
2. **Experimental** - randomly assign subjects to treatment conditions and observe the response of interest. Results can be deemed to be "causal," that is the treatment caused any effect observed.

Examples

- Earth worm presence as a function of soil characteristics
- Vole fidelity as a function of certain allele characteristics
- Behavior of various organisms (fish, mammals, etc) in the presence of varying concentrations of chemicals

- Community study where elderly receive care using either a consumer-directed system or a traditional case manager system.
- Perch growth in the presence of gobi and/or zebra mussels
- Species presence as a function of stream characteristics
- Web weight as a function of temperature and species of spider
- Contaminant level in effluent as a function of temperature and pressure

Research Plan ingredients (OL 14, p. 831)

1. Objectives
2. Study Factors (ie "Treatments")
3. Extraneous Factors (ie Block Factors)
4. Response
5. Randomization method
6. Protocol for recording responses
7. Replications needed
8. Resources

Principles to consider when designing an experiment

1. Randomization - create groups as similar as possible prior to an experiment
2. Control - comparison group (concurrently conducted with the study)
3. Replication - how many experimental units? sensitivity to detect important differences (power of analysis).

How might you do a RANDOMIZATION?

Example: The Meat Packaging method and amount of bacteria.

Suppose we are back at the "start" of this experiment. We have already decided that:

1. The EU is a certain size or quantity of meat (eg a pound of filet!)
2. The "treatments" are the four packaging conditions: CO₂, plastic, mixed, and vacuum
3. There will be three pieces of meat for each packaging condition

The next step will be to actually randomly assign the twelve pieces of meat to the conditions.

To randomly assign 12 experimental units (here pieces of meat) to one of four packaging conditions with 3 units assigned to each condition, here's one way:

- Step 1: Assign a unique number (label) to each experimental unit - say "1" to "12" ($=n_T$)
- Step 2: Randomly permute the labels (put slips of paper with the labels 1 - 12) in a hat and randomly select slips from the hat
- Step 3: Assign the units corresponding to the first $n_1=3$ permuted labels to group 1, the next $n_2=3$ permuted labels to group 2, etc.

Let's make this concrete by having R and/or SAS do the actually do the "randomization." We assume the labels are just the numbers 1 to 12. We need R or SAS to simply randomly arrange these 12 numbers.

Using R

```
> labels = c(1:12) * note that labels = 1:12 also works!  
> labels  
[1] 1 2 3 4 5 6 7 8 9 10 11 12  
> sample(labels, size=12, replace=FALSE)  
[1] 2 6 11 3 12 1 7 4 10 9 8 5
```

We assign meat pieces 2, 6, and 11 to the CO2 package condition; 3, 12, and 1 to plastic, etc.

Using SAS

```
proc plan;  
title "generate randomization/allocation scheme for 12 steaks";  
factors meat=12;  
run;
```

generate randomization/allocation scheme for 12 steaks

2

The PLAN Procedure

Factor	Select	Levels	Order
meat	12	12	Random

-----meat-----

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

We assign meat pieces 3, 1, and 10 to the CO2 package condition; 2, 5, and 11 to plastic, etc.

As an aside, you can also use this to generate a **random sample of n items** from a population with **N items**. Suppose we have 40 pieces of meat and want to randomly select 12 for our study.

Using R

```
> labels = 1:40
> sample(labels, size=12) * note that replace = FALSE is the default
[1] 35  6 12 26 24 37 32  1 27 39 16 10
```

Our random sample of 12 meat pieces are those labeled: 35 6 12 26 24 37 32 1 27 39 16 10

Using SAS

```
proc plan;
title "generate randomization/allocation scheme for 12 steaks";
  factors meat=12;
run;
```

NOTE: At the start of processing, random number seed=581671001.

```
proc plan;
title "generate random sample of 10 from 40 in sampling frame";
  factors n=10 of 40;
```

generate random sample of 10 from 40 in sampling frame

3

The PLAN Procedure

Factor	Select	Levels	Order
n	10	40	Random

-----n-----

2 20 13 25 11 23 30 31 4 29

The SAS log noted that

```
10  proc plan;
NOTE: At the start of processing, random number seed=364282001.
11  title "generate random sample of 10 from 40 in sampling frame";
12  factors n=10 of 40;
13  run;
```

(in case you wanted to replicate this stream of random numbers).

Lastly, you CAN use random number tables or other devices to do a randomization.

Why have control groups? Does control group = untreated group?

Guaranteed treatment for the common cold - I call it "chicken soup" - You will be cured after 3 days or your money back! Justification for guarantee? I did a study where I gave soup to 25 people with colds and they all felt better when I asked them 3 days later. Reaction?

LOTS OF TYPES OF CONTROL GROUPS:

1. Untreated
2. Placebo
3. Sham (often in surgery or neuroanatomy studies)
4. Standard treatment (may not be ethical to have an untreated group)
5. Vehicle (sometimes you have to give a treatment in some medium)
6. Historical (can be problematic)

WHAT DOES "BLINDING" MEAN IN AN EXPERIMENT?

Single-blind study (subject doesn't know treatment)

Double-blind study (subject & physician/experimenter don't know treatment)

Triple-blind study (subject, experimenter & analyst don't know treatment)

[code is broken after completion of the study]

Absolutely blind study (like a Triple-blind study with the "code" lost!)

Treatment Structure

Factor = manipulation/population of interest (analogous to independent variable in regression)

Level = unique value of factor

Treatment = (single factor study) level of factor

Treatment = (multiple factor study) unique combination of factor levels

SINGLE FACTOR

Factor = packaging condition with Levels = CO₂, plastic, mixed, vacuum

Packaging Condition (Factor)			
CO ₂	plastic	mixed	vacuum
1	2	3	4
Treatment			

MULTIPLE FACTORS

Factor A = gobi (levels = present/absent)

Factor B = Zebra mussel (levels = present/absent)

		B:Zebra mussel	
		present	absent
A:gobi	present	1	2
	absent	3	4

	Treatment			
	1	2	3	4
A: gobi	present	present	absent	absent
B: Zebra mussel	present	absent	present	absent

Experimental Units (EU) and Measurement Unit (MU)

EXPERIMENTAL UNITS (EU) = entity to which treatment is randomly assigned or is randomly sampled from one of the "treatment" population [OL 14, p. 833]

MEASUREMENT UNIT (MU) = entity on which a measurement is taken

Example: Meat packaging study: EU = MU = piece of meat

Example: Teratology study: EU=dam/litter; MU=pup

NOTE: Sometimes, MU called "pseudoreplicate" if MU doesn't equal the EU (in ecology literature)

Experimental Error

EXPERIMENTAL ERROR = variation among EUs assigned to the same treatment and observed under the "same" experimental conditions

SOURCES OF EXPERIMENTAL ERROR?

1. Natural differences in EUs
2. Variation in devices/people that record the MUs (thermometer/student repeatability?)
3. Variation in the treatment conditions (doses/concentrations EXACTLY the same?)
4. Variation in extraneous/uncontrolled factors

Controlling Experimental Error? OL 14.4

1. Procedures for conducting study standardized ("local control" of Kuehl) - train data collectors/lab technicians, standard protocol for recording data and conducting experiment, etc.
2. Choice of EU/MU (e.g. same age/size class, same level of disability, etc.) - randomly select from population and then randomly assign treatments, select EUs that are similar (although if too similar that generalizability may be questioned) - e.g. treatment for a specific level of Alzheimer's impairment
3. Measurement procedure
4. Blocking (a "design" control - before conducting study) - EUs placed in groups
5. Covariates (an "analysis" control - possible after study conducted)

Blocking Designs (and Blocking Variables)

A **BLOCKING DESIGN** imposes a **CONSTRAINT** on the randomization of experimental units to treatments

- EUs are placed in groups (**BLOCK**) that are similar with respect to some important characteristic that may affect the response.
- EUs are randomly assigned to treatments **WITHIN** each group/block

A **BLOCKING VARIABLE** is a variable that accounts for another source of variation among experimental units and usually occurs in observational studies in which another source of variation is accounted for in the analysis/model

- EUs in a **BLOCK** are similar with respect to some characteristic that may affect the response

NOTE: **BLOCKS** or **BLOCKING VARIABLES** are **not usually of interest** in the context of the experiment, but is included to **ACCOUNT FOR THE UNEXPLAINED VARIATION!**

SOME CRITERIA FOR DETERMINING BLOCKS/BLOCK VARIABLE (OL 14, P. 839)

- i. physical characteristic (e.g. age, weight, size class, sex, health, education)
- ii. related units (e.g. twins, animals from the same litter)
- iii. spatial location (e.g. neighboring plots of land, oven, table, rack)
- iv. time (e.g. day of week, time of day)
- v. person conducting study (e.g. technician, operator, rater)
- vi. same unit (e.g. same EU "gets" ALL of treatments)

Examples

- Temp in Great Miami at different sites example from STA 671

1 st Site	29.02	28.72	29.10	28.09
2 nd Site	29.57	30.71	31.00	29.86
3 rd Site	41.77	41.99	41.82	37.30
4 th Site	38.27	38.01	37.85	35.61
6 th Site	32.74	33.92	34.21	33.20

- Typing Scores and Background Music example from HW (OL Exerc 15.5)

No Music	20	17	24	20	22	25	18
Hard Rock	20	18	23	18	21	22	19
Classical	24	20	27	22	24	28	16

Using a covariate (aka Analysis of Covariance ANOCVA)

A **COVARIATE** is a numeric (!) variable related to the response variable (you might have blocked on this covariate as an alternative to controlling the covariate, for example, age of the EU could be a covariate or a block)

DEPTH could be used as a covariate in an analysis comparing DO of two lakes.

ANOCVA is essentially a REGRESSION problem where the **TREATMENT** could be represented by one or more indicator variables and the **COVARIATE** is some numeric variable.

Example: Tahoe DO Study

In the Lake Tahoe region of Nevada a study was done investigating whether the dissolved oxygen was different at "Tahoe Keys" (a high human activity site) from "Eagle Lake" (low human activity site)? Note that we will use the log of DO for reasons to be seen later.

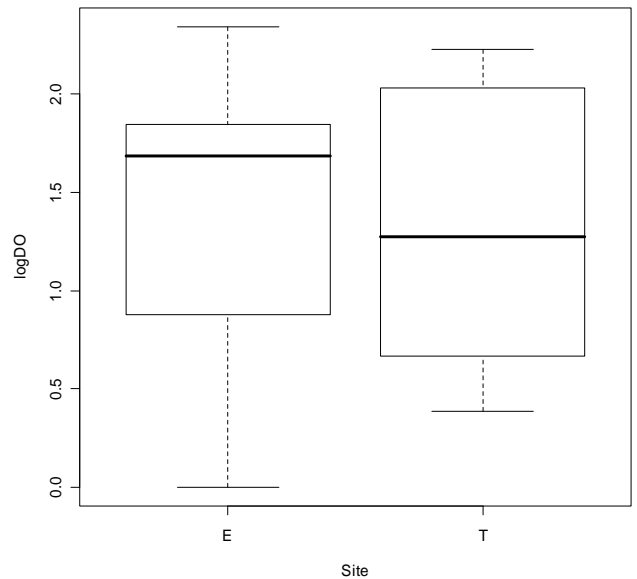
Depth	0	1	2	3	4	5	6	11	16	21	26	0	1	2	3	4	5
DO	10.4	7.5	6.6	6.1	5.7	5.4	5.1	2.9	2	1.2	1	9.26	7.63	5.05	2.52	1.95	1.47
Site	E	E	E	E	E	E	E	E	E	E	E	T	T	T	T	T	T

Our model is: $\log DO_{ij} = \mu + \alpha_i + \varepsilon_{ij}$, where α_i is the "effect" of the i^{th} site and $\varepsilon_{ij} = \text{NID}(0, \sigma^2)$.

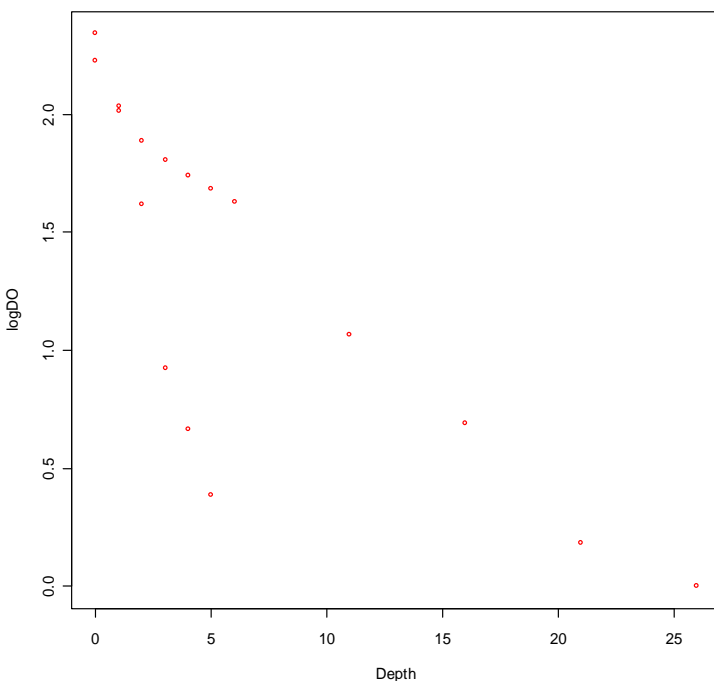
Note that the mean logDO does not appear to be different at the two sites because of:

Notice we have "ignored" the DEPTH variable (and hence effect) in our model!

But notice what we see if we plot logDO versus Depth.



```
> scatterplot(logDO~Depth, reg.line=FALSE, smooth=FALSE, labels=FALSE,
boxplots=FALSE, span=0.5, data=DepthDO) * from Rcmdr
```



We see ANOTHER source of variation! DEPTH!!!! As depth increases, logDO goes down.

Could this be affecting the comparison of the mean logDO at the two sites?

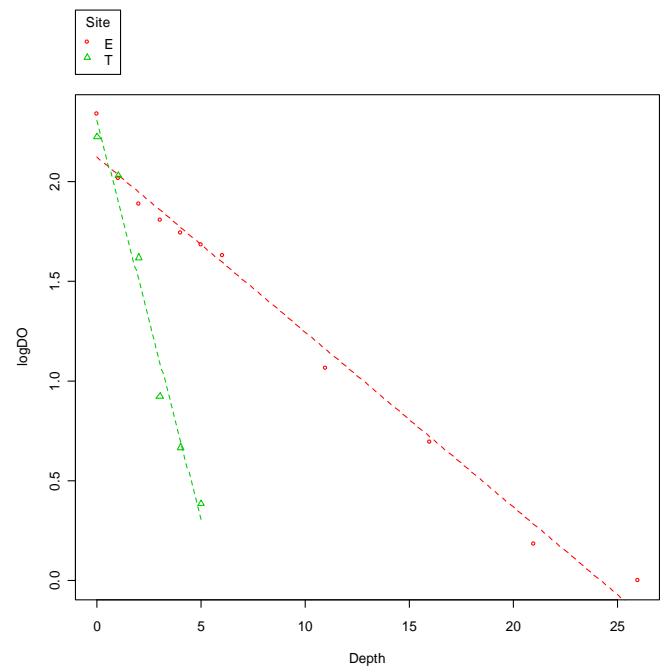
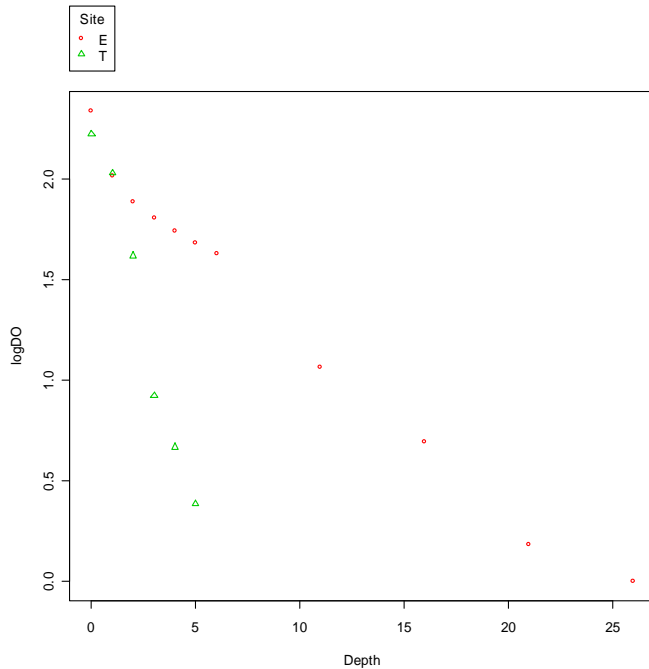
Let's "add" this effect to our model:

What ASSUMPTION does this model make about the effect of Depth on DO?

How could we check this assumption? Visually?

Let's identify which points in this graph correspond to the different sites.

```
> scatterplot(logDO~Depth | Site, reg.line=FALSE, smooth=FALSE, labels=FALSE,
boxplots=FALSE, span=0.5, by.groups=FALSE, data=DepthDO) * from Rcmdr
```



And now with the least squares lines for each group separately.

```
> scatterplot(logDO~Depth | Site, reg.line=lm, smooth=FALSE, labels=FALSE,
boxplots=FALSE, span=0.5, by.groups=TRUE, data=DepthDO) * from Rcmdr
```

We initially included Depth in our model as:

$$\log DO_{ij} = \mu + \alpha_i + \beta * \text{Depth}_{ij} + \varepsilon_{ij}, \text{ where } \varepsilon_{ij} = \text{NID}(0, \sigma^2)$$

but now what do you conclude?

What's the remedy?

ANALYSIS OF COVARIANCE (ANOCVA) MODELS

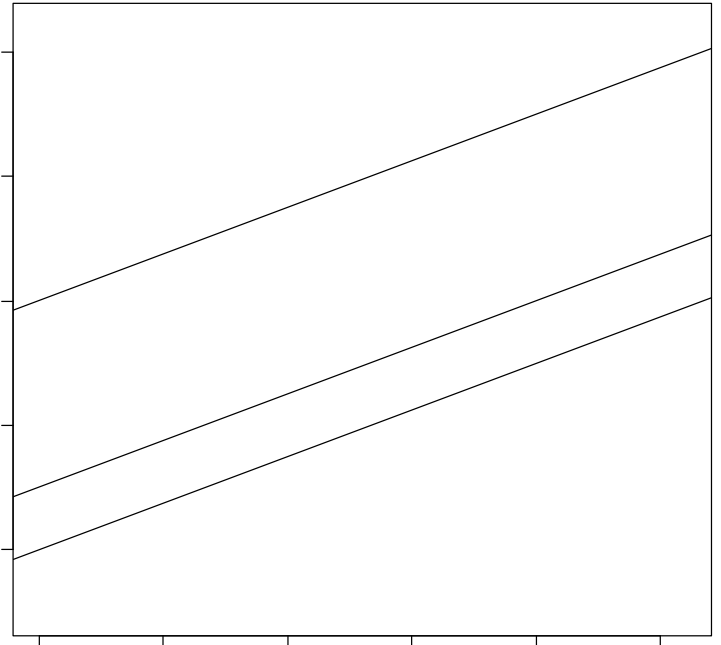
Usual ANOCVA Model—The Equal Slope Version

The most common ANOCVA model is one that ASSUMES equal "slopes" as indicated in the following picture.

What would the model be for this case?

Some implications of this model:

1. The **slope** of these lines which is the **effect of the Covariate**, the X variable, is the same for all Treatments.
2. The effect of the Treatment is the vertical differences in the lines and is reflected in the different **intercepts**.



Hypotheses of Interest and How would we test these Hypotheses?

1.

2.

CAUTION: Most statistically uneducated will assume that this is the model you use if say an ANOCVA was done!

THE ANOCVA Model—The UNEQUAL Slope Version

To those in the "know," the first question one SHOULD ask is:

"IS THE EFFECT OF THE COVARIATE THE SAME FOR ALL OF THE TREATMENTS?"

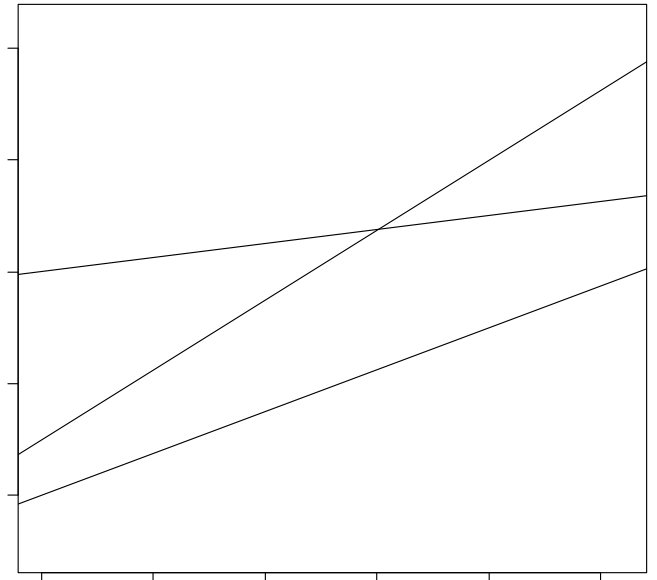
or put another way

"ARE THE SLOPES OF THE LINES THE SAME?"

In fact a more general version of the Equal Slope ANOCVA model is one in which the slopes differ. For example, the following is an illustration:

What are the implications?

1. Do the Treatments differ?



2. Does the covariate have the same effect for each of the treatments? What does this mean?

3. How do we investigate the "Treatment Effect?"

What would the model be for this *General ANOCVA* case?

What hypotheses could we test in this model *AND* what do they imply?

Null Hypothesis	What is investigating?

Could we use a *Regression Model* to the same end?

What would this model look like?

What would the hypotheses of interest be?

Notes and Comments

1. If you fit the *Unequal Slope ANOCVA* model *AND FAIL TO REJECT* the slopes are different, then *YOU MUST FIT* the *Equal Slope ANOCVA* model in order to obtain a statistically valid test of the *Treatment effect!!!!*
2. *Could/Should* you do an *ANCOVA* if there was a *nonlinear* relationship with the *covariate*?

Example: Tahoe logDO Study again

Using R

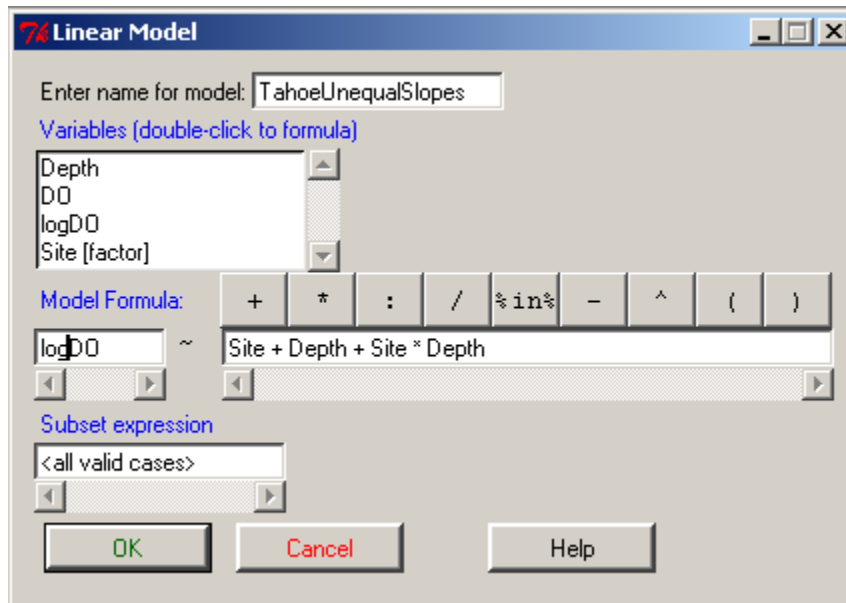
Here are the questions we need addressed:

1. Is the effect of Depth the same at the two sites?
2. If the effect of Depth is the same at the two sites, how does the mean DO compare at the two sites?

Using RCommander, with *DepthDO* the active data set:

```
> DepthDO
  Depth  DO Site  logDO
1     0 10.40  E 2.3418058
2     1  7.50  E 2.0149030
3     2  6.60  E 1.8870696
...
12    0  9.26  T 2.2257040
13    1  7.63  T 2.0320878
14    2  5.05  T 1.6193882
15    3  2.52  T 0.9242589
16    4  1.95  T 0.6678294
17    5  1.47  T 0.3852624
```

Statistics > Fit Models > Linear model ...



Our model: $Y_{ij} = \mu + \alpha_i + \beta X_{ij} + \beta_i X_{ij} + \varepsilon_{ij}$ are $NID(0, \sigma^2)$ or

$$\log DO_{ij} = \mu + SITE_i + \beta Depth_{ij} + \beta_i Depth_{ij} + \varepsilon_{ij} \quad \text{are } NID(0, \sigma^2)$$

Site	Model
Eagle Lake	$\log\text{DO}_{Ej} = \mu + \text{SITE}_E + \beta\text{Depth}_{Ej} + \beta_E\text{Depth}_{Ej} + \varepsilon_{Ej} = \mu + \text{SITE}_E + (\beta + \beta_E)\text{Depth}_{Ej} + \varepsilon_{Ej}$ $= \beta_{0E} + (\beta + \beta_E)\text{Depth}_{Ej} + \varepsilon_{Ej}$
Tahoe Keys	$\log\text{DO}_{Tj} = \mu + \text{SITE}_T + \beta\text{Depth}_{Tj} + \beta_T\text{Depth}_{Tj} + \varepsilon_{Tj} = \mu + \text{SITE}_T + (\beta + \beta_T)\text{Depth}_{Tj} + \varepsilon_{Tj}$ $= \beta_{0T} + (\beta + \beta_T)\text{Depth}_{Tj} + \varepsilon_{Tj}$

But note that we have:

1. Only two intercepts, β_{0E} and β_{0T} , but our model has three parameters associated with these two parameters, μ , SITE_E , and SITE_T .
2. Likewise, there are just two slopes, but our model uses three parameters to model these two slopes, β , β_E , and β_T .

What do stat packages, such as R and SAS, do in this case?

What does R do?

AND so how do the models above change AND how would you interpret the resulting parameters?

- $\text{SITE}_E =$
- $\mu =$
- $\text{SITE}_T =$
- $\beta_E =$
- $\beta =$
- $\beta_T =$

Finally, what hypothesis would be used to test whether the slopes are equal (ie the covariate, Depth, has the same effect on logDO at the two sites)?

$H_0:$ versus $H_A:$

```
> TahoeUnequalSlopes <- lm(logDO ~ Site + Depth + Site * Depth, data=DepthDO)
> summary(TahoeUnequalSlopes)
```

Call:

```
lm(formula = logDO ~ Site + Depth + Site * Depth, data = DepthDO)
```

Residuals:

```
      Min       1Q   Median       3Q      Max
-0.18497 -0.06209 -0.03007  0.07547  0.21751
```

Coefficients:

```
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    2.124292   0.050584  41.996 2.86e-15 ***
Site[T.T]       0.184090   0.098337   1.872  0.0839 .
Depth          -0.087567   0.004214 -20.780 2.34e-11 ***
Site[T.T]:Depth -0.312151   0.028170 -11.081 5.40e-08 ***
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Residual standard error: 0.1165 on 13 degrees of freedom

Multiple R-squared: 0.9801, Adjusted R-squared: 0.9755

F-statistic: 212.9 on 3 and 13 DF, p-value: 2.680e-11

What about the fit of the model?

What are the estimates of the model that we fit and what do correspond to?

- estimate of $\mu =$
- $SITE_T =$
- $\beta =$
- $\beta_T =$

So what do we conclude about:

$H_0: \beta_T = 0$ [Depth has same effect at both sites] versus $H_A: \beta_T \neq 0$ [Depth effect differs]

Conclusion:

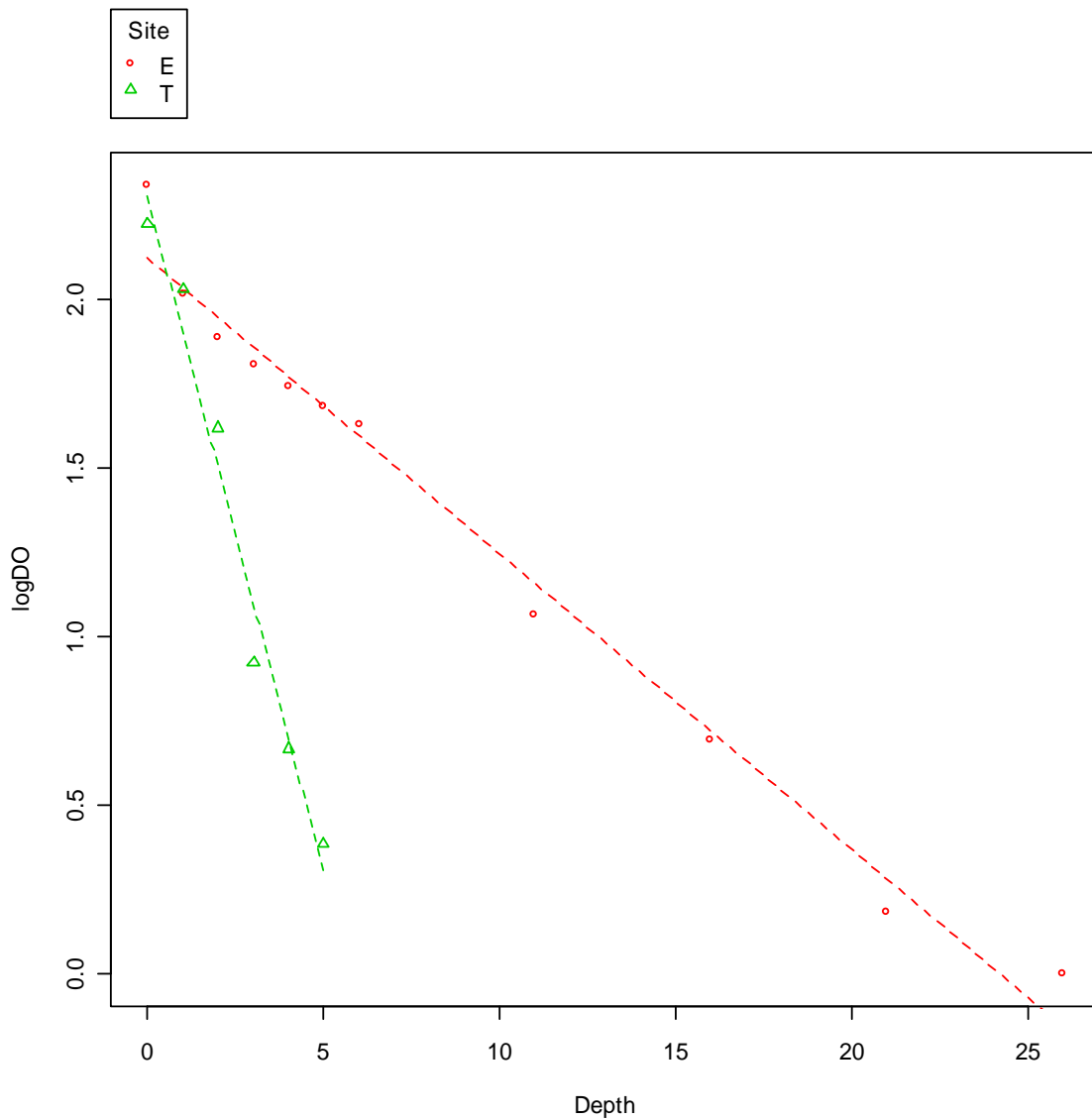
Interpretation:

What else should we do before moving on?

What do we conclude?

Would a picture help?

```
> scatterplot(logDO~Depth | Site, reg.line=lm, smooth=FALSE, labels=FALSE,  
boxplots=FALSE, span=0.5, by.groups=TRUE, data=DepthDO) *Rcmdr script
```



How would you summarize the logDO at the two sites?

So, evidence that the relationship between log(DO) and depth differs between Tahoe Keys and Eagle Lake. Additional analyses would likely focus on slope differences versus intercept differences.

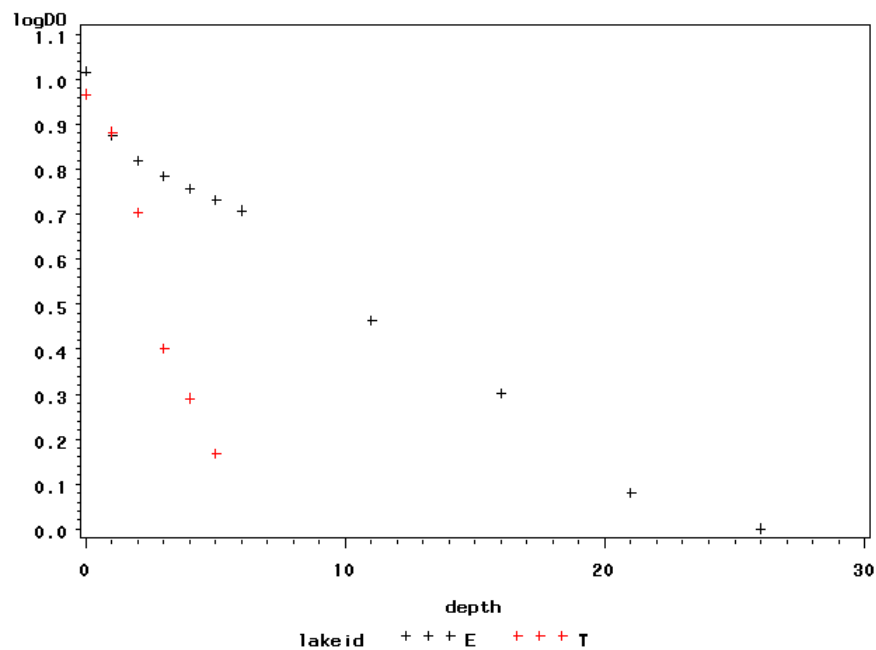
```

data lake;
  input obs depth do lakeid $;
  logDO = log10(do);
  datalines;
1      0 10.40      E
2      1  7.50      E
3      2  6.60      E
4      3  6.10      E
5      4  5.70      E
6      5  5.40      E
7      6  5.10      E
8     11  2.90      E
9     16  2.00      E
10    21  1.20      E
11    26  1.00      E
12     0  9.26      T
13     1  7.63      T
14     2  5.05      T
15     3  2.52      T
16     4  1.95      T
17     5  1.47      T
;

ods html;
proc gplot data=lake;
  plot logDO*depth=lakeid;
  run;

proc glm;
  class lakeid;
  model logDo = depth lakeid depth*lakeid;
  run;

```



The GLM Procedure

Dependent Variable: logDO

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	1.63564934	0.54521645	212.92	<.0001
Error	13	0.03328897	0.00256069		
Corrected Total	16	1.66893830			

R-Square	Coeff Var	Root MSE	logDO Mean
0.980054	8.648742	0.050603	0.585094

Source	DF	Type I SS	Mean Square	F Value	Pr > F
depth	1	1.06920821	1.06920821	417.55	<.0001
lakeid	1	0.25202396	0.25202396	98.42	<.0001
depth*lakeid	1	0.31441717	0.31441717	122.79	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
depth	1	0.76620168	0.76620168	299.22	<.0001
lakeid	1	0.00897392	0.00897392	3.50	0.0839
depth*lakeid	1	0.31441717	0.31441717	122.79	<.0001

Parameter	Estimate	Standard Error	t Value	Pr > t
Intercept	1.002517675 B	0.03662395	27.37	<.0001
depth	-0.173595111 B	0.01209649	-14.35	<.0001
lakeid E	-0.079949392 B	0.04270734	-1.87	0.0839
lakeid T	0.000000000 B	.	.	.
depth*lakeid E	0.135565272 B	0.01223415	11.08	<.0001
depth*lakeid T	0.000000000 B	.	.	.

NOTE: The X'X matrix has been found to be singular, and a generalized inverse was used to solve the normal equations. Terms whose estimates are followed by the letter 'B' are not uniquely estimable.

An alternative analysis using PROC REG ...

```

data lake;
  input obs depth do lakeid $ @@;
  iTK = (lakeid="T"); * defining the indicator variable;
  logDO = log10(do);
  iTK_depth = iTK*depth;
  datalines;
1      0 10.40      E 2      1 7.50      E 3      2 6.60      E
4      3 6.10      E 5      4 5.70      E 6      5 5.40      E
7      6 5.10      E 8      11 2.90      E 9      16 2.00      E
10     21 1.20      E 11     26 1.00      E 12     0 9.26      T
13     1 7.63      T 14     2 5.05      T 15     3 2.52      T
16     4 1.95      T 17     5 1.47      T
;
proc reg data=lake;
  model logDO = depth iTK iTK_depth;
  test iTK=0, iTK_depth=0;
run;

```

Sample Size Determination: How many observations should be used?

Warning: don't confuse the determination of a replicate with the replication of an entire experiment. This is the cause of much confusion.

Sample Size based on Margin of Error (OL 14, p. 831)

Goal: A specified margin of error of "E" for estimating μ_i assume equal replication $r = n_1 = \dots = n_t$

The sample size required to be $100(1-\alpha)\%$ confident that the estimator is within "E" units of the treatment mean μ_i is:

$$r = \frac{(z_{\alpha/2})^2 \hat{\sigma}^2}{E^2}$$

where α (specified by experimenter) & E = desired accuracy (specified by experimenter).

Estimated standard deviation?

Sources: Pilot study; Similar past experiments (literature); range/4?

Example: Meat Packaging study (revisited)

Sample size required to estimate the mean log(bacterial count) with a margin of error = 0.20 and confidence level 95%. Use the MSE = 0.12 from current experiment as an estimate of the variance.

$$r = \frac{(z_{\alpha/2})^2 \hat{\sigma}^2}{E^2} = \frac{(1.96)^2(0.12)}{(0.20)^2} = 11.52 \sim 12$$

Specify Power to Detect Specified Difference Among Means (OL p. 846)

$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_t$ versus $H_a: \mu_i \neq \mu_j$ [at least two population means differ]

Need to specify

1. α , significance level
2. Power = $1 - \beta$
3. Variance σ^2 estimate
4. assume equal replication $r = n_1 = n_2 = \dots = n_t$
5. D = pairwise difference that represents an important difference = $|\mu_i - \mu_j|$ or $\sum (\mu_i - \mu_j)^2$ which can be used to determine λ as indicated to the right \Rightarrow

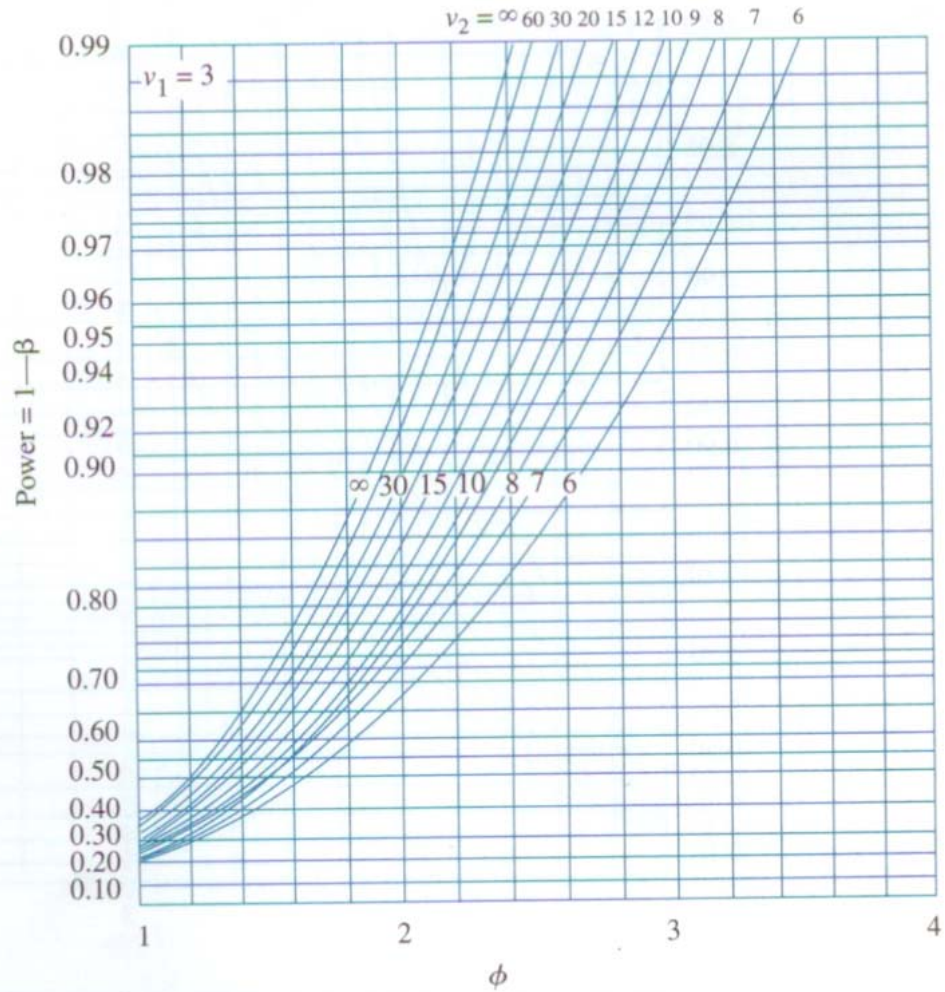
$$\lambda = \frac{r D^2}{2 \sigma^2}$$

or

$$\lambda = \frac{r \sum_{i=q}^t (\mu_i - \mu_j)^2}{\sigma^2}$$

Table 14 in the Appendix has power calculations indexed by $\phi = \sqrt{\lambda/t}$

Power of the analysis
of variance test
($\alpha = .05, t = 4$)



Example: Meat Packaging study (revisited)

Sample size required to detect a "0.5" difference in log(bacterial counts) with a power of say 0.8. Use Type I error rate of 5% and the MSE = 0.12 from this study as an estimate of the variance.

1. α , significance level, = 0.05
2. Power = $1 - \beta = 0.80$
3. Variance σ^2 estimate is MSE = 0.12
4. assume equal replication $r = n_1 = n_2 = \dots = n_t$
5. $D = |\mu_i - \mu_j| = 0.5$

$$\lambda = \frac{r D^2}{2 \sigma^2}$$

or

$$\lambda = \frac{r \sum_{i=q}^t (\mu_i - \mu_{.})^2}{\sigma^2}$$

or

$$\sum (\mu_i - \mu_{.})^2 = 0.1875 \text{ assuming } \mu_1 = \mu_2 = \mu_3 = 3.0 \text{ and } \mu_4 = 3.5$$

r	$v_2 = 4(r-1) = n_T - 4$	$\lambda = \frac{rD^2}{2\sigma^2}$	$\phi = \sqrt{\lambda/t}$	$* \lambda = \frac{r \sum_{i=q}^t (\mu_i - \mu.)^2}{\sigma^2}$	$\phi = \sqrt{\lambda/t}$	Power (Table 14)
5	16	5.21	1.14	7.81	1.40	~0.50
6	20	6.25	1.25	9.38	1.53	~0.60
7	24	7.29	1.35	10.94	1.65	
8	28	8.33	1.44	12.50	1.77	~0.81
9	32	9.38	1.53	14.06	1.88	~0.84
10	36	10.42	1.61	15.63	1.98	
11	40	11.46	1.69	17.19	2.07	
12	44	12.50	1.77	18.75	2.17	
13	48	13.54	1.84	20.31	2.25	~0.96

This calculation can also be done using software. R, as we see below, allows you to specify five of the six values and it calculates the last.

```
power.anova.test(groups = NULL, n = NULL,
                 between.var = NULL, within.var = NULL,
                 sig.level = 0.05, power = NULL)
```

SAS has a power and sample size application that is a web application and it also has sample size procedures - PROC GLMPower and PROC Power.

Using R

```
> power.anova.test( groups=4, sig.level=0.05, within.var=0.12,
between.var=var(c(3,3,3,3.5)), power=0.80)
Balanced one-way analysis of variance power calculation

  groups = 4
    n = 8.02505
between.var = 0.0625
within.var = 0.12
 sig.level = 0.05
  power = 0.8
NOTE: n is number in each group

> power.anova.test( groups=4, sig.level=0.05, within.var=0.12,
between.var=var(c(3,3,3,3.5)), n=9)
Balanced one-way analysis of variance power calculation

  groups = 4
    n = 9
between.var = 0.0625
within.var = 0.12
 sig.level = 0.05
  power = 0.8541396
NOTE: n is number in each group
```

Using SAS PROC GLMPower

```
data meats;
  input condition $ logbact CellWgt;
  datalines;
Plastic    3.5    1
Mixed      3      1
CO2        3      1
Vacuum     3      1
;
proc glmpower data=meats;
  class condition;
  model logbact = condition;
  weight CellWgt;
  power
  stddev = .3464
  alpha = 0.05
  ntotal = .
  power = 0.8;
run;
```

The SAS System

1

The GLMPower Procedure

Fixed Scenario Elements

Dependent Variable	logbact
Source	condition
Weight Variable	CellWgt
Alpha	0.05
Error Standard Deviation	0.3464
Nominal Power	0.8
Test Degrees of Freedom	3

Computed N Total

Error	Actual	N
DF	Power	Total
32	0.854	36

Using SAS PROC POWER

```
proc power;  
  onewayanova  
  npergroup = .  
  power = 0.80  
  stddev = 0.3464  
  groupmeans = 3 | 3 | 3 | 3.5;  
  plot x=power min=.5 max=.95;  
run;
```

The SAS System

1

The POWER Procedure

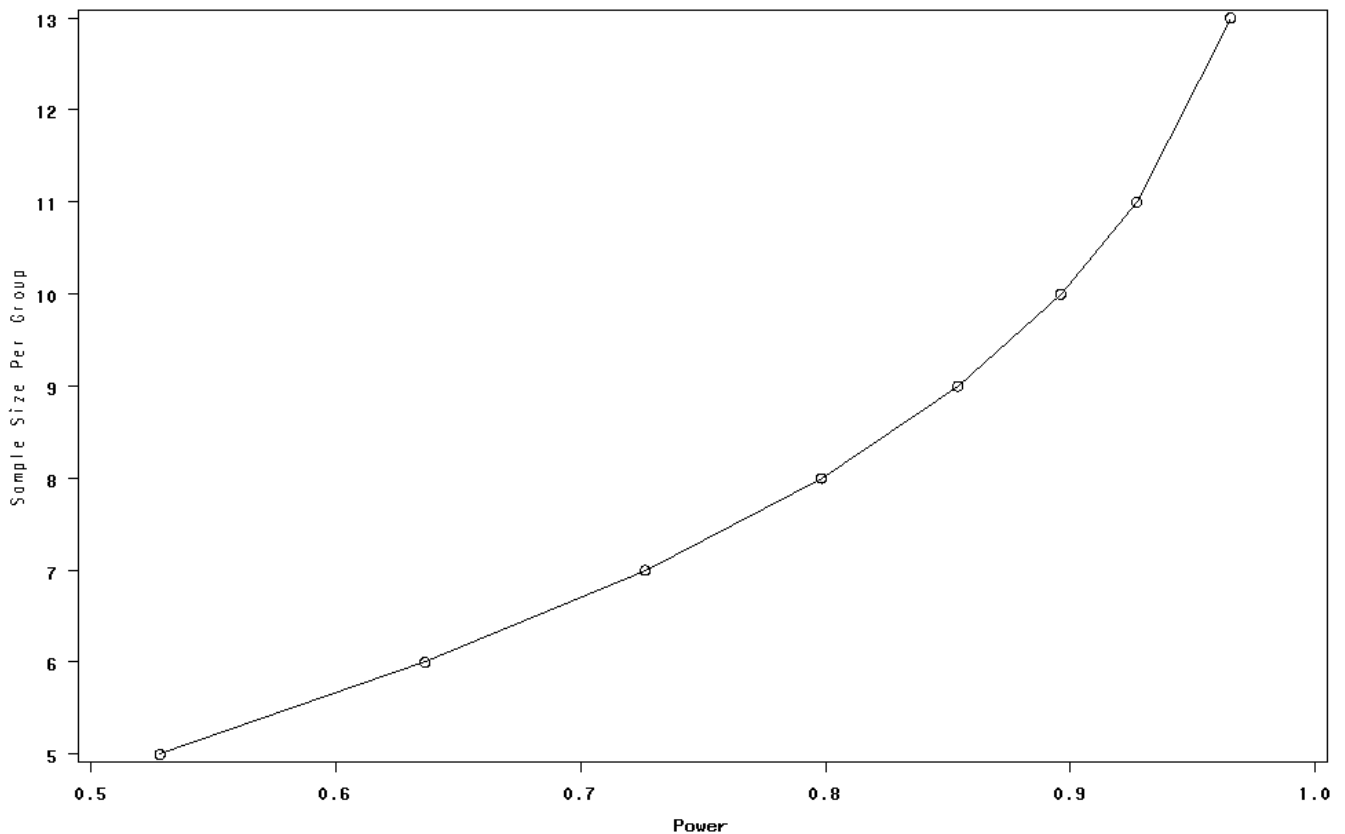
Overall F Test for One-Way ANOVA

Fixed Scenario Elements

Method	Exact
Group Means	3 3 3 3.5
Standard Deviation	0.3464
Nominal Power	0.8
Alpha	0.05

Computed N Per Group

Actual Power	N Per Group
0.854	9



An alternative perspective where you look at power achieved by different sample sizes

```
proc power;  
  onewayanova  
  npergroup = 5 to 15 by 1  
  power = .  
  stddev = 0.3464  
  groupmeans = 3 | 3 | 3 | 3.5;  
  plot x=n min=5 max=15;  
run;
```

The SAS System

3

The POWER Procedure

Overall F Test for One-Way ANOVA

Fixed Scenario Elements

Method	Exact
Group Means	3 3 3 3.5
Standard Deviation	0.3464
Alpha	0.05

Computed Power

Index	N Per Group	Power
1	5	0.528
2	6	0.637
3	7	0.727
4	8	0.798
5	9	0.854
6	10	0.896
7	11	0.927
8	12	0.950
9	13	0.966
10	14	0.977
11	15	0.984

