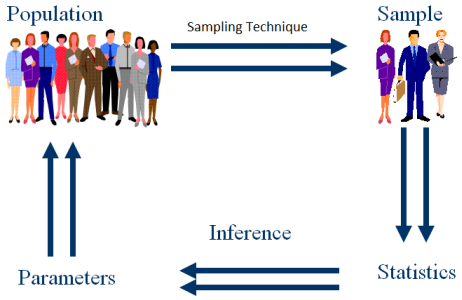
Contents

Chapter 1

ST 512 - Review

Readings: Chapters 1-8 as needed

- _____ - all the values, items, or individuals of interest
- _____ - a summary value about the population
- _____ - a subset of the population we observe data on
- _____ - a summary value calculated from the sample



Examples of parameters - (true) mean μ, (true) variance σ^2.
Examples of statistics - sample mean $\bar{Y}$, sample variance $S^2 = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n-1}$

_____ - Making claims about the population using sample data.

Scales (Types) of Data:

- _____ - A variable that is described by attributes or labels
  Subscales:
  Nominal - categories have no ordering (Male, Female)
  Ordinal - can order categories (Lickert scale data)

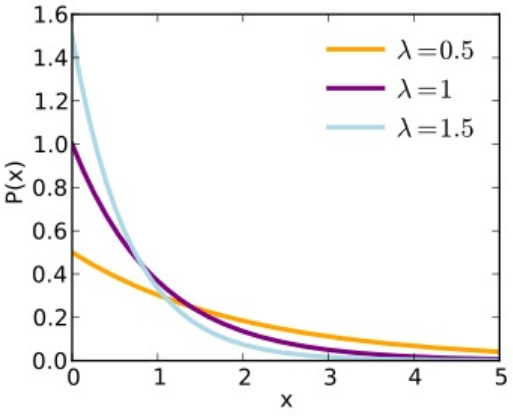
- _____ - A variable that is described by numerical measurements where arithmetic can be performed
  Subscales:
  Discrete - finite or countably infinite # of values (# of flowers, 0, 1, 2, ...)
  Continuous - any value in an interval is possible (Temperature, $(-459.67\,\text{deg}\,F, \infty)$)

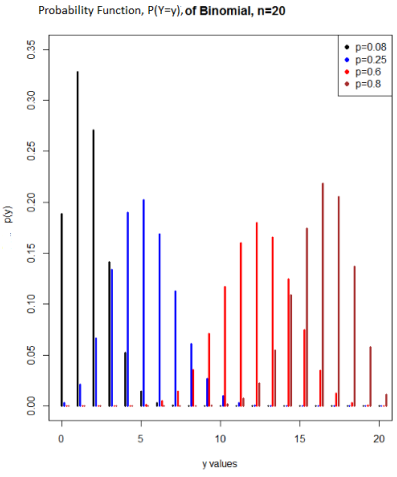
Random Variables and Things of Interest:

- _____ - Numeric outcome to a random process

  Things of interest

  – _____ - pattern and frequency of observable values
  For continuous RVs, a smooth curve. For discrete a 'probability histogram.'





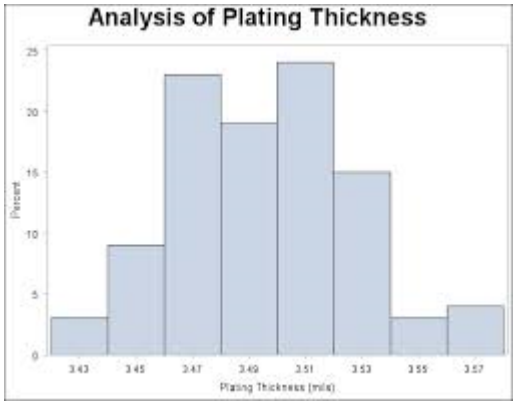
  – _____ - measures of center of the distribution
  Main focus often on mean: true mean μ, RV sample mean $\bar{Y}$, observed sample mean $\bar{y}$

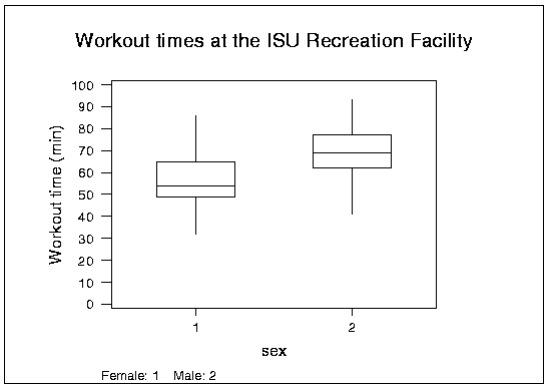
– _____ - measures of spread for the distribution
For our purposes, SD and Variance are most important: true variance σ^2, true SD σ, observed sample variance s^2, observed SD s

Graphical Descriptions of RV's:

- _____ - Graphs the frequencies or relative frequencies of realizations of a RV

- _____ - Uses the Five Number Summary (min, Q_1, M, Q_3, max) to display the realizations of a RV





Statistics are RVs!

The distribution of a statistic is called a _____

We almost always require that our sample is a **random sample** (RS), or equivalently that our random variables are **iid** (independent and identically distributed).

Central Limit Theorem (CLT):

If a RV Y has a (true) mean μ and (true) variance σ^2, and a random sample is of size $n \geq 30$ is taken then

If $Y \sim^{iid} N(\mu, \sigma^2)$ then $\bar{Y} \sim N(\mu, \sigma^2/n)$ for any n.

Two general methods for inference:

1. _____ - range of values we believe contain the parameter with some level of confidence

   For an observed $(1-\alpha)100\%$ confidence interval (c_L, c_U) we can say

   We are $(1-\alpha)100\%$ confident the true parameter value is contained in the interval. (***Do not say probability or chance!)

   The idea of Confidence means

   The procedure used to create the interval has a $(1-\alpha)100\%$ probability of producing an interval that contains the parameter.

   i.e. If the experiment were done repeatedly and an interval made for each sample, $(1-\alpha)100\%$ of the intervals would contain the parameter value.

2. _____ - test to determine if a given value is reasonable for a parameter

   For a hypothesis test, the p-value means

   the probability of observing a test statistic as extreme or more extreme than the one observed, assuming the null hypothesis is true.

   Statistical significance implies

   the observed value was unlikely to have occurred by random chance alone (assuming the null hypothesis is true).

Common ways make inference about μ (true mean):

One sample Z-test

When the true SD, σ, is known and $\bar{Y}$ has a normal distribution the sampling distribution of the statistic

$$Z = \frac{\bar{Y} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

100(1-α)%Ci for μ is

$$\bar{Y} \pm z_{\alpha/2}\sigma/\sqrt{n}$$

HT: for $H_0 : \mu = \mu_0 \;\; vs \;\; H_A : \mu > \mu_0 \;\; or \;\; \mu < \mu_0 \;\; or \;\; \mu \neq \mu_0$

$$\text{Test Statistic: } Z = \frac{\bar{Y} - \mu_0}{\sigma/\sqrt{n}}$$

$$RR : \{z_{obs} : z_{obs} > z_\alpha\} \;\; or \;\; \{z_{obs} : z_{obs} < -z_\alpha\} \;\; or \;\; \left\{z_{obs} : |z_{obs}| > z_{\alpha/2}\right\}$$
$$P - value : P(Z > z_{obs}) \;\; or \;\; P(Z < z_{obs}) \;\; or \;\; 2 * P(Z > |z_{obs}|)$$

One sample T-test

When the true SD, σ, is unknown we looked at the sampling distribution of the statistic

$$T = \frac{\bar{Y} - \mu}{s/\sqrt{n}} \sim t_{n-1} \;\; \text{(valid if RS and Y is normal)}$$

100(1-α)% CI for μ is

$$\bar{Y} \pm t_{\alpha/2,n-1}S/\sqrt{n}$$

HT: for $H_0 : \mu = \mu_0 \;\; vs \;\; H_a : \mu > \mu_0 \;\; or \;\; \mu < \mu_0 \;\; or \;\; \mu \neq \mu_0$

$$\text{Test Statistic: } T = \frac{\bar{Y} - \mu_0}{S/\sqrt{n}}$$

$$RR : \{t_{obs} : t_{obs} > t_{\alpha,n-1}\} \;\; or \;\; \{t_{obs} : t_{obs} < -t_{\alpha,n-1}\} \;\; or \;\; \left\{t_{obs} : |t_{obs}| > t_{\alpha/2,n-1}\right\}$$
$$P - value : P(T_{n-1} > t_{obs}) \;\; or \;\; P(T_{n-1} < t_{obs}) \;\; or \;\; 2 * P(T_{n-1} > |t_{obs}|)$$

Inference about two (true) means, μ_1 and μ_2 or $\mu_d = \mu_1 - \mu_2$:

Paired Data: (Paired t-test) Assume **differences** are a RS and normally distributed
100(1-α)% CI for μ_d is

$$\overline{D} \pm t_{\alpha/2,n-1} S_D/\sqrt{n} = \bar{Y} - \bar{X} \pm t_{\alpha/2,n-1} S_{\bar{Y}-\bar{X}}/\sqrt{n}$$

HT: for $H_0 : \mu_d = \Delta_0$ *vs* $H_a : \mu_d > \Delta_0$ *or* $\mu_d < \Delta_0$ *or* $\mu_d \neq \Delta_0$

$$\text{Test Statistic: } T = \frac{\bar{Y} - \bar{X} - \Delta_0}{S_d/\sqrt{n}}$$

$$RR : \{t_{obs} : t_{obs} > t_{\alpha,n-1}\} \quad or \quad \{t_{obs} : t_{obs} < -t_{\alpha,n-1}\} \quad or \quad \left\{t_{obs} : |t_{obs}| > t_{\alpha/2,n-1}\right\}$$
$$P - value : P(T_{n-1} > t_{obs}) \quad or \quad P(T_{n-1} < t_{obs}) \quad or \quad 2 * P(T_{n-1} > |t_{obs}|)$$

Independent Samples: Assume populations are independent RS's with each population having a normal distribution
Equal Variance (Two-sample pooled t-test):
100(1-α)% CI for μ_d is

$$\bar{Y} - \bar{X} \pm t_{\alpha/2,n_1+n_2-2} \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

HT: for $H_0 : \mu_d = \Delta_0$ *vs* $H_a : \mu_d > \Delta_0$ *or* $\mu_d < \Delta_0$ *or* $\mu_d \neq \Delta_0$

$$\text{Test Statistic: } T = \frac{\bar{Y} - \bar{X} - \Delta_0}{\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$RR : \{t_{obs} : t_{obs} > t_{\alpha,n_1+n_2-2}\} \quad or \quad \{t_{obs} : t_{obs} < -t_{\alpha,n_1+n_2-2}\} \quad or \quad \left\{t_{obs} : |t_{obs}| > t_{\alpha/2,n_1+n_2-2}\right\}$$
$$P - value : P(T_{n_1+n_2-2} > t_{obs}) \quad or \quad P(T_{n_1+n_2-2} < t_{obs}) \quad or \quad 2 * P(T_{n_1+n_2-2} > |t_{obs}|)$$

Unequal Variance (Two-sample t-test)
100(1-α)% CI for μ_d is

$$\bar{Y} - \bar{X} \pm t_{\alpha/2,\widehat{df}} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

HT: for $H_0 : \mu_d = \Delta_0$ *vs* $H_a : \mu_d > \Delta_0$ *or* $\mu_d < \Delta_0$ *or* $\mu_d \neq \Delta_0$

$$\text{Test Statistic: } T = \frac{\bar{Y} - \bar{X} - \Delta_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$RR : \left\{t_{obs} : t_{obs} > t_{\alpha,\widehat{df}}\right\} \quad or \quad \left\{t_{obs} : t_{obs} < -t_{\alpha,\widehat{df}}\right\} \quad or \quad \left\{t_{obs} : |t_{obs}| > t_{\alpha/2,\widehat{df}}\right\}$$
$$P - value : P(T_{\widehat{df}} > t_{obs}) \quad or \quad P(T_{\widehat{df}} < t_{obs}) \quad or \quad 2 * P(T_{\widehat{df}} > |t_{obs}|)$$

$$\widehat{df} = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\left(\frac{s_1^2}{n_1}\right)^2/(n_1-1) + \left(\frac{s_2^2}{n_2}\right)^2/(n_2-1)}$$

Extension of two-sample pooled t-test to inference about t (true) means, $\mu_1, \mu_2, ..., \mu_t$:

Analysis used for a completely randomized design.

One Way ANOVA model:

$$Y_{ij} = \mu_i + E_{ij}$$

where $E_{ij} \sim^{iid} N(0, \sigma^2)$, $i = 1, 2, ...t$, and $j = 1, 2, ..., n$ (total sample size $= nt = N$)

μ_i is the mean for group i and σ^2 is the common variance for each population.

An alternative model:

$$Y_{ij} = \mu + \tau_i + E_{ij}$$

where μ is the overall (grand) mean, τ_i is the effect for the i^{th} treatment, and $E_{ij} \sim^{iid} N(0, \sigma^2)$
Data and labeling:

Corn Syrup	Replicate #	'L' measurement	Label
26	1	51.89	y_{11}
26	2	51.52	y_{12}
26	3	52.69	y_{13}
42	1	47.21	y_{21}
42	2	48.57	y_{22}
42	3	47.57	y_{23}
55	1	41.43	y_{31}
55	2	42.31	y_{32}
55	3	42.31	y_{33}

Balanced One-way ANOVA table (same number of replicates per group)

Source	DF	SS	MS	F-stat	P-value
Treatment	$t-1$	$n\sum_{i=1}^{t}(\bar{Y}_{i\bullet} - \bar{Y}_{\bullet\bullet})^2$	$\frac{SS(Trt)}{t-1}$	$\frac{MS(Trt)}{MS(E)}$	Use $F(t-1, t(n-1))$
Error	$t(n-1)$	$\sum_{i=1}^{t}\sum_{j=1}^{n}(Y_{ij} - \bar{Y}_{i\bullet})^2$	$\frac{SS(E)}{t(n-1)}$		
Total	$nt-1$	$\sum_{i=1}^{t}\sum_{j=1}^{n}(Y_{ij} - \bar{Y}_{\bullet\bullet})^2$			

P-value from ANOVA table tests

$$H_0 : \mu_1 = \mu_2 = ... = \mu_t \text{ vs } H_A : \text{at least 1 mean differs}$$

Chapter 2

ST 512 - Contrasts and Multiple Comparisons

Readings 9.1-9.5

For the next section or two we will continue to consider experiments that have a quantitative response and categorical predictors only (i.e. experiments that can be analyzed using One-Way ANOVA). To start with let's just look at a completely randomized experimental design (i.e. all treatments are randomly assigned to the experimental units). We will look at contrasts of the parameters that will lead directly into the way to analyze Multi-Way ANOVA models where a 'factorial' treatment structure is used.

Consider the traditional balanced One-Way ANOVA model (**ANOVA (Analysis of Variance, i.e. comparing mean squares)**). That is,

- we have a *continuous* response, Y
- *qualitative or categorical* predictor(s) which we call our **factor(s)** (often we will use a continuous factor that is only observed at a few values as our categorical factor)
- if a single factor then the **levels** of the factor are our **treatments**, if multiple factors the combinations of levels from the factors is the treatment. (Either way, t = total number of treatments)

(One-Way corresponds to having only one factor of interest.)

One form of the One-Way ANOVA model is

$$Y_{ij} = \mu_i + E_{ij}$$

- E_{ij} are iid $N(0, \sigma^2)$
- $i = 1, ..., t$ describes the treatment group
- $j = 1, ..., n_i$ represents the number of replications we have in treatment group i.

We will consider 'balanced' designs for now, where $n_i = n$, same number of replicates for each treatment. Total number of observations $= N = nt$

Unknown parameters:

- μ_i - (true) mean response for all members of population i
  Estimate:

- σ^2 - (true) variance within a given treatment group (assumed constant across groups)
  Estimate:

Goals of One-Way ANOVA: Determine

1. if all treatment group means are equal.

2. if treatment group means are not equal, which means differ from each other.

One-Way ANOVA example:
An experiment was done to determine if there was a difference between antibiotic types in terms of their mean binding fraction in bovines.
There were N=20 bovines that were randomly assigned to one of t=5 types of antibiotics (the levels of the factor, since only one factor these levels are also the treatments), yielding n=4 replicates for each treatment.
The data given here, labeled in terms of the One-Way ANOVA format:

Binding Fraction (Y)	Antibiotic	True Trt Mean	Sample Mean
$y_{11} = 29.2$	Chloramphenicol	μ_1	$\bar{y}_{1\bullet} = 27.8$
$y_{12} = 32.8$	Chloramphenicol		
$y_{13} = 25.0$	Chloramphenicol		
$y_{14} = 24.2$	Chloramphenicol		
$y_{21} = 21.6$	Erythromycin	μ_2	$\bar{y}_{2\bullet} = 19.1$
$y_{22} = 17.4$	Erythromycin		
$y_{23} = 18.3$	Erythromycin		
$y_{24} = 19.0$	Erythromycin		
$y_{31} = 29.6$	Penicillin G	μ_3	$\bar{y}_{3\bullet} = 28.6$
$y_{32} = 24.3$	Penicillin G		
$y_{33} = 28.5$	Penicillin G		
$y_{34} = 32.0$	Penicillin G		
$y_{41} = 5.8$	Streptomycin	μ_4	$\bar{y}_{4\bullet} = 7.8$
$y_{42} = 6.2$	Streptomycin		
$y_{43} = 11.0$	Streptomycin		
$y_{44} = 8.3$	Streptomycin		
$y_{51} = 27.3$	Tetracyclin	μ_5	$\bar{y}_{5\bullet} = 31.4$
$y_{52} = 32.6$	Tetracyclin		
$y_{53} = 30.8$	Tetracyclin		
$y_{54} = 34.8$	Tetracyclin		
$\bar{y}_{\bullet\bullet} = 22.9$			

Recall: Notation for means in One-Way ANOVA -
Overall sample mean = $\bar{y}_{\bullet\bullet}$ or $\bar{y}_{++}$ (mean over index i and index j)

Treatment i sample mean = $\bar{y}_{i\bullet}$ or $\bar{y}_{i+}$ (mean over index i)

Goal: Test if the population means for these 5 treatments are plausibly equal.
If so, which treatment means differ significantly?

Modeling the binding fraction experiment:
One-Way ANOVA model is appropriate:

$$Y_{ij} = \mu_i + E_{ij}$$

for $i = 1, \ldots, 5$ and $j = 1, \ldots, 4$, where E_{ij} are i.i.d. $N(0, \sigma^2)$ errors.

$$\begin{aligned} \mu_1 &= \text{mean of Cholramphenicol treatment} \\ \mu_2 &= \text{mean of Erythromycin treatment} \\ &\vdots \\ \mu_5 &= \text{mean of Tetracyclin treatment} \end{aligned}$$

To test H_0 :

vs

H_A :

we use software to compute the One-Way ANOVA table and look at 'global' p-value from the table.

Table for balanced one-way ANOVA:

Source	DF	SS	MS	F	P-value
Treatments	$t-1$	$SS(T)$	$MS(T) = \frac{SS(T)}{(t-1)}$	$F = \frac{MS(T)}{MS(E)}$	$P(F_{t-1,t(n-1)} > F_{obs})$
Error	$t(n-1)$	$SS(E)$	$MS(E) = \frac{SS(E)}{(N-t)}$		
Total	$nt-1$	$SS(Tot)$			

where

$$\begin{aligned} SS(T) &= \sum_{i=1}^{t}\sum_{j=1}^{n}(\bar{y}_{i\bullet} - \bar{y}_{\bullet\bullet})^2 = n\sum_{i=1}^{t}(\bar{y}_{i\bullet} - \bar{y}_{\bullet\bullet})^2 \\ SS(E) &= \sum_{i=1}^{t}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{i\bullet})^2 \\ SS(Tot) &= \sum_{i=1}^{t}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{\bullet\bullet})^2 \end{aligned}$$

Note: SS(T) is also called SS(Between) and SS(E) is also called SS(Within).

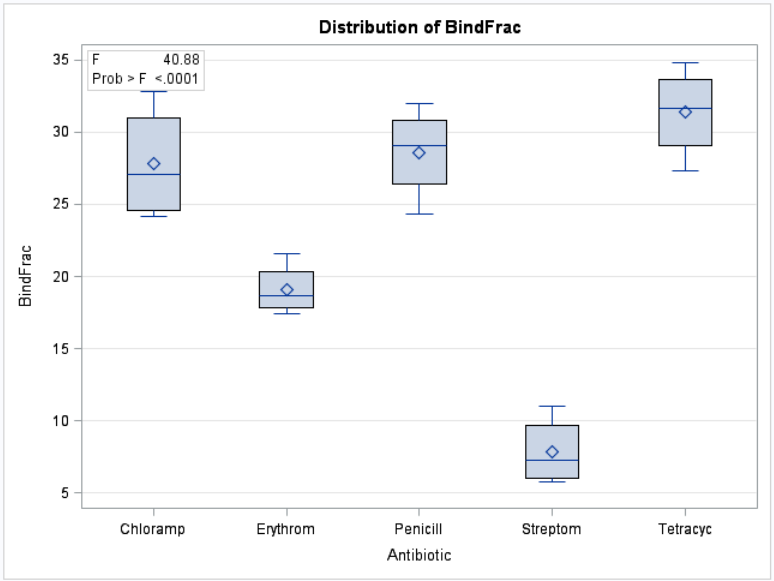
In SAS,

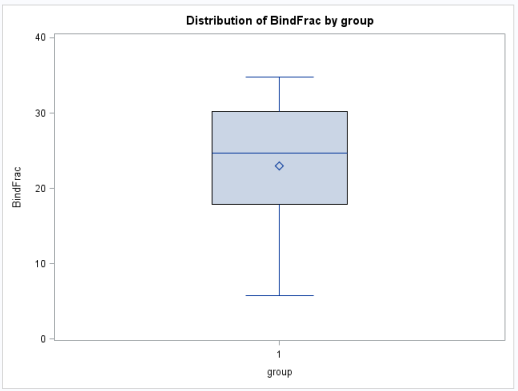
```
proc glm data=binding;
class antibiotic;
model bindfrac=antibiotic;
means antibiotic;
run;
```

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	1480.823000	370.205750	40.88	<.0001
Error	15	135.822500	9.054833		
Corrected Total	19	1616.645500			

R-Square	Coeff Var	Root MSE	BindFrac Mean
0.915985	13.12023	3.009125	22.93500

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Antibiotic	4	1480.823000	370.205750	40.88	<.0001





Level of Antibiotic	N	BindFrac	
		Mean	Std Dev
Chloramp	4	27.8000000	3.98998747
Erythrom	4	19.0750000	1.80623919
Penicill	4	28.6000000	3.21765960
Streptom	4	7.8250000	2.38379949
Tetracyc	4	31.3750000	3.17109340

Conclusion about treatment means begin equal?

Given the answer to the previous question, the next logical question to answer is: 'Which treatment means are different?'

Suppose we want first inspect the difference between the Cholramphenicol (μ_1) and Erythromycin treatment means (μ_2). In terms of μ_1 and μ_2, how can we write this question as a null and alternative hypotheses?

We get an estimator this quantity with the corresponding sample means

The standard error of this quantity can be found by taking the square root of the variance (recall we assume our samples are independent so covariance is 0)

For a balanced design we have

yielding a standard error of

By the normality assumption on the data we then have a case similar to the two-sample t test with pooled variance!

$$\bar{Y}_{1\bullet} - \bar{Y}_{2\bullet} \sim N(\mu_1 - \mu_2, \sigma^2(1/n_1 + 1/n_2)) = N(\mu_1 - \mu_2, 2\sigma^2/n)$$

We estimate σ^2 by the common pooled estimate (over all the samples, not just these two)

$$MS(E) = S_w^2 = \frac{SS(E)}{t(n-1)} = \frac{\sum_{i=1}^{t}\sum_{j=1}^{n}(Y_{ij} - \bar{Y}_{i\bullet})^2}{N-t}$$

Thus, we can for a t-test for testing $H_0 : \mu_1 = \mu_2 \quad vs \quad H_A : \mu_1 \neq \mu_2$ using

$$T = \frac{\bar{Y}_{1\bullet} - \bar{Y}_{2\bullet}}{\sqrt{MS(E)(1/n_1 + 1/n_2)}} = \frac{\bar{Y}_{1\bullet} - \bar{Y}_{2\bullet}}{\sqrt{2MS(E)/n}} \sim t_{N-t} = t_{t(n-1)}$$

We can form a $(1-\alpha)100\%$ CI for $\mu_1 - \mu_2$ using

$$\bar{Y}_{1\bullet} - \bar{Y}_{2\bullet} \pm t_{\alpha/2,N-t}\sqrt{MS(E)(1/n_1 + 1/n_2)} \quad or \quad \bar{Y}_{1\bullet} - \bar{Y}_{2\bullet} \pm t_{\alpha/2,N-t}\sqrt{2MS(E)/n}$$

and check if 0 is in the interval.

1. Create a 95% CI for $\mu_1 - \mu_2$ ($P(T_{15} > 2.13) = 0.025$) and also find the test statistic for testing the above hypotheses. What is your conclusion?

More generally than just wanting to see which means differ, we may want to compare different functions of means.

The functions of the means will be in the form of a **linear combination** of the treatment means.
In general, a linear combination of treatment means takes the form

For the bovine experiment, which of the following are linear combinations of treatment means?

- $\theta_6 = 4\mu_1 + 3\mu_2 - 7\mu_5$
- $\theta_7 = 3\mu_1\mu_4 + 2\mu_3$
- $\theta_8 = \mu_1 + \mu_2 + \mu_3 + \mu_4 + \mu_5$
- $\theta_9 = \mu_1^2 + 3\mu_2 + 1$

If the coefficients of the linear combination sum to zero the linear combination is called a **contrast**.

Is our linear combination $\mu_1 - \mu_2 = 0$ a contrast? How about any of θ_6 through θ_9?

If we want to do inference about a linear combination of treatment means we need an estimator of θ, call it $\hat{\theta}$ and we will also need a measure of variability, say $\hat{SE}(\hat{\theta})$.

An estimator is given by substitution of the sample means:

$$\hat{\theta} = c_1\bar{Y}_{1+} + c_2\bar{Y}_{2+} + ... + c_t\bar{Y}_{t+}$$

Use the output from previous to estimate the following linear combinations:

$\theta_1 = \mu_1 \quad \hat{\theta}_1 =$ _____

$\theta_2 = \mu_2 \quad \hat{\theta}_2 =$ _____

$\theta_5 = \mu_5 \quad \hat{\theta}_5 =$ _____

$\theta_6 = 4\mu_1 + 3\mu_2 - 7\mu_5 \quad \hat{\theta}_6 =$ _____

$\theta_8 = \mu_1 + \mu_2 + \mu_3 + \mu_4 + \mu_5 \quad \hat{\theta}_8 =$ _____

We now have a point estimate of the quantity in our null hypothesis, in order to conduct our test we must also know about the variability of this estimate, i.e. What is $\hat{Var}(\hat{\theta})$ or $\hat{SE}(\hat{\theta})$?

The variance of a linear combination of means in One-way ANOVA has a very nice form:

Making inference about the linear combination: (includes contrasts)

Due to the normality assumption each mean has a normal distribution and further, the linear combination will also have a normal distribution. Therefore, once we estimate the variance, we can use a t-test and a t-interval.

Let θ_0 be a value of interest for our contrast (often 0).
To test $H_0 : \theta = \theta_0$ vs $H_A : \theta \neq \theta_0$ we can use

$$t = \frac{\hat{\theta} - \theta_0}{\hat{SE}(\hat{\theta})} \sim t_{t(n-1)} \text{ under } H_0$$

A $(1-\alpha)100\%$ CI for θ is

$$\hat{\theta} \pm t_{\alpha/2,t(n-1)}\sqrt{\hat{SE}(\hat{\theta})} = \sum_{i=1}^{t} c_i \bar{y}_{i+} \pm t_{\alpha/2,t(n-1)}\sqrt{MS(E)\sum_{i=1}^{t}\frac{c_i^2}{n_i}}$$

What is the value of this test statistic and a 95% CI for the θ_8? Compare the test stat to $t_{0.025,15} = 2.13$ to make a conclusion. What is your interpretation?

Getting the answers from SAS

Pairwise Contrasts in SAS

Note: A contrast that has only two nonzero c's is called a **pairwise contrast** (as it looks at only two means).
These can be had easily in proc glm using the code below.

Recall, if our global p-value is significant then our secondary goal is to find which means differ. This question is answered via looking at all pairwise comparisons (contrasts) of means.

A **complex** contrast is a contrast that involves more than two non-zero coefficients. For example, $\theta_{10} = \frac{\mu_1+\mu_2}{2} - \frac{\mu_3+\mu_4}{2}$ is a complex contrast.

```
proc glm data=binding;

class antibiotic;

model bindfrac=antibiotic;

means antibiotic/ lsd cldiff lines;

lsmeans antibiotic/stderr pdiff;

run;
```

*Generally we'll want to use lsmeans not means, but ok here since no covariates involved and a balanced design was done. (To be discussed more later.)

The GLM Procedure

t Tests (LSD) for BindFrac

Note: This test controls the Type I comparisonwise error rate, not the experimentwise

Alpha	0.05
Error Degrees of Freedom	15
Error Mean Square	9.054833
Critical Value of t	2.13145
Least Significant Difference	4.5352

Comparisons significant at the 0.05 level are indicated by *.**

Antibiotic Comparison	Difference Between Means	95% Confidence Limits		
Tetracyc - Penicill	2.775	-1.760	7.310	
Tetracyc - Chloramp	3.575	-0.960	8.110	
Tetracyc - Erythrom	12.300	7.765	16.835	***
Tetracyc - Streptom	23.550	19.015	28.085	***
Penicill - Tetracyc	-2.775	-7.310	1.760	
Penicill - Chloramp	0.800	-3.735	5.335	
Penicill - Erythrom	9.525	4.990	14.060	***
Penicill - Streptom	20.775	16.240	25.310	***
Chloramp - Tetracyc	-3.575	-8.110	0.960	
Chloramp - Penicill	-0.800	-5.335	3.735	
Chloramp - Erythrom	8.725	4.190	13.260	***
Chloramp - Streptom	19.975	15.440	24.510	***
Erythrom - Tetracyc	-12.300	-16.835	-7.765	***
Erythrom - Penicill	-9.525	-14.060	-4.990	***
Erythrom - Chloramp	-8.725	-13.260	-4.190	***
Erythrom - Streptom	11.250	6.715	15.785	***
Streptom - Tetracyc	-23.550	-28.085	-19.015	***
Streptom - Penicill	-20.775	-25.310	-16.240	***
Streptom - Chloramp	-19.975	-24.510	-15.440	***
Streptom - Erythrom	-11.250	-15.785	-6.715	***

The SAS System 12:34 Sunday, February 16,

The GLM Procedure

t Tests (LSD) for BindFrac

Note: This test controls the Type I comparisonwise error rate, not the experimentwise error rate.

Alpha	0.05
Error Degrees of Freedom	15
Error Mean Square	9.054833
Critical Value of t	2.13145
Least Significant Difference	4.5352

Means with the same letter are not significantly different.

t Grouping	Mean	N	Antibiotic
A	31.375	4	Tetracyc
A			
A	28.600	4	Penicill
A			
A	27.800	4	Chloramp
B	19.075	4	Erythrom
C	7.825	4	Streptom

The GLM Procedure
Least Squares Means

Antibiotic	BindFrac LSMEAN	Standard Error	Pr > |t|	LSMEAN Number
Chloramp	27.8000000	1.5045625	<.0001	1
Erythrom	19.0750000	1.5045625	<.0001	2
Penicill	28.6000000	1.5045625	<.0001	3
Streptom	7.8250000	1.5045625	0.0001	4
Tetracyc	31.3750000	1.5045625	<.0001	5

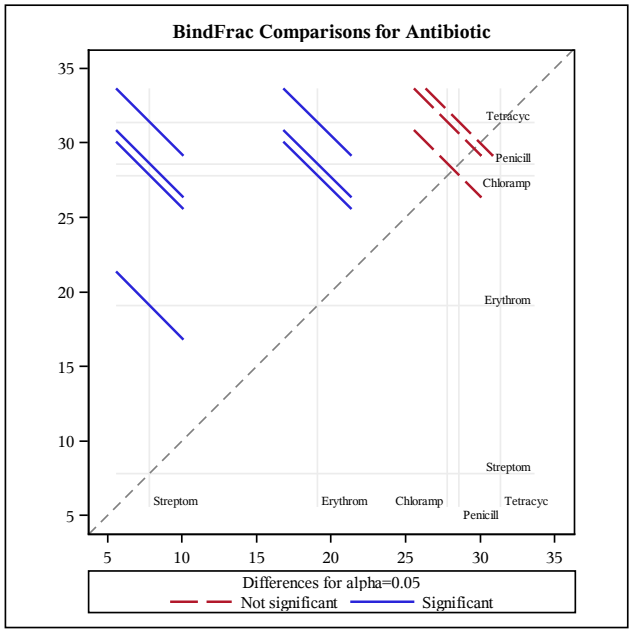
Least Squares Means for effect Antibiotic
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: BindFrac

i/j	1	2	3	4	5
1		0.0009	0.7122	<.0001	0.1136
2	0.0009		0.0004	<.0001	<.0001
3	0.7122	0.0004		<.0001	0.2118
4	<.0001	<.0001	<.0001		<.0001
5	0.1136	<.0001	0.2118	<.0001	

The SAS System 12:34 Sunday, February 16,

The GLM Procedure
Least Squares Means



Note: To ensure overall protection level, only probabilities associated with pre-planned comparisons should be used.

Use the output to construct a 95% CI for θ_{10}.

Any linear combination in SAS.
To get the estimate for θ_{10} (and a few other linear combinations of means) in SAS we can use proc glm or use proc mixed.

Using the estimate and contrast statements in One-Way ANOVA:
We need to write our contrast in terms of the model parameters μ, τ_1, τ_2, τ_3, τ_4, and τ_5 (the alternative parameterization of the model, $Y_{ij} = \mu + \tau_i + E_{ij}$).

For instance,

$$\theta_{10} = \frac{1}{2}(\mu_1 + \mu_2) - \frac{1}{2}(\mu_3 + \mu_4) = \frac{1}{2}(\mu + \tau_1 + \mu + \tau_2 - (\mu + \tau_3 + \mu + \tau_4))$$

$$= 0\mu + \frac{1}{2}\tau_1 + \frac{1}{2}\tau_2 - \frac{1}{2}\tau_3 - \frac{1}{2}\tau_4$$

In terms of syntax, we write contrast or estimate followed by a name to distinguish it. Then we do

intercept coef on μ treatment coef on τ_1 coef on τ_2 coef on τ_3 coef on τ_4 coef on τ_5

A contrast statements will give you the contrast sum of squares and a p-value.

An estimate statement will estimate any 'estimable' function of parameters and a p-value.

```
proc glm data=binding;
class antibiotic;
model bindfrac=antibiotic/clparm;
estimate 'lsmean for trt 2'                intercept 1 antibiotic 0 1 0 0 0;
estimate 'avg of trt 2 and trt 3 mean'     intercept 2 antibiotic 0 1 1 0 0/divisor=2;
estimate 'trt 1 vs trt 5'                  intercept 0 antibiotic 1 0 0 0 -1;
estimate 'avg of 1 and 2 vs avg of 3 and 4' intercept 0 antibiotic 1 1 -1 -1 0/divisor=2;
run;
```

```
proc mixed data=binding;
class antibiotic;
model bindfrac=antibiotic;
lsmestimate antibiotic 'lsmean for trt 2'                [1,2]/cl;
lsmestimate antibiotic 'avg of trt 2 mean and trt 3 mean' [0.5,2] [0.5,3]/cl;
lsmestimate antibiotic 'trt 1 vs trt 5'                  [1,1] [-1,5]/cl;
lsmestimate antibiotic 'avg of 1 and 2 vs avg of 3 and 4' [1,1] [1,2] [-1,3] [-1,4]/divisor=2 cl;
run;
```

Least Squares Means Estimate									
Effect	Label	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
Antibiotic	lsmean for trt 2	19.0750	1.5046	15	12.68	<.0001	0.05	15.8681	22.2819

Least Squares Means Estimate									
Effect	Label	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
Antibiotic	avg of trt 2 mean and trt 3 mean	23.8375	1.0639	15	22.41	<.0001	0.05	21.5699	26.1051

Least Squares Means Estimate									
Effect	Label	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
Antibiotic	trt 1 vs trt 5	-3.5750	2.1278	15	-1.68	0.1136	0.05	-8.1102	0.9602

Least Squares Means Estimate									
Effect	Label	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
Antibiotic	avg of 1 and 2 vs avg of 3 and 4	5.2250	1.5046	15	3.47	0.0034	0.05	2.0181	8.4319

Parameter	Estimate	Standard Error	t Value	Pr > |t|	95% Confidence Limits	
lsmean for trt 2	19.0750000	1.50456251	12.68	<.0001	15.8681009	22.2818991
avg of trt 2 mean and trt 3 mean	23.8375000	1.06388635	22.41	<.0001	21.5698799	26.1051201
trt 1 vs trt 5	-3.5750000	2.12777270	-1.68	0.1136	-8.1102402	0.9602402
avg of 1 and 2 vs avg of 3 and 4	5.2250000	1.50456251	3.47	0.0034	2.0181009	8.4318991

Multiple Comparisons Corrections

It is not safe to go carrying out many many significance tests suggested by the data all willy-nilly. If we do, our *experiment-wise* type I error rate will not be controlled.

For instance, we only look at all pairwise comparisons of treatment means if the global p-value is significant. Thus, these are data driven hypotheses we are testing.

Recall: $\alpha = P(\text{Type I Error})$

Decision	H_0 true	H_0 false
Reject H_0	Type I Error	Correct!
Fail to Reject H_0	Correct!	Type II Error

For a given test, we fix the probability of a type I error to be small (often 0.05) as it is usually considered worse than a type II error.

Consider the case with $t = 5$ (antibiotic treatments):

- the number of pairwise contrasts of the form $\theta = \mu_i - \mu_j$ is $\begin{pmatrix} 5 \\ 2 \end{pmatrix} = 10$
- each test has type I error $\alpha = 0.05$, but overall what is our experiment-wise type I error rate?

  i.e. P(rejecting at least one null hypothesis that is true)
- This is called the *experiment-wise* or *family-wise* (fwe) type I error rate. We should really control this in addition to the type I error for each test!

Example: Is a certain type of coin fair (equal probability of flipping a head and a tail)?

$$H_0 : \text{Coin fair}, p = 0.5 \qquad H_A : \text{Coin biased}, p \neq 0.5$$

Experiment - flip one of these coins 10 times, if 9 or 10 heads appear or if 9 or 10 tails appear then declare coin biased.

Assuming the coin is fair, what is the significance level of this test?

Now suppose we have 100 coins of this type and we test each in the same manner. If the coins were truly fair, how many of the experiments would we expect to conclude we have a biased coin?

For a particular coin to come up heads or tails 9-10 times is very unlikely, but seeing any 1 coin of the 100 behave this way would be more likely than not.
In fact,

Under the assumption of independence of all tests,

$$\text{fwe rate} = P(\text{At least 1 type I error}) = \alpha^* = 1 - (1 - \alpha)^k$$

where k is the number of tests being done and α is the significance level used for each test.

Comparisons can be categorized as *a priori* or *post-hoc*:

- *A priori*: Significance tests which will be carried out without regard to the observed outcome of the experiment.
- *Post-hoc* or data-driven: Significance tests which are suggested by the observed outcome of the experiment.

Methods for simultaneous inference for multiple comparisons include (but there are many many of these)

- Bonferroni (section 9.3)
- Tukey (section 9.5)
- Fisher's LSD (don't use this, section 9.4, you should still read this section)
- Duncan (Not required, used when there is a control treatment, section 9.7)
- Scheffé (Not required, section 9.8)

Bonferroni Correction

Suppose interest lies in exactly k linear combinations of means. **Bonferroni correction** is to replace the usual α with

$$\alpha' = \frac{\alpha}{k}$$

By doing so our fwe rate will be less than α.

We can now create **simultaneous** CIs. These are a group of CIs we are $(1-\alpha)\%$ confident will <u>all</u> contain their true values.

Simultaneous 95% confidence intervals for the k linear combinations of means is given by

$$a_1\bar{y}_{1\bullet} + a_2\bar{y}_{2\bullet} + \cdots + a_t\bar{y}_{t\bullet} \pm t_{\alpha'/2,t(n-1)}\sqrt{MS(E)\sum_{i=1}^{t}\frac{a_i^2}{n_i}}$$

$$\vdots$$

$$k_1\bar{y}_{1\bullet} + k_2\bar{y}_{2\bullet} + \cdots + k_t\bar{y}_{t\bullet} \pm t_{\alpha'/2,t(n-1)}\sqrt{MS(E)\sum_{i=1}^{t}\frac{k_i^2}{n_i}}$$

Note: $t_{\alpha'/2,t(n-1)}$ might have to be obtained using software.

For the binding fraction example, consider only pairwise comparisons of Chloramphenicol (μ_1):

$$\theta_1 = \mu_1 - \mu_2, \quad \theta_2 = \mu_1 - \mu_3, \quad \theta_3 = \mu_1 - \mu_4, \quad \theta_4 = \mu_1 - \mu_5$$

We have $k = 4$, $\alpha' = 0.05/k = 0.0125$, and $t_{\alpha'/2,15} = 2.84$.

What is the Margin of Error for one of these contrasts? Find the simultaneous 95% intervals for the four contrasts. Which of these antibiotic means differ significantly from the chloramphenicol mean?

Tukey-Kramer Correction (or just Tukey)

Tukey's correction is the best method when making inference on **all pairwise** contrasts in balanced designs. That is, for simple contrasts of the form

$$\theta = \mu_j - \mu_k$$

it will tend to have a lower type II error rate in these cases than other multiple comparison corrections. (It has a greater chance of detecting differences i.e. is more powerful.)

- Uses multipliers from a distribution called the 'studentized range distribution'
- Denoted $q(\alpha, t, t(n-1))$

For a balanced design, **simultaneous** 95% confidence intervals for $\theta = \mu_j - \mu_k$ are given by

$$\begin{aligned}
\hat{\theta} &\pm \frac{q(\alpha, t, t(n-1))}{\sqrt{2}} \hat{SE}(\hat{\theta}) \\
\bar{y}_{j+} - \bar{y}_{k+} &\pm \frac{q(\alpha, t, t(n-1))}{\sqrt{2}} \sqrt{MS(E)\left(\frac{1}{n} + \frac{1}{n}\right)} \\
\bar{y}_{j+} - \bar{y}_{k+} &\pm q(\alpha, t, t(n-1)) \sqrt{\frac{MS(E)}{n}}
\end{aligned}$$

How to do these multiple comparison corrections in SAS?

- Bonferonni can be done by manually changing the α level SAS uses. For the example above, $\alpha' = 0.05/4 = 0.0125$:

```
proc glm data=binding;  class antibiotic;
model bindfrac=antibiotic/clparm alpha=0.0125;  *Can drop intercept since it has 0 coefficient;
estimate 'theta 1' antibiotic 1 -1;             *Note these are different thetas than previous;
estimate 'theta 2' antibiotic 1 0 -1;
estimate 'theta 3' antibiotic 1 0 0 -1 0;
estimate 'theta 4' antibiotic 1 0 0 0 -1;   run;

proc mixed data=binding; class antibiotic;
model bindfrac=antibiotic;
lsmestimate antibiotic 'theta 1' [1,1] [-1,2],
                       'theta 2' [1,1] [-1,3],
                       'theta 3' [1,1] [-1,4],
                       'theta 4' [1,1] [-1,5]/cl adjust=bon;  run;
```

- Tukey correction is an option you ask for:

```
proc glm data=binding; class antibiotic;
model bindfrac=antibiotic;
lsmeans antibiotic/pdiff adjust=tukey cl;  run;

proc mixed data=binding; class antibiotic;
model bindfrac=antibiotic;
lsmeans antibiotic/pdiff adjust=tukey cl;  run;
```

Bonferroni GLM output:

Parameter	Estimate	Standard Error	t Value	Pr > |t|	98.75% Confidence Limits	
theta 1	8.7250000	2.12777270	4.10	0.0009	2.6893015	14.7606985
theta 2	-0.8000000	2.12777270	-0.38	0.7122	-6.8356985	5.2356985
theta 3	19.9750000	2.12777270	9.39	<.0001	13.9393015	26.0106985
theta 4	-3.5750000	2.12777270	-1.68	0.1136	-9.6106985	2.4606985

Bonferroni Mixed output:

Least Squares Means Estimates Adjustment for Multiplicity: Bonferroni												
Effect	Label	Estimate	Standard Error	DF	t Value	Pr > |t|	Adj P	Alpha	Lower	Upper	Adj Lower	Adj Upper
Antibiotic	theta 1	8.7250	2.1278	15	4.10	0.0009	0.0038	0.05	4.1898	13.2602	2.6893	14.7607
Antibiotic	theta 2	-0.8000	2.1278	15	-0.38	0.7122	1.0000	0.05	-5.3352	3.7352	-6.8357	5.2357
Antibiotic	theta 3	19.9750	2.1278	15	9.39	<.0001	<.0001	0.05	15.4398	24.5102	13.9393	26.0107
Antibiotic	theta 4	-3.5750	2.1278	15	-1.68	0.1136	0.4545	0.05	-8.1102	0.9602	-9.6107	2.4607

Tukey Mixed output:

Differences of Least Squares Means													
Effect	Antibiotic	_Antibiotic	Estimate	Standard Error	DF	t Value	Pr > |t|	Adjustment	Adj P	Alpha	Lower	Upper	Adj Lower	Adj Upper
Antibiotic	Chloramp	Erythrom	8.7250	2.1278	15	4.10	0.0009	Tukey	0.0072	0.05	4.1898	13.2602	2.1546	15.2954
Antibiotic	Chloramp	Penicill	-0.8000	2.1278	15	-0.38	0.7122	Tukey	0.9953	0.05	-5.3352	3.7352	-7.3704	5.7704
Antibiotic	Chloramp	Streptom	19.9750	2.1278	15	9.39	<.0001	Tukey	<.0001	0.05	15.4398	24.5102	13.4046	26.5454
Antibiotic	Chloramp	Tetracyc	-3.5750	2.1278	15	-1.68	0.1136	Tukey	0.4738	0.05	-8.1102	0.9602	-10.1454	2.9954
Antibiotic	Erythrom	Penicill	-9.5250	2.1278	15	-4.48	0.0004	Tukey	0.0035	0.05	-14.0602	-4.9898	-16.0954	-2.9546
Antibiotic	Erythrom	Streptom	11.2500	2.1278	15	5.29	<.0001	Tukey	0.0007	0.05	6.7148	15.7852	4.6796	17.8204
Antibiotic	Erythrom	Tetracyc	-12.3000	2.1278	15	-5.78	<.0001	Tukey	0.0003	0.05	-16.8352	-7.7648	-18.8704	-5.7296
Antibiotic	Penicill	Streptom	20.7750	2.1278	15	9.76	<.0001	Tukey	<.0001	0.05	16.2398	25.3102	14.2046	27.3454
Antibiotic	Penicill	Tetracyc	-2.7750	2.1278	15	-1.30	0.2118	Tukey	0.6928	0.05	-7.3102	1.7602	-9.3454	3.7954
Antibiotic	Streptom	Tetracyc	-23.5500	2.1278	15	-11.07	<.0001	Tukey	<.0001	0.05	-28.0852	-19.0148	-30.1204	-16.9796

Tukey GLM output:

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey

Antibiotic	BindFrac LSMEAN	LSMEAN Number
Chloramp	27.8000000	1
Erythrom	19.0750000	2
Penicill	28.6000000	3
Streptom	7.8250000	4
Tetracyc	31.3750000	5

Least Squares Means for effect Antibiotic
Pr > |t| for H0: LSMean(i)=LSMean(j)
Dependent Variable: BindFrac

i/j	1	2	3	4	5
1		0.0072	0.9953	<.0001	0.4738
2	0.0072		0.0035	0.0007	0.0003
3	0.9953	0.0035		<.0001	0.6928
4	<.0001	0.0007	<.0001		<.0001
5	0.4738	0.0003	0.6928	<.0001	

Antibiotic	BindFrac LSMEAN	95% Confidence Limits	
Chloramp	27.800000	24.593101	31.006899
Erythrom	19.075000	15.868101	22.281899
Penicill	28.600000	25.393101	31.806899
Streptom	7.825000	4.618101	11.031899
Tetracyc	31.375000	28.168101	34.581899

Least Squares Means for Effect Antibiotic

i	j	Difference Between Means	Simultaneous 95% Confidence Limits for LSMean(i)-LSMean(j)	
1	2	8.725000	2.154598	15.295402
1	3	-0.800000	-7.370402	5.770402
1	4	19.975000	13.404598	26.545402
1	5	-3.575000	-10.145402	2.995402
2	3	-9.525000	-16.095402	-2.954598
2	4	11.250000	4.679598	17.820402
2	5	-12.300000	-18.870402	-5.729598
3	4	20.775000	14.204598	27.345402
3	5	-2.775000	-9.345402	3.795402
4	5	-23.550000	-30.120402	-16.979598

Independent Contrasts

So, what else can we do with contrasts?

Consider a contrast θ, then

$$\theta = c_1\mu_1 + c_2\mu_2 + ... + c_t\mu_t$$

where $\sum_{i=1}^{t} c_i = 0$. The point estimate of a contrast is

$$\hat{\theta} = c_1\bar{y}_{1+} + c_2\bar{y}_{2+} + ... + c_t\bar{y}_{t+}$$

and the estimated variance is given by

$$\hat{Var}(\hat{\theta}) = MS(E)\sum_{i=1}^{t}\frac{c_i^2}{n_i}$$

Recall: The idea behind ANOVA is that we partition $SS(Tot)$ (with df=N-1) into independent components $SS(Trt)$ and $SS(E)$ (whose df add up to N-1).

Similarly, we can take $SS(Trt)$ and partition it into $t-1$ independent contrasts each with 1 df. How you ask??

Orthogonal contrasts:
Let

$$\theta_1 = \sum_{i=1}^{t} c_i\mu_i \text{ and } \theta_2 = \sum_{i=1}^{t} d_i\mu_i$$

be two contrasts. θ_1 and θ_2 are **orthogonal** if

$$c_1d_1 + c_2d_2 + ... + c_td_t = \sum_{i=1}^{t} c_id_i = 0$$

(for balanced designs).

If two contrasts are orthogonal, then one contrast conveys no information about the other contrast. Hence we can break up $SS(Trt)$ into completely separate sources.

A set of k contrasts is mutually orthogonal if all pairs are orthogonal. As we have $t-1$ df, we can break $SS(Trt)$ into at most $t-1$ orthogonal contrasts.

This will allow us to attribute a certain amount of variation to given contrasts, which in turn will represent something of interest to you the researcher.

Examples:
$(-1, 1, 0, 0, 0)$ and $(0, 0, -1, 1, 0)$ orthogonal ?

$(1, -1/2, -1/2, 0, 0)$ and $(0, 0, 0, -1, 1)$ orthogonal ?

$(-1, 1, 0, 0, 0)$ and $(0, -1, 1, 0, 0)$ orthogonal ?

Due to the joint normality of our data, **orthogonality implies independence**!

Sums of squares for contrasts
Recall we are going to partition $SS(Trt)$ into $t-1$ independent contrasts, each will have its own sum of squares. The sums of squares for a contrast are defined as

$$SS(\hat{\theta}) = \frac{\hat{\theta}^2}{\left(\frac{c_1^2}{n_1} + \cdots + \frac{c_t^2}{n_t}\right)} = \frac{\hat{\theta}^2}{\left(\sum_{i=1}^{t} \frac{c_i^2}{n_i}\right)}$$

This contrast has 1 df associated with it.

Alternative (but equivalent) test for a contrast
Previously, we saw that we could test a linear combination (and hence a contrast) via a t-test. Now, we can define

$$MS(\hat{\theta}) = SS(\hat{\theta})/1 = SS(\hat{\theta})$$

and can then test

$$H_0 : \theta = 0 \quad vs \quad H_A : \theta \neq 0$$

using the F-statistic

$$F = \frac{MS(\hat{\theta})}{MS(E)} \sim F_{1,t(n-1)}$$

Compare this to the t-test done earlier for testing a contrast equal to 0

$$t = \frac{\hat{\theta} - 0}{\hat{SE}(\hat{\theta})} \sim t_{t(n-1)}$$

If we square a t stat we get an F stat!

Simultaneous test for contrasts
We can also test multiple orthogonal contrasts all equal to 0 at once

$$H_0 : \theta_1 = \theta_2 = ... = \theta_k = 0 \quad vs \quad H_A : \text{At least 1 } \theta \neq 0$$

using the F-statistic

$$F = \frac{\frac{SS(\hat{\theta_1})+SS(\hat{\theta_2})+...+SS(\hat{\theta_k})}{k}}{MS(E)} \sim F_{k,t(n-1)}$$

How to relate this to $SS(Trt)$?
Generally, if $\theta_1, \theta_2, ..., \theta_{t-1}$ are $t-1$ mutually orthogonal contrasts then

$$SS(Trt) = SS(\hat{\theta}_1) + SS(\hat{\theta}_2) + \cdots + SS(\hat{\theta}_{t-1})$$

and $df_{Trt} = df_{\hat{\theta}_1} + ... + df_{\hat{\theta}_{t-1}} = 1 + ... + 1 = t - 1$

Notice, testing all $t-1$ contrasts equal to 0 is equivalent to testing our global F-test!

Again consider the Binding Fraction data. In this case we have $5 - 1 = 4$ df for treatment. Consider the following set of 4 mutually orthogonal contrasts:

$$\begin{aligned}
\theta_1 &= (-2 \quad -1 \quad 0 \quad 1 \quad 2) \\
\theta_2 &= (2 \quad -1 \quad -2 \quad -1 \quad 2) \\
\theta_3 &= (-1 \quad 2 \quad 0 \quad -2 \quad 1) \\
\theta_4 &= (1 \quad -4 \quad 6 \quad -4 \quad 1)
\end{aligned}$$

Since these are mutually orthogonal, they are all independent.

Let's use SAS to get estimates. Test for $\theta_4 = 0$ using both the t and F tests. Then show that $SS(Trt) = SS(\theta_1) + SS(\theta_2) + SS(\theta_3) + SS(\theta_4)$ and conduct the global F test.

```
proc glm data=binding; class antibiotic; model bindfrac=antibiotic;
contrast 'theta 1' antibiotic -2 -1 0 1 2;
contrast 'theta 2' antibiotic 2 -1 -2 -1 2;
contrast 'theta 3' antibiotic -1 2 0 -2 1;
contrast 'theta 4' antibiotic 1 -4 6 -4 1; run;
```

```
proc mixed data=binding; class antibiotic; model bindfrac=antibiotic;
lsmestimate antibiotic 'theta 1' [-2,1] [-1,2] [0,3] [1,4] [2,5],
'theta 2' [2,1] [-1,2] [-2,3] [-1,4] [2,5],
'theta 3' [-1,1] [2,2] [0,3] [-2,4] [1,5],
'theta 4' [1,1] [-4,2] [6,3] [-4,4] [1,5]; run;

*I've done the estimate statements in proc mixed, but you could do them in glm instead;
*Also, we may want to do a multiple comparison correction depending on
```

Note the equivalence of the p-value for a contrast and for an estimate statement.

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Antibiotic	4	1480.823000	370.205750	40.88	<.0001

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
theta 1	1	6.7240000	6.7240000	0.74	0.4024
theta 2	1	335.1607143	335.1607143	37.01	<.0001
theta 3	1	271.9622500	271.9622500	30.04	<.0001
theta 4	1	866.9760357	866.9760357	95.75	<.0001

Least Squares Means Estimates						
Effect	Label	Estimate	Standard Error	DF	t Value	Pr > |t|
Antibiotic	theta 1	-4.1000	4.7578	15	-0.86	0.4024
Antibiotic	theta 2	34.2500	5.6296	15	6.08	<.0001
Antibiotic	theta 3	26.0750	4.7578	15	5.48	<.0001
Antibiotic	theta 4	123.18	12.5881	15	9.79	<.0001

Another example
Data consists of the number of contaminants in IV fluids made by $t = 3$ pharmaceutical companies

	Cutter	Abbott	McGaw
y_{i1}	255	105	577
y_{i2}	264	288	515
y_{i3}	342	98	214
y_{i4}	331	275	413
y_{i5}	234	221	401
y_{i6}	217	240	260
$\bar{y}_{i\bullet}$	273.8	204.5	396.7

Source	d.f.	Sum of squares	Mean Square	F
Treatments (pharmacy)	2	113646	56823	5.81
Error	15	146753	9784	
Total	17	260400		

Consider the following 2 contrasts:

$$\theta_1 = \mu_M - \mu_A \quad \text{and} \quad \theta_2 = \mu_C - \frac{\mu_M + \mu_A}{2}$$

Which levels of the factor will each of these be in SAS?

Rewrite these contrasts in terms of μ_1, μ_2, and μ_3.

Are these contrasts orthogonal (and thus independent)?

Use the output to compute $SS(\hat{\theta}_1)$ and $SS(\hat{\theta}_2)$. What should these add up to and why?

```
proc glm data=pharm; class company; model contam=company;
contrast 'McGaw vs Abbot' company -1 0 1;
estimate 'McGaw vs Abbot'  company -1 0 1;
contrast 'Cutter vs avg of McGaw and Abbot' company -1 2 -1;
estimate 'Cutter vs avg of McGaw and Abbot' company -1 2 -1/divisor=2; run;
```

Note: We should really do a multiple comparison correction for our two contrasts. Bonferonni is easiest, compare our p-values to $0.05/2 = 0.025$.

The GLM Procedure

Class Level Information		
Class	Levels	Values
Company	3	Abbott Cutter McGaw

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	113646.3333	56823.1667	5.81	0.0136
Error	15	146753.6667	9783.5778		
Corrected Total	17	260400.0000			

R-Square	Coeff Var	Root MSE	contam Mean
0.436430	33.91268	98.91197	291.6667

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Company	2	113646.3333	56823.1667	5.81	0.0136

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Company	2	113646.3333	56823.1667	5.81	0.0136

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
McGaw vs Abbot	1	110784.0833	110784.0833	11.32	0.0043
Cutter vs avg of McGaw and Abbot	1	2862.2500	2862.2500	0.29	0.5965

Parameter	Estimate	Standard Error	t Value	Pr > |t|
McGaw vs Abbot	192.166667	57.1068524	3.37	0.0043
Cutter vs avg of McGaw and Abbot	-26.750000	49.4559849	-0.54	0.5965

Company	contam LSMEAN
Abbott	204.500000
Cutter	273.833333
McGaw	396.666667

For another good example, see page 457, example 9.3.

Chapter 3

ST 512 - Analysis of Factorial Designs (Multiway ANOVA)

Readings: 14.1-14.6, especially pg 886-887,891,904

We've looked at one-way ANOVA so far. This models the t treatments using $t-1$ degrees of freedom. However, the treatments may actually be combinations of more than 1 factor of interest. What we'll be able to do now is answer the question, which factor(s) (not treatments) are important!

A multiway ANOVA example
An observational study was done to investigate the Cholesterol levels of different groups of people. There were two factors in this experiment

- Age - levels = Younger than 50 (Young), Older than 50 (Old)
- Gender - levels = Male, Female

Therefore, we have 2x2=4 treatments. Label the treatments as

- OF = Old Female (group 1)
- OM = Old Male (group 2)
- YF = Young Female (group 3)
- YM = Young Male (group 4)

Within each treatment group we have $n_j = n = 7$ observations (a balanced design). To investigate if any treatment means differ, we can do a One-Way ANOVA analysis by fitting the model

$$Y_{ij} = \mu + \tau_i + E_{ij}$$

where $i = 1(OF), 2(OM), 3(YF), 4(YM)$, $j = 1, 2, \ldots, 7$ and E_{ij} i.i.d. $N(0, \sigma^2)$. We constrain that $\sum_{i=1}^{t} \tau_i = 0$.

The data and One-Way ANOVA output are given below:

Treatment	Cholesterol level							avg	std. dev.
OF (i=1)	262	193	224	201	161	178	265	$\bar{y}_{1\bullet} = 212.0$	$s_1 = 40$
OM (i=2)	192	253	248	278	232	267	289	$\bar{y}_{2\bullet} = 251.3$	$s_2 = 32$
YF (i=3)	221	213	202	183	185	197	162	$\bar{y}_{3\bullet} = 194.7$	$s_3 = 20$
YM (i=4)	271	192	189	209	227	236	142	$\bar{y}_{4\bullet} = 209.4$	$s_4 = 41$

```
proc glm data=cholesterol;
class Treatment;
model Chol=Treatment;
lsmeans Treatment/cl pdiff adjust=tukey;
run;
```

Treatment	Chol LSMEAN	LSMEAN Number
OF	212.000000	1
OM	251.285714	2
YF	194.714286	3
YM	209.428571	4

Least Squares Means for effect Treatment
Pr > |t| for H0: LSMean(i)=LSMean(j)
Dependent Variable: Chol

i/j	1	2	3	4
1		0.1707	0.7841	0.9990
2	0.1707		0.0250	0.1322
3	0.7841	0.0250		0.8538
4	0.9990	0.1322	0.8538	

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	12280.85714	4093.61905	3.46	0.0323
Error	24	28434.57143	1184.77381		
Corrected Total	27	40715.42857			

Treatment	Chol LSMEAN	95% Confidence Limits	
OF	212.000000	185.149211	238.850789
OM	251.285714	224.434925	278.136503
YF	194.714286	167.863497	221.565075
YM	209.428571	182.577782	236.279360

R-Square	Coeff Var	Root MSE	Chol Mean
0.301627	15.87245	34.42054	216.8571

Least Squares Means for Effect Treatment

i	j	Difference Between Means	Simultaneous 95% Confidence Limits for LSMean(i)-LSMean(j)	
1	2	-39.285714	-90.040133	11.468704
1	3	17.285714	-33.468704	68.040133
1	4	2.571429	-48.182990	53.325847
2	3	56.571429	5.817010	107.325847
2	4	41.857143	-8.897276	92.611561
3	4	-14.714286	-65.468704	36.040133

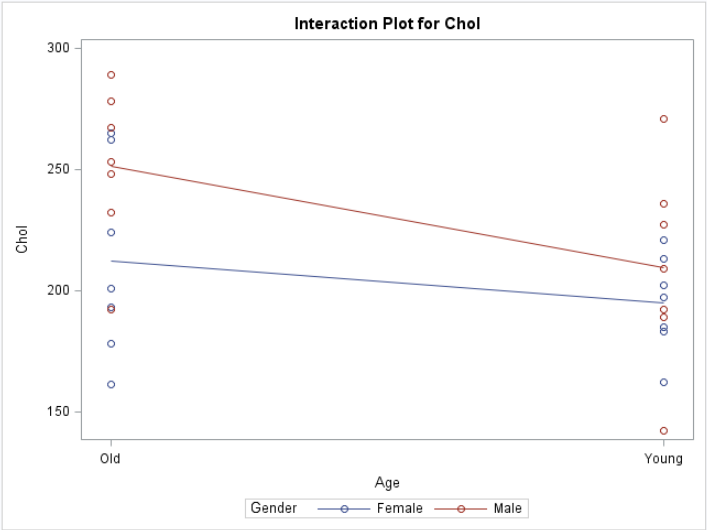
Source	DF	Type I SS	Mean Square	F Value	Pr > F
Treatment	3	12280.85714	4093.61905	3.46	0.0323

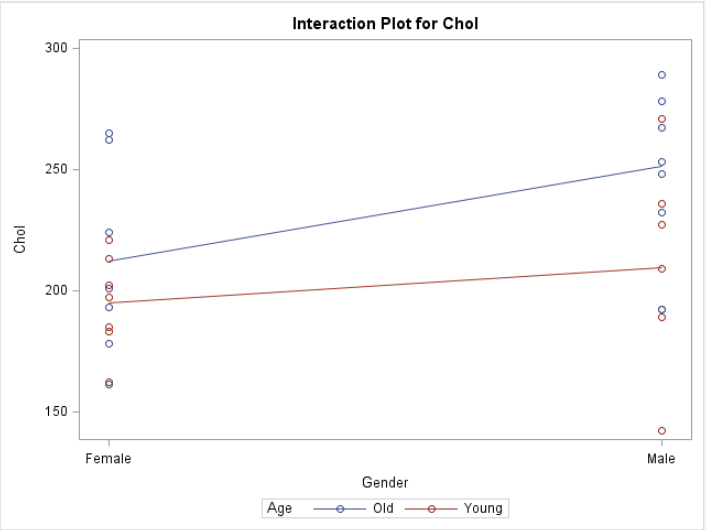
Source	DF	Type III SS	Mean Square	F Value	Pr > F
Treatment	3	12280.85714	4093.61905	3.46	0.0323

Conclusion from ANOVA table p-value is that the treatment means, $\mu + \tau_i$ or equivalently μ_i, are not plausibly equal (using $\alpha = 0.05$). From the lsmeans statement we can see that Young Females and Old Males are the only groups that differ significantly.

Now suppose we want to decide what effects the Age factor and the Gender factor have on the response. That is, rather than just inspect treatment mean differences, can we say that Age is significant for predicting Blood Pressure? What about Gender?

We can investigate these by looking at contrasts! Consider the (profile or interaction) plots below:





Define the **main effect** of factor A is the change in the response for switching levels of factor A (averaged over all other factors).

What contrast would test for the *main effect* of Age?

$\theta_{Age} =$

What contrast would test for the *main effect* of Gender?

$\theta_{Gender} =$

If the main effect contrast for factor A is significantly different from 0, then that **factor** is important for predicting the response.

There is a third contrast of interest. This contrast represents the **interaction** between Age and Gender.

An **interaction** between factor A and factor B implies that the effect of factor A on the response depends on the level of factor B (and vice-versa).

In terms of the plots above, Age and Gender interact if the lines are not parallel (can look at either plot). What contrast can we write to investigate this?

$\theta_{Age*Gender} =$

1. Check that θ_{Age}, θ_{Gender}, and $\theta_{Age*Gender}$ are mutually orthogonal.

2. Use the previous output to find estimates for each contrast and provide standard errors. Recall: $\hat{\theta} = \sum c_i \bar{y}_{i+}$ and $\hat{SE}(\hat{\theta}) = \sqrt{MS(E) \sum_{i=1}^{t} \frac{c_i^2}{n_i}}$

   $\hat{\theta}_{Age} =$

   $\hat{\theta}_{Gender} =$

   $\hat{\theta}_{Age*Gender} =$

3. Find the sums of squares for each contrast. How many degrees of freedom are associated with each contrast? Recall: $SS(\hat{\theta}) = \frac{\hat{\theta}^2}{\sum \frac{c_i^2}{n_i}}$.

   $SS(\hat{\theta}_{Age}) =$

   $SS(\hat{\theta}_{Gender}) =$

   $SS(\hat{\theta}_{Age*Gender}) =$

4. Formulate a test of $H_0 : \theta_i = 0$ for each of these three contrasts. Obtain the F-ratio for each of these tests. Compare them to the F-critical value ($F_{0.05,1,24} = 4.26$) and make a decision about the importance of each effect.

$F_{Age} =$ _____ $num\ df =$ _____ $den\ df =$ _____

$F_{Gender} =$ _____ $num\ df =$ _____ $den\ df =$ _____

$F_{Age*Gender} =$ _____ $num\ df =$ _____ $den\ df =$ _____

5. What do you notice about the sum of these sums of squares? Find the F-statistic for a test of $H_0 : \theta_{Age} = \theta_{Gender} = \theta_{Age*Gender} = 0$ vs H_A : at least one differs. What do you notice about this test and the overall F-test from the ANOVA table in the One-Way model?

Notice what we've done:
Very similar to partitioning the $SS(Tot)$ into $SS(Trt)$ and $SS(E)$, we've partitioned $SS(Trt)$, which has $t-1 = 4-1 = 3$ degrees of freedom into 3 independent components that represent different effects of interest!

We can test for each effect to learn more about our factors rather than just the treatment means. This allows for much more insight!

This is the idea of Multi-Way ANOVA! (Although it gets a little bit more complicated when a factor has more than 2 levels.)

Let's look at how we could get these contrasts in SAS. Recall: We need to write our contrast in terms of the model parameters μ, τ_1, τ_2, τ_3, and τ_4.

For instance,

$$\theta_{Gender} = \frac{1}{2}(\mu_1 + \mu_3) - \frac{1}{2}(\mu_2 + \mu_4) = \frac{1}{2}(\mu + \tau_1 + \mu + \tau_3 - \mu - \tau_2 - \mu - \tau_4)$$
$$= 0\mu + \frac{1}{2}\tau_1 - \frac{1}{2}\tau_2 + \frac{1}{2}\tau_3 - \frac{1}{2}\tau_4$$

In terms of syntax, we write contrast or estimate followed by a name to distinguish it. Then we do

intercept coef on μ treatment coef on τ_1 coef on τ_2 coef on τ_3 coef on τ_4

A contrast statements will give you the contrast sum of squares and a p-value.

An estimate statement will estimate any 'estimable' function of parameters (coefficients don't have to sum to 0).

Tests of each contrast and estimate individually:

```
proc glm data=cholesterol;  class Treatment;  model Chol=Treatment;
contrast 'Age Main Effect Contrast' intercept 0 treatment 0.5 0.5 -0.5 -0.5;
contrast 'Gender Main Effect Contrast' intercept 0 treatment 0.5 -0.5 0.5 -0.5;
contrast 'Age*Gender Interaction Effect Contrast' intercept 0 treatment 0.5 -0.5 -0.5 0.5;
estimate 'Age Main Effect Estimate' intercept 0 treatment 0.5 0.5 -0.5 -0.5;
estimate 'Gender Main Effect Estimate' intercept 0 treatment 0.5 -0.5 0.5 -0.5;
estimate 'Age*Gender Interaction Effect Estimate' intercept 0 treatment 0.5 -0.5 -0.5 0.5;  run;
```

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
Age Main Effect Contrast	1	6121.285714	6121.285714	5.17	0.0323
Gender Main Effect Contrast	1	5103.000000	5103.000000	4.31	0.0488
Age*Gender Interaction Effect Contrast	1	1056.571429	1056.571429	0.89	0.3544

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Age Main Effect Estimate	29.5714286	13.0097426	2.27	0.0323
Gender Main Effect Estimate	-27.0000000	13.0097426	-2.08	0.0488
Age*Gender Interaction Effect Estimate	-12.2857143	13.0097426	-0.94	0.3544

Test of the contrasts simultaneously, **same as global f-test here:

```
proc glm data=cholesterol;  class Treatment;  model Chol=Treatment;
contrast 'All three at once' intercept 0 Age 1 -1,
 intercept 0 Age 0 0 Gender 1 -1,
 intercept 0 Age 0 0 Gender 0 0 Age*Gender 0.5 -0.5 -0.5 0.5; run;
```

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
All three at once	3	12280.85714	4093.61905	3.46	0.0323

Two-Way ANOVA example

Rather than fit a One-Way ANOVA model, we can use a different parameterization of that model called a Two-Way ANOVA model that will **automatically test for these contrasts of interest!**

The Two-Way ANOVA model for the cholesterol measurements is:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + E_{ijk}$$

$i = 1, 2(old, young) \quad j = 1, 2(female, male)$ and $k = 1, 2, \ldots, 7$.

- Y_{ijk} is the response for replicate k at level i of Age and level j of Gender
- μ represents the overall mean of cholesterol,

$$\text{estimate is } \hat{\mu} = \bar{Y}_{\bullet\bullet\bullet}$$

- α_i represents the 'effect' for being at level i of Age,

$$\text{estimate is } \hat{\alpha}_i = \bar{Y}_{i\bullet\bullet} - \bar{Y}_{\bullet\bullet\bullet}$$

- β_i represents the 'effect' for being at level j of Gender,

$$\text{estimate is } \hat{\beta}_j = \bar{Y}_{\bullet j\bullet} - \bar{Y}_{\bullet\bullet\bullet}$$

- $(\alpha\beta)_{ij}$ represents the 'joint effect' for being at level i of Age and level j of Gender,

$$\text{estimate is } \hat{(\alpha\beta)}_{ij} = \bar{Y}_{ij\bullet} - \bar{Y}_{i\bullet\bullet} - \bar{Y}_{\bullet j\bullet} + \bar{Y}_{\bullet\bullet\bullet}$$

- E_{ijk} is a random error

We still assume $E_{ijk} \overset{iid}{\sim} N(0, \sigma^2)$. With parameter constraints:

$$\alpha_1 + \alpha_2 = 0, \beta_1 + \beta_2 = 0, \text{ and}$$
$$(\alpha\beta)_{11} + (\alpha\beta)_{12} = 0, (\alpha\beta)_{21} + (\alpha\beta)_{22} = 0, (\alpha\beta)_{11} + (\alpha\beta)_{21} = 0, (\alpha\beta)_{12} + (\alpha\beta)_{22} = 0$$

This is called a 2×2 **factorial design** since we have 2 factors with 2 levels each and the treatments are found by *crossing* the levels of the factors. (A three factor design with 2, 3, and 5 levels crossed would be a 2×3×5 factorial design.)

For level i of Age and level j of Gender we are model the mean cholesterol as

$$\mu_{ij} = E(Y_{ijk}) = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}$$

and, as you might expect, the estimate for level i of Age and level j of Gender is

$$\hat{\mu_{ij}} = \bar{Y}_{\bullet\bullet\bullet} + (\bar{Y}_{i\bullet\bullet} - \bar{Y}_{\bullet\bullet\bullet}) + (\bar{Y}_{\bullet j\bullet} - \bar{Y}_{\bullet\bullet\bullet}) + (\bar{Y}_{ij\bullet} - \bar{Y}_{i\bullet\bullet} - \bar{Y}_{\bullet j\bullet} + \bar{Y}_{\bullet\bullet\bullet}) = \bar{Y}_{ij\bullet}$$

The two-way ANOVA model can be fit easily in proc glm using the following code:

```
proc glm data=cholesterol;
class Age Gender;
model Chol=Age Gender Age*Gender;
run;
```

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	12280.85714	4093.61905	3.46	0.0323
Error	24	28434.57143	1184.77381		
Corrected Total	27	40715.42857			

R-Square	Coeff Var	Root MSE	Chol Mean
0.301627	15.87245	34.42054	216.8571

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Age	1	6121.285714	6121.285714	5.17	0.0323
Gender	1	5103.000000	5103.000000	4.31	0.0488
Age*Gender	1	1056.571429	1056.571429	0.89	0.3544

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Age	1	6121.285714	6121.285714	5.17	0.0323
Gender	1	5103.000000	5103.000000	4.31	0.0488
Age*Gender	1	1056.571429	1056.571429	0.89	0.3544

The sums of squares for each effect are equal to the sums of squares for our contrasts in the one-way ANOVA model.

When looking at Two-Way ANOVA output,

1. Inspect the overall ANOVA table p-value. If significant,
2. Inspect the interaction p-value.
   (a) If significant, both factors are important for predicting the response, look at *simple effects*.
   (b) If not significant, investigate the main effect p-values for significance to determine which factors are important for predicting the response - look at those *main effects*.

Types of Effects Investigated

Use this table of means from the cholesterol example in the following questions:

	Gender	
	female(j=1)	male(j=2)
Age Older (i=1)	$\hat{\mu}_{11}$=212.0 (previously $\hat{\mu}_1$)	$\hat{\mu}_{12}$=251.3 (previously $\hat{\mu}_2$)
Young (i=2)	$\hat{\mu}_{21}$=194.7 (previously $\hat{\mu}_3$)	$\hat{\mu}_{22}$=209.4 (previously $\hat{\mu}_4$)

Simple Effects

1. In a Two-way ANOVA problem, a simple effect for factor A is the difference in the level means of factor A *at a given level of factor B*. That is,

   Simple effect of A at level 1 of B is defined as $\mu_{21} - \mu_{11}$,

   Simple effect of A at level 2 of B is defined as $\mu_{22} - \mu_{12}$,

2. Define the simple effects of factor B.

3. For the Cholesterol example, estimate these four simple effects and explain what each measures.

Main Effects

1. Main effects in a 2x2 experiment are the averages of the simple effects. Define the main effect of factor A as

$$\mu_A = \frac{1}{2}\left((\mu_{22} - \mu_{12}) + (\mu_{21} - \mu_{11})\right) = \frac{1}{2}(\mu_{22} + \mu_{21}) - \frac{1}{2}(\mu_{12} + \mu_{11})$$

   Our θ_{Age} contrast from before!

2. Define the main effect for factor B.

   *****Main effects should (usually) only be looked at when interaction effects are not significant.

3. For the Cholesterol example, estimate these two main effects and explain what each measures.

Interaction Effects

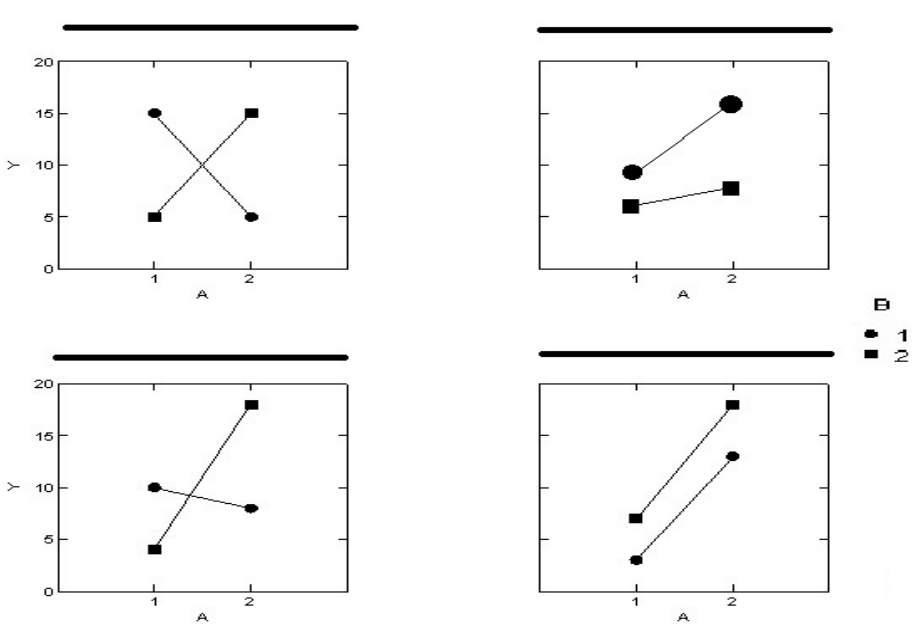
Interaction effects in a 2x2 experiment are the average of the *difference* of simple effects. The value of this effect is not usually of interest, mostly we just want to find out if the interaction is significant.

If an interaction is present, both factors are important and we look at simple effects only. Why?

Visually we can investigate interactions via a 'profile' or 'interaction' plot.

1. If an interaction effect causes the relationship between the levels of a factor to change, this is called a **qualitative** interaction.
2. If the interaction effect just changes the magnitude of the relationship between the levels of a factor, this is called a **quantitative** interaction.

Label the plots below accordingly:



If interactions exist, looking at main effects is usually not necessary. For example, approximately what would the main effect for factor B be equal to for the top left plot?

To test for the Age main effect notice,

$$\begin{aligned}
\text{A main effect} &= \frac{1}{2}(\hat{\mu}_{22} - \hat{\mu}_{12} + \hat{\mu}_{21} - \hat{\mu}_{11}) \\
&= \frac{1}{2}(\bar{Y}_{22+} - \bar{Y}_{12+} + \bar{Y}_{21+} - \bar{Y}_{11+}) \\
&= \frac{1}{2}(\bar{Y}_{22+} + \bar{Y}_{21+}) - \frac{1}{2}(\bar{Y}_{12+} + \bar{Y}_{11+}) \\
&= \bar{Y}_{2++} - \bar{Y}_{1++} \\
&= \hat{\alpha}_2 - \hat{\alpha}_1
\end{aligned}$$

Our test for the main effect of Age can be written as

$$H_0 : \alpha_1 = \alpha_2 = 0 \text{ vs } H_A : \text{ At least 1 is not 0}$$

Similarly, we can test for the Gender main effect by

$$H_0 : \beta_1 = \beta_2 = 0 \text{ vs } H_A : \text{ At least 1 is not 0}$$

And we can test the interaction using

$$H_0 : (\alpha\beta)_{11} = (\alpha\beta)_{12} = (\alpha\beta)_{21} = (\alpha\beta)_{22} = 0 \text{ vs } H_A : \text{ At least 1 is not 0}$$

Thus, this parameterization of the model gives a very nice way to test for these different effects.

Again the point of this section is that once we have a significant global p-value, we then want to attribute the significance to our factors and/or their interaction. This model allows us to do so by testing groups of parameters!

The general Two-Way ANOVA model:

Suppose we have a *continuous* response, Y, and two factors, A and B from a CRD.

Most experiments with multiple factors we will look at will have a **factorial** treatment structure. This implies 'treatments' are combinations of the levels of different factors (also called a crossed design). For the most part we will have **complete** experiments (i.e. observations at each level combination). Later in the semester we'll look at 'nested designs.'

1. Factor A has a levels and factor B has b levels.
2. There are n replicates for each treatment.
3. Total of $N = abn$ EUs.
4. Our main interest lies in whether or not the response differs due to the factors.

The parametrization of the Two-way ANOVA model we will use is

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + E_{ijk}, \quad i = 1, ..., a, \quad j = 1, ..., b, k = 1, ..., n \text{ balanced design}$$

where $E_{ijk} \sim^{iid} N(0, \sigma^2)$ and we have sum to zero constraints on the parameters (to get a unique solution).

- α_i is an effect due to level i of factor A
- β_j is an effect due to level j of factor B
- $(\alpha\beta)_{ij}$ is an joint effect due to level i of factor A and level j of factor B

To construct the ANOVA table, we will still take the total sum of squares and split it up. Now we split it into a few more parts than previous (using contrasts as before, just a little more complicated now):

ANOVA table for a×b (balanced) factorial experiment

Source	df	Sum of Squares (SS)	Mean Square (MS)	F-stat
A	$a-1$	$SS(A)$	MS(A)=SS(A)/(a-1)	MS(A)/MS(E)
B	$b-1$	$SS(B)$	MS(B)=SS(B)/(b-1)	MS(B)/MS(E)
AB	$(a-1)(b-1)$	$SS(AB)$	MS(AB)=SS(AB)/((a-1)(b-1))	MS(AB)/MS(E)
Error	$ab(n-1)$	$SS(E)$	MS(E)=SS(E)/(ab(n-1))	
Total	$N-1$	$SS(Tot)$		

$$SS(A) = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} (\bar{Y}_{i\bullet\bullet} - \bar{Y}_{\bullet\bullet\bullet})^2$$

$$SS(B) = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} (\bar{Y}_{\bullet j\bullet} - \bar{Y}_{\bullet\bullet\bullet})^2$$

$$SS(AB) = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} (\bar{Y}_{ij\bullet} - \bar{Y}_{i\bullet\bullet} - \bar{Y}_{\bullet j\bullet} + \bar{Y}_{\bullet\bullet\bullet})^2$$

$$SS(E) = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} (Y_{ijk} - \bar{Y}_{ij\bullet})^2$$

$$SS(Tot) = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} (Y_{ijk} - \bar{Y}_{\bullet\bullet\bullet})^2$$

Note:
$SS(A) + SS(B) + SS(AB) = SS(Trt)$ in One-Way ANOVA

$(a-1) + (b-1) + (a-1)(b-1) = t - 1$ (degrees of freedom in One-Way ANOVA)

$SS(A) + SS(B) + SS(AB) + SS(E) = SS(Tot)$

$df_A + df_B + df_{AB} + df_E = df_{Tot}$.

An a×b example

An entomologist records energy expended (Y) by $N = 18$ honeybees at $a = 3$ temperature (A) levels ($20, 30, 40^oC$) consuming liquids with $b = 2$ levels of sucrose concentration (B)($20\%, 40\%$) in a balanced, completely randomized crossed 3×2 design. The data are given below:

Temp	Suc	Sample		
20	20	y_{111}=3.1	y_{112}=3.7	y_{113}=4.7
20	40	y_{121}=5.5	y_{122}=6.7	y_{123}=7.3
30	20	y_{211}=6	y_{212}=6.9	y_{213}=7.5
30	40	y_{221}=11.5	y_{222}=12.9	y_{223}=13.4
40	20	y_{311}=7.7	y_{312}=8.3	y_{313}=9.5
40	40	y_{321}=15.7	y_{322}=14.3	y_{323}=15.9

```
proc glm data=ent;
class Temp Suc;
model Energy=Temp|Suc; *Vertical Bar fits all combinations of Temp and Suc (main effects and interactions);
run;
```

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	271.9444444	54.3888889	70.43	<.0001
Error	12	9.2666667	0.7722222		
Corrected Total	17	281.2111111			

R-Square	Coeff Var	Root MSE	energy Mean
0.967047	9.849135	0.878762	8.922222

Source	DF	Type I SS	Mean Square	F Value	Pr > F
temp	2	141.4577778	70.7288889	91.59	<.0001
Suc	1	116.5355556	116.5355556	150.91	<.0001
temp*Suc	2	13.9511111	6.9755556	9.03	0.0040

Source	DF	Type III SS	Mean Square	F Value	Pr > F
temp	2	141.4577778	70.7288889	91.59	<.0001
Suc	1	116.5355556	116.5355556	150.91	<.0001
temp*Suc	2	13.9511111	6.9755556	9.03	0.0040

Answer the following:

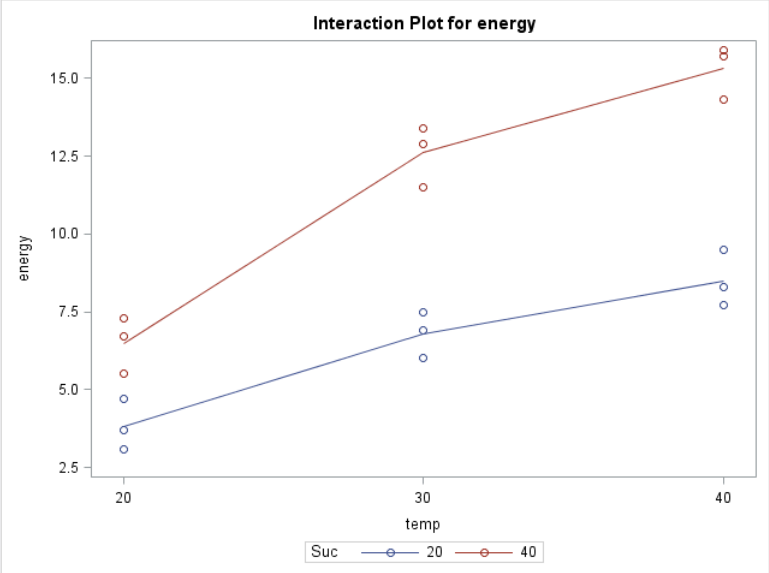
1. Is anything in the model useful?
2. If so, what effects should we investigate and why?

Unlike a 2×2 study, it is not possible to express interaction between Temp and Suc using 1 contrast.

```
title 'Means for each treatment group'; proc tabulate data=ent;
class Temp Suc; var Energy;
table Temp*mean, Energy*Suc; run;
```

Means for each treatment group

		energy	
		Suc	
		20	40
temp			
20	Mean	3.83	6.50
30	Mean	6.80	12.60
40	Mean	8.50	15.30



Testing Interaction in an $a \times b$ experiment
Here we have (3-1)(2-1)=2 degrees of freedom for interaction. This is because no interaction implies changing the level of Temp doesn't change the effect of Suc on Energy, i.e.

$$\mu_{12} - \mu_{11} = \mu_{22} - \mu_{21} \text{ or } \mu_{Temp20,Suc40} - \mu_{Temp20,Suc20} = \mu_{Temp30,Suc40} - \mu_{Temp30,Suc20}$$

$$\mu_{12} - \mu_{11} = \mu_{32} - \mu_{31} \text{ or } \mu_{Temp20,Suc40} - \mu_{Temp20,Suc20} = \mu_{Temp40,Suc40} - \mu_{Temp40,Suc20}$$

$$\mu_{32} - \mu_{31} = \mu_{22} - \mu_{21} \text{ or } \mu_{Temp40,Suc40} - \mu_{Temp40,Suc20} = \mu_{Temp30,Suc40} - \mu_{Temp30,Suc20}$$

Notice that given any two of these contrasts (move all means to one side to see these as contrasts), we can get the third contrast. So we have 3-1=2 contrasts that are needed for testing interaction.

In terms of the plot, no interaction would imply **piecewise parallel lines** across all the levels of the factor on the axis.

Test for interaction effect generalizes as:

$$H_0 : (\alpha\beta)_{ij} \equiv 0 \text{ vs. } H_1 : (\alpha\beta)_{ij} \neq 0 \text{ for some } i, j$$

$$F = \frac{MS(AB)}{MS(E)}$$

on $(a-1)(b-1)$ numerator and $ab(n-1)$ denominator df.

For honeybee data,

$$F = SS(AB)/SS(E) = 6.976/0.7722 = 9.03$$

which is significant ($p = 0.0040$) at the 0.05 S.L. on 2 and 12 degrees of freedom.

As interaction is significant, both factors are important! Our next step would be to analyze the **simple effects** of each factor.

That is, we would investigate the **effects of Temperature at given levels of Sucrose** and also look at the **effect of Sucrose at given levels of Temperature**.

Inspection of main effects in not appropriate here.

To SAS!

```
proc glm data=ent; class Temp Suc;
model Energy=Temp|Suc;
lsmeans Temp*Suc/adjust=tukey pdiff cl; run;
```

temp	Suc	energy LSMEAN	95% Confidence Limits	
20	20	3.833333	2.727905	4.938761
20	40	6.500000	5.394572	7.605428
30	20	6.800000	5.694572	7.905428
30	40	12.600000	11.494572	13.705428
40	20	8.500000	7.394572	9.605428
40	40	15.300000	14.194572	16.405428

Least Squares Means for Effect temp*Suc

i	j	Difference Between Means	Simultaneous 95% Confidence Limits for LSMean(i)-LSMean(j)	
1	2	-2.666667	-5.076699	-0.256635
1	3	-2.966667	-5.376699	-0.556635
1	4	-8.766667	-11.176699	-6.356635
1	5	-4.666667	-7.076699	-2.256635
1	6	-11.466667	-13.876699	-9.056635
2	3	-0.300000	-2.710032	2.110032
2	4	-6.100000	-8.510032	-3.689968
2	5	-2.000000	-4.410032	0.410032
2	6	-8.800000	-11.210032	-6.389968
3	4	-5.800000	-8.210032	-3.389968
3	5	-1.700000	-4.110032	0.710032
3	6	-8.500000	-10.910032	-6.089968
4	5	4.100000	1.689968	6.510032
4	6	-2.700000	-5.110032	-0.289968
5	6	-6.800000	-9.210032	-4.389968

In reality, we would probably only be interested in simple effects such as Temp 20, Suc 40 vs Temp 20, Suc 20 (i.e. the effect of Suc, holding temperature at 20). Here we are given way more than that!

We could write estimate or contrast to get just those, but this is easier. However, as we are correcting for multiple comparisons, we may be correcting too much! In real life, you'd take the time to write the contrast and estimate statements of interest.

Another $a \times b$ Design - Interaction Not Significant

Yields on 36 tomato crops from balanced, complete, crossed design with $a = 3$ varieties (A) at $b = 4$ planting densities (B) :

Variety	Density $k/hectare$	Sample		
1	10	7.9	9.2	10.5
2	10	8.1	8.6	10.1
3	10	15.3	16.1	17.5
1	20	11.2	12.8	13.3
2	20	11.5	12.7	13.7
3	20	16.6	18.5	19.2
1	30	12.1	12.6	14.0
2	30	13.7	14.4	15.4
3	30	18.0	20.8	21.0
1	40	9.1	10.8	12.5
2	40	11.3	12.5	14.5
3	40	17.2	18.4	18.9

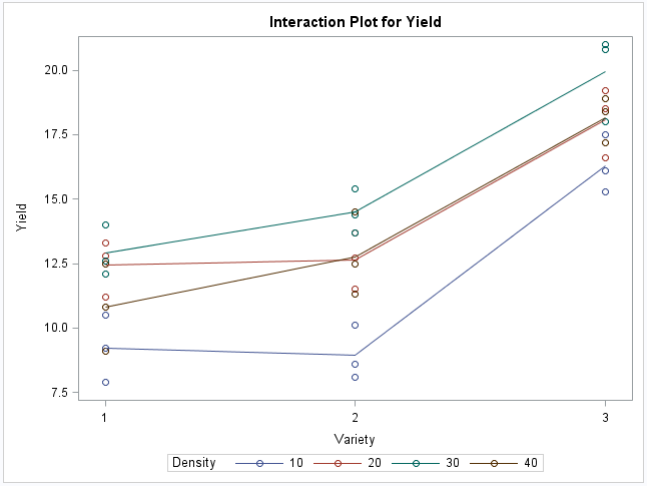
```
proc glm data=tomato; class variety density;
model Yield=Variety|Density;
lsmeans Variety Density/adjust=tukey cl; *adjust=tukey tells sas to do pdiff; run;
```

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	422.3155556	38.3923232	24.22	<.0001
Error	24	38.0400000	1.5850000		
Corrected Total	35	460.3555556			

R-Square	Coeff Var	Root MSE	Yield Mean
0.917368	9.064568	1.258968	13.88889

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Variety	2	327.5972222	163.7986111	103.34	<.0001
Density	3	86.6866667	28.8955556	18.23	<.0001
Variety*Density	6	8.0316667	1.3386111	0.84	0.5484

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Variety	2	327.5972222	163.7986111	103.34	<.0001
Density	3	86.6866667	28.8955556	18.23	<.0001
Variety*Density	6	8.0316667	1.3386111	0.84	0.5484



We proceed with our usual approach:

1. Inspect the overall ANOVA table p-value. If significant,
2. Inspect the interaction p-value.
   (a) If significant, both factors are important for predicting the response, look at *simple effects*.
   (b) If not significant, investigate the main effect p-values for significance to determine which factors are important for predicting the response - look at those *main effects*.

1. Global p-value is signficant.
2. As the interaction is not significant we want to look at main effects p-values.
3. These are significant for both factors, thus both are important. We should now look at the main effect differences for both factors (we would not look at differences for a factor deemed non-significant).

Ideas for the main effect in an axb experiment:
Note that the variety main effect has 2 df. That implies that this effect comes from testing two contrasts simultaneously.

Main effects are averages of simple effects, so we must define the simple effects of A.

We have ones corresponding to switching from variety 2 to variety 1:

$$\mu_{21}-\mu_{11} \quad \mu_{22}-\mu_{12} \quad \mu_{23}-\mu_{13} \quad \mu_{24}-\mu_{14}$$

$\mu_{Var2,Den10}-\mu_{Var1,Den10}$ $\mu_{Var2,Den20}-\mu_{Var1,Den20}$ $\mu_{Var2,Den30}-\mu_{Var1,Den30}$ $\mu_{Var2,Den40}-\mu_{Var1,Den40}$
We have ones corresponding to switching from variety 3 to variety 1:

$$\mu_{31}-\mu_{11} \quad \mu_{32}-\mu_{12} \quad \mu_{33}-\mu_{13} \quad \mu_{34}-\mu_{14}$$

$\mu_{Var3,Den10}-\mu_{Var1,Den10}$ $\mu_{Var3,Den20}-\mu_{Var1,Den20}$ $\mu_{Var3,Den30}-\mu_{Var1,Den30}$ $\mu_{Var3,Den40}-\mu_{Var1,Den40}$
We have ones corresponding to switching from variety 3 to variety 2:
Notice that these could be found by subtracting above simple effects. For example,

$$(\mu_{31}-\mu_{11})-(\mu_{21}-\mu_{11})=\mu_{31}-\mu_{32}$$

So these are redundant and not needed.

Now the main effects are the averages of these groups of simple effects. That is, one component of the main effect is

$$\frac{1}{4}\left((\mu_{21}-\mu_{11})+(\mu_{22}-\mu_{12})+(\mu_{23}-\mu_{13})+(\mu_{24}-\mu_{14})\right)$$
$$=\frac{1}{4}\left(\mu_{21}+\mu_{22}+\mu_{23}+\mu_{24}\right)-\frac{1}{4}\left(\mu_{11}+\mu_{12}+\mu_{13}+\mu_{14}\right)$$

the average of the responses at level 2 of variety against the average of the responses at level 1 of variety. (Which should make sense intuitively.)

Similarly, averaging the other group of simple effects gives the average of the responses at level 3 of variety against the average of the responses at level 1 of variety.

The average of 3 vs 2 could be found by subtracting and so it is not needed. Thus, we have 2 df for testing this main effect.

First let's look at the Variety main effects (lsmeans Variety/adjust=tukey cl; output):

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey

Variety	Yield LSMEAN	LSMEAN Number
1	11.3333333	1
2	12.2083333	2
3	18.1250000	3

Least Squares Means for effect Variety
Pr > |t| for H0: LSMean(i)=LSMean(j)
Dependent Variable: Yield

i/j	1	2	3
1		0.2249	<.0001
2	0.2249		<.0001
3	<.0001	<.0001	

Variety	Yield LSMEAN	95% Confidence Limits	
1	11.333333	10.583245	12.083422
2	12.208333	11.458245	12.958422
3	18.125000	17.374912	18.875088

Least Squares Means for Effect Variety

i	j	Difference Between Means	Simultaneous 95% Confidence Limits for LSMean(i)-LSMean(j)	
1	2	-0.875000	-2.158534	0.408534
1	3	-6.791667	-8.075201	-5.508132
2	3	-5.916667	-7.200201	-4.633132

Now let's look at the Density main effects (lsmeans Density/adjust=tukey cl; output):

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey

Density	Yield LSMEAN	LSMEAN Number
10	11.4777778	1
20	14.3888889	2
30	15.7777778	3
40	13.9111111	4

Least Squares Means for effect Density
Pr > |t| for H0: LSMean(i)=LSMean(j)
Dependent Variable: Yield

i/j	1	2	3	4
1		0.0003	<.0001	0.0022
2	0.0003		0.1169	0.8514
3	<.0001	0.1169		0.0213
4	0.0022	0.8514	0.0213	

Density	Yield LSMEAN	95% Confidence Limits	
10	11.477778	10.611650	12.343905
20	14.388889	13.522762	15.255016
30	15.777778	14.911650	16.643905
40	13.911111	13.044984	14.777238

Least Squares Means for Effect Density

i	j	Difference Between Means	Simultaneous 95% Confidence Limits for LSMean(i)-LSMean(j)	
1	2	-2.911111	-4.548299	-1.273923
1	3	-4.300000	-5.937188	-2.662812
1	4	-2.433333	-4.070521	-0.796145
2	3	-1.388889	-3.026077	0.248299
2	4	0.477778	-1.159410	2.114966
3	4	1.866667	0.229479	3.503855

Ideas can be easily extended to an arbitrary number of factors

A three-factor example: In a balanced, complete, crossed design, $N = 36$ shrimp were randomized to $abc = 12$ treatment combinations from the factors below:

A1: Temperature at 25^o C
A2: Temperature at 35^o C
B1: Density of shrimp population at 80 shrimp/$40l$
B2: Density of shrimp population at 160 shrimp/$40l$
C1: Salinity at 10 units
C2: Salinity at 25 units
C3: Salinity at 40 units

Thus, this is a 2×2×3 experiment. The response variable of interest is weight gain Y_{ijkl} after four weeks.

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + E_{ijkl}$$

where

$$i = 1, 2, \quad j = 1, 2, \quad k = 1, 2, 3, \quad l = 1, 2, 3$$

$E_{ijkl} \overset{iid}{\sim} N(0, \sigma^2)$. Note: Many constraints are required on the parameters.

Analysis of a Multi-Way ANOVA model starts by investigating the highest order interactions and working down from there, just as in the Two-Way model.

```
proc glm data=shrimp; class Temp Density Salinity;
model y=Temp|Density|Salinity; run;
```

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	467636.3333	42512.3939	14.64	<.0001
Error	24	69690.6667	2903.7778		
Corrected Total	35	537327.0000			

R-Square	Coeff Var	Root MSE	y Mean
0.870301	19.30270	53.88671	279.1667

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Temp	1	15376.0000	15376.0000	5.30	0.0304
Density	1	21218.7778	21218.7778	7.31	0.0124
Temp*Density	1	8711.1111	8711.1111	3.00	0.0961
Salinity	2	96762.5000	48381.2500	16.66	<.0001
Temp*Salinity	2	300855.1667	150427.5833	51.80	<.0001
Density*Salinity	2	674.3889	337.1944	0.12	0.8909
Temp*Density*Salinit	2	24038.3889	12019.1944	4.14	0.0285

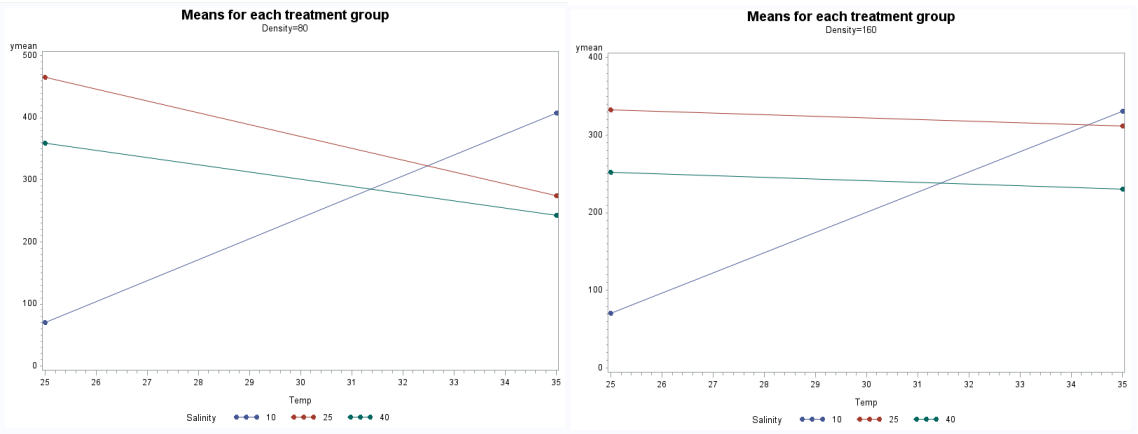
Source	DF	Type III SS	Mean Square	F Value	Pr > F
Temp	1	15376.0000	15376.0000	5.30	0.0304
Density	1	21218.7778	21218.7778	7.31	0.0124
Temp*Density	1	8711.1111	8711.1111	3.00	0.0961
Salinity	2	96762.5000	48381.2500	16.66	<.0001
Temp*Salinity	2	300855.1667	150427.5833	51.80	<.0001
Density*Salinity	2	674.3889	337.1944	0.12	0.8909
Temp*Density*Salinit	2	24038.3889	12019.1944	4.14	0.0285

Three-way interaction is significant implying that all three factors are important. We would now look at simple effects.

For example,

- the effect of temperature (35 degrees vs 25 degrees) at density 80 and salinity 10.
- the effect of density (160 vs 80) at temperature 25 and salinity 40.
- the effect of salinity (40 vs 10) at temperature 35 and density 80.
- Any effects like these that are of interest to the researcher. Key is that we aren't averaging over any of the variables as the three-way interaction is significant. This implies that, for instance, the way temperature effects the response depends on both density and salinity together. Therefore, it doesn't make sense to average across density, salinity, or both when looking at temperature.

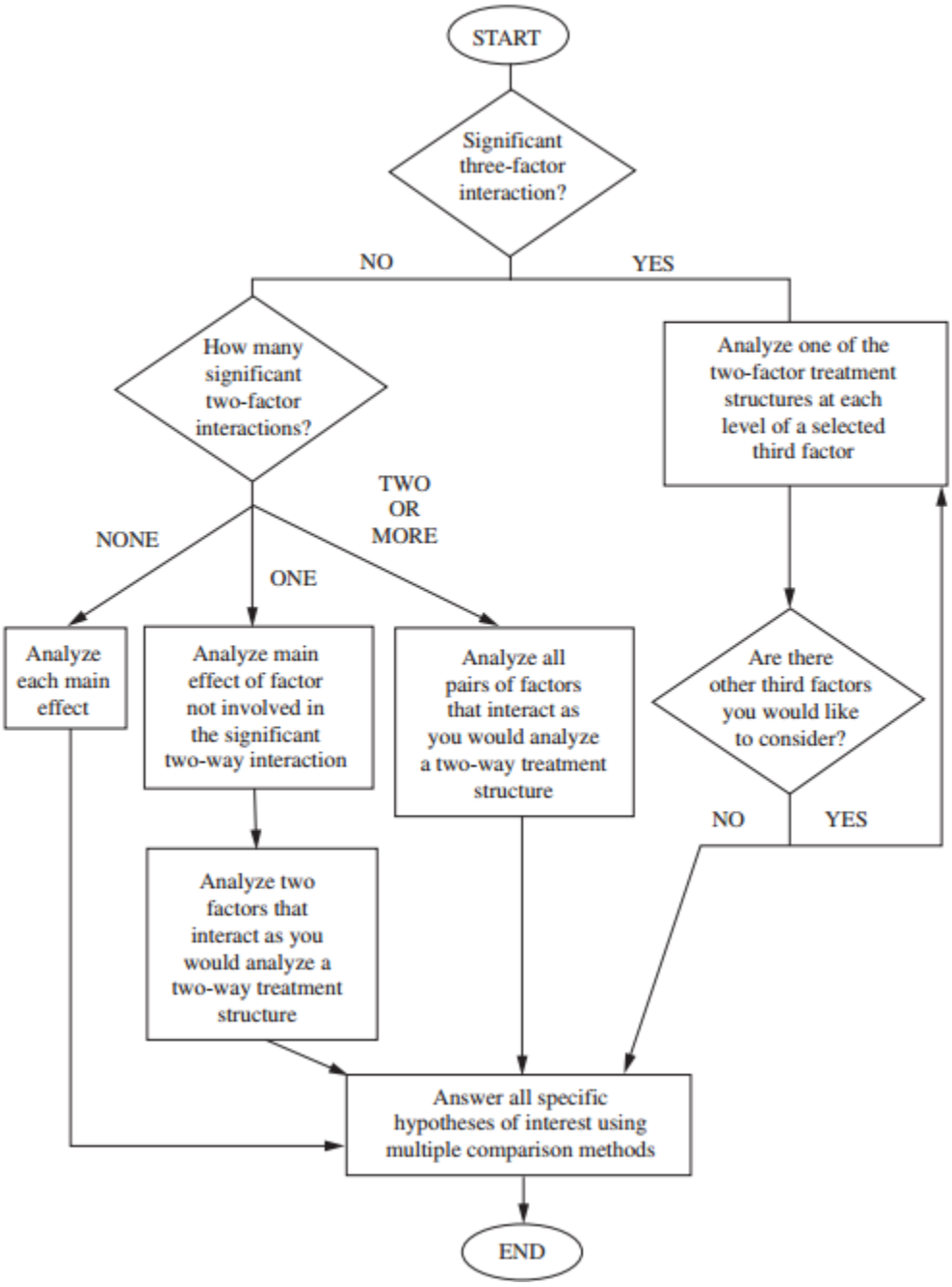
Interpretation of three-way (second order) interaction



Suppose the 3-way interaction was not significant. Then we would proceed to the 2-way interaction p-values and, if those were not significant, the main effect p-values.

For example, we might find the only significant effects were the Temp*Salinity, Salinity, and Density. What should we do here?
We should investigate the simple effects of temperature and salinity (averaged over density). That is, effects such as Temp 35 vs Temp 25 at Salinity 10 (averaged over Density). This could be found by
lsmeans Temp*Salinity/pdiff;
or by writing the appropriate contrast/estimate statements.

We would also look at the main effects of Density. That is, the effect of Density 160 vs Density 80. This could be found by
lsmeans Density/pdiff; or by writing the appropriate contrast/estimate statements.


START
Significant three-factor interaction?
NO
YES
How many significant two-factor interactions?
Analyze one of the two-factor treatment structures at each level of a selected third factor
NONE
ONE
TWO OR MORE
Analyze each main effect
Analyze main effect of factor not involved in the significant two-way interaction
Analyze all pairs of factors that interact as you would analyze a two-way treatment structure
Are there other third factors you would like to consider?
NO
YES
Analyze two factors that interact as you would analyze a two-way treatment structure
Answer all specific hypotheses of interest using multiple comparison methods
END