

## ✧ RNR 613 — Multi-factor ANOVAs

### Review...

### Analysis of Variance

Models to assess means from studies with  $\geq 1$  *factors* that have  $\geq 2$  *levels* that are often qualitative.

### One-factor ANOVA

Appropriate model for completely randomized designs (CRD) where one treatment is assigned to experimental units at random.

### Multi-factor ANOVA

More than one factor examined simultaneously in a single study. Appropriate for completely randomized designs with multiple factors, randomized block designs, etc.

Sleuth views ANOVA as a special case of Regression.

*Note:* For designs with  $>1$  factor, there are no nonparametric equivalents that work well; hence, model checking and appropriate transformations are more important than ever.

*Example:* Response is *plasma calcium concentration* in birds; *factors* considered simultaneously are *hormone trt* and *sex*.

No Hormone Treatment		Hormone Treatment	
Female	Male	Female	Male
16.5	14.5	39.1	32.0
18.4	11.0	26.2	23.8
12.7	10.8	21.3	28.8
14.0	14.3	35.8	25.0
12.8	10.0	40.2	29.3
Cell $\bar{y} = 14.88$	12.12	32.52	27.78

In ANOVAs, responses from each treatment combination can be envisioned as a *cell* in a table (e.g., Male, Not Trt).

		Factor B	
		B1	B2
Factor A	A1	A1, B1	A1, B2
	A2	A2, B1	A2, B2

		Hormone Treatment	
		Not Treated	Treated
Sex	Male	Male, Not Trt	Male, Trt
	Fem	Fem, Not Trt	Fem, Trt

A treatment set is *factorial* if all levels of all treatments are tested with all levels of all other treatments. For example: *Factorial treatment sets* are designated by the number of levels of each factor. Above, each of 2 factors has 2 levels, yielding a  $2 \times 2$  factorial.

With 3 factors: factor A has 3 levels, factor B has 3 levels, factor C has 2 levels, the design is  $3 \times 3 \times 2$ .

When all levels of any two factors are examined in all combinations, then these factors are *crossed*.

A design is *full factorial* when all factors are *crossed* with all other factors at all combinations of factor levels.

Factorials are powerful designs because they allow you to precisely estimate *interactions* among factors.

Multiply the number of levels of each factor to determine the total number of *cells* in a study. A 3 x 3 x 2 factorial has 18 cells.

A treatment set is *balanced* if there is an equal number of replicates for each treatment combination.

If a design is balanced then  $N$  for the study = *cells* x *replicates*; e.g., with 3 replicates per cell and 18 cells,  $N = 18 \times 3 = 54$ .

### Additive Models

ANOVAs *without interactions*, where the effect of one factor has the same effect on the response regardless of the levels of the other factors.

$$\mu \{ \text{plasma} \mid \text{hormone}, \text{sex} \} = \text{hormone} + \text{sex}$$

### Nonadditive Models

ANOVAs *with interactions*, where the effect of one factor on the response depends on the levels of the other factors.

$$\mu \{ \text{plasma} \mid \text{hormone}, \text{sex} \} = \text{hormone} + \text{sex} + \text{hormone} \times \text{sex}$$

### Notation

Uniquely identify every experimental unit in a 2-way design with 3 subscripts:

Let  $y_{ijl}$  denotes the  $l^{\text{th}}$  replicate of the  $i^{\text{th}}$  level of factor  $A$  and the  $j^{\text{th}}$  level of factor  $B$ .

e.g., from the plasma calcium experiment, if *hormone trt* is factor  $A$  and *sex* factor  $B$ , then

$y_{213} = 21.3$  mg/100 ml,  $y_{115} = 12.8$  mg/100 ml, etc.

Each cell formed by the combination of level  $i$  of factor  $A$  and level  $j$  of factor  $B$  has a mean for all of its replicates:  $\bar{y}_{ij}$ .

In the example,  $\bar{y}_{11} = 14.88$ ,  $\bar{y}_{12} = 12.12$ , etc.

The mean of all replicates for level  $i$  of factor  $A$  is  $\bar{y}_{i\cdot}$ , and for level  $j$  of factor  $B$  is  $\bar{y}_{\cdot j}$

e.g., mean for all non-treated birds is  $\bar{y}_1$ .

mean for all treated birds is  $\bar{y}_2$ .

mean for all females is  $\bar{y}_{\cdot 1}$

mean for all males is  $\bar{y}_{\cdot 2}$

### Kinds of Questions

Factor  $A$  effect? Factor  $B$  effect? (*Main effects*)

$A \times B$  interaction? (*Interactive effect*)

**Null hypotheses for these questions are:**

- $H_0$ : No effect of hormone treatment on mean plasma calcium concentration of birds (i.e.,  $\mu_{\text{no hormone}} = \mu_{\text{hormone}}$  or  $\mu_1 = \mu_2$ .)
- $H_a$ : Effect of hormone treatment on mean plasma calcium concentration of birds (i.e.,  $\mu_{\text{no hormone}} \neq \mu_{\text{hormone}}$  or  $\mu_1 \neq \mu_2$ .)
- $H_0$ : No effect of sex on the mean plasma calcium concentration of birds (i.e.,  $\mu_{\text{female}} = \mu_{\text{male}}$  or  $\mu_{\cdot 1} = \mu_{\cdot 2}$ .)
- $H_a$ : Effect of sex on mean plasma calcium concentration of birds (i.e.,  $\mu_{\text{female}} \neq \mu_{\text{male}}$  or  $\mu_{\cdot 1} \neq \mu_{\cdot 2}$ .)
- $H_0$ : No interaction of sex and hormone treatment on the mean plasma calcium concentration of birds
- $H_a$ : Interaction of sex and hormone treatment on the mean plasma calcium concentration of birds.

To test null hypotheses, partition *Total SS* into appropriate sources.

- *Total SS* — variability among all data.

Squared difference between every observation (each factor, level, replicate) and the grand mean with  $abn - 1$  df.

$$\text{Total SS} = \sum_{i=1}^a \sum_{j=1}^b \sum_{l=1}^n (y_{ijl} - \bar{y}_{\cdot})^2$$

- *Among-group, Model or Treatment SS* — variability among cells, with  $ab - 1$  df ( $a \times b$  total cells).

$$\text{Model SS} = \sum_{i=1}^a \sum_{j=1}^b n(\bar{y}_{ij} - \bar{y}_{\cdot})^2$$

$n$  is the no. replicates in a cell.

- *Within-group or Error SS* — variability within cells, with  $ab(n - 1)$  df.

Because there are more than 2 factors, *Model SS* contains the effects of *both* factors which must be separated.

Partition the *Model SS* into SS attributable to *factor A* by considering factor A as the sole factor in a one-factor ANOVA (i.e., ignore factor B), then do the same for factor B (i.e., ignore factor A).

$$\text{within-cell or Error SS} = \sum_{i=1}^a \sum_{j=1}^b \sum_{l=1}^n (y_{ijl} - \bar{y}_{ij})^2$$

- For *factor A* SS, determine the squared difference between the mean for each level  $i$  of *factor A* and the grand mean:

$$\text{Factor A SS} = bn \sum_{i=1}^a (\bar{y}_{i\cdot} - \bar{y}_{\cdot})^2, \text{ with } a - 1 \text{ df.}$$

- For *factor B*:

$$\text{Factor B SS} = an \sum_{j=1}^b (\bar{y}_{\cdot j} - \bar{y}_{\cdot})^2, \text{ with } b - 1 \text{ df.}$$

## Interactions

Two factors interact if the effect that one factor has on the mean response depends on the value of the other factor.

*Does the effect of hormone treatment on plasma calcium levels in birds depend on the sex of the bird?*

There is evidence of interaction when variability among cells  $\neq$  variability among levels of *factor A* plus the variability among levels of *factor B*. That is

$$\text{Model SS} \neq \text{factor A SS} + \text{factor B SS}.$$

This difference in variability is that due to any interaction between *factors A* and *B*.

*Interaction effects* influence the mean response *in addition* to the sum of the effects of each factor considered separately.

$$A \times B \text{ interaction SS} = \text{Model SS} - \text{factor A SS} - \text{factor B SS}$$

$$A \times B \text{ interaction df} = \text{Model df} - \text{factor A df} - \text{factor B df}$$

$$A \times B \text{ interact df} = (\text{factor A df}) (\text{factor B df}) = (a - 1) (b - 1)$$

Whole-Model Test				
Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	3	1461.3255	487.109	21.2727
Error	16	366.3720	22.898	Prob>F
C Total	19	1827.6975		<.0001

JMP output for bird calcium plasma levels...

Need SS's and df's to determine Mean Squares to generate F-statistics for null hypotheses of interest, specifically, *Total SS*, *Cell SS (or Model SS)*, *Sex SS*, *Hormone SS*, *Sex x Hormone SS*, and *Error SS (Within-cells SS)* and their respective df's.

Effect Test					
Source	Nparm	DF	Sum of Squares	F Ratio	Prob>F
Sex	1	1	70.3125	3.0706	0.0989
Hormone Trt	1	1	1386.1125	60.5336	<.0001
Sex*Hormone Trt	1	1	4.9005	0.2140	0.6499

Sex				
Effect Test				
	Sum of Squares	F Ratio	DF	Prob>F
	70.312500	3.0706	1	0.0989

Least Squares Means			
Level	Least Sq Mean	Std Error	Mean
Female	23.70000000	1.513216772	23.7000
Male	19.95000000	1.513216772	19.9500

Hormone Trt				
Effect Test				
	Sum of Squares	F Ratio	DF	Prob>F
	1386.1125	60.5336	1	<.0001

Least Squares Means			
Level	Least Sq Mean	Std Error	Mean
No	13.50000000	1.513216772	13.5000
Yes	30.15000000	1.513216772	30.1500

Sex*Hormone Trt				
Effect Test				
	Sum of Squares	F Ratio	DF	Prob>F
	4.9005000	0.2140	1	0.6499

Least Squares Means		
Level	Least Sq Mean	Std Error
Female,No	14.88000000	2.140011682
Female,Yes	32.52000000	2.140011682
Male,No	12.12000000	2.140011682
Male,Yes	27.78000000	2.140011682

*Sex* and *Hormone* are called *Main Effects* to distinguish them from *interaction* effects

*Effect* is a more general term than *factor* because effects can include other quantities, such as interactions, that are not actual factors specified in an experiment.

**ANOVA F-tests**

Total SS (1827.70, 19 df)

Error SS (366.37, 16 df, MSE = 22.89)

Model SS is SS attributable to **all** treatment effects

Partition *Model* SS into components, including main effects and interaction; note SS's for these 3 effects sum to the *Model* SS.

**Hormone Effect?**

$$H_o: \mu_{\text{no hormone}} = \mu_{\text{hormone}} \text{ OR } \mu_{1\cdot} = \mu_{2\cdot}$$

$$H_a: \mu_{\text{no hormone}} \neq \mu_{\text{hormone}} \text{ OR } \mu_{1\cdot} \neq \mu_{2\cdot}$$

$$F_{1,16} = \text{MST}_{\text{hormone}} \div \text{MSE} = (1386.11/1) \div (366.37/16) = 1386.11 \div 22.898 = 60.53, \text{ with } P < 0.0001$$

Conclude that hormone-treated animals, *regardless of sex*, had a higher mean ( $\bar{y} \pm \text{SE}$ ) plasma calcium level ( $30.15 \pm 1.51$ ) compared to those not treated ( $13.5 \pm 1.51$ ).

**Sex Effect?**

$$H_o: \mu_{\text{female}} = \mu_{\text{male}} \text{ OR } \mu_{\cdot 1} = \mu_{\cdot 2}$$

$$H_a: \mu_{\text{female}} \neq \mu_{\text{male}} \text{ OR } \mu_{\cdot 1} \neq \mu_{\cdot 2}$$

$$F = \text{MST}_{\text{sex}} \div \text{MSE} = (70.31/1) \div (366.37/16) = 70.31 \div 22.898 = 3.07, \\ P = 0.09 \text{ (marginal)}$$

Conclude that females had slightly higher plasma calcium levels ( $23.7 \pm 1.51$ ), *regardless of hormone treatment*, than males ( $20.0 \pm 1.51$ ).

**Interaction?**

$$H_o: \text{No interaction of Sex and Hormone treatment on the mean plasma calcium}$$

$$H_a: \text{There is interaction of sex and hormone treatment on the mean plasma calcium}$$

$$F = \text{MST}_{\text{sex} \times \text{hormone}} \div \text{MSE} = (4.90/1) \div (366.37/16) =$$

$$4.90 \div 22.898 = 0.21, P = 0.65$$

Conclude no obvious *Sex x Hormone* interaction.

**Summary ANOVA Table**

Source of Variation	SS	df	MS	F	P
Cells (model)	1461.33	3			
Hormone (Factor A)	1386.11	1	1386.11	60.50	0.0001
Sex (factor B)	70.31	1	70.31	3.07	0.099
Hormone x Sex	4.90	1	4.90	0.21	0.65
Error (within-cells)	366.37	16	22.90		
Total	1827.70	19			

Interaction is a test of additive versus non-additive models.

The  $F$ -test for interaction is equivalent to an Extra SS  $F$ -test comparing full (main effects and interaction) and reduced models (main effects only).

$$F = (\text{error SS reduced} - \text{error SS full} / \text{diff in df}) / \text{MSE full} = F_{1,16} = (371.37 - 366.27 / 1) 22.90 = 0.21, P = 0.65$$

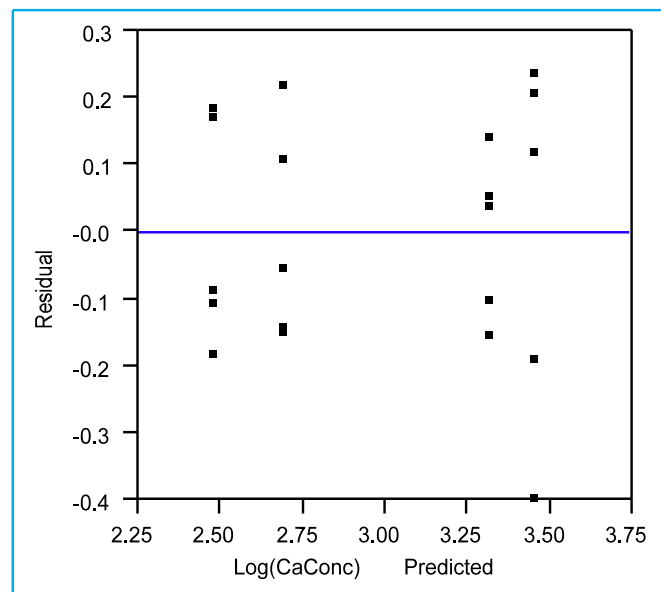
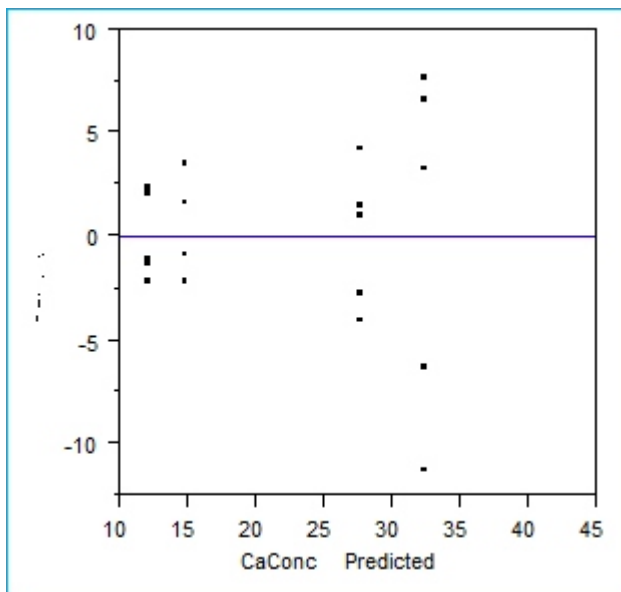
### What if interaction was significant?

Then it is not appropriate to consider the effect of *hormone treatment* on *plasma calcium* without also knowing sex of the bird.

Tests for main effects are no longer meaningful and should not be considered independently.

Assess by examining Cell, Row, and Column means graphically.

	Female	Male	
No hormone	14.9	12.1	13.5
Hormone	32.5	27.8	30.2
	23.7	20.0	



### Check Residuals

To assess model fit and assumptions. Examine (1) the centers, (2) relative spreads, (3) general shapes of each group's distribution, and (4) outliers.

Residuals are the original observations with their cell means subtracted out.

## ✧ RNR 613 — Randomized Block Designs

---

ANOVA analyses and Experimental Design are linked closely.

- **Experimental design** — the way in which treatments are assigned to experimental units.

To determine the type of design, consider the restrictions placed on the randomization...

In the pig-feed experiment we covered previously, animals were assigned treatment groups completely at random.

This describes a **Completely Randomized Design**.

### • Randomized Block Design

When possible, choose experimental units that are as similar as possible.

Sometimes, more experimental units are needed than like-units are available so it is often effective to group similar experimental units into **blocks**, each of which contains the same number of units as there are factor levels.

**Block** — a group of homogeneous experimental units.

Each block usually contains one complete replication of the complete experiment (sometimes more than one).

E.g., Evaluate the effects of *burning intensity* (intense, light, no burning = control) on density of a bird (1 factor with 3 levels).

Need large 25-ha plots, so employ a randomized block experimental design (RBD) and perform experiment in different areas across a region. Each area is a *block*.

**Randomized Block Designs** replicate the entire treatment structure *within* a set of predefined blocks.

Blocks are then be identified as a factor in the analysis so variability among blocks is removed.

*Blocking* is an extension of *pairing* used in paired *t*-test framework, and identifies experimental units that are related to one another.

*Example:* Guinea pigs separated into “blocks” based on the lab where they are kept. 20 guinea pigs randomly placed into 5 groups (blocks) of 4 individuals each; individuals within each block randomly assigned 1 of 4 diets.

Animals are raised on respective diets and the ending body masses:

Block/Lab	Diet				Block $\bar{y}$
	1	2	3	4	
1	7.0	5.3	4.9	8.8	6.50
2	9.9	5.7	7.6	8.9	8.03
3	8.5	4.7	5.5	8.1	6.70
4	5.1	3.5	2.8	3.3	3.68
5	10.3	7.7	8.4	9.1	8.88
Diet $\bar{y}$	8.16	5.38	5.84	7.64	$\bar{y}_{..} = 6.76$

Null hypothesis of interest is one of no *diet effect*.

Note the similarity between the above table and the one we examined for a two-way ANOVA.

Determine necessary SS:

$SS_{\text{Total}}$  and  $SS_{\text{diet}}$  but also  $SS_{\text{block}}$

If block is ignored, differences among blocks then become part of *error SS*.

### Interaction in RBD?

With two factors and a Completely Randomized Design, we assessed interactions between factors.

In RBD, the *block x treatment interaction* provides the estimate of experimental error.

In statistics programs, you must specify *treatment* and *block* effects but *not* the *block x treatment* interaction because it is the error term (when there is only one replicate per block).

In the Guinea pig-diet example, we find *Diet* effect highly significant ( $F_{3,19} = 11.82$ ,  $P = 0.0007$ ) after removing the variation explained by blocks (that due to differences in lab conditions) from experimental error.

### Correct RBD Analysis

Whole-Model Test				
Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	7	90.072500	12.8675	16.6444
Error	12	9.277000	0.7731	Prob>F
C Total	19	99.349500		<.0001

Block				
Effect Test				
Sum of Squares	F Ratio	DF	Prob>F	
62.647000	20.2588	4	<.0001	

Least Squares Means				
---------------------	--	--	--	--

Diet				
Effect Test				
Sum of Squares	F Ratio	DF	Prob>F	
27.425500	11.8252	3	0.0007	

Least Squares Means				
---------------------	--	--	--	--

### Incorrect One-way Analysis

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	3	27.425500	9.14183	2.0337
Error	16	71.924000	4.49525	Prob>F
C Total	19	99.349500		0.1497

Diet				
Effect Test				
Sum of Squares	F Ratio	DF	Prob>F	
27.425500	2.0337	3	0.1497	

Least Squares Means				
---------------------	--	--	--	--

If analyzed incorrectly without accounting for the block effect, *diet* is no longer significant ( $P = 0.15$ ).

Remember, *SS* and *df* like energy: they can never be created or lost, they can only change form.



$SS$  and  $df$  that were partitioned into *Block* in RBD analysis are included as error in the one-way analysis, so *error MS* increased from 0.77 to 4.49 when we did not account for block-block variation.

The probability of finding a treatment effect (power) was reduced.

### ✧ RNR 613 — Nested ANOVAs

When every level of factor one is examined in combination with every level of factor two, the two factors are *crossed*.

Drug	1		2		3	
Source	A	B	A	B	A	B

When some levels of factor one are examined in combination with some levels of a factor two, and the remaining levels of factor one are examined with different levels of factor two, the factors are *nested (or hierarchical)*.

Drug	1		2		3	
Source	A	B	C	D	E	F

In both schematics above, each of the 3 *Drug* levels are examined from 2 levels of supply *Source*.

In the first, the supply *Source* is the same for all levels of *Drug*. These factors are *crossed*.

In the second, the supply *Source* is different for each level of *Drug*. These factors are *nested*. Therefore, *Source* is *nested within Drug*.

For example, if 4 tests were administered to 4 high school classes (i.e., a factor with 4 levels), and 2 of those 4 classes are in high school A, whereas the other two classes are in high school B

The levels of the first factor (4 different tests) are nested in the second factor (2 different high schools).

Levels of nested factors (here *Source*) are often determined at *random* rather than being *fixed*.

In other cases, experiments are designed with nesting so that hypotheses about individual samples can be tested.

Typically, however, inclusion of a random-effects "nested" factor are included to broaden inferences. In the example above, primary interest is in the effect of *Drug*, a fixed-effect factor.

By recognizing and accounting for the effect of supply *Source*, we account for some of the random variation due to *Source* to increase the strength of the hypothesis tested about *Drug*.

ANOVA table if *Source* not identified

Source	df	SS	MS	F	P
Drug	2	61.16	30.58	26.21	0.0002
Error	9	10.50	1.166		
C Total	11	71.66			

ANOVA table with *Source(Drug)* estimated incorrectly as a fixed effect

Source	df	SS	MS	F	P
Model	5	62.66			
Drug	2	61.16	30.58	20.38	0.0021
Source[Drug]	3	1.50	0.33	0.33	0.8022
Error	6	9.00	1.50		
C Total	11	71.66			

ANOVA table with Source(Drug) estimated correctly as a random effect

Source	df	SS	MS	F	P
Model	5	62.66			
Drug	2	61.16	30.58	61.16	0.0037
Source[Drug]	3	1.50	0.50	0.50	0.8022
Error	6	9.00	1.50		
C Total	11	71.66			

In the last table, *Source[Drug]* is used as the error term (denominator for *F*-test) by which the *Drug* effect is tested, not the MS error for the full model. Notice the change in the P-value for *Drug* between fixed and random effect models.

Tests with Random Effects

Source	SS	MS Num	DF Num	F Ratio	Prob>F
Drug	61.1667	30.5833	2	61.1667	0.0037
Source[Drug]	1.5	0.5	3	0.3333	0.8022

### Random Effects Models

In most cases, the particular levels of a factor studied are chosen because they are of direct interest.

E.g., we wish to study the effect of a herbicide treatment at three levels: control, label application rate, and twice label rate.

Groups means are considered fixed, so this is a *fixed-effects model*.

In some cases the levels of a factor are selected as representative of a broader population.

E.g., we wish to study the effect of degree of invasion by a nonnative grass on small mammal density.

If we were interested in particular levels of invasion we could choose a fixed-effect approach.

Instead we could study the degree of invasion in general by choosing samples (and therefore the particular factor levels studied) from among all of those available.

Because factor levels are chosen at random, and the groups means are a random sample from a population of means, this is a *random-effects model*.

When is a random-effects model appropriate?

1. If inference to a larger set is desired
2. If the groups represent a random sample of a larger set

The first question trigger a random-effects approach and the second question justifies the broader inference.

In the fixed-effect model, sample *i* is considered to be a random sample from a normal population with mean  $\mu$  and variance  $\sigma^2$ .

The model has *I* + 1 parameters: *I* means + variance  $\sigma^2$

In the random-effects model, the  $\mu_i$  means are thought to be a random sample from a normal population with mean  $\mu$  and variance  $\sigma_\mu^2$ .

The model has 3 parameters:

- the observed mean  $\mu$ ,
- the within-group variance  $\sigma^2$
- the among-group variance  $\sigma_{\mu}^2$ .

Analysis of a one-way random-effects model involves a test of whether  $\sigma_{\mu}^2$  is zero and an estimate of the ratio  $\sigma_{\mu}^2/(\sigma_{\mu}^2 + \sigma^2)$ .

This ratio is between 0 and 1; it = 0 when there is no among-group variance and it = 1 when there is no within-group variance.

Because the denominator describes the total variance of the measurements, the ratio is the proportion of the total variance of the measurements that is explained by among-group variance.

## ✧ RNR 613 — Repeated-Measures Experimental Designs

---

Appropriate when multiple measurements made on the same experimental unit.

Key: experimental units or sampling units measured more than once.

Note: a) a single characteristic of each unit is measured on more than one occasion

b) multiple characteristics of the same unit are measured

*Repeated measures* are designs where the same characteristic (response variable) is measured on multiple occasions, often through time or under different conditions.

Experimental units often called “subjects” in these designs.

In essence, each experimental unit is treated as its own block.

Like blocks, effects about specific subjects may or may not be of interest to the researcher, but they must be included in the model so that SSs are partitioned appropriately.

Other names for RM designs:

✧ *within-subject* designs

✧ *treatment-by-subject* designs

### Types of Repeated Measures Studies

*Longitudinal Studies* — Repeated measurements are of the same variable on the same subject at different points in time.

*Crossover Experiments* — Each subject receives more than one treatment; the order in which the treatments are received by each subject is randomized.

*Split-Plot Experiments with Repeated Measurements* — Examine two separate treatments, with two different randomizations; have features of both *longitudinal* and *crossover* experiments.

*First* randomization allocates experimental units to different levels of the first treatment.

*Second* randomization determines the temporal sequence or spatial location for *all* levels of a second treatment.

*Why might these designs be beneficial?*

When all subjects can receive all treatments, then the benefits parallel those of randomized blocks.

Specifically, when there is substantial variation among subjects, more precise estimates of treatment effects are achieved by within-subject comparisons (removes between-subject variation).

*Oat-bran example in Sleuth...*

Assess the effect of oat bran on cholesterol. For each subject a baseline cholesterol level was measured, then they received a high-fiber and low-fiber diet, each for 6 weeks.

### Univariate Approach: One Experimental Factor

*Example:* Assess the effects of 3 drugs on blood cholesterol (mg) in animals. All 7 subject were treated with all 3 drugs for some period of time; the order was randomized.

Null and alternate hypotheses for this 1-factor repeated-measures experiment are:

- $H_o$ : Mean cholesterol is the same on all 3 drugs ( $\mu_1 = \mu_2 = \mu_3$ ).
- $H_a$ : Mean cholesterol differs among drugs.

Data can be tabulated where  $i$  columns represent  $i$  different treatments, and  $j$  rows represent  $j$  subjects:

Subject	Drug 1	Drug 2	Drug 3	Total ( $S_j$ )
1	164	152	178	494
2	202	181	222	605
3	143	136	132	411
4	210	194	216	620
5	228	219	245	692
6	173	159	182	514
7	161	157	165	483
Total ( $G_i$ )	1281	1198	1340	3819

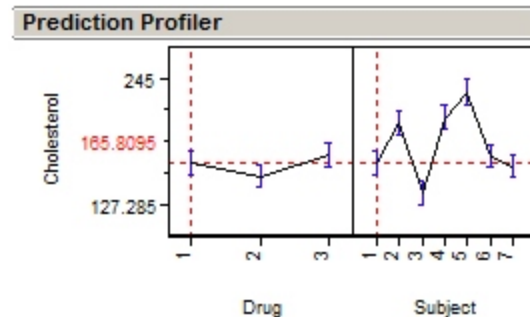
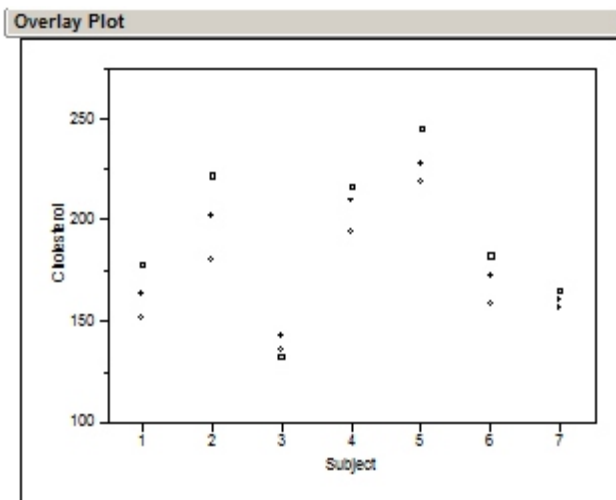
$S$  = sum of responses from all treatments  $i$  for each Subject  $j$ ,  
 $G$  = sum of responses from all subjects  $j$  for each treatment Group  $i$

Grand total of all data  $\sum\sum y_{ij}$  equals both  $\sum G_i$  and  $\sum S_j$

Number of subjects ( $j$ ) is the number of experimental units ( $n$ ) and the number of total observations in the study is  $N$  ( $N = n \times i$ ).

**Profile Plot**

Profile plots illustrate variation *within subjects*...



Drug	$n$	mean	SE	
1	7	183		11.61
2	7	171.14	10.76	
3	7	191.43	14.56	

## Partition the Sum of Squares

Calculate the *SS Total*, which is the sum of the squared differences between  $y_{ij}$  and  $\bar{y}_{..}$ .

Next, *SS Total* into *SS among-subjects* and *SS within-subjects*:

$$\text{SS among - subjects} = \sum_{j=1}^n \frac{s_j^2}{i} - C$$

$C$  is a correction factor ( $=\sum\sum y_{ij}^2/N$ ); bookkeeping to subtract a quantity included in the calculation of both *within-subject* and *among-subject* *SS*.

*SS within-subject* is found by subtraction:  $\text{SS within-subject} = \text{SS total} - \text{SS among-subject}$

*Total df*,  $N - 1$ , are partitioned between

$$\begin{aligned} \text{among-subject df, which} &= n - 1 \\ \text{within-subject df} &= \text{total df} - \text{among-subject df} = n(i - 1). \end{aligned}$$

Lastly, partition the *within-subject* variability into *treatment SS*

$$\text{treatment SS} = \sum_{i=1}^i \frac{G_j^2}{n} - C$$

and that unexplained by both *treatments* and *subjects*:

$$\text{SS remainder} = \text{SS within-subjects} - \text{SS treatment}$$

*SS remainder* is the *treatment x subject interaction* (conceptually identical to a *block x treatment interaction* in a RBD), which is the all-important estimate of experimental error, or *SS error*.

*SS remainder* (or *SS error*) measures how similarly *subjects* responded to *treatments*.

The associated *df* are:

$$\begin{aligned} \text{treatment df} &= i - 1 \\ \text{remainder df} &= \text{within-subject df} - \text{treatment df} = (i-1)(n-1). \end{aligned}$$

The appropriate *F*-statistic for *treatment* effect is:

$$F = \text{MS treatment} / \text{MS error}.$$

The test for a *subject* effect is usually not relevant to the null hypothesis, but is:  $F = \text{MS among-subject} / \text{MS error}$ .

To use JMP (or any other package) to calculate *F*-tests with the appropriate *SS*'s and *df*'s:

- identify the response ( $Y$ ),
- identify *subjects* (*among-subjects*) as a blocking factor
- identify *treatment* (*within-subjects*).

By identifying *subjects* as a factor, the *SS* attributable to differences *among-subjects* is removed from experimental error.

The *within-subjects* *SS* then contains both the *treatment* effect and the remainder (the *drug x subject interaction* which is estimate of error used for testing).

**Analysis of Variance**

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	8	20185.238	2523.15	43.5444
Error	12	695.333	57.94	Prob > F
C. Total	20	20880.571		<.0001

**Effect Tests**

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Subject	6	6	18731.238	53.8770	<.0001
Drug	2	2	1454.000	12.5465	0.0011

With these components, build the following ANOVA table:

<i>Source of Variation</i>	SS	df	MS	<i>F</i>	<i>P</i>
Among-subjects	18731.24	6			
Within-subjects	2149.33	14			
Drugs	1454.00	2	727.00	12.6	0.0011
Error (Subject x Drugs)	695.33	12	57.94		
Total	20880.57	20			

Conclude there is strong evidence for a drug effect on cholesterol ( $F_{2,20} = 12.6$ ,  $P = 0.0011$ ).

Had we not identified subjects in the model, the treatment effect would not have been judged significant at any reasonable  $\alpha$  ( $F_{2,20} = 0.67$ ,  $P = 0.52$ ).

Note that using the univariate approach above, data are entered as one measurement per row... such as:

<i>Subject</i>	<i>Choles</i>	<i>Drug</i>
1	164	1
2	202	1
3	143	1
4	210	1
5	228	1
6	173	1
7	161	1
1	152	2
2	181	2
3	136	2
4	194	2
5	219	2

...

**Multivariate Approach: Same Example**

In multivariate analyses, multiple measurements of the response ( $Y_1, Y_2, Y_3$ ) for each subject are considered simultaneously and are entered on the same row in a data table.

In JMP, choose all responses, select the "MANOVA" personality, then "Run Model."

In the subsequent *Manova Fit* dialog, you must specify the methods to assess interrelationships among  $Y$ 's (*Response Specification*), which are sets of *linear contrasts* among  $Y$ 's.

Common response designs (*M*-matrices) include:

<i>Repeat. Meas.</i>	longitudinal designs; it is both <i>Sum</i> and <i>Contrast</i> responses.
<i>Sum</i>	sum of all responses, one value
<i>Identity</i>	each response, the identity matrix
<i>Contrast</i>	each response compared to the first
<i>Helmert</i>	each response versus the ones after it, except the last
<i>Profile</i>	each response versus all others, except the last
<i>Mean</i>	each response versus the mean of the others, except the last
<i>Custom</i>	any set of contrasts you choose

In an example with 3 measurements per subject:

*Sum* yields:

<i>Drug1</i>	<i>Drug2</i>	<i>Drug3</i>
1	1	1

*Contrast* yields:

<i>Drug1</i>	<i>Drug2</i>	<i>Drug3</i>
-1	1	0
-1	0	1

*Identity* yields:

<i>Drug1</i>	<i>Drug2</i>	<i>Drug3</i>
1	0	0
0	1	0
0	0	1

*Repeated*, which is equivalent to specifying is both *Sum* [between-subject effects] and *Contrast* [within-subject effects]) yields both:

*Between subjects*

<i>Drug1</i>	<i>Drug2</i>	<i>Drug3</i>
1	1	1

*Within subjects*

<i>Drug1</i>	<i>Drug2</i>	<i>Drug3</i>
-1	1	0
-1	0	1

The appropriate response specification depends on your question and will affect your results.

In the Drug example,

Results for a *repeated measures* specification are:

*Drugs (within subjects)*

Test	Value	Exact F	NumDF	DenDF	Prob>F
F Test	5.1620662	12.9052	2	5	0.0106



Results for an *identity* specification are:

*Intercept*

Test	Value	Exact F	NumDF	DenDF	Prob>F
F Test	65.658749	87.5450	3	4	0.0004

The M-transformed parameter estimates, which in this case are mean responses for each drug, depend on the response specification, which is why the result depends on the M-matrix specified:

	<i>Drug1</i>	<i>Drug2</i>	<i>Drug3</i>
Intercept	183	171.14	191.43

### Univariate or Multivariate Approaches?

Univariate repeated-measures analyses assume that treatment responses are equally correlated among pairs of subjects in the study.

That is, responses of subject 1 and 2 should be correlated with those of subject 2 and 3, and so forth. This is called *compound symmetry*, which when coupled with the usual ANOVA assumption of homogeneity of variances among groups is called *sphericity*.

The *Sphericity* assumption is necessary for using univariate *F*-tests for *within-subject* effects.

If violated, and errors are confirmed to be nonspherical, then an *F*-test with *Greenhouse-Geisser* or *Huynh-Feldt* adjusted degrees of freedom can be used.

In a multivariate setting, all responses from the same subject are considered simultaneously, and the correlation among responses accounted for by the test statistic. In the two-response case:

$$\text{Hotelling's } T^2 = \frac{t_1^2 + t_2^2 - 2Rt_1t_2}{1 - R^2}$$

where  $t_1$  and  $t_2$  are Student's *t*-ratios for each response, and  $R$  is the sample correlation coefficient between responses. When multiplied by a degrees-of-freedom factor, Hotelling's  $T^2$  has an *F*-distribution.

Hence, multivariate repeated-measures analyses *do not* have the same restrictive assumptions as univariate analyses. Plus, multivariate tests are the only type that are appropriate for true multivariate responses.