Multivariate ANOVA: Advanced Designs

Multivariate ANOVA: Advanced Designs

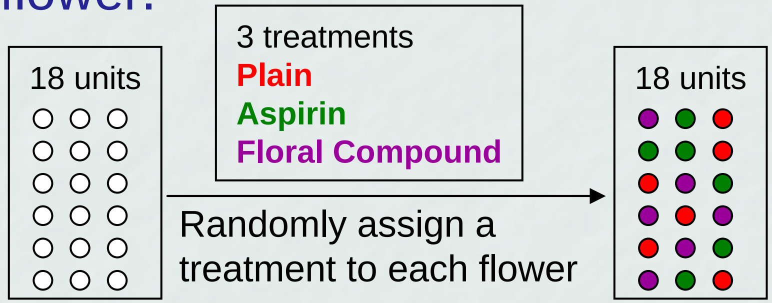
The previous tutorial introduced **Analysis of Variance (ANOVA)** by discussing one type of design structure, the **randomized basic factorial design**. This tutorial will discuss more advanced design structures. To develop the appropriate ANOVA for advanced designs, it is necessary to answer the following questions:

1) Is structure of the design:
   - a Complete Randomized Design (CR) also called the Randomized Basic Factorial Design or Factorial Design
   - a Block Design, or
   - a Split Plot/Repeated Measure (SP/RM) Design?
2) Is each factor **crossed** or **nested**?
3) Is each factor **fixed** or **random**?

Three Types of Design Structures

In **Completely Randomized/Factorial Designs**, each **condition** (specific combination of factor levels) is randomly assigned to an **Experimental Unit (unit)**.

Flower Example 1: Students in an introductory statistics class tested the impact of water solutions on the longevity of cut flowers. They purchased 18 carnations and randomly assigned one of three treatments (plain water, one aspirin added to the water, and a floral compound provided by the flower shop) to each flower.



Factor: water solution with **Levels**: plain, aspirin, and floral compound

Unit or EU (experimental unit): each of the 18 flowers

Response: longevity (i.e. number of days until flower starts to wilt)

Null hypothesis: Solution of water makes no difference in the longevity of carnations.

Three Types of Design Structures

This CR design could be analyzed with a 1-factor ANOVA, with factor A, water solution, having three levels. An F-statistic would be used to determine if $\mathbf{MS_A}$, the variability between the level means of factor A, is large compared to the mean square error (**MSE**), the general flower-to-flower variability. Before this study can extend to the general population of all carnations, the students need to also address other issues:

1) What care did they take in the experimental process to ensure that other external factors did not influence the longevity? For example, were the carnation stems cut at the same time and in the same way? Were the carnations exposed to the same amount of sunlight, temperature, and humidity levels?

2) Is their sample truly representative of the the entire population? If their sample was all the same color or from the same store, it is very likely that the MSE they calculate from their sample will show less variability than a true random sample from the entire population. If the MSE is not accurate, the ANOVA is not valid.

In order to draw conclusions from any experiment, it is essential to realize that the design and procedures are just as important (if not more important) than any statistical calculations.

Three Types of Design Structures

To address these issues, it is important to write very clear procedures before the experiment begins. For example, since flowers absorb moisture through the stem, the angle at which a stem is cut is known to impact flower longevity. Unless cut angle is another variable in their experimental design, it is important that this potential nuisance variable is kept as consistent as possible.

These students also chose to restrict their experiment to only white carnations. Since it is impossible to select a true random sample from all carnations sold on a particular date, the students had purchased 6 white carnations from three different stores. While not perfect, this is a very practical approach to account for population variability of white carnations. Store type clearly was not a factor of interest, but it could impact the results. Thus it is appropriate to include the nuisance factor, Store, in the model and analyze the data using a (Randomized) Block Design instead of a Completely Randomized Design.

Three Types of Design Structures

Blocking is the process of grouping units based on some pre-existing similarity that might impact the results. *Units can be sorted, reused, or subdivided to create a block.*

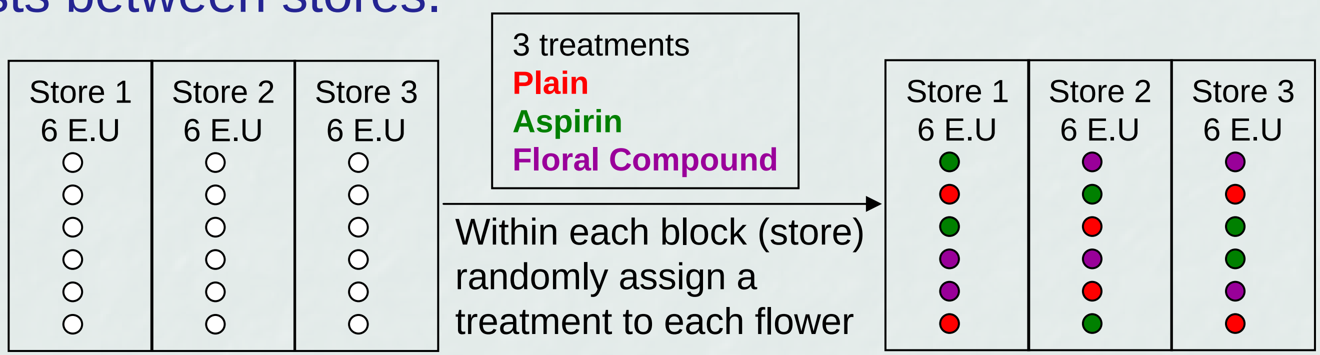
In factorial designs, a treatment is a specific combination of predetermined factor levels that is assigned to an EU. However, there are many situations in which a study also includes **nuisance factors** (factors that may impact the results but not be of specific interest in the study). Blocking incorporates nuisance factors into the design in order to provide more accurate results.

Block effects can be of interest in a study, however since blocks are pre-existing conditions and thus not assigned to EU, there is no causation. Even without proving causation, blocking is beneficial because it can increase the efficiency of a design by accounting for some of the model variability. The Mathematical Calculations section will describe that including the blocking factor may reduce the experimental error (MSE) and thus help identify other factors of interest as significant.

Three Types of Design Structures

Block Designs restrict the way in which the conditions are assigned. Units are placed into groups (or blocks) of similar units. Units within each group are assumed to have some similarity that may impact the results. ***Within each block***, treatments are randomly assigned to one unit.

Flower Example (continued): The random assignment of treatments to units was done within each block. Since there are an equal number of treatments for each store, the effect of water solutions is not biased by store type. In addition, the students are able to measure the variability that exists between stores.

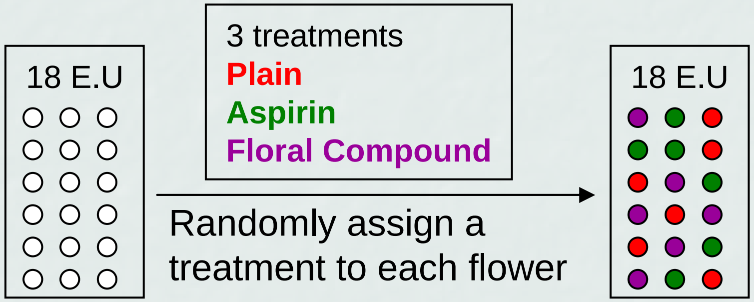


*The Mathematical Calculations section will show that while this is a block design, the ANOVA is identical to a 2-factor ANOVA with no interaction term. One F-test for store effect and another for water solution effect.

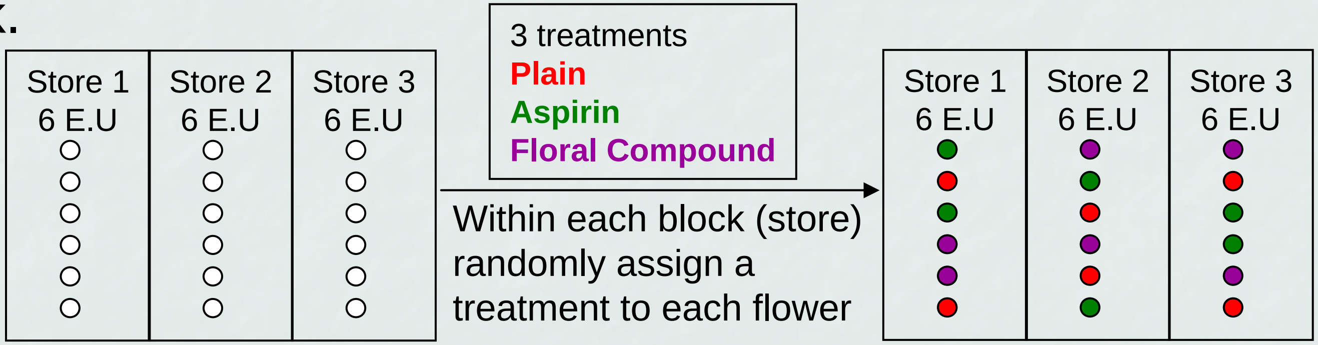
Three Types of Design Structures

Before the third design structure is discussed, it is important to understand the difference between replications and repeated measures. **Replications** occur when each condition is assigned to more than one unit.

In the factorial design in Example 1, each condition had 6 replicates (6 flowers units) that was assigned to each level of the factor solution.



In the block design in Example 1, each condition had 2 replicates (2 flower units) assigned to each level within each block.

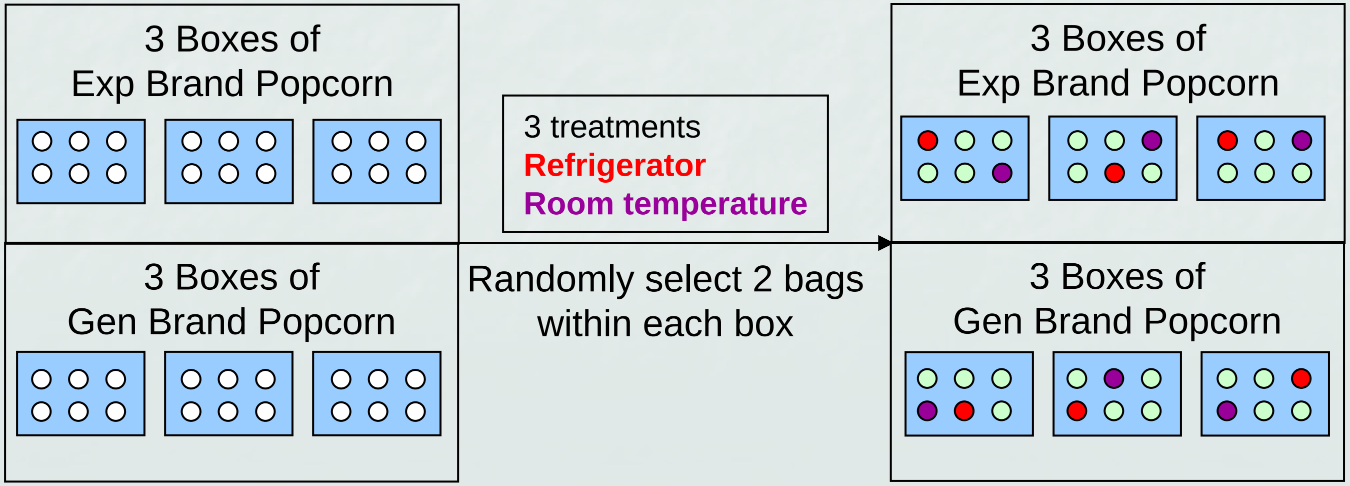


Repeated Measures occur when multiple conditions are assigned to one unit. Thus replications have one measurement for each unit while repeated measurements have multiple measurements on one unit.

Three Types of Design Structures

Split Plot/Repeated Measures Designs have at least two sizes of units in one design. A condition is assigned to a whole plot unit and then the whole plot unit is reused or subdivided into subgroups (split plot units) which also receive a condition. The whole plot units act as blocks for the split plot units.

Popcorn Example 2: To test the effect of storage temperature and brand on the percentage of popped kernels, a student purchased three boxes of both an expensive (exp) and generic (gen) popcorn brand. Each box contained six microwavable bags. Two bags were randomly selected from each box and stored for one week, one in the refrigerator (frig) and the other at room (room) temperature. The bags were popped in random order and the popped and un-popped kernels were counted.



Three Types of Design Structures

Whole Plot Factor: brand **Whole Plot Unit (Blocks)**: Box

Split Plot Factor: storage temperature **Split Plot Unit**: Bag

Response: % popped kernels in a bag

Since boxes were randomly selected from each brand population, the box-to-box variation should be measured by the experimental error (whole plot MSE) within brands. There are three replicates (three boxes) for each brand.

The popcorn bags *within a box* are repeated measures (not replicates) because the bag-to-bag variability with a box is not representative of the population variability. The bags within each box are likely to be handled by the same person, at the same time and at the same location. The bags within a box are considered as sub plot units, because the box effect is likely to have an impact on the bags selected within the box.

To test the effect of storage temperature, bags were randomly assigned to a treatment (frig or room) so bags are the appropriate experimental error (split plot MSE) to measure the temperature effect.

Crossed Vs. Nested Effects

Factors A and B are **crossed** if every level of A can occur in every level of B. Factor B is **nested** in factor A if levels of B only have meaning within specific levels of A.

In *Factorial Designs,* all factors of interest are crossed and there are no repeated measures. *Block Designs* can have either crossed or nested factors. Units are always nested within blocks.

Flower Example (continued): Store and water solution are crossed factors. Since the same water solution is assigned to flowers from each store, the effects of water, aspirin, and floral compound have meaning across stores and each store effect can also be calculated.

Six flowers (units) are nested within each of the three stores. The first flower purchased from Store 1 is not expected to have any relation to the first flower purchased from Store 2. So finding a Flower 1 effect across all three stores is meaningless.

Crossed Vs. Nested Effects

Split Plot Designs typically have both crossed and nested effects.

Popcorn Example (continued): The whole plot unit (boxes) are nested in brand. The three boxes (B1, B2, B3) appear only under the expensive level of factor A (brand) and the next three boxes (B4, B5, B6) appear only under the generic level of factor A. In many texts, B4 ,B5, and B6 are also labeled B1 ,B2, and B3, but it is understood that the B1 occurring in expensive is different from the B1 occurring in generic.

Bags are nested within boxes. Each bag can only come from one box.

Storage Temperatures are crossed with brand. Each temp [room (T1) and frig (T2)] occurs in each brand. Since these factors are crossed, room (T1) is the same under both the expensive and the generic brands.

Storage Temperatures are also crossed with box, but this interaction effect is of no interest and typically not shown in an ANOVA table.

Calculating Crossed Vs. Nested Effects

The calculations for effect size depend on whether a factor is crossed or nested. These calculations do not depend on whether the factor is a factor of interest or a nuisance factor.

As shown in the ANOVA: Full Factorial Design tutorial, all **crossed effects are calculated by finding the appropriate average and subtracting the partial fit**. All nested effects are also calculated by finding the appropriate average and subtracting the partial fit. However, partial fits in nested factors include the factor level in which it is nested.

Flower Example (continued): The effect of "aspirin solution" is the average result of all flowers treated with the aspirin water solution minus the grand mean. The "store 3" effect is the average result of all flowers purchased from store 3 minus the grand mean.

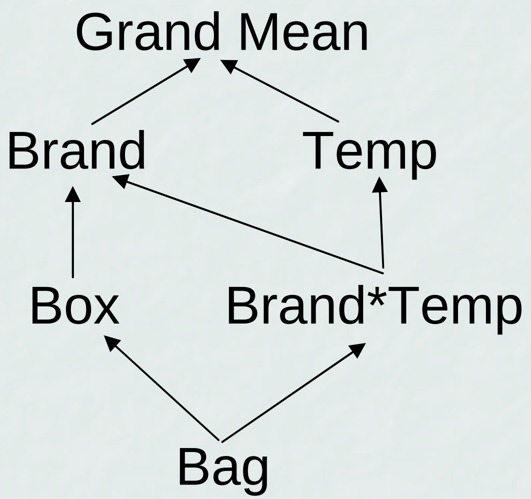
Flower (unit) is nested within store. The effect of the 1st flower from Store1 is calculated:

(average of flower 1 within store1) - (store1 effect + grand mean)

Note that the flower effect is used to calculate MSE. In this study, each flower is the unit and the average is just the observed result for that single flower.

Calculating Crossed Vs. Nested Effects

Popcorn Example (continued): We can visualize the design structure of any balanced model with a Hasse (pronounced hahs) diagram. Interaction terms are listed below the main effects and arrows point from the interaction to the main effects. Arrows also point from nested factors up to the factors in which they are nested.



Temperatures and brand are crossed: only the grand mean is included in their partial fit.

Boxes are nested in brand: brand and grand mean are included in their partial fit.

The brand by temp interaction includes both brand and temp. Thus the partial fit for this term includes the brand, temp and grand mean effects.

Bags are nested within boxes (and so also is necessarily nested within **brand**). Bags are also randomly selected within temp. The partial fit for bag includes the box, brand, brand*temp, temp and grand mean effects.

Hasse Diagrams will be explained in more detail later in this tutorial. This example simply is used to visualize the relationship between all factors in the experiment.

Calculating Crossed Vs. Nested Effects

The tables shows a slightly modified data set for the Popcorn Example. **Brand** and **temp** effects are found by subtracting the grand mean from the appropriate averages.

Brand	Box	Temp	%Popped
exp	1	room	84
exp	1	frig	76
exp	2	room	86
exp	2	frig	86
exp	3	room	91
exp	3	frig	84
gen	1	room	74
gen	1	frig	87
gen	2	room	84
gen	2	frig	83
gen	3	room	83
gen	3	frig	90

Averages			
Brand	%Popped	Temp	%Popped
exp	**84.5**	room	83.67
gen	83.5	frig	84.33

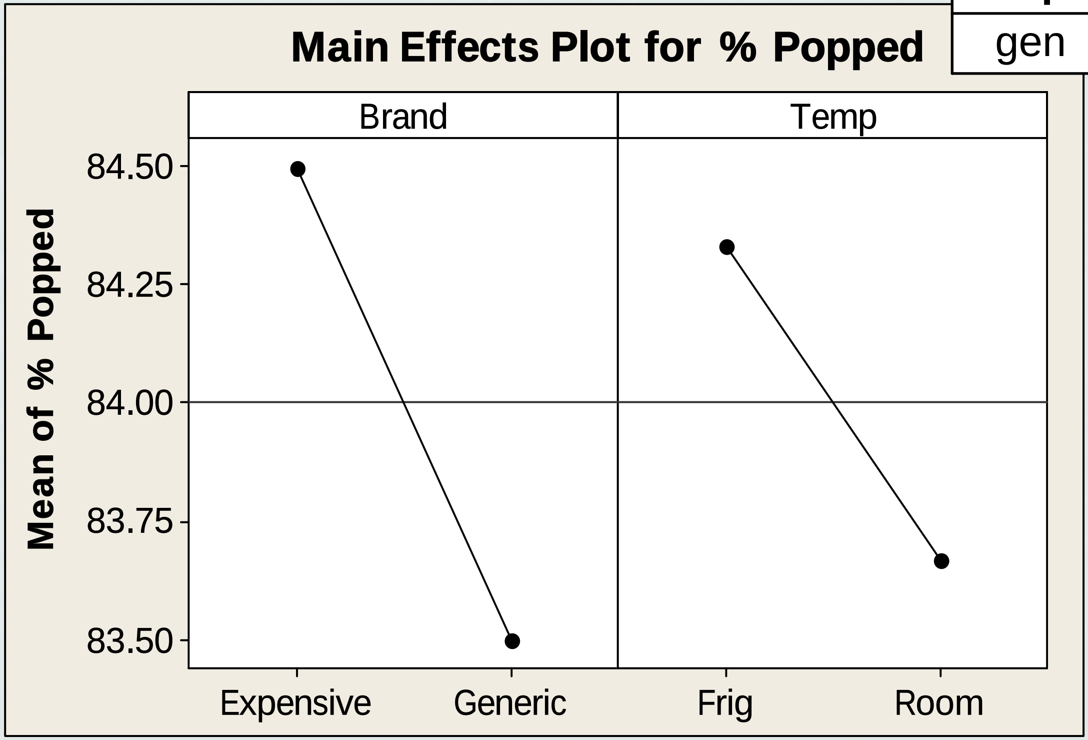
Grand Mean = 84

Main Effects			
Brand	%Popped	Temp	%Popped
exp	**.5**	room	-.33
gen	-.5	frig	.33

Calculating Crossed Vs. Nested Effects

The Main Effects plot shows that the effect of brand is larger than the effect of temp. In this sample, the expensive brand did better than generic and refrigerated bags did better than room temperature bags.

Main Effects			
Brand	%Popped	Temp	%Popped
exp	**.5**	room	-.33
gen	-.5	frig	.33

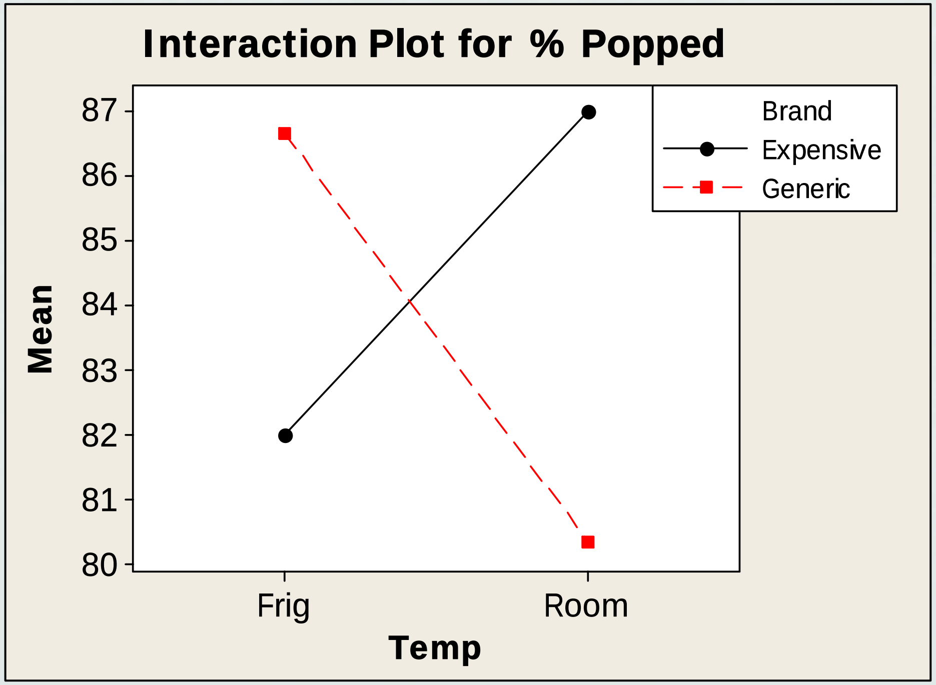


Even though the graph appears to show a difference between levels, we do not know at this time whether these differences are significant. In other words, if there really is no difference in brand, how often would we expect effects this large in a random sample?

Calculating Crossed Vs. Nested Effects

The brand by temp interaction is also found with the following formula: Level average - (brand effect + temp effect + grand mean).

Factor	Level Average	Brand Effect	Temp Effect	Grand Mean	Interaction Effect
exp room	**87**	**.5**	-.333	84	**2.83**
exp frig	**82**	**.5**	.333	84	**- 2.83**
gen room	80.3333	-.5	-.333	84	- 2.83
gen frig	86.6667	-.5	.333	84	2.83



Calculating Crossed Vs. Nested Effects

Referring back to the Hasse diagram, the effects of box and bag factors still need to be calculated. Since each box only has meaning within a brand, there are 6 box averages that need to be calculated

Brand	Box	Temp	%Popped
exp	1	room	84
exp	1	frig	76
exp	2	room	86
exp	2	frig	86
exp	3	room	91
exp	3	frig	84
gen	1	room	74
gen	1	frig	87
gen	2	room	84
gen	2	frig	83
gen	3	room	83
gen	3	frig	90

Brand	Box	Box Average	Brand effect	Grand Mean	Box Effect
exp	**1**	**80**	**.5**	**84**	**- 4.5**
exp	**2**	**86**	**.5**	**84**	**1.5**
exp	**3**	**87.5**	**.5**	**84**	**3**
gen	1	80.5	-.5	84	-3
gen	2	83.5	-.5	84	0
gen	3	86.5	-.5	84	3

Crossed effects always sum to zero. Nested effects (box) ***also sum to zero within each appropriate factor level*** (brand). Box B1, B2, and B3 effects sum to zero within the exp brand. Box B1, B2, and B3 effects sum to zero within the gen brand.

Crossed Vs. Nested Effects

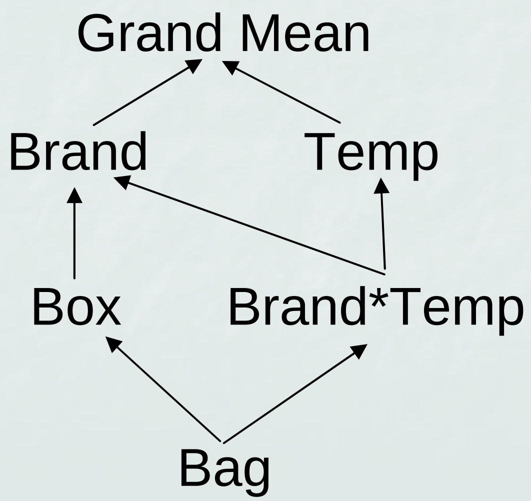
Since **bags** are the units in this study, the **bag** effect is the same as a residual effect. To calculate the ***residual effect***, subtract all other effects from the bag average (observed % popped from each bag).

Brand	Box	Temp	% Popped	Brand Effect	Temp Effect	Brand*Temp Effect	Box Effect	Grand Mean	Bag Effect
exp	1	room	84	**.5**	-.333	**2.833**	**- 4.5**	84	1.5
exp	1	frig	76	**.5**	.333	**- 2.833**	**- 4.5**	84	-1.5
exp	2	room	86	**.5**	-.333	**2.833**	**1.5**	84	-2.5
exp	2	frig	86	**.5**	.333	**- 2.833**	**1.5**	84	2.5
exp	3	room	91	**.5**	-.333	**2.833**	**3**	84	1
exp	3	frig	84	**.5**	.333	**- 2.833**	**3**	84	-1
gen	1	room	74	-.5	-.333	- 2.833	-3	84	-3.333
gen	1	frig	87	-.5	.333	2.833	-3	84	3.333
gen	2	room	84	-.5	-.333	- 2.833	0	84	3.667
gen	2	frig	83	-.5	.333	2.833	0	84	-3.667
gen	3	room	83	-.5	-.333	- 2.833	3	84	-0.333
gen	3	frig	90	-.5	.333	2.833	3	84	0.333
Effect sizes still sum to 0				0	0	0	0	0	0

Calculating Crossed Vs. Nested Effects

In Summary:

- All effect sizes are calculated by finding the appropriate average and subtracting the partial fit.
- Partial fits depend on whether a factors are crossed or nested.
- Hasse diagrams are helpful in visualizing complex design structures.



Mathematical Calculations

Effects show the impact of each factor combination and identify which factors are most influential in our sample. However, a statistical hypotheses test is needed in order to determine if any of these effects are **significant**. Each row corresponding to a factor of interest in the Analysis of Variance (ANOVA) consists of hypothesis tests to determine if there is statistical evidence that the effects are non-zero.

While effect size calculations vary depending on whether the factor is crossed or nested, the following calculations are used for all terms in all **balanced** designs:

- Sum of Squares (SS) = sum of all the squared effects
- Degrees of Freedom (df) = number of free units of information
- Mean Square (MS) = SS/df for each factor

In SP/RM designs, there are multiple unit sizes and each unit size has an experimental error (residual) term. The appropriate denominator (MSE) in the F tests will depend on the three initial questions: 1) design structure, 2) crossed vs. nested factors, and 3) fixed vs. random factors.

- Mean Square Error (MSE) = pooled variance of sample units within each level
- F statistic = (MS for each factor)/(appropriate MSE)

Mathematical Calculations

Sum of Squares (SS) is calculated by summing the squared factor effect for each run, $SS = \sum_{i=1}^{N} (effect)^2$. For Example 2:

Brand Effect	Temp Effect	B*T Effect	Box Effect	Bag Effect
.5	-.33	**2.83**	**- 4.5**	1.5
.5	.33	**- 2.83**	**- 4.5**	-1.5
.5	-.33	**2.83**	**1.5**	-2.5
.5	.33	**- 2.83**	**1.5**	2.5
.5	-.33	**2.83**	**3**	1
.5	.33	**- 2.83**	**3**	-1
-.5	-.33	- 2.84	-3	-3.33
-.5	.33	2.84	-3	3.33
-.5	-.33	- 2.84	0	3.67
-.5	.33	2.84	0	-3.67
-.5	-.33	- 2.84	3	-0.33
-.5	.33	2.84	3	0.33

	Brand Effect Squared	Temp Effect Squared	B*T Effect Squared	Box Effect Squared	Bag Effect Squared
	0.25	0.11	8.03	20.25	2.25
	0.25	0.11	8.03	20.25	2.25
	0.25	0.11	8.03	2.25	6.25
	0.25	0.11	8.03	2.25	6.25
	0.25	0.11	8.03	9.00	1.00
	0.25	0.11	8.03	9.00	1.00
	0.25	0.11	8.03	9.00	11.11
	0.25	0.11	8.03	9.00	11.11
	0.25	0.11	8.03	0.00	13.44
	0.25	0.11	8.03	0.00	13.44
	0.25	0.11	8.03	9.00	0.11
	0.25	0.11	8.03	9.00	0.11
Sum of Squares	3.00	1.33	96.33	99.00	68.33

Mathematical Calculations

Degrees of Freedom (df) = number of free units of information. In the popcorn example, there are 2 levels of **brand** and the ANOVA assumptions require that the effects sum to 0. Knowing the effect of the generic brand automatically forces a known expensive brand effect.

a = # of levels in brand, b = # of levels in temp, c = # of levels of bags within each brand

Brand Effect	Temp Effect	B*T Effect	Box Effect	Bag Effect
.5	-.33	**2.83**	**- 4.5**	**1.5**
.5	.33	**- 2.83**	**- 4.5**	**-1.5**
.5	-.33	**2.83**	**1.5**	**-2.5**
.5	.33	**- 2.83**	**1.5**	**2.5**
.5	-.33	**2.83**	**3**	**1**
.5	.33	**- 2.83**	**3**	**-1**
-.5	-.33	- 2.83	-3	-3.33
-.5	.33	2.83	-3	3.33
-.5	-.33	- 2.83	0	3.67
-.5	.33	2.83	0	-3.67
-.5	-.33	- 2.83	3	-0.33
-.5	.33	2.83	3	0.33

For factors not nested in any other factors, the df is the number of levels minus one.

$df_{Brand} = df_A = a - 1 = 2\text{-}1 = 1$

$df_{Temp} = df_B = b - 1 = 2\text{-}1 = 1$

For nested factors, restrictions in ANOVA require that ***all nested effects sum to zero within each level*** of the factor it is nested in. **Box**, factor C, is nested in **brand**. The three boxes in the expensive brand need to sum to 0. If two box effects in expensive are known, the third box effect is fixed. There are c-1 pieces of free information for every level of brand.

$df_{Box} = df_C = a * (c - 1) = 2(3\text{-}1) = 4$

Mathematical Calculations

For the brand*temp (AB) factor interaction, there are a*b effects that are calculated. Restrictions in ANOVA require:

Brand	Temp	Interaction Effect
exp	room	**2.83**
exp	frig	**- 2.83**
exp	room	**2.83**
exp	frig	**- 2.83**
exp	room	**2.83**
exp	frig	**- 2.83**
gen	room	- 2.83
gen	frig	2.83
gen	room	- 2.83
gen	frig	2.83
gen	room	- 2.83
gen	frig	2.83

1) AB interaction factor effects sum to 0. This requires 1 piece of information to be fixed.

2) The interaction effects within the exp Brand level sum to 0. The same is true for the gen Brand level. This requires 1 piece of information to be fixed in each Brand level. Since 1 value is already used in restriction 1), this requires a-1 pieces of information.

3) The AB effects also sum to 0 within each Temp level. This requires b-1 pieces of information.

Thus, general rules for a factorial ANOVA:

$$df_{Brand*Temp} = df_{AB} = ab - [(a-1) + (b-1) + 1] = (a-1)(b-1)$$
$$= 4 - [1+1+1] = 1$$

Similarly, the df for residuals (bags in our example) also fits these restrictions.

$$\begin{aligned} df_{Bag} &= \text{\# of effects} - [\text{pieces of information already accounted for}] \\ &= \text{\# of effects} - [df_{Box} + df_{AB} + df_{Brand} + df_{Temp} + 1] \\ &= abc - [a(c-1) + (a-1)(b-1) + (a-1) + (b-1) + 1] \\ &= 12 - [2(3-1) + (2-1)*(2-1) + (2-1) + (2-1) + 1] = 4 \end{aligned}$$

where abc = number of units (bags)

Mathematical Calculations

Mean Squares (MS) = SS/df for each factor. MS is a measure of variability for each factor. Below is the ANOVA for Example 2):

Source	DF	SS	MS	F	P
Brand	1	3.00	3.00	0.12	0.745
Box(Brand)	4	99.00	24.75	1.45	0.364
Temp	1	1.33	1.33	0.08	0.794
Brand*Temp	1	96.33	96.33	5.64	0.076
Error	4	68.33	17.08		
Total	11	268.00			

F-statistic = MS for each factor/MSE. Since there are 2 unit sizes, boxes and bags, there are two error (MSE) terms. The brand F-test uses box(brand) [stated box nested within brand] in the denominator. Since box best represents the variation within brand, it is the whole plot error.

To test the effect of temperature, we have two bags that are as similar as possible (from the same box) and randomly assign a bag to either room or frig. The F-statistic for temp is MS_{Temp}/MS_{Bag}. MS_{Bag} is called the split plot error and is the best measure of variability between bags.

In addition to the design structure and crossed vs. nested factors, each factor needs to be classified as fixed or random in order to determine what error term should be used in the denominator for every F-test.

Fixed vs. Random Effects

Fixed factors: the levels tested represent all levels of interest

Random factors: the levels tested represent a random sample from some population of possible levels of interest.

Flower Example (continued): The levels of water solution (plain, aspirin, and floral compound) are all of specific interest. They are not just a random selection of all possible items that could be added to water. Thus, solution is a fixed factor.

The students did not want to compare three specific stores to determine which store had the best flowers. Instead three stores were randomly selected from all possible stores to better understand the variability that exists within the population. Store and flower (units) are random factors.

Popcorn Example (continued): Bags and boxes are random factors. There were random selections from a population of boxes and a population of bags. Brand and temp are fixed factors. *If* we were not interested in finding the effect of expensive and generic brands, but instead simply randomly selected two solutions of brands from all possible brands, then brand would be a random factor. Determination of whether effects are fixed or random can vary, and the choice can greatly impact the ANOVA analysis.

Fixed vs. Random Effects

Fixed factors have meaning only at the levels that were included in the experimental design. The same levels of that factor would be used if the experiment was repeated.

Since the levels of random factors were randomly selected, the results have meaning for the levels selected in the study as well as any levels not included in the study. Different levels would be randomly selected if the experiment was repeated. Blocks and units are typically classified as random effects.

Now that the key questions have been answered:

1) Is structure of the design:
   - a Complete Randomized Design/Factorial Design
   - a Block Design, or
   - a Split Plot/Repeated Measure Design?
2) Is each factor is crossed or nested?
3) Is each factor is fixed or random?

Hasse diagrams can be used to determine what error term should be used in the denominator for every F test. Hasse diagrams are effective for all **balanced designs** (i.e. same number of units in every condition).

Rules for Developing Hasse Diagrams

1) Start row 1 with node M for the grand mean
2) Put a node on row 2 for each factor that is not nested in any term. Add arrows from each node on row 2 to the grand mean. Place parentheses around any random factor.
3) Add a node on row 3 for any factor nested in row 2, and draw arrows to the row 2 nodes. Add a node for any 2-way interaction and draw arrows to the individual factors in row 2. Place parentheses around any random factor or any factor that is nested in a random factor. If an interaction term contains at least 1 random effect, the entire interaction is considered random.
4) On each successive row, say row "i", add a node for any factor nesting in row "i-1". Add a node for any "i-way interaction". Draw appropriate arrows to the "i-1" nodes and place parentheses around any random factor or any factor that is nested in a random factor.
5) When all interactions or nested factors are exhausted, add a node for error on the bottom line, and draw arrows to nodes in the row above.
6) For each node, add a superscript that indicates the number of effects for each term (# of interaction effects are always products of the # of main effects).
7) For each node, add a subscript that indicates the degrees of freedom for that term. Degrees of freedom for a term are found by starting with the superscript for that particular node and subtracting out the degrees of freedom for all terms connected with arrows above it.

Mathematical Calculations: Hasse Diagram

If the Hasse Diagram is developed, the denominator for the appropriate F test is typically straightforward:

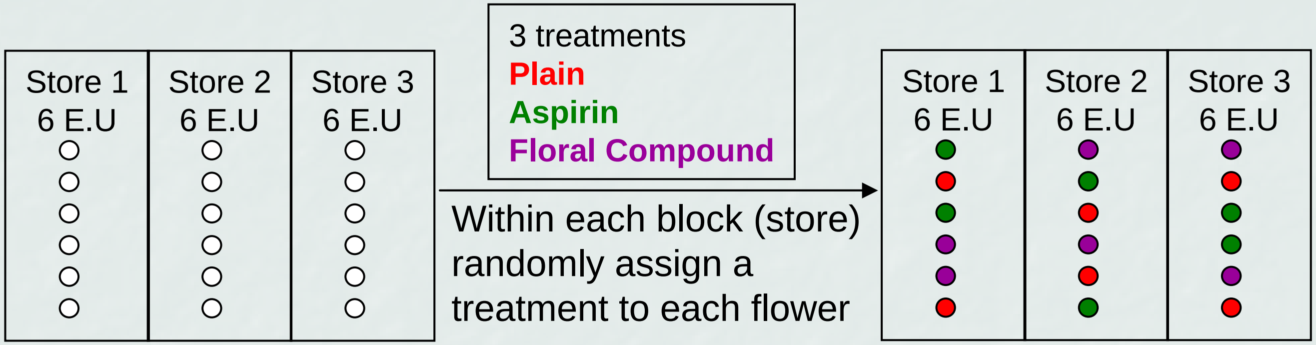
The denominator for testing node A is the next eligible random term below A in the Hasse diagram.

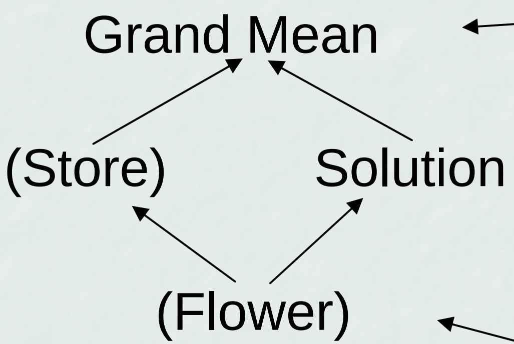
If there are 2 or more "next eligible random terms" then use an *approximate test*. Approximate tests usually are a combination of existing MS values. Most software packages do this automatically.

In more complex models that include **mixed interaction terms** (there are both fixed and random factors in the interaction), it is necessary to determine whether the effects are **Restricted** or **Unrestricted.** In general, restricted effects sum to zero while unrestricted effects do not. However, this classification tends to be rather complex and the analysis is best done with a statistician. Texts listed at the end of this tutorial all discuss restricted and unrestricted effects in more detail.

Mathematical Calculations

Hasse Diagram for the **Flower Example**;





1) Start row 1 with a node for the grand mean

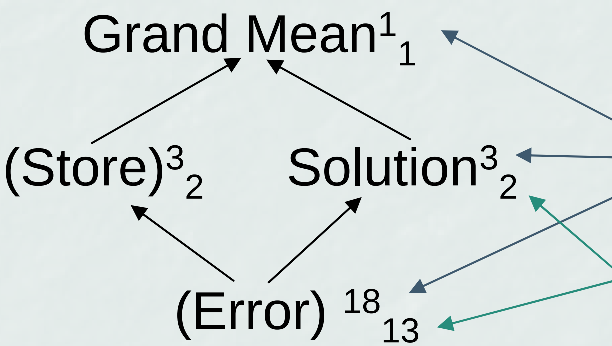
2) Put a node on row 2 for each factor that is not nested in any term. Add arrows from each node on row 2 to the grand mean. Place parentheses around any random factor.

3) Add a node on row 3 for any factor nested in row 2, and draw arrows to the row 2 nodes. Place parentheses around any random factor or any factor that is nested in a random factor. No 2-factor interactions are used.

4) No successive rows needed.

Mathematical Calculations

Hasse Diagram for the **Flower Example (continued):**



5) When all interactions or nested factors are exhausted, add a node for error on the bottom line, and draw arrows to nodes in the row above.
6) For each node, add a superscript that indicates the number of levels for each term. (# of interaction effects are always products of the # of main effects)
7) For each node, add a subscript that indicates the degrees of freedom for that term. Degrees of freedom for a term are found by starting with the superscript for that particular node and subtracting out the degrees of freedom for all terms connected with arrows above it (i.e. subtracting df from all terms included in the partial fit).

Flowers are the units, thus the flower effect is identical to the residual effect (i.e. the error term).

For this study, there is only one error term (thus only one MSE). Error is the first random term following the store effect and the water solution effect. Thus error (i.e. flower or unit) is the denominator in both F-tests.

Mathematical Calculations

The correct ANOVA for the **Flower Example:**

Source	df	SS	MS	F	P
Store	2	13	6.5	8.05	0.005
Solution	2	9	4.5	5.57	0.018
Error	13	10.5	0.8		
Total	17	32.5			

solution	Store 1 days	Store 2 days	Store 3 days
water	7	6	6
water	9	7	5
aspirin	6	6	4
aspirin	5	5	5
compound	8	7	6
compound	7	8	4

Note that if the students had ignored the store effect. The ANOVA would look like:

Source	df	SS	MS	F	P
Solution	2	9.0	4.50	2.87	0.088
Error	15	23.5	1.57		
Total	17	32.5			

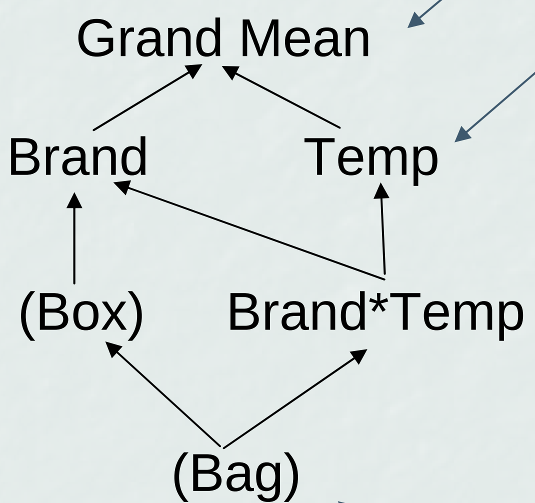
Note that the solution df and SS are identical in both ANOVAs. However, the error term in the 2nd ANOVA combines the store and the flower (error) effects.

If store variability is ignored (just assumed to be flower-to-flower variability), it overshadows the effects of solution, and the students would incorrectly conclude that solution is not significant.

Students could have decided to eliminate store variability by only buying flowers from Store 1. However, then conclusions from this study could only extend to carnations purchased from that particular store.

Mathematical Calculations

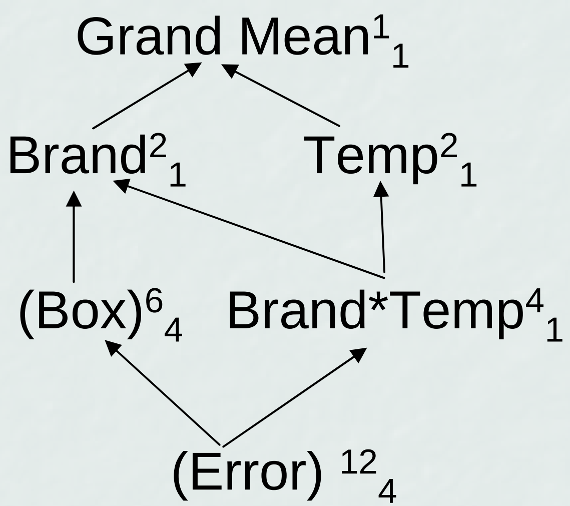
Hasse Diagram for **Popcorn Example**:



1) Start row 1 with a node for the grand mean
2) Put a node on row 2 for each factor that is not nested in any term. Add arrows from each node on row 2 to the grand mean.
3) Add a node on row 3 for any factor nested in row 2, and draw arrows to the appropriate row 2 nodes. Add a node for any 2-way interaction and draw arrows to the individual factors in row 2. Place parentheses around any random factor or any factor that is nested in a random factor. If an interaction term contains at least 1 random effect, the entire interaction is considered random.
4) On each successive row, say row “i”, add a node for any factor nesting in row “i-1”. Add a node for any “i-way interaction”. Draw appropriate arrows to the “i-1” nodes and place parentheses around any random factor or any factor that is nested in a random factor.

Mathematical Calculations

Hasse Diagram for **Popcorn Example (continued):**



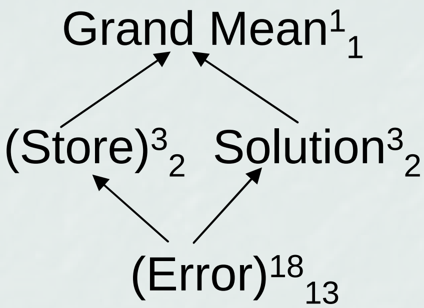
5) When all interactions or nested factors are exhausted, add a node for error on the bottom line, and draw arrows to nodes in the row above.
6) For each node, add a superscript that indicates the number of levels for each term. (# of interaction effects are always products of the # of main effects)
7) For each node, add a subscript that indicates the degrees of freedom for that term. Degrees of freedom for a term are found by starting with the superscript for that particular node and subtracting out the degrees of freedom for all terms connected with arrows above it.

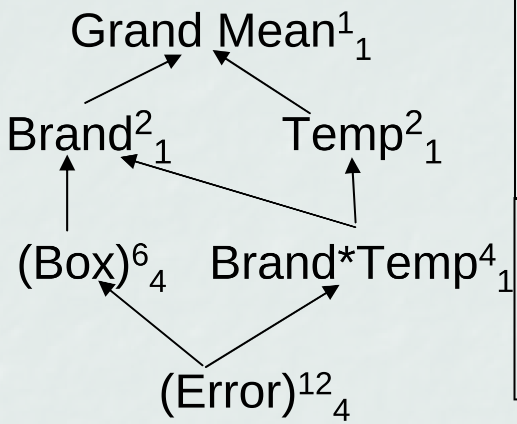
For this study, **box** (whole plot error) is the first random term below **brand** and thus is used in the denominator for the brand F-test.

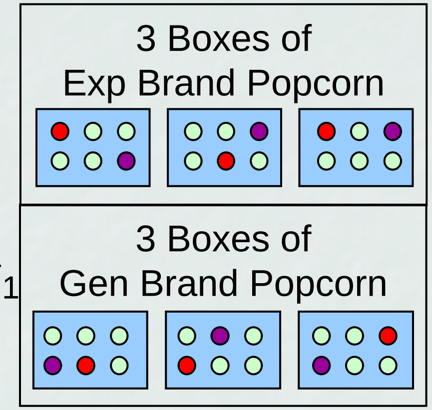
Bag (error or sub plot error) is the first random term below **temp** and **brand*temp**, thus error (i.e. bag or split plot unit) is the denominator in both temp and brand*temp F-tests.

Mathematical Calculations

Comparison of the Flower and Popcorn Examples:







Both the Flower and Popcorn Examples have units assigned within larger groups (flowers within stores and boxes within brand). The key difference between Block designs and SP/RM designs are the units. In the Flower Example, only one measurement (longevity) is taken on each unit. SP/RM designs take more than one measurement on the same unit (2 bags were measured within each box). Thus the Popcorn SP/RM Example has two sizes of units (boxes – the whole plot unit and bag-the split plot unit).

While graphics are very useful in determining what condition is applied to each unit, the Hasse diagram can also identify whether the factors are crossed/nested or fixed/random.

Assignment

Comparison of 2 factor models: Remember that df, SS and MS depend upon 1) the design structure and 2) whether the factors are crossed or nested. Also, F-ratios cannot be calculated unless you also know 3) whether factors are fixed or random. The following table shows data from an experiment with 2 factors, A and B. For each of the four situations, use the same data to create Hasse diagrams and ANOVA tables.

trial	A	B	Results
1	a1	b1	80
2	a1	b1	70
3	a2	b1	30
4	a2	b1	36
5	a3	b1	46
6	a3	b1	38
7	a1	b2	46
8	a1	b2	44
9	a2	b2	16
10	a2	b2	18
11	a3	b2	32
12	a3	b2	24

1) A and B crossed [written A x B or AB], any combination of fixed and random effects.

2) A nested in B [this is written A(B)], A fixed B fixed

3) A nested in B [this is written A(B)], A random B fixed

4) A nested in B [this is written A(B)], A random B random

*It typically doesn't make sense to have a fixed factor nested within a random factor. For example, if you randomly select 10 trees, factor B, and then choose 5 leaves nested within each tree, factor A, it is reasonable to assume that since trees are random, so are the leaves.

Assignment

Two students designed an experiment to test for the effects of several factors on memorization. 13 students were selected from 4 majors (52 students total). Each student was asked to read 4 lists of 20 words, (ABS/M, ABS/E, CONC/M, and CONC/E), in random order.

Three Factors of interest:

- Major, 4 levels: Mathematics, Computer Science, History, and English
- List-Type, 2 levels: Abstract Words (harder to remember) and Concrete Words
- Distraction-Type, 2 levels: Math [Count down from 262 by 3's for 30 seconds] and English [read Robert Frost's *The Road Less Traveled* and describe the theme of the poem]

Nuisance Factor: Student 6 students were randomly selected from each major.

Response: number of words a student could recall.

List, Distraction form a 2-way factorial design within each student, this is called a compound within-blocks treatment with students as the block/whole plot unit. There are a total of 96 split plot units [4 majors*6 students*2 list-types*2 distractions]. Each list of 20 words is a split plot unit used to test the effect of List-Type and Distraction.

Assignment

Create a Hasse Diagram for this study and complete the following ANOVA.

Source	df	SS	MS	F	P
Major		40.792			.314
Subject(Major)		215.167			.003
Distraction		5.042			.277
List-Type		112.667			.000
Distraction*List-Type		0.667			.691
Major*Distraction		41.125			
Major*List-Type		9.000			
Major*Distraction*List-Type		13.000			
Error		251.500			
Total		688.958			

Suggested Reading

Hunter, W. G., “Some Ideas about Teaching and Design of Experiments, with 2^5 Examples of Experiments Conducted by Students”, *The American Statistician,* Vol. 31, No. 1 (Feb., 1977), 12-17.

Suggested Textbooks

Design and Analysis of Experiments by George Cobb,
Prerequisites: none
Examples: wide variety

Design and Analysis of Experiments by Gary Oehlert,
Prerequisites: prior statistics courses are beneficial
Examples: primarily from science and engineering

Design and Analysis of Experiments by David Montgomery,
Prerequisites: prior statistics courses are beneficial
Examples: primarily engineering

Suggested Reading

The Hasse diagram provides a visual display of the relationships between factors for balanced complete experimental designs. Using the Hasse diagram, rules exist for determining the appropriate linear model, ANOVA table, expected means squares, and F-tests.

Iverson, P.W. and Marasinghe, M.G.(2005). Visualizing Experimental Designs for Balanced ANOVA Models Using Lisp-Stat. *Journal of Statistical Software*, 18, 3. http://www.jstatsoft.org/v13/i03/v13i03.pdf

Kempthorne, O. (1982). Classificatory Data Structures and Associated Linear Models, in *Essays in Honor of C. R. Rao*, G. Killianpur, P. R. Krishnaiah, J. K. Ghosh,eds. New York: North Holland, 397-410.

Lohr, S. L. (1995). Hasse Diagrams in Statistical Consulting and Teaching. *The American Statistician*, 49, 4, 376-381.

Marasinghe, M. G. and Darius P. L. (1990). A Structure-Based Approach for Model Determination in Experimental Designs. *Proc. Stat. Comp. Section, American Statistical Association*, 143-150.

Oehlert, G., “A First Course in Design and Analysis of Experiments”, Freeman Publishers, 2000

Searle, S. R. (1971). *Linear Models*. New York: Wiley.

Taylor, W. H. and Hilton, H. G. (1981). A Structure Diagram Symbolization for Analysis of Variance. *The American Statistican*, 35, 2, 85-93.

Additional Terminology

The following are additional terms and techniques that can be used in developing advanced ANOVA models. While many of these are beyond the scope of this tutorial, they are all discussed in more detail in the suggested textbooks.

Often factors or designs are described based on their relationship to units. When units are nested within factors, they are called ***between group factors***.
If some combination of factors are allocated within units, they are called ***within group factors***.

Latin Square Designs are designs with two blocking factors (two nuisance factors) and one factor of interest.

Nested designs are also called **hierarchical designs**. These designs can have 3 or more levels of nesting. For example, trees could be nested within fields and leaves nested in trees.