STATISTICS
TEXAS A&M UNIVERSITY

Handout03:

Design of Experiments

Completely Randomized Design (CRD)

Objectives

- Illustrate the basic terms of experimental design using an example.
- Follow the experimental design process to set up an experiment and determine the appropriate design.
- Generate and analyze a completely randomized design.

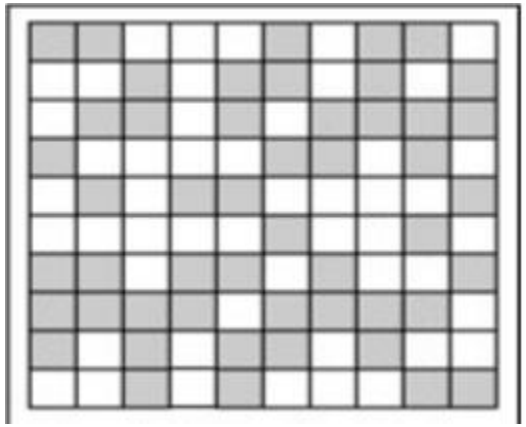
Introduction

A single factor is varied over two or more levels.

Levels of the factor (treatments) are completely randomly assigned to experimental units, and a response variable is measured from each unit.

Interest lies in determining if the mean response differs among the treatment levels in the respective populations.

The randomization for completely randomized design with 2 treatments and 50 replications



Completely Randomized Design (CRD)

- Bio-availability of dietary zinc fed to sheep:

  Sheep were randomly assigned to diets containing supplemental zinc. Three sheep were assigned to each of eight diets. This is a completely randomized design with treatments given by the diets and sheep being the experimental units. Zinc uptake in bone was measured on each sheep.

- Serum zinc in dogs at the College of Veterinary Medicine:

  Blood samples were obtained from dogs that were taken to the clinic. The dogs were diagnosed according to five classifications as related to skin diseases; allergic, non-allergic, sick for non-skin disease, immune deficient, and healthy. Numbers of dogs varied in the various diagnoses. Serum zinc was measured in each of the serum samples.

Completely Randomized Design (CRD)

- Wear due to friction applied to samples of wood veneer material:

  Five brands of synthetic wood veneer material that are used for counter tops were compared for their durability. Four samples were used from each brand. The samples were subjected to a friction test in a randomly assigned order. Amounts of wear resulting from the friction test were measured on each sample. Although there are no "treatments" per se in this experiment, the brands are treated as such. The random assignment of samples to the friction test avoids systematic bias that might result from the first to the last test.

- Muzzle velocity of bullets:

  Rifle cartridges were made using three types of gunpowder. Four cartridges of each powder type were fired from a rifle in random order and muzzle velocities were measured.

CRD is not appropriate

- Animals are given a drug and measured at several time points in the future. Each animal is measured more than once which violates the assumption of a simple CRD.

  Use Repeated Measures Design.

- Large plots of land are prepared using different fertilizers. Each large plot is divided into smaller plots which receive different varieties of wheat. There are two sizes of experimental units - large plots receiving fertilizers and smaller plots receiving variety. This violates the assumption of a CRD that there is only one size of experimental unit.

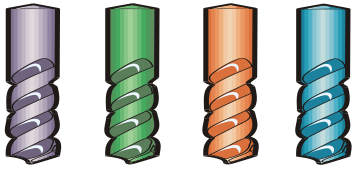
  Use Split-Plot Design.

- Honey bees colonies are arranged on pallets, three per pallet. Interest lies in comparing a method of killing bee mites. Three methods are used, and each pallet receives all three methods. There was not complete randomization because each pallet has to receive all three treatments which violates one of the assumptions of a CRD.

  Use Randomized Complete Block Design.

Define the Purpose if the Drill Experiment

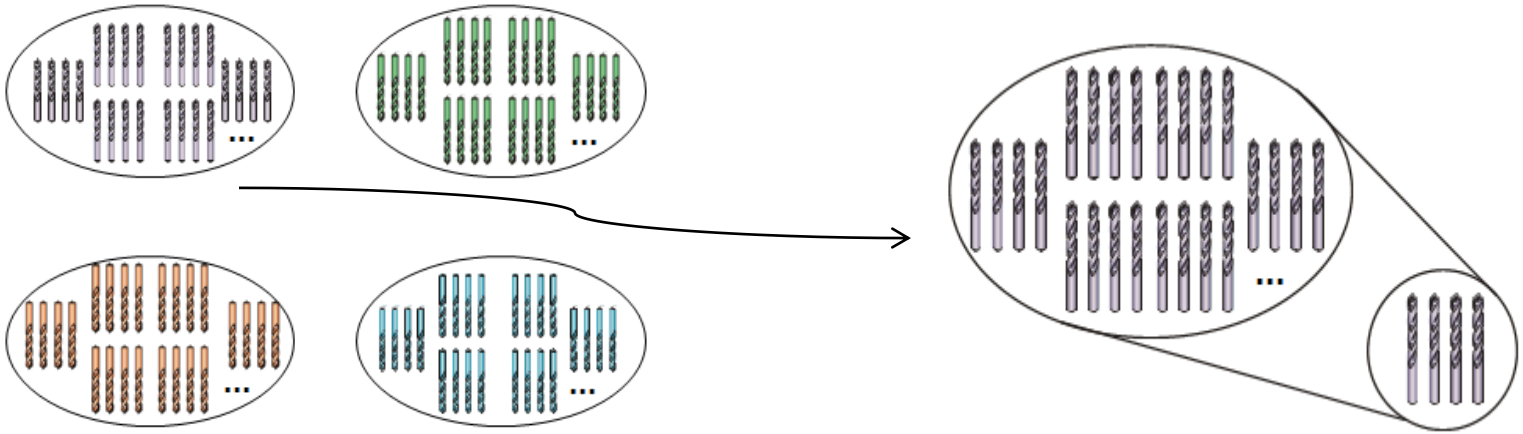
The company wants to determine whether **Hardness** readings from four types of drill tips are different.



Specific Question for this experiment: Are the average **Hardness** readings from the four types of drill tips significantly different from each other?

Define the Populations of Interest and the Need for Sampling

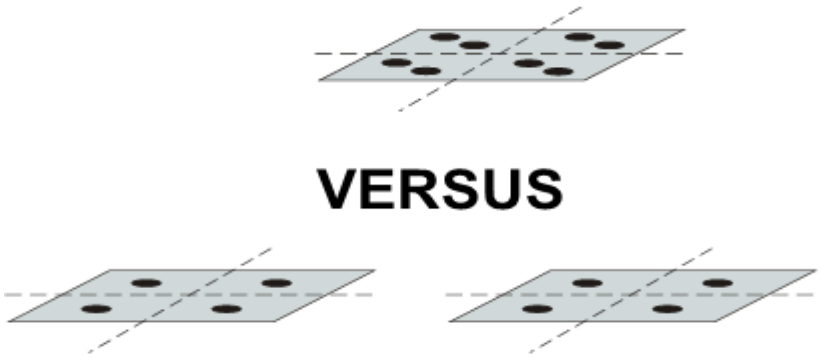
- The company is interested in all drill tips of these four types produced by its supplier, the XYZ Corporation.
- It is physically impossible to collect information on all of the drill tips made by the XYZ Corporation; therefore, you need to use a sample of each of the populations of drill tips. A sample of the **`Purple`** tips is shown below. Each of the other three populations would be sampled the same way.



Define the Data Collection Protocol – Determine the "Best" Design for the Experiment

- The experiment has one factor, **`Tip Type`**, with four levels, **`Purple`**, **`Green`**, **`Orange`**, **`and Blue`**. These factor levels are easy to change from run to run.
- The experimental unit is a quadrant of a metal sheet.
- The experimental units are believed to be homogeneous.
- The completely randomized design is appropriate.
- **Steps:**
  - Select all experimental units (metal sheets) from the same lot.
  - Randomly assign treatments (drill tips) to the experimental units.
  - Use a single drill at one setting for the entire experiment.
  - Use a single instrument to measure and record each indentation depth.

Experimental Units and Replication

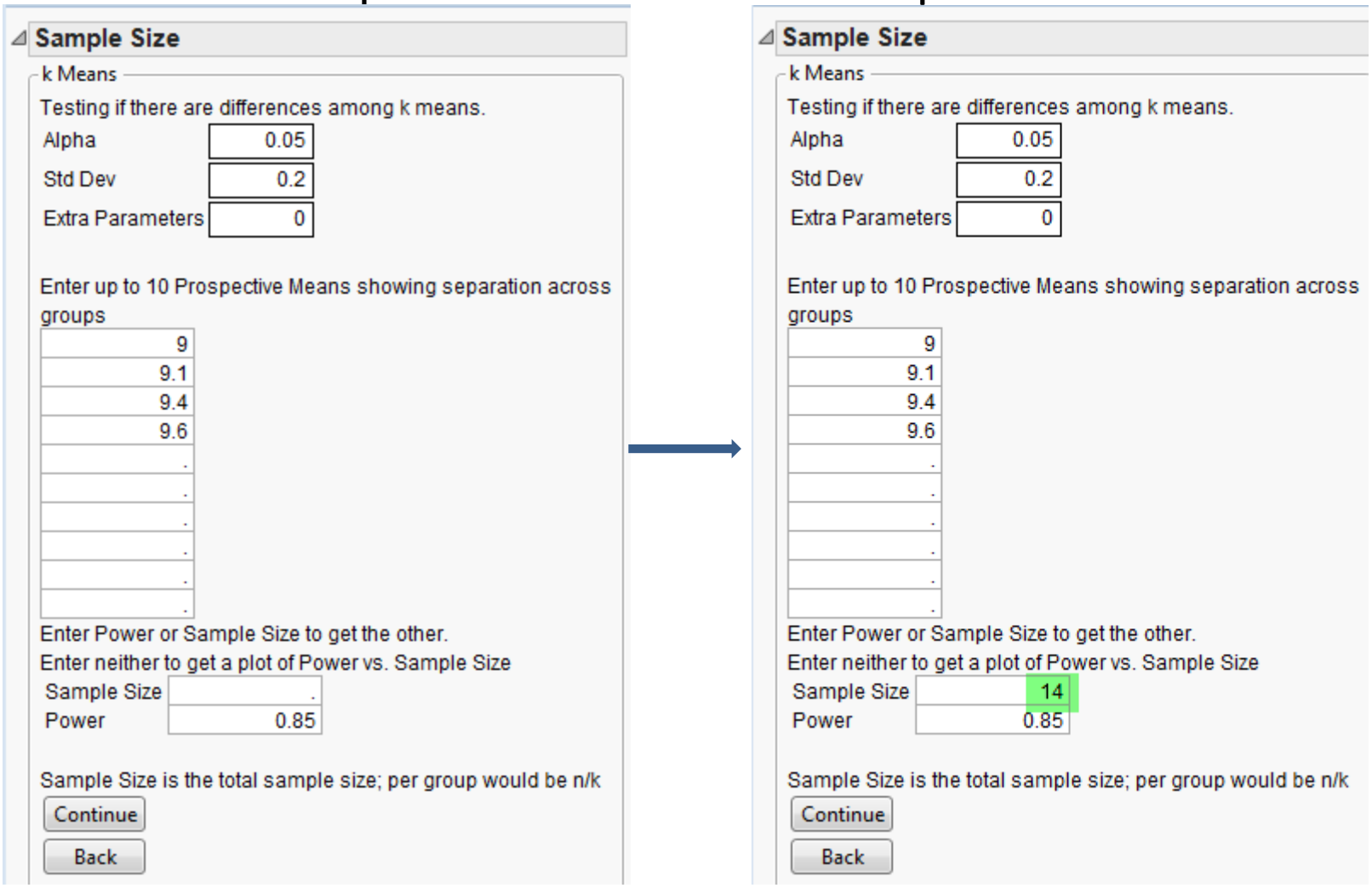


You also want to determine the power and sample size. In preparation for this experiment, you have consulted with industry experts and reviewed previous experiments on drill tips. The following information is determined:

- The expected standard deviation for each treatment group is approximately .2.
- Alpha=.05.
- Power needs to be at least 85%.
- The estimates for the means for each **`Tip Type`** are given by 9.0, 9.1, 9.4, and 9.6.

JMP:Determining Power and Sample Size

Select DOE-Sample Size and Power-k sample means



Determining Power and Sample Size

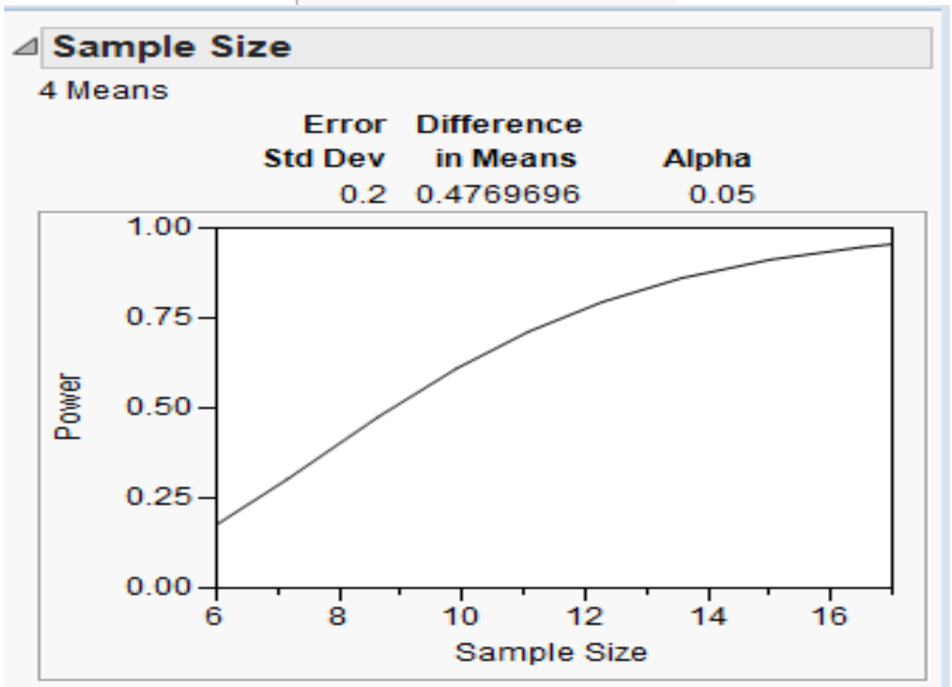
The necessary sample size is 14. Since 14 is not divisible by 4 (number of drill tips), you must prepare 16 pieces.

Sample Size	16
Power	0.9373688283

Sample Size

4 Means

Error Std Dev	Difference in Means	Alpha
0.2	0.4769696	0.05



SAS:Determining Power and Sample Size

```
proc power;
    onewayanova
      groupmeans = 9 | 9.1 | 9.4 | 9.6
      stddev = 0.2
      alpha = 0.05
      ntotal = .
      power = 0.85
      contrast = (3 -1 -1 -1) (1  1 -1 -1)
                 (1 -1  0  0) (0  0  1 -1);
  run;
```

The SAS System

The POWER Procedure
Overall F Test for One-Way ANOVA

Fixed Scenario Elements	
Method	Exact
Alpha	0.05
Group Means	9 9.1 9.4 9.6
Standard Deviation	0.2
Nominal Power	0.85
Group Weights	1 1 1 1

Computed N Total	
Actual Power	N Total
0.937	16

Computed N Total						
Index	Contrast				Actual Power	N Total
1	3	-1	-1	-1	0.915	20
2	1	1	-1	-1	0.925	12
3	1	-1	0	0	0.853	292
4	0	0	1	-1	0.860	76

```
proc power;
onewayanova alpha=.05 test=overall
groupmeans=(9 9.1 9.4 9.6) npergroup=.
stddev=0.2
power=.85;
run;
```

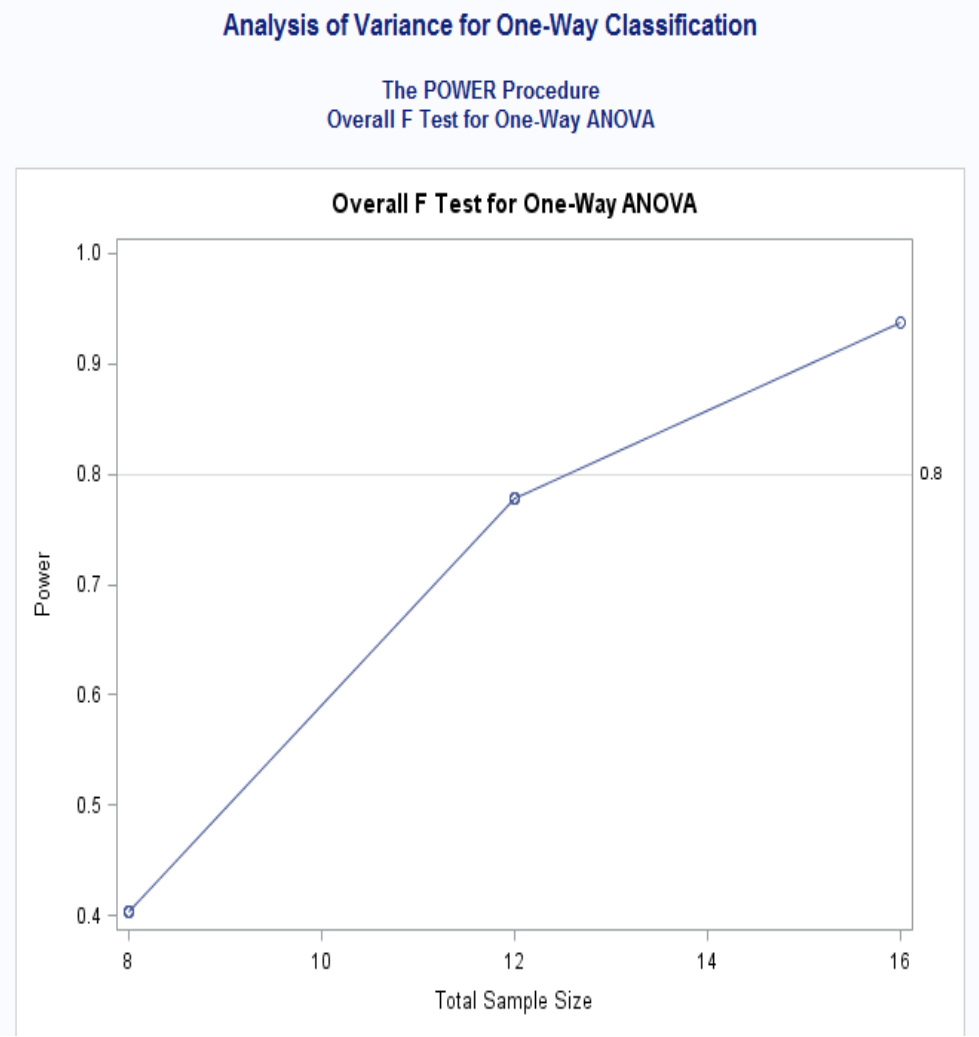
The POWER Procedure
Overall F Test for One-Way ANOVA

Fixed Scenario Elements	
Method	Exact
Alpha	0.05
Group Means	9 9.1 9.4 9.6
Standard Deviation	0.2
Nominal Power	0.85

Computed N per Group	
Actual Power	N per Group
0.937	4

SAS:Determining Power and Sample Size

```
proc power;
onewayanova alpha=.05 test=overall
groupmeans=(9 9.1 9.4 9.6) ntotal=8 to 16
stddev=0.2
power=.;
plot interpol=join yopts=(ref=.80);
run;
```



Generating a Completely Randomized Design in JMP

Select DOE-Full Factorial Design.

Under Responses,

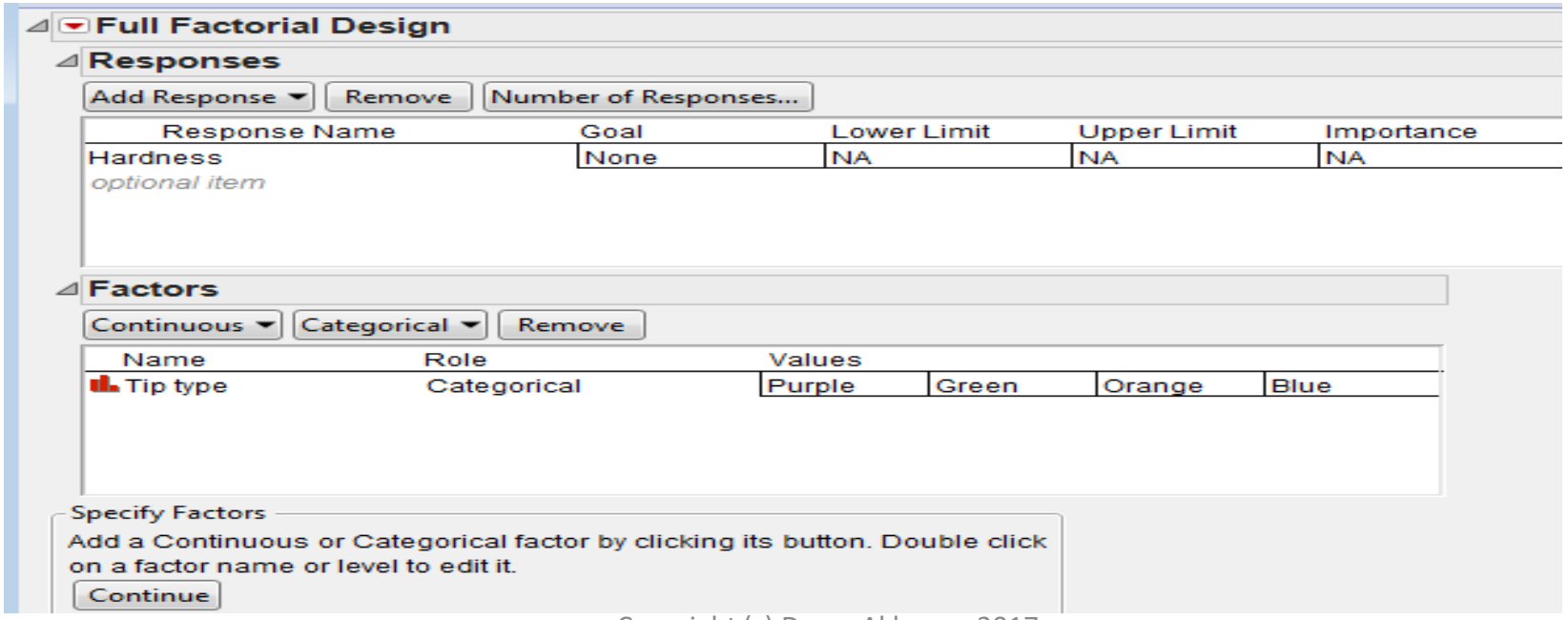
the default response name is Y. Double click Y and type Hardness to change it.

the default Goal is Maximize. Change it to None

Under Factors, select categorical - 4 level

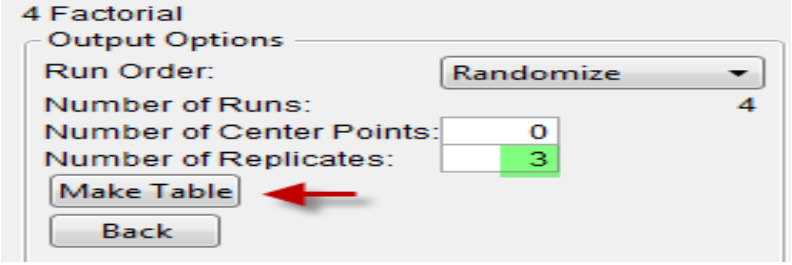
the default name is X1. Change it to Tip type.

Press TAB to move the cursor to the values for Tip type then type Purple, Green, Orange, Blue



DOE (continue)

Change number of Replicates from 0 to 3 then click on Make a Table.



4 Factorial 2

Design 4 Factorial

Model

Columns (3/0)

Pattern

Tip type *

Hardness *

Rows

	All rows	Selected	Excluded	Hidden	Labelled
	16	0	0	0	0

	Pattern	Tip type	Hardness
1	2	Green	•
2	2	Green	•
3	1	Purple	•
4	1	Purple	•
5	3	Orange	•
6	4	Blue	•
7	3	Orange	•
8	1	Purple	•
9	2	Green	•
10	2	Green	•
11	4	Blue	•
12	4	Blue	•
13	4	Blue	•
14	3	Orange	•
15	1	Purple	•
16	3	Orange	•

SAS :DOE

The PROC PLAN statement starts the PLAN procedure and, optionally, specifies a random number seed or a default method for selecting levels of factors. By default, the procedure uses a random number seed generated from reading the time of day from the computer's clock and randomly selects levels of factors.

```
proc plan seed=27371;
factors Replicate=3 ordered tiptype=4;
run;
```

Analysis of Variance for One-Way Classification

The PLAN Procedure

Factor	Select	Levels	Order
Replicate	3	3	Ordered
tiptype	4	4	Random

Replicate	tiptype			
1	3	2	4	1
2	1	2	4	3
3	4	1	2	3

SAS :DOE

```
proc factex;
  factors tiptype / nlev=4;
  size design=16;
  output out=Experiment randomize
    tiptype cvals=('Green' 'Purple' 'Blue' 'Orange');
run;
proc print data=Experiment;
run;
```

Analysis of Variance for One-Way Classification

Obs	tiptype
1	Blue
2	Orange
3	Green
4	Green
5	Blue
6	Purple
7	Blue
8	Orange
9	Blue
10	Purple
11	Green
12	Green
13	Purple
14	Orange
15	Purple
16	Orange

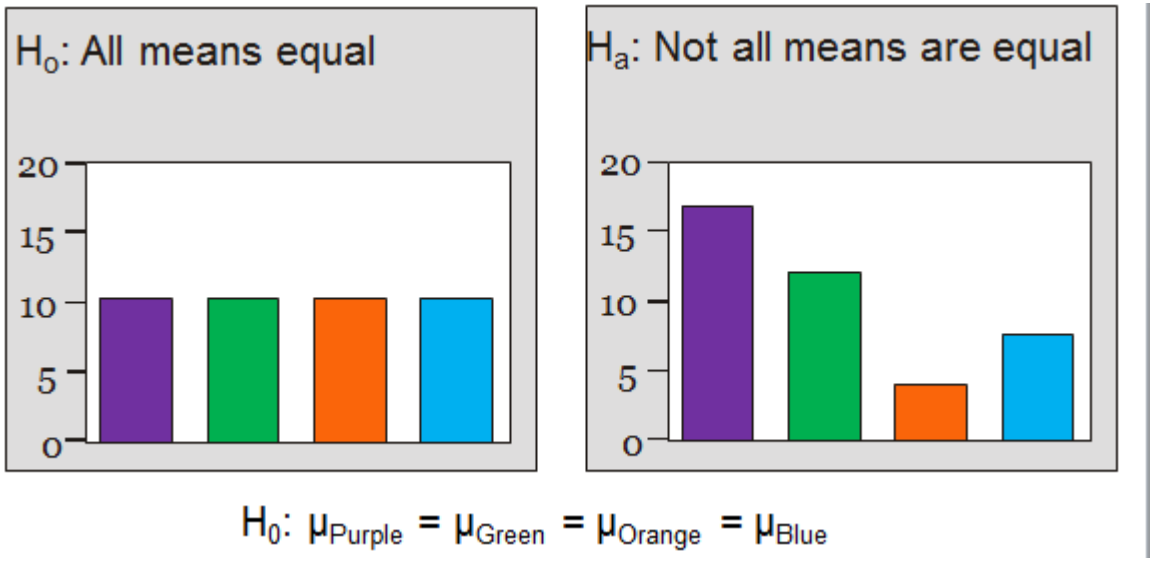
Completely Randomized Design

$Y_{ij} = \mu_i + \varepsilon_{ij} = \mu + \tau_i + \varepsilon_{ij}$, i=1,…,t and j=1,…..,$n_i$

- each population (each value of Tip type, i) has a unique mean ($\mu_i=\mu+\tau_i$)
- Each population is normally distributed but there is a shift in the mean for each population

Commonly used Restriction: τ_t =0, then τ_i =μ_i-μ and μ= μ_t for i=1,…,t-1

Old Restriction : $\sum_{i=1}^{t} n_i\tau_i$ =0, then τ_i =μ_i-μ and $\mu=\frac{1}{n}\sum_{i=1}^{t} n_i\mu_i$ for i=1,…,t-1



Completely Randomized Design

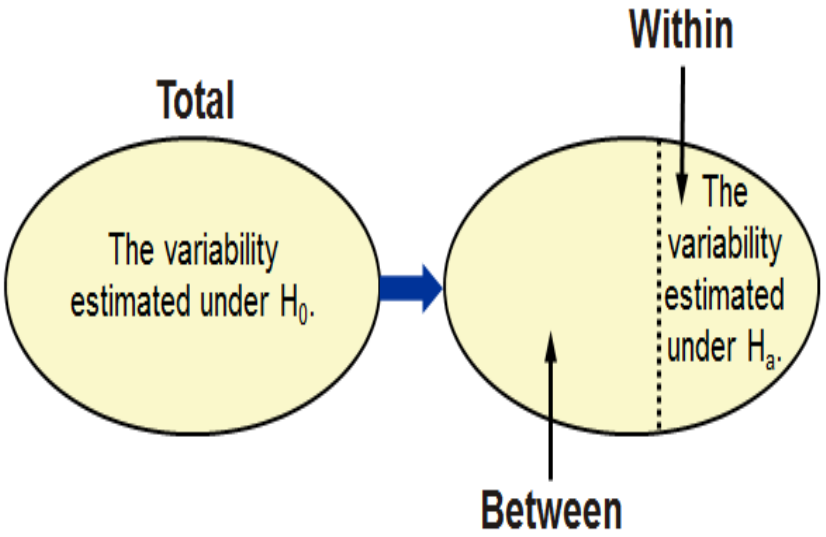
Assumptions:

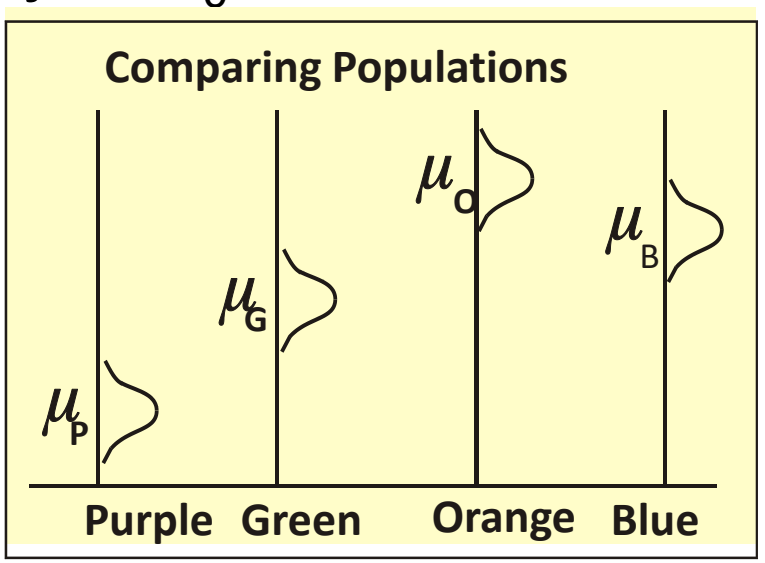
(i) independent observations,

(ii) normally distributed residuals,

(iii) equal variances for each population.

Decision: If the between-group variability is larger than the within-group variability, reject H_0.





STATISTICS
TEXAS A&M UNIVERSITY

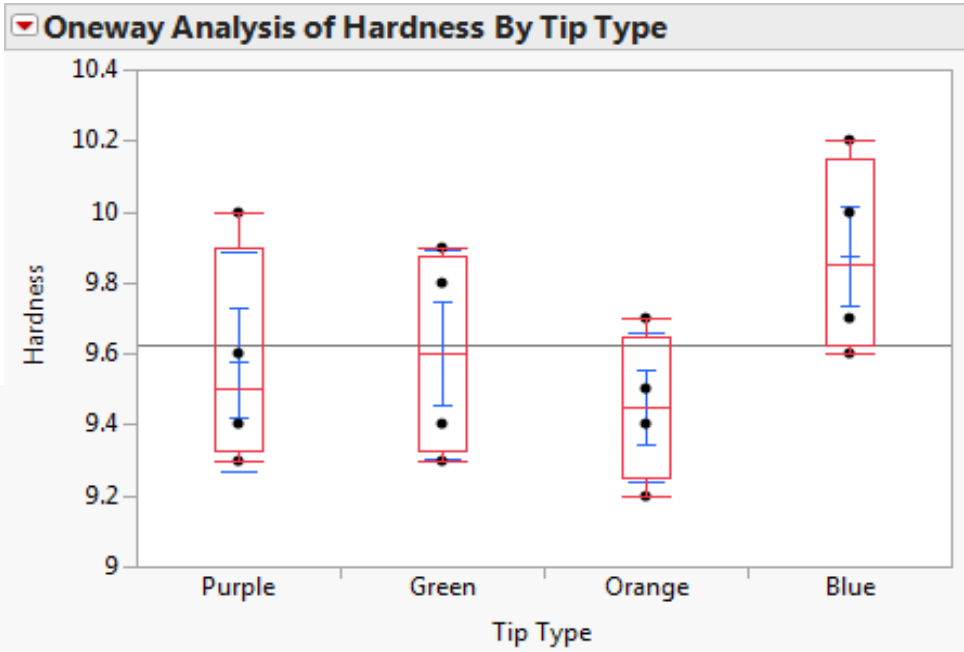
Example (Tips.JMP)

(i) lower power than expected

(ii) Means are closer than expected

(iii) RMSE is higher than expected

NEED LARGER SAMPLE SIZE



Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
Purple	4	9.57500	0.309570	0.15478	9.0824	10.068
Green	4	9.60000	0.294392	0.14720	9.1316	10.068
Orange	4	9.45000	0.208167	0.10408	9.1188	9.781
Blue	4	9.87500	0.275379	0.13769	9.4368	10.313

Test	F Ratio	DFNum	DFDen	Prob > F
O'Brien[.5]	0.3036	3	12	0.8223
Brown-Forsythe	0.4865	3	12	0.6980
Levene	0.5625	3	12	0.6500
Bartlett	0.1492	3	.	0.9302

Warning: Small sample sizes. Use Caution.

Power Details

Test Tip Type

Power

α	σ	δ	Number	Power
0.0500	0.274621	0.155121	16	0.3357

Oneway Anova

Summary of Fit

Rsquare	0.29845
Adj Rsquare	0.123062
Root Mean Square Error	0.274621
Mean of Response	9.625
Observations (or Sum Wgts)	16

Analysis of Variance

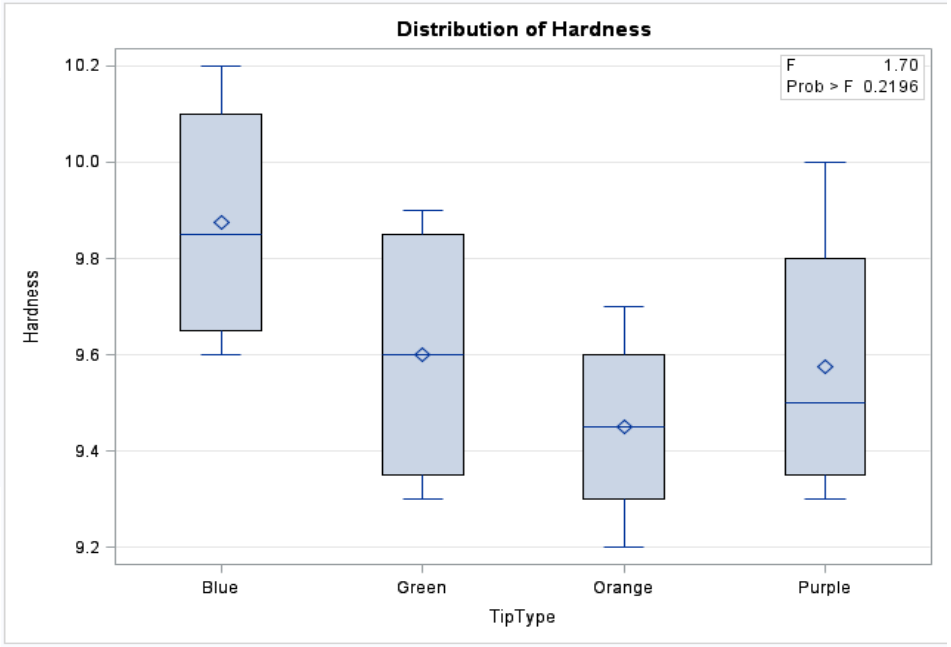
Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Tip Type	3	0.3850000	0.128333	1.7017	0.2196
Error	12	0.9050000	0.075417		
C. Total	15	1.2900000			

Example (SASHardness)

Analysis of Variance for One-Way Classification

The GLM Procedure

Brown and Forsythe's Test for Homogeneity of Hardness Variance ANOVA of Absolute Deviations from Group Medians					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
TipType	3	0.0225	0.00750	0.49	0.6980
Error	12	0.1850	0.0154		



Analysis of Variance for One-Way Classification

The POWER Procedure
Overall F Test for One-Way ANOVA

Fixed Scenario Elements	
Method	Exact
Alpha	0.05
Group Means	9.575 9.6 9.45 9.875
Standard Deviation	0.274621
Sample Size per Group	4

Computed Power
Power
0.336

Dependent Variable: Hardness

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.38500000	0.12833333	1.70	0.2196
Error	12	0.90500000	0.07541667		
Corrected Total	15	1.29000000			

R-Square	Coeff Var	Root MSE	Hardness Mean
0.298450	2.853205	0.274621	9.625000

Prospective versus Retrospective Power

	Prospective	Retrospective
Alpha (α)	.05	.05
Error Standard Deviation	.2	.274
Mean for Orange	9.0	9.450
Mean for Purple	9.1	9.575
Mean for Green	9.4	9.600
Mean for Blue	9.6	9.875
Sample Size (n)	16	16
Power (1-β)	.9374	.3357

Differences in Means

Contrast is the linear combination of means.

Contrast=$\sum_{i=1}^{t} a_i \mu_i$ such that $\sum_{i=1}^{t} a_i = 0$

Suppose there are t=5 treatments

(i) Want to test $\mu_2 = \mu_4$ then contrast= $\mu_2 - \mu_4$ and testing H_0:$\mu_2 - \mu_4$ =0

(ii) Want to test if the average of treatments 1 and 3 is different from the average of treatments 2, 4 and 5 then contrast= $\frac{1}{2}(\mu_1 + \mu_3) - \frac{1}{3}(\mu_2 + \mu_4 + \mu_5)$ and testing H_a:$3\mu_1 - 2\mu_2 + 3\mu_3 - 2\mu_4 - 2\mu_5 \neq 0$

Inferences for Contrasts

The point estimator for the Contrast, C=$\sum_{i=1}^{t} a_i \mu_i$ is $\hat{C} = \sum_{i=1}^{t} a_i \bar{y}_i$.

The point estimator is an unbiased estimator

The variance for the point estimator is $\sum_{i=1}^{t} a_i^2 \frac{MSE}{n_i}$

The 100(1-α)% confidence interval for the point estimator is

$$\hat{C} \pm t_{\frac{\alpha}{2};errordf} \sqrt{MSE} \sqrt{\sum_{i=1}^{t} \frac{a_i^2}{n_i}}$$

The test statistics for testing H_0:C=0 versus H_a:C ≠ 0 is

F=$\frac{\hat{C}^2 / \sqrt{\sum_{i=1}^{t} \frac{a_i^2}{n_i}}}{MSE}$ with df numerator=1 and df denominator=error df

Orthogonal Contrasts

Two contrasts, $C_1=\sum_{i=1}^{t} a_i\mu_i$ and $C_2=\sum_{i=1}^{t} b_i\mu_i$ are said to be orthogonal if $\sum_{i=1}^{t} a_i b_i = 0$

If there are t treatments and t-1 contrasts where all t-1 pairwise comparisons are orthogonal then t-1 contrasts are mutually orthogonal.

SSTreatment can be written as the sum of t-1 independent sums of squares.

If all sample sizes are equal then t-1 mutually orthogonal contrasts satisfy $SSTreatment=\sum_{i=1}^{t-1} SSC_i$ where SSC_i is the sums of squares for the i^{th} contrast.

CONTRAST and ESTIMATES Statements in SAS

They are similar in syntax.

For a single comparison, the ESTIMATE statement provides the estimate of the difference, as well as the significance of the difference; the CONTRAST statement only provides the significance of the difference.

For two or more simultaneous comparisons, only the CONTRAST statement can be used.

In some situations, the ESTIMATE statement requires the DIVISOR= option, whereas the CONTRAST statement does not.

You can write CONTRAST and/or ESTIMATE statements to reproduce any hypothesis tests produced by default.

CONTRAST and ESTIMATES Statements in SAS

$H_0: L'\beta = 0$

CONTRAST *'label' effect values / options*;
ESTIMATE *'label' effect values / options*;

} PROC GLM

$H_0: L'\phi = 0$

CONTRAST *'label' fixed-effect values |*
random-effect values / options;
ESTIMATE *'label' fixed-effect values |*
random-effect values / options;

} PROC MIXED

WRITING CONTRASTS

Write the general linear model for your design.

Write the hypothesis of interest in terms of cell means.

Rewrite the hypothesis as a linear combination of model parameters.

Compute the coefficients for the CONTRAST statement.

Translate the results into an appropriate CONTRAST statement.

EXAMPLE: WRITING CONTRASTS

Suppose you want to know whether there is a difference between the mean hardness of tiptype a and b

The hypothesis can be written as
$H_0: \mu_a = \mu_b$ or $H_0: \mu_a - \mu_b = 0$

Since $E(y_{ij}) = \mu + \alpha_i = \mu_i$, $\mu_a - \mu_b = (\mu + \alpha_a) - (\mu + \alpha_b) = \alpha_a - \alpha_b$

Therefore, the coefficients or the tiptype effect are (1, -1, 0, 0)

In SAS, you would write the following in the model statement:

Contrast ‘a versus b’ tiptype 1 -1 0 0;

EXAMPLE: WRITING ESTIMATES

Write an ESTIMATE statement to estimate the mean hardness value for tiptype c.
(Hint: Use the keyword intercept or int for μ.)

$E(y_{cj}) = \mu + \alpha_c = \mu_c$

In SAS, you would write the following in the model statement:

estimate 'mean pressure for tiptype c' int 1 tiptype 0 0 1 0;

Types of Single Step Multiple Comparisons

1. Compare the set of treatments to a control or standard.

   Use Dunnett's.

2. Make all pairwise comparisons among a set of t means.

   If equal number of observations per group, use Tukey's

   If unbalanced, use a method which simulates percentage point

3. Construct a set of simultaneous C.Is or test of hypotheses.

   Use Bonferroni.

   For large number of comparisons (≥20), scheffe's

4. Explanatory experiments where numerous tests being conducted.

   Use t-test or unadjusted C.I based on LSD

5. Data snooping where the comparisons are possibly data driven.

   Use Scheffe's.

Evaluation of Normality

- Use Shapiro Wilk test
- Look at the normal probability plot of residuals
- Look at the boxplot of residuals to see if at least they are symmetric
- Unequal sample sizes and unequal variances minimally affect the results
- Correlated residuals affect the Shapiro Wilk results .

  Standardized or studentized residuals need to be looked at.
- If the absolute value of the standardized residual is larger than 3, it calls for attention for this outlier

Tests for Homogeneity in Variances

Hartley's Fmax test: requires same sample size and normality.

Bartlett's test: requires normality

Levene's test: uses $|y_{ij}-\bar{y}_{i.}|$ where $\bar{y}_{i.}$ is the mean of i^{th} group.

Brown and Forsythe's test: uses $|y_{ij}-\tilde{y}_{i.}|$ where $\tilde{y}_{i.}$ is the median of i^{th} group. More powerful than Levene's.

O'Brien's test: uses a weight parameter for the scores

Tests for Homogeneity in Variances

- Use **Brown-Forsythe** if the distributions have heavy tails. If this is not available, use Levene's.
- Use O'Brien if the distributions are somewhat skewed
- Use any test including, Bartlett's and Hartley's, if the data are nearly normally distributed.

Recommendation:

(i) No heterogeneity at 1%, use usual ANOVA

(ii) Heterogeneity at 1%, use the mixed models with appropriate denominator df (Satterthwaite approximation,…will be discussed later) OR use AIC to determine a simpler or fewer number of variances can be used to increase the power of the tests concerning means.

Example (Pulserate_CRD.jmp)

Experiment: 78 male in 20s assigned at random to 6 work tasks.

Question: How six different kinds of work tasks affect worker's pulse rate (number of heart pulsations) ?

(1) Design this experiment.

(2) Record the responses. See if you still have total 78 pulse rates recorded?

(3) Compare LSD and Tukey's HSD and also Dunnett's (task2 is control)

(4) Test to see if the mean pulse rate in task3 is 30

(5) Test to see if the mean pulse rate in task4 and task5 are different.

(6) Test to see if the mean pulse rate in task1 is different from the average pulse rate of task2 to task6.

(7) Test to see if the mean pulse rate in task1 is different from the average pulse rate of task2 to task4.

(8) Test to see if the mean pulse rate in task1 is different from the average pulse rate of task3 to task6.

Example

A steel company uses 5 different alloys, A,B,C,D, and E in the production of metal rods. The company wants to determine whether the breaking strength of rods made with different alloys is the same. Other than the different alloy used, the company believes that the rods are identical.

(1) Does CRD appear to be appropriate?

(2) The initial research indicated that the average breaking strength will be approximately 275. the expected standard deviation for each treatment is approximately 5.5. The company has determined the significance level of 10% and power of at least 85% is desired. Suppose your effect size is 3 units of breaking strength. Determine the sample size to detect such and effect. Also if power is at least 70% or 90%.

(3) Use the full factorial design platform to generate a CRD for this problem with a total of 25 runs. How many runs for each value of Alloy are in the design?

(4) Use Rods.jmp data. Are there differences in the mean breaking strength between the different values of Alloy? Compute the power of this analysis. Use Tukey's HSD if you find differences being 90% confident.

Random and Fixed Effects

Some factors have fixed effects and other factors are random effects.

Fixed Effects

All levels of interest are selected by a nonrandom process and are included in the study.

Inferences are to be made only to those level included in the study.

Random Effects

Levels consist of a random sample of levels from a population of possible levels.

Inference is about the population of levels, not just the subset of levels included in the study.

Mixed Models

Models in which some factors are fixed and other factors are random effects.

Random and Fixed Effects

Fixed effects are constant across individuals, and random effects vary. For example, in a growth study, a model with random intercepts, a_i and fixed slope, b corresponds to parallel lines for different individuals, or the model $y_{it}=a_i+bt$. Kreft and De Leeuw (1998) thus distinguish between fixed and random coefficients.

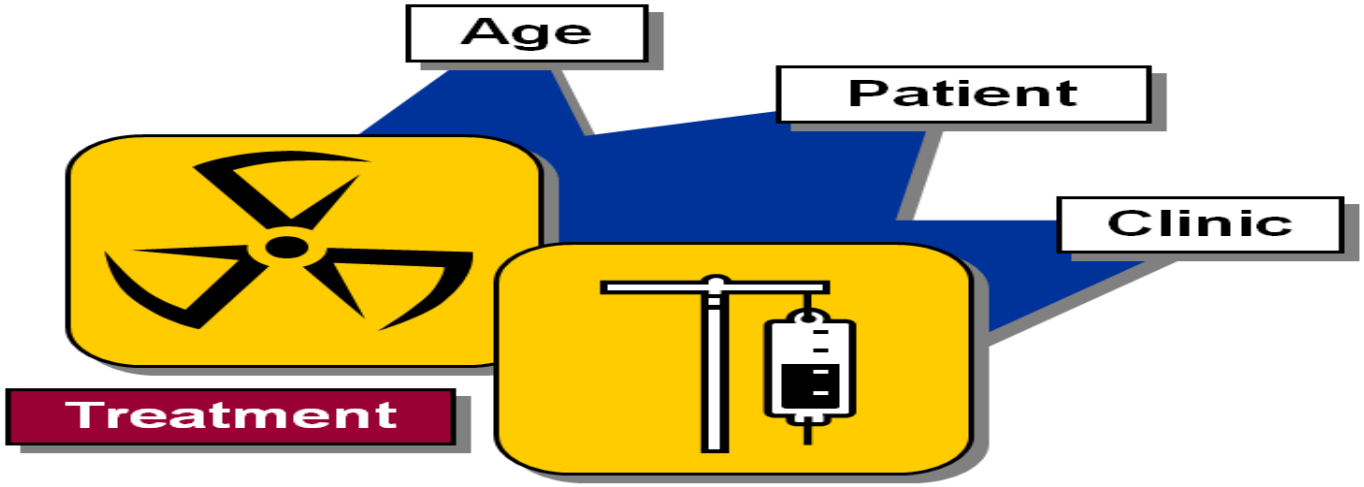
Effects are fixed if they are interesting in themselves or random if there is interest in the underlying population. Searle, Casella, and McCulloch (1992, Section 1.4) explore this distinction in depth.

"When a sample exhausts the population, the corresponding variable is fixed; when the sample is a small (i.e., negligible) part of the population the corresponding variable is random." (Green and Tukey, 1960)

"If an effect is assumed to be a realized value of a random variable, it is called a random effect." (LaMotte, 1983)

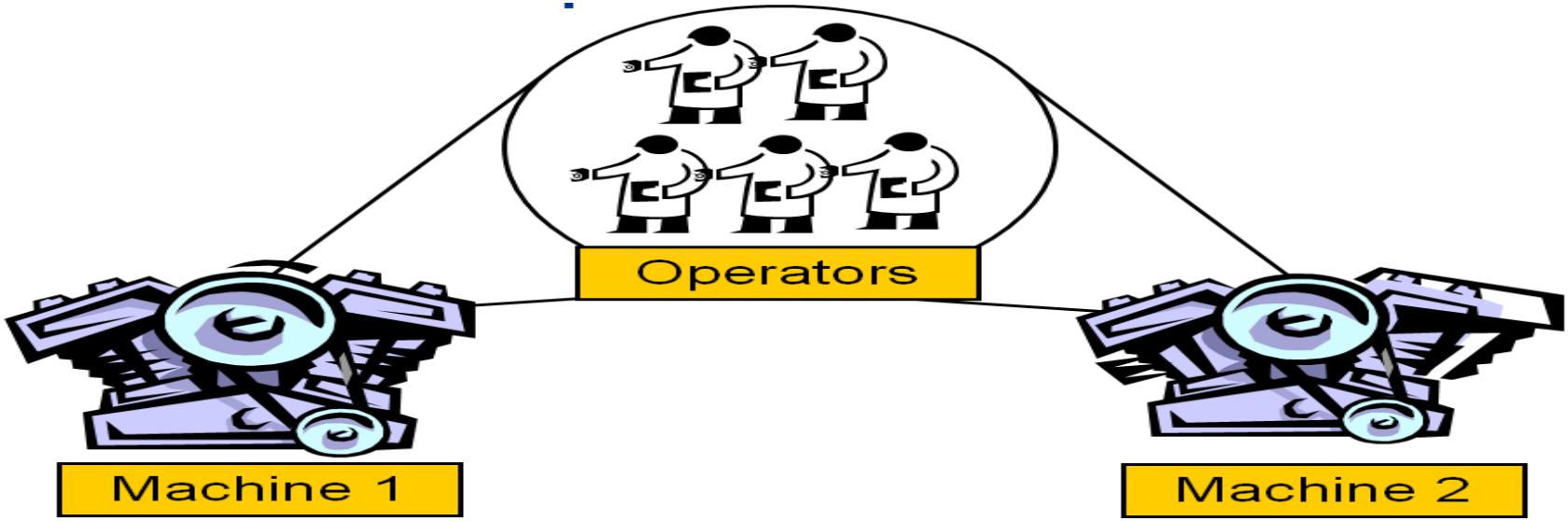
Fixed effects are estimated using least squares (or, more generally, maximum likelihood) and random effects are estimated with shrinkage ("linear unbiased prediction" in the terminology of Robinson, 1991). This definition is standard in the multilevel modeling literature (see, for example, Snijders and Bosker, 1999, Section 4.2) and in econometrics.

Cancer Example



- A physician studies the effect of chemotherapy treatment and a radiotherapy treatment on a certain form of cancer.
- The physician also wants to determine which treatment is more effective with children.
- 10 adults and 10 children are selected from each of 5 cancer treatment clinics.
- The reduction in the size of the tumor is measured after a two-month treatment period.

Machine Example



- The manager of an automotive plant must replace the machines producing a certain component used in automatic transmissions.
- 2 different machines are available.
- The manager wants to evaluate the productivity of the machines when operated by the plant's employees.
- 5 employees are randomly selected from the workforce to operate each machine at 3 time periods.

School Example

- Gains (end-beginning) in scores on a standardized test were recorded for 1,515 fourth-grade students in all school in a district.
- The students' genders and ethnicities, as well as the identification numbers of the students' teachers, were recorded.
- The primary objective was to evaluate and compare the schools in the gain scores.
- A secondary objective was to assess the effects of gender and ethnicity.

Fixed and Random effects

A consumer group studied the variations in coffee prices in U.S. cities with population of at least 20,000.

Ten states were selected at random for the study.

Within each state, five cities were selected randomly. Within each city, ten stores were chosen at random.

The price of a particular grade of coffee at each store was recorded.

Identify the correct statements:

a. State is a random effect

b. City is a random effect

c. Store is a random effect

d. State and City are both fixed effects

e. State, City and Store are both fixed effects

Completely Randomized Design

- Management Policy:

  if a company is interested in the effects of implementing a management policy at its stores and the experiment includes all 5 of its existing stores, it might consider "store" to be a fixed factor, because the levels are not a random sample. But if the company has 100 stores and picks 5 for the experiment, or if the company is considering a rapid expansion and is planning to implement the selected policy at the new locations as well, then "store" would be considered a random factor.

- Bio-availability of dietary zinc fed to sheep:

  Sheep were randomly assigned to diets containing supplemental zinc. Three sheep were assigned to each of randomly selected eight diets.

- Looms in a textile weaving company:

  Four looms have been chosen randomly from a population of looms within a weaving shed and four observations of fabric strength were made on each loom.

Completely Randomized Design

If the tip types are randomly selected from the population of tip types.

$Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$, i=1,…,t and j=1,…,$n_i$ where

- τ_i ~ IID Normal(0,σ_τ^2)
- ε_{ij} ~ IID Normal(0, σ_ε^2)
- τ_i and ε_{ij} are independent
- The variance components are σ_τ^2 and σ_ε^2
- The appropriate test for the random effect is H_0: σ_τ^2=0 versus H_a: σ_τ^2>0

PROC MIXED versus PROC GLM in SAS

They are very similar in syntax.

PROC GLM uses the ordinary least squares (OLS) method to make inferences about fixed effects. Therefore, the inferences are based on a fixed-effects- only model, even when you specify a RANDOM statement.

PROC MIXED uses the generalized least squares (GLS) method to make inferences about fixed effects. Therefore, the inferences directly incorporate the variance-covariance structure that you specify.

PROC GLM computes analysis of variance to assess variations. PROC MIXED uses the maximum-likelihood approach to estimate the variance components. (REML is the default.)

The MIXED Procedure in SAS

Since the GLM is a fixed effects procedure, PROC GLM is not recommended for mixed model analysis and PROC MIXED is recommended.

```
PROC MIXED < options > ;
    BY variables ;
    CLASS variables ;
    ID variables ;
    MODEL dependent = < fixed-effects > < / options > ;
    RANDOM random-effects < / options > ;
    REPEATED < repeated-effect >< / options > ;
    PARMS (value-list) . . . < / options > ;
    PRIOR < distribution >< / options > ;
    CONTRAST 'label' < fixed-effect values . . . >
                < | random-effect values . . . >, . . . < / options > ;
    ESTIMATE 'label' < fixed-effect values . . . >
                < | random-effect values . . . >< / options > ;
    LSMEANS fixed-effects < / options > ;
    WEIGHT variable ;
```

Items within angle brackets (< >) are optional. The CONTRAST, ESTIMATE, LSMEANS, and RANDOM statements can appear multiple times; all other statements can appear only once.

The PROC MIXED and MODEL statements are required, and the MODEL statement must appear after the CLASS statement if a CLASS statement is included. The CONTRAST, ESTIMATE, LSMEANS, RANDOM, and REPEATED statements must follow the MODEL statement. The CONTRAST and ESTIMATE statements must also follow any RANDOM statements.

The MIXED Procedure in SAS

Statement	Description	Important Options
PROC MIXED	invokes the procedure	DATA= specifies input data set, METHOD= specifies estimation method
BY	performs multiple PROC MIXED analyses in one invocation	none
CLASS	declares qualitative variables that create indicator variables in design matrices	none
ID	lists additional variables to be included in predicted values tables	none
MODEL	specifies dependent variable and fixed effects, setting up **X**	S requests solution for fixed-effects parameters, DDFM= specifies denominator degrees of freedom method, OUTP= outputs predicted values to a data set, INFLUENCE computes influence diagnostics
RANDOM	specifies random effects, setting up **Z** and **G**	SUBJECT= creates block-diagonality, TYPE= specifies covariance structure, S requests solution for random-effects parameters, G displays estimated **G**
REPEATED	sets up **R**	SUBJECT= creates block-diagonality, TYPE= specifies covariance structure, R displays estimated blocks of **R**, GROUP= enables between-subject heterogeneity, LOCAL adds a diagonal matrix to **R**
PARMS	specifies a grid of initial values for the covariance parameters	HOLD= and NOITER hold the covariance parameters or their ratios constant, PARMSDATA= reads the initial values from a SAS data set
PRIOR	performs a sampling-based Bayesian analysis for variance component models	NSAMPLE= specifies the sample size, SEED= specifies the starting seed
CONTRAST	constructs custom hypothesis tests	E displays the **L** matrix coefficients
ESTIMATE	constructs custom scalar estimates	CL produces confidence limits
LSMEANS	computes least squares means for classification fixed effects	DIFF computes differences of the least squares means, ADJUST= performs multiple comparisons adjustments, AT changes covariates, OM changes weighting, CL produces confidence limits, SLICE= tests simple effects

Completely Randomized Design

If the tip types are randomly selected from the population of tip types.

The appropriate test for the random effect is H_0: σ_τ^2=0 versus H_a: σ_τ^2>0

Response Hardness

Summary of Fit

RSquare	0.195381
RSquare Adj	0.195381
Root Mean Square Error	0.274621
Mean of Response	9.625
Observations (or Sum Wgts)	16

Parameter Estimates

Term	Estimate	Std Error	DFDen	t Ratio	Prob>|t|
Intercept	9.625	0.089559	3	107.47	<.0001*

Random Effect Predictions

REML Variance Component Estimates

Random Effect	Var Ratio	Var Component	Std Error	95% Lower	95% Upper	Pct of Total
Tip Type	0.1754144	0.0132292	0.0273034	0.0018747	31002.736	14.924
Residual		0.0754167	0.0307887	0.0387802	0.2055049	85.076
Total		0.0886458	0.0349206	0.04648	0.2312591	100.000

-2 LogLikelihood = 8.1646481975

Note: Total is the sum of the positive variance components.

Total including negative estimates = 0.0886458

Class Level Information

Class	Levels	Values
TipType	4	Blue Green Orange Purple

Dimensions	
Covariance Parameters	2
Columns in X	1
Columns in Z	4
Subjects	1
Max Obs per Subject	16

Number of Observations	
Number of Observations Read	16
Number of Observations Used	16
Number of Observations Not Used	0

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	8.53962498	
1	1	8.16464820	0.00000000

Convergence criteria met.

Covariance Parameter Estimates

Cov Parm	Estimate	Alpha	Lower	Upper
TipType	0.01323	0.05	0.001875	31003
Residual	0.07542	0.05	0.03878	0.2055

Fit Statistics	
-2 Res Log Likelihood	8.2
AIC (Smaller is Better)	12.2
AICC (Smaller is Better)	13.2
BIC (Smaller is Better)	10.9