BAN 602: Quantitative Fundamentals

Spring, 2020 Lecture Slides – Week 6

CAL STATE EAST BAY

Agenda

- Analysis of Variance: A Conceptual Overview
- Analysis of Variance and Completely Randomized Design
- Testing for the Equality of k Population Means: A Completely Randomized Design
- Testing for the Equality of k Population Means: An Observational Study
- Multiple Comparison: Fisher's LSD
- Randomized Block Design
- Factorial Experiment

Introduction to Experimental Design and Analysis of Variance

- Statistical studies can be classified as being either experimental or observational.
- In an experimental study, one or more factors are controlled so that data can be obtained about how the factors influence the variables of interest.
- In an observational study, no attempt is made to control the factors.
- Cause-and-effect relationships are easier to establish in experimental studies than in observational studies.
- Analysis of variance (ANOVA) can be used to analyze the data obtained from experimental or observational studies.

Introduction to Experimental Design and Analysis of Variance

In this class three types of experimental designs are introduced.

- A completely randomized design
- A randomized block design
- A factorial experiment

- A <u>factor</u> is a variable that the experimenter has selected for investigation.
- A <u>treatment</u> is a level of a factor.
- <u>Experimental units</u> are the objects of interest in the experiment.
- A <u>completely randomized design</u> is an experimental design in which the treatments are randomly assigned to the experimental units.

Analysis of Variance: A Conceptual Overview

- Analysis of Variance (ANOVA) can be used to test for the equality of three or more population means.
- Data obtained from observational or experimental studies can be used for the analysis.
- We want to use the sample results to test the following hypotheses:

$$H_0\colon \mu_1 = \mu_2 = \mu_3 = \cdots = \mu_k$$

$$H_a\text{: Not all population means are equal.}$$

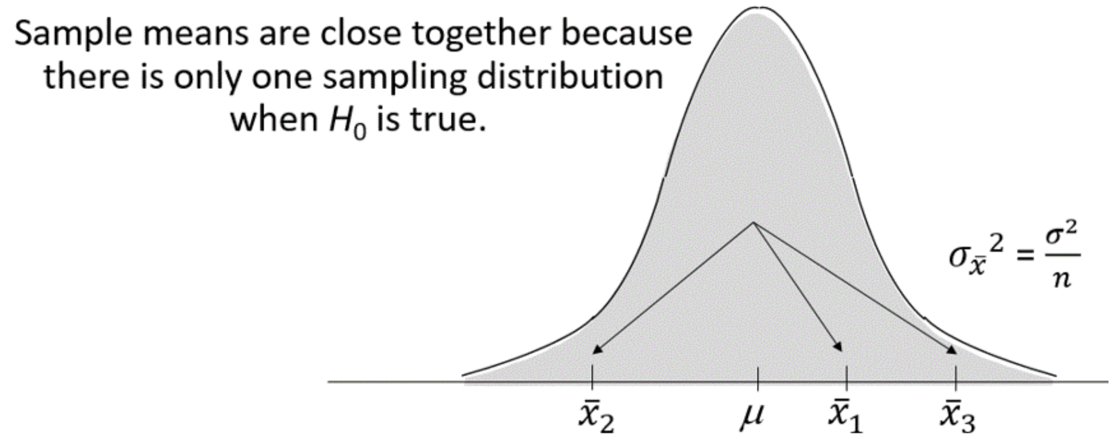
If H_0 is rejected, we cannot conclude that all population means are different. Rejecting H_0 means that at least two population means have different values.

Assumptions for Analysis of Variance:

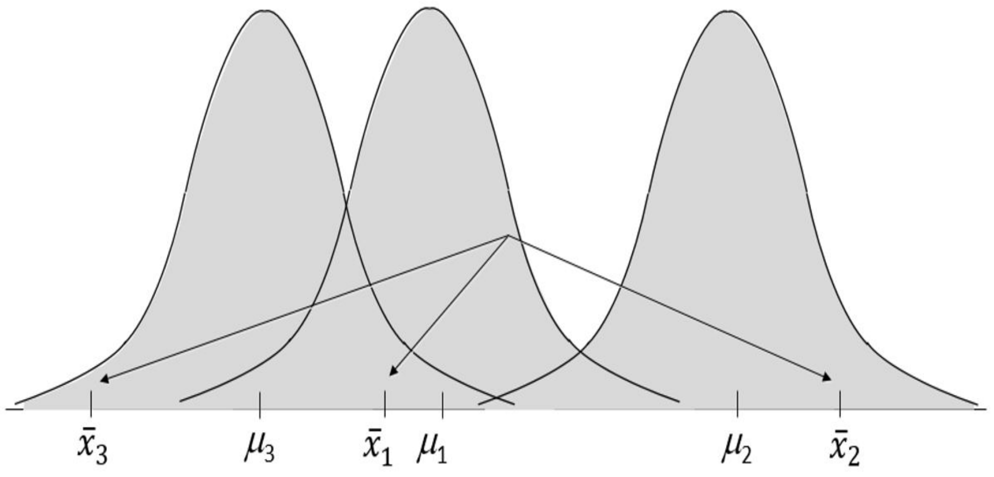
1. For each population, the response (dependent) variable is normally distributed.
2. The variance of the response variable, denoted σ^2, is the same for all of the populations.
3. The observations must be independent.

Analysis of Variance: A Conceptual Overview

- Sampling distribution of $\bar{x}$, given H_0 is true.



- Sampling distribution of $\bar{x}$, given H_0 is false.



Sample means come from different sampling distributions and are not as close together when H_0 is false.

Analysis of Variance and Completely Randomized Design

- Between-Treatments Estimate of Population Variance
- Within-Treatments Estimate of Population Variance
- Comparing the Variance Estimates: The F Test
- ANOVA Table

Between-Treatments Estimate of Population Variance σ^2

The estimate of σ^2 based on the variation of the sample means is called the mean square due to treatments and is denoted by MSTR.

$$\text{MSTR} = \frac{\sum_{j=1}^{k} n_j \left(\bar{x}_j - \bar{\bar{x}}\right)^2}{k-1}$$

Numerator is called the sum of squares due to treatments (SSTR).

Denominator is the degrees of freedom associated with SSTR.

The estimate of σ^2 based on the variation of the sample observations within each sample is called the mean square error and is denoted by MSE.

$$\text{MSE} = \frac{\sum_{j=1}^{k} (n_j - 1) {s_j}^2}{n_T - k}$$

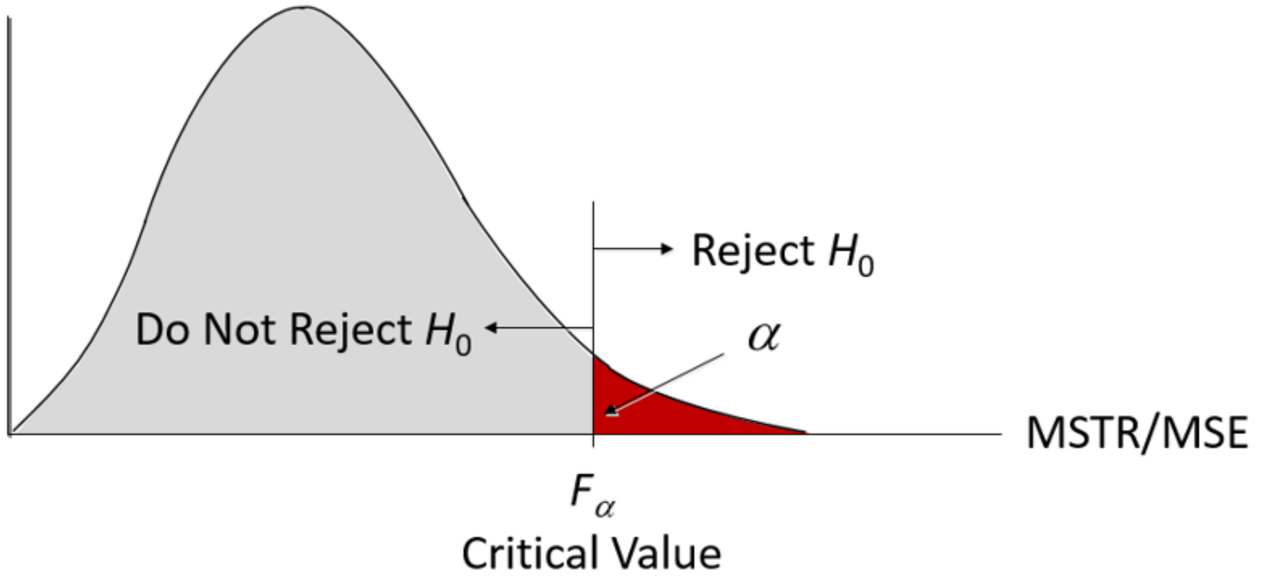
Numerator is called the sum of squares due to error (SSE).

Denominator is the degrees of freedom associated with SSE.

Comparing the Variance Estimates: The F Test

- If the null hypothesis is true and the ANOVA assumptions are valid, the sampling distribution of MSTR/MSE is an F distribution with MSTR degrees of freedom equal to k – 1 and MSE degrees of freedom equal to $n_T - k$.
- If the means of the k populations are not equal, the value of MSTR/MSE will be inflated because MSTR overestimates σ^2.
- Hence, we will reject H_0 if the resulting value of MSTR/MSE appears to be too large to have been selected at random from the appropriate F distribution.

Sampling Distribution of MSTR/MSE



ANOVA Table for a Completely Randomized Design

SST is partitioned into SSTR and SSE.

SST's degrees of freedom (df) are partitioned into SSTR's df and SSE's df.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	p-Value
Treatments	SSTR	k - 1	$\text{MSTR} = \dfrac{\text{SSTR}}{k-1}$	$\dfrac{\text{MSTR}}{\text{MSE}}$	
Error	SSE	n_T - k	$\text{MSE} = \dfrac{\text{SSE}}{n_T - k}$		
Total	SST	n_T - 1			

CAL STATE EAST BAY

ANOVA Table for a Completely Randomized Design

- SST divided by its degrees of freedom $n_T - 1$ is the overall sample variance that would be obtained if we treated the entire set of observations as one data set.
- With the entire data set as one sample, the formula for computing the total sum of squares, SST, is:

$$SST = \sum_{j=1}^{k} \sum_{i=1}^{n_j} (x_{ij} - \bar{\bar{x}})^2 = SSTR + SSE$$

- ANOVA can be viewed as the process of partitioning the total sum of squares and the degrees of freedom into their corresponding sources: treatments and error.
- Dividing the sum of squares by the appropriate degrees of freedom provides the variance estimates, the F value and the p-value used to test the hypothesis of equal population means.

Test for the Equality of k Population Means

- Hypotheses

$$H_0: \mu_1 = \mu_2 = \mu_3 = \cdots = \mu_k$$

H_a: Not all population means are equal.

- Test Statistic $F = \dfrac{MSTR}{MSE}$

Rejection Rule:

p-value approach: Reject H_0 if the *p*-value $\leq \alpha$

Critical value approach: Reject H_0 if $F \geq F_\alpha$

Where the value of F_α is based on an F distribution with $k - 1$ numerator degrees of freedom and $n_T - k$ denominator degrees of freedom.

Testing for the Equality of k Population Means: A Completely Randomized Design

AutoShine, Inc. is considering marketing a long- lasting car wax. Three different waxes (Type 1, Type 2, and Type 3) have been developed. In order to test the durability of these waxes, 5 new cars were waxed with Type 1, 5 with Type 2, and 5 with Type 3. Each car was then repeatedly run through an automatic carwash until the wax coating showed signs of deterioration.

The number of times each car went through the carwash before its wax deteriorated is shown on the next slide. AutoShine, Inc. must decide which wax to market. Are the three waxes equally effective?

Factor . . . Car wax

Treatments . . . Type 1, Type 2, Type 3

Experimental units . . . Cars

Response variable . . . Number of washes

Testing for the Equality of k Population Means: A Completely Randomized Design

Observation	Wax Type 1	Wax Type 2	Wax Type 3
1	27	33	29
2	30	28	28
3	29	31	30
4	28	30	32
5	31	30	31
Sample Mean	29.0	30.4	30.0
Sample Variance	2.5	3.3	2.5

$$H_0\colon \mu_1 = \mu_2 = \mu_3$$

H_a: Not all population means are equal.

Where:

μ_1 = the mean number of washes using Type 1 wax

μ_2 = the mean number of washes using Type 2 wax

μ_3 = the mean number of washes using Type 3 wax

Testing for the Equality of k Population Means: A Completely Randomized Design

Mean Square Between Treatments: (Because the sample sizes are all equal)

$$\bar{\bar{x}} = \frac{(\bar{x}_1 + \bar{x}_2 + \bar{x}_3)}{3} = \frac{(29 + 30.4 + 30)}{3} = 29.8$$

$$SSTR = 5(29 - 29.8)^2 + 5(30.4 - 29.8)^2 + 5(30 - 29.8)^2 = 5.2$$

$$MSTR = \frac{5.2}{(3-1)} = 2.6$$

Mean Square Error: $SSE = 4(2.5) + 4(3.3) + 4(2.5) = 33.2$

$$MSE = \frac{33.2}{(15-3)} = 2.77$$

Rejection Rule: p-value approach: Reject H_0 if the p-value ≤ 0.05

Critical value approach: Reject H_0 if $F \geq 3.89$

Where the value of $F_{0.05} = 3.89$ is based on an F distribution with 2 numerator degrees of freedom and 12 denominator degrees of freedom.

CAL STATE EAST BAY

Testing for the Equality of k Population Means: A Completely Randomized Design

Test Statistic:

$$F = \frac{MSTR}{MSE} = \frac{2.60}{2.77} = 0.939$$

The p-value is greater than 0.10, where $F = 2.81$. R provides a p-value of 0.42. Therefore, we cannot reject H_0.

Conclusion:

There is insufficient evidence to conclude that the mean number of washes for the three wax types are not all the same.

ANOVA Table:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	F	p-Value
Treatments	5.2	2	2.60	0.939	0.42
Error	33.2	12	2.77		
Total	38.4	14			

Testing for the Equality of k Population Means: An Observational Study

Example: Reed Manufacturing

Janet Reed would like to know if there is any significant difference in the mean number of hours worked per week for the department managers at her three manufacturing plants (in Buffalo, Pittsburgh, and Detroit). An F test will be conducted using α =0.05.

A simple random sample of five managers from each of the three plants was taken and the number of hours worked by each manager in the previous week is shown on the next slide.

Factor . . . Manufacturing plant

Treatments . . . Buffalo, Pittsburgh, Detroit

Experimental units . . . Managers

Response variable . . . Number of hours worked

Testing for the Equality of k Population Means: An Observational Study

Observation	Plant 1 Buffalo	Plant 2 Pittsburgh	Plant 3 Detroit
1	48	73	51
2	54	63	63
3	57	66	61
4	54	64	54
5	62	74	56
Sample Mean	55	68	57
Sample Variance	26.0	26.5	24.5

1. Develop the hypotheses:

H_0: $\mu_1 = \mu_2 = \mu_3$

H_a: Not all population means are equal.

Where:

μ_1 = the mean number of hours worked per week by the managers at Plant 1.

μ_2 = the mean number of hours worked per week by the managers at Plant 2.

μ_3 = the mean number of hours worked per week by the managers at Plant 3.

Testing for the Equality of k Population Means: An Observational Study

2. Specify the level of significance. $\alpha = 0.05$
3. Compute the value of the test statistic.

Mean Square Due to Treatments (all sample sizes are equal):

$$\bar{\bar{x}} = \frac{(\bar{x}_1 + \bar{x}_2 + \bar{x}_3)}{3} = \frac{(55 + 68 + 57)}{3} = 60$$

$$SSTR = 5(55 - 60)^2 + 5(68 - 60)^2 + 5(57 - 60)^2 = 490$$

$$MSTR = \frac{490}{(3 - 1)} = 245$$

Mean Square Due to Error

$$SSE = 4(26) + 4(26.5) + 4(24.5) = 308$$

$$MSE = \frac{308}{(15 - 3)} = 25.667$$

$$F = \frac{MSTR}{MSE} = \frac{245}{25.667} = 9.55$$

Testing for the Equality of k Population Means: An Observational Study

ANOVA Table

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	p-Value
Treatment	490	2	245	9.55	.0033
Error	308	12	25.667		
Total	798	14			

p-value approach

4. Compute the p –value.

   With 2 numerator *df* and 12 denominator *df*, the *p*-value is 0.01 for $F = 6.93$.

   Therefore, the *p*-value is less than 0.01 for $F = 9.55$.

5. Determine whether to reject H_0. The p-value ≤ 0.05, so we reject H_0.

We can conclude that the mean number of hours worked per week by department managers is not the same at all 3 plants.

Testing for the Equality of k Population Means: An Observational Study

Critical Value Approach:

4. Determine the critical value and rejection rule.

   Based on an F distribution with 2 numerator *df* and 12 denominator *df*, $F_{0.05} = 3.89$.
   We will reject H_0 if $F \geq 3.89$.

5. Determine whether to reject H_0.

   Because $F = 9.55 \geq 3.89$, we reject H_0.
   We can conclude that the mean number of hours worked per week by department managers is not the same at all 3 plants.

Multiple Comparison Procedures

- Suppose that analysis of variance has provided statistical evidence to reject the null hypothesis of equal population means.
- Fisher's least significant difference (LSD) procedure can be used to determine where the differences occur.

Hypotheses:

$$H_0\colon \mu_i = \mu_j$$
$$H_a\colon \mu_i \neq \mu_j$$

Test Statistic:

$$t = \frac{\bar{x}_i - \bar{x}_j}{\sqrt{MSE\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}}$$

Rejection Rule:

p-value approach: Reject H_0 if the *p*-value ≤ 0.05

Critical value approach: Reject H_0 if $t \leq -t_{\alpha/2}$ or $t \geq t_{\alpha/2}$

Where the value of $t_{\alpha/2}$ is based on a t distribution with $n_T - k$ degrees of freedom.

Fisher's LSD Procedure Based on the Test Statistic $\bar{x}_i - \bar{x}_j$

- Hypotheses:

$$H_0\colon \mu_i = \mu_j$$
$$H_a\colon \mu_i \neq \mu_j$$

- Test Statistic:

$$\bar{x}_i - \bar{x}_j$$

- Rejection Rule:

$$\text{Reject } H_0 \text{ if } \left|\bar{x}_i - \bar{x}_j\right| \geq \text{LSD}$$

$$\text{Where LSD} = t_{\alpha/2}\sqrt{MSE\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$$

Fisher's LSD Procedure Based on the Test Statistic $\bar{x}_i - \bar{x}_j$

Example: Reed Manufacturing

Recall that Janet Reed wants to know if there is any significant difference in the mean number of hours worked per week for the department managers at her three manufacturing plants.

Analysis of variance has provided statistical evidence to reject the null hypothesis of equal population means. Fisher's least significant difference (LSD) procedure can be used to determine where the differences occur.

For $\alpha = 0.05$ and $n_T - k = 15 - 3 = 12$ degrees of freedom $t_{0.025} = 2.179$.

$$\text{LSD} = t_{\alpha/2}\sqrt{MSE\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$$

$$\text{LSD} = 2.179\sqrt{25.667\left(\frac{1}{5} + \frac{1}{5}\right)} = 6.98$$

Fisher's LSD Procedure Based on the Test Statistic $\bar{x}_i - \bar{x}_j$

LSD for Plants 1 and 2

- Hypotheses (A): $H_0\colon \mu_1 = \mu_2$

  $H_a\colon \mu_1 \neq \mu_2$

- Rejection Rule: Reject H_0 if $|\bar{x}_1 - \bar{x}_2| \geq 6.98$

- Test Statistic: $|\bar{x}_1 - \bar{x}_2| = |55 - 68| = 13$

- Conclusion: The mean number of hours worked at Plant 1 is <u>not equal</u> to the mean number worked at Plant 2.

Fisher's LSD Procedure Based on the Test Statistic $\bar{x}_i - \bar{x}_j$

LSD for Plants 1 and 3

- Hypotheses (B): $H_0: \mu_1 = \mu_3$
  $H_a: \mu_1 \neq \mu_3$

- Rejection Rule: Reject H_0 if $|\bar{x}_1 - \bar{x}_3| \geq 6.98$

- Test Statistic: $|\bar{x}_1 - \bar{x}_3| = |55 - 57| = 2$

- Conclusion: There is <u>no significant difference</u> between the mean number of hours worked at Plant 1 and the mean number of hours worked at Plant 3.

Fisher's LSD Procedure Based on the Test Statistic $\bar{x}_i - \bar{x}_j$

LSD for Plants 2 and 3

- Hypotheses (C): $H_0\colon \mu_2 = \mu_3$

  $H_a\colon \mu_2 \neq \mu_3$

- Rejection Rule: Reject H_0 if $|\bar{x}_2 - \bar{x}_3| \geq 6.98$

- Test Statistic: $|\bar{x}_2 - \bar{x}_3| = |68 - 57| = 11$

- Conclusion: The mean number of hours worked at Plant 2 is <u>not equal</u> to the mean number worked at Plant 3.

CAL STATE EAST BAY

Type I Error Rates

The comparison-wise Type I error rate α indicates the level of significance associated with a single pairwise comparison.

The experimentwise Type I error rate α_{EW} is the probability of making a Type I error on at least one of the $(k-1)!$ pairwise comparisons.

$$\alpha_{EW} = 1 - (1-\alpha)^{(k-1)!}$$

The experiment-wise Type I error rate gets larger for problems with more populations (larger k).

Randomized Block Design

- Experimental units are the objects of interest in the experiment.
- A completely randomized design is an experimental design in which the treatments are randomly assigned to the experimental units.
- If the experimental units are heterogeneous, blocking can be used to form homogeneous groups, resulting in a randomized block design.

ANOVA Procedure

For a randomized block design the sum of squares total (SST) is partitioned into three groups: sum of squares due to treatments, sum of squares due to blocks, and sum of squares due to error.

$$SST = SSTR + SSBL + SSE$$

- The total degrees of freedom, $n_T - 1$, are partitioned such that $k - 1$ degrees of freedom go to treatments, $b - 1$ go to blocks, and $(k - 1)(b - 1)$ go to the error term.

Randomized Block Design

ANOVA Table:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	p-Value
Treatments	SSTR	k - 1	$\text{MSTR} = \dfrac{\text{SSTR}}{k-1}$	$\dfrac{\text{MSTR}}{\text{MSE}}$	
Blocks	SSBL	b - 1	$\text{MSBL} = \dfrac{\text{SSBL}}{b-1}$		
Error	SSE	$(k-1)(b-1)$	$\text{MSE} = \dfrac{\text{SSE}}{(k-1)(b-1)}$		
Total	SST	n_T - 1			

Randomized Block Design

Example: Crescent Oil Co.

Crescent Oil has developed three new blends of gasoline and must decide which blend or blends to produce and distribute. A study of the miles per gallon ratings of the three blends is being conducted to determine if the mean ratings are the same for the three blends.

Five automobiles have been tested using each of the three gasoline blends and the miles per gallon ratings are shown on the next slide.

Factor … Gasoline blend

Treatments … Blend X, Blend Y, Blend Z

Blocks … Automobiles

Response variable … Miles per gallon

Randomized Block Design

Automobile (Block)	Type of Gasoline (Treatment) Blend X	Type of Gasoline (Treatment) Blend Y	Type of Gasoline (Treatment) Blend Z	Block Means
1	31	30	30	30.333
2	30	29	29	29.333
3	29	29	28	28.667
4	33	31	29	31.000
5	26	25	26	25.667
Treatment Means	29.8	28.8	28.4	29

- Mean Square Due to Treatments $SSTR = 5[(29.8 - 29)^2 + (28.8 - 29)^2 + (28.4 - 29)^2] = 5.2$

$$MSTR = \frac{5.2}{(3-1)} = 2.6$$

Randomized Block Design

- Mean Square Due to Blocks $SSBL = 3[(30.333 - 29)^2 + \cdots + (25.667 - 29)^2] = 51.3$

$$MSTR = \frac{51.33}{(5-1)} = 12.8$$

- Mean Square Due to Error $SSE = 62 - 5.2 - 5\,133 = 5.47$

$$MSE = \frac{5.47}{(3-1)(5-1)} = 0.68$$

ANOVA Table:

Source of variation	Sum of squares	Degrees of Freedom	Mean Square	*F*	*p*-value
Treatment	5.20	2	2.60	3.82	.07
Blocks	51.33	4	12.80		
Error	5.47	8	.68		
Total	62.00	14			

Randomized Block Design

Rejection Rule: p-value approach: Reject H_0 if the p-value ≤ 0.05

Critical value approach: Reject H_0 if $F \geq 4.46$

For $\alpha = 0.05, F_{0.05} = 4.46$

(2 numerator *df* and 8 denominator *df*)

Test Statistic: $F = \dfrac{MSTR}{MSE} = \dfrac{2.6}{0.68} = 3.82$

Conclusion:

The p-value is greater than 0.05 (where $F = 4.46$) and less than 0.10 (where $F = 3.11$). Therefore, we cannot reject H_0.

There is insufficient evidence to conclude that the miles per gallon ratings differ for the three gasoline blends.

Factorial Experiment

- In some experiments we want to draw conclusions about more than one variable or factor.
- Factorial experiments and their corresponding ANOVA computations are valuable designs when simultaneous conclusions about two or more factors are required.
- The term factorial is used because the experimental conditions include all possible combinations of the factors.
- For example, for a levels of factor A and b levels of factor B, the experiment will involve collecting data on ab treatment combinations.

Two-Factor Factorial Experiment

ANOVA Procedure

- The ANOVA procedure for the two-factor factorial experiment is similar to the completely randomized experiment and the randomized block experiment.
- We again partition the sum of squares total (SST) into its sources.

$$\text{SST} = \text{SSA} + \text{SSB} + \text{SSAB} + \text{SSE}$$

- The total degrees of freedom, $n_T - 1$, are partitioned such that $(a-1)$ degrees of freedom go to Factor A, $(b-1)$ degrees of freedom go to Factor B, $(a-1)(b-1)$ degrees of freedom go to interaction, and $ab(r-1)$ degrees of freedom go to error.

Two-Factor Factorial Experiment

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	p-Value
Factor A	SSA	a - 1	$\text{MSA} = \dfrac{\text{SSA}}{a-1}$	$\dfrac{\text{MSA}}{\text{MSE}}$	
Factor B	SSB	b - 1	$\text{MSB} = \dfrac{\text{SSB}}{b-1}$	$\dfrac{\text{MSB}}{\text{MSE}}$	
Interaction	SSAB	(a–1)(b–1)	$\text{MSAB} = \dfrac{\text{SSAB}}{(a-1)(b-1)}$	$\dfrac{\text{MSAB}}{\text{MSE}}$	
Error	SSE	$ab(r-1)$	$\text{MSE} = \dfrac{\text{SSE}}{ab(r-1)}$		
Total	SST	n_T - 1			

Two-Factor Factorial Experiment

Step 1: Compute the total sum of squares.

$$SST = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}\left(x_{ijk} - \bar{\bar{x}}\right)^2$$

Step 2: Compute the sum of squares for factor A.

$$SSA = br\sum_{i=1}^{a}(\bar{x}_{i.} - \bar{\bar{x}})^2$$

Step 3: Compute the sum of squares for factor B.

$$SSB = ar\sum_{j=1}^{b}\left(\bar{x}_{.j} - \bar{\bar{x}}\right)^2$$

Step 4: Compute the sum of squares for interaction.

$$SSAB = r\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\bar{x}_{ij} - \bar{x}_{i.} - \bar{x}_{.j} - \bar{\bar{x}}\right)^2$$

Step 5: Compute the sum of squares due to error.

SSE = SST – SSA – SSB – SSAB

CAL STATE EAST BAY

Two-Factor Factorial Experiment

Example: State of Ohio Wage Survey. A survey was conducted of hourly wages for a sample of workers in two industries at three locations in Ohio. Part of the purpose of the survey was to determine if differences exist in both industry type and location. The sample data are shown here.

Industry	Cincinnati	Cleveland	Columbus
I	$12.10	$11.80	$12.90
I	11.80	11.20	12.70
I	12.10	12.00	12.20
II	12.40	12.00	12.10
II	12.50	12.00	12.10
II	12.00	12.50	12.70

Factors

Factor A: Industry Type (2 levels)

Factor B: Location (3 levels)

Replications

Each experimental condition is repeated 3 times

Two-Factor Factorial Experiment

ANOVA Table:

Source of Variation	Sum of Squares	Degree of Freedom	Mean Square	F	P-value
Factor A	.50	1	.50	4.19	.06
Factor B	1.12	2	.56	4.69	.03
Interaction	.37	2	.19	1.55	.25
Error	1.43	12	.12		
Total	3.42	17			

Two-Factor Factorial Experiment

- **Conclusions using the critical value approach**
  - Industries:

    $F = 4.19 \leq F_\alpha = 4.75$

    Mean wages do not differ by industry type.
  - Locations:

    $F = 4.69 \geq F_\alpha = 3.89$

    Mean wages differ by location.
  - Interactions:

    $F = 1.55 \leq F_\alpha = 3.89$

    Interaction is not significant.

- **Conclusions using the *p*-value approach**
  - Industries:

    p-value = $0.06 > \alpha = 0.05$

    Mean wages do not differ by industry type.
  - Locations:

    p-value = $0.03 < \alpha = 0.05$

    Mean wages differ by location.
  - Interactions:

    p-value = $0.25 > \alpha = 0.05$

    Interaction is not significant.