Part A
Experimental Design and Analysis of Variance

- An Introduction to Experimental Design and Analysis of Variance
- Analysis of Variance and the Completely Randomized Design
- Multiple Comparison Procedures

An Introduction to Experimental Design and Analysis of Variance

- Statistical studies can be classified as being either experimental or observational.
- In an <u>experimental study</u>, one or more factors are controlled so that data can be obtained about how the factors influence the variables of interest.
- In an <u>observational study</u>, no attempt is made to control the factors.
- <u>Cause-and-effect relationships</u> 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.

An Introduction to Experimental Design and Analysis of Variance

- Three types of experimental designs are introduced.
  - a completely randomized design
  - a randomized block design
  - a factorial experiment

An Introduction to Experimental Design and Analysis of Variance

- A factor is a variable that the experimenter has selected for investigation.
- A treatment is a level of a factor.
- 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.

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:

$$H_0\!: \; \mu_1 = \mu_2 = \mu_3 = \cdot \; \cdot \; \cdot = \mu_k$$

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

Analysis of Variance: A Conceptual Overview

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

H_a: 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.

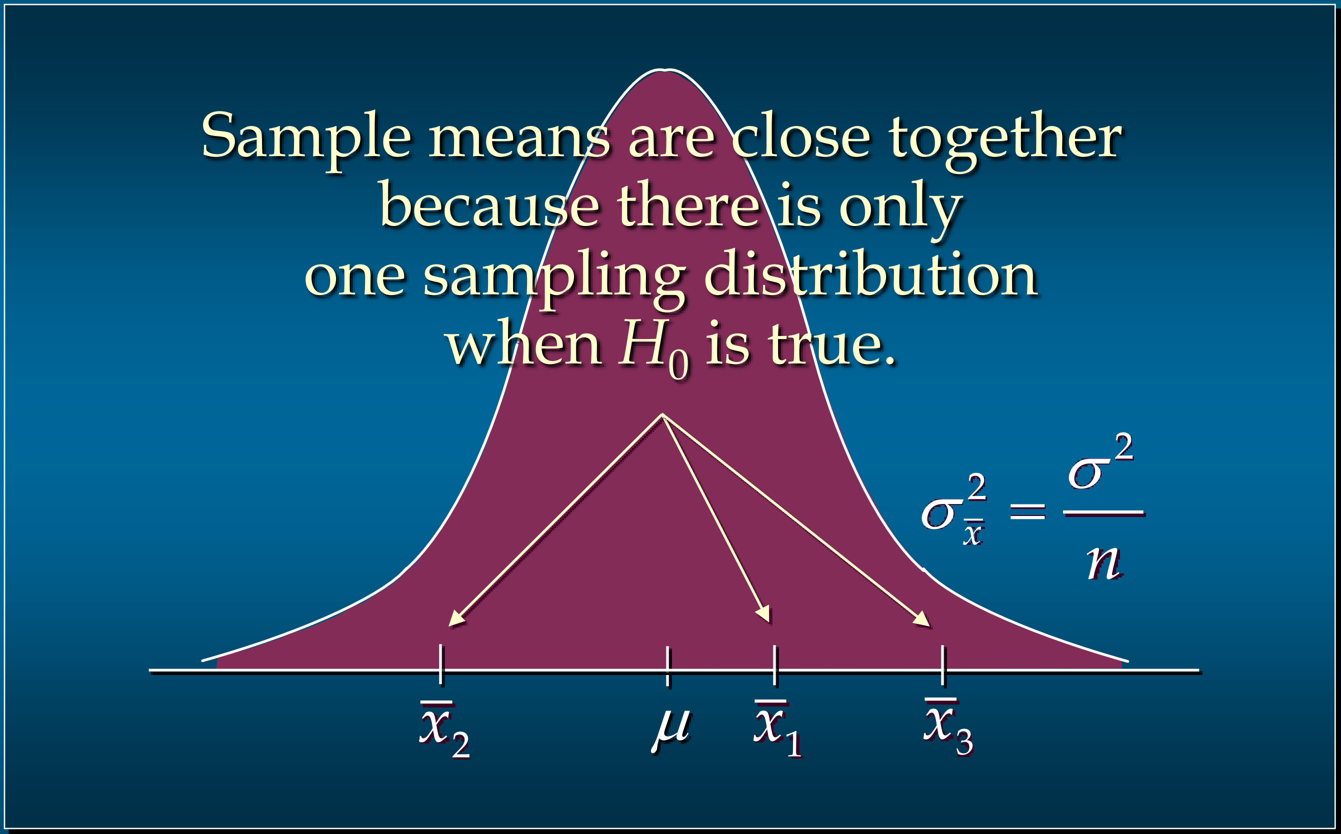
Analysis of Variance: A Conceptual Overview

- Assumptions for Analysis of Variance

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

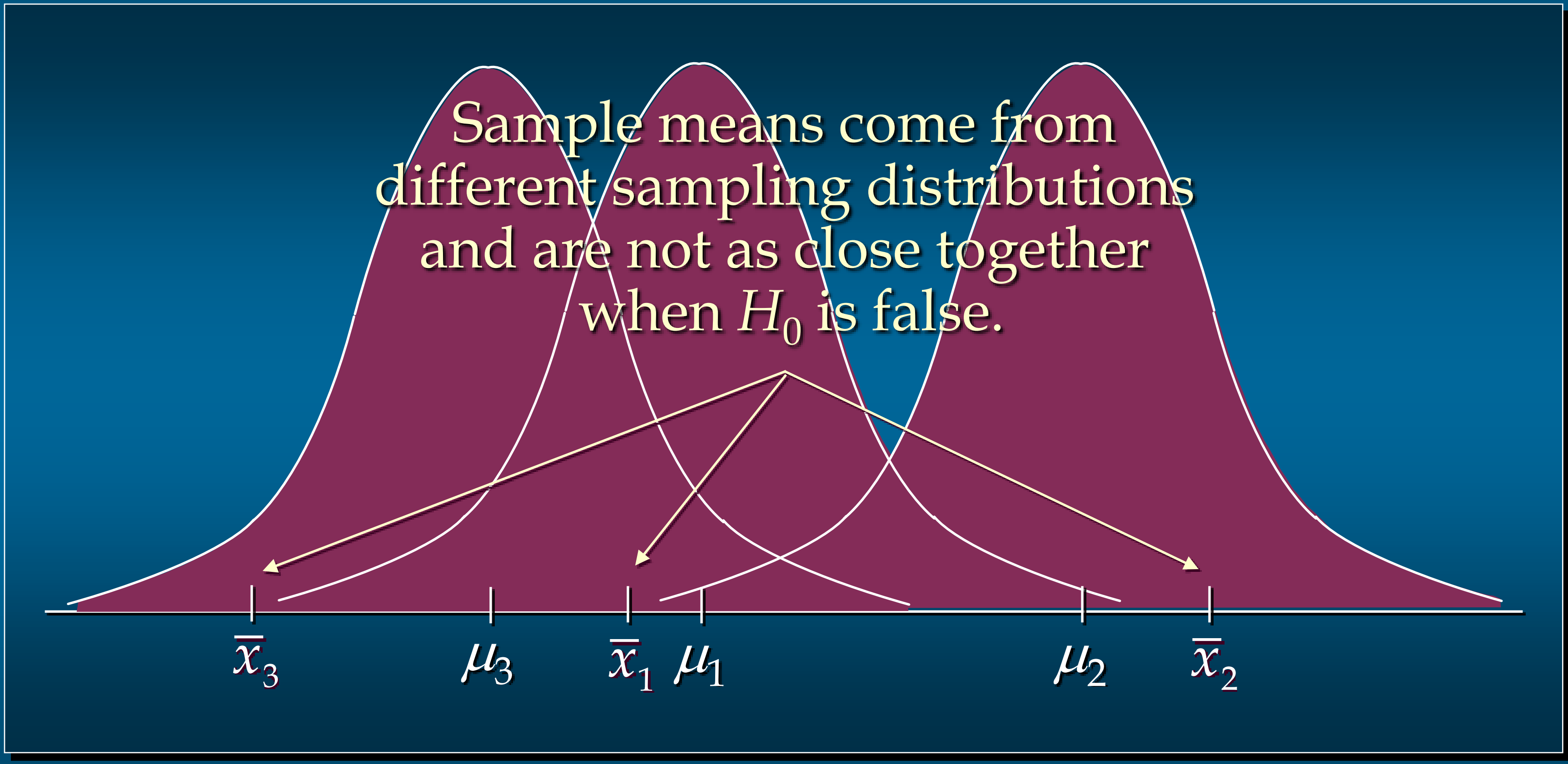
Analysis of Variance: A Conceptual Overview

- Sampling Distribution of $\bar{x}$ Given H_0 is True



Analysis of Variance: A Conceptual Overview

- Sampling Distribution of $\overline{x}$ Given H_0 is False



Analysis of Variance and the 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.

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

Denominator is the degrees of freedom associated with SSTR

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

Within-Treatments Estimate of Population Variance σ^2

- 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.

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

Denominator is the degrees of freedom associated with SSE

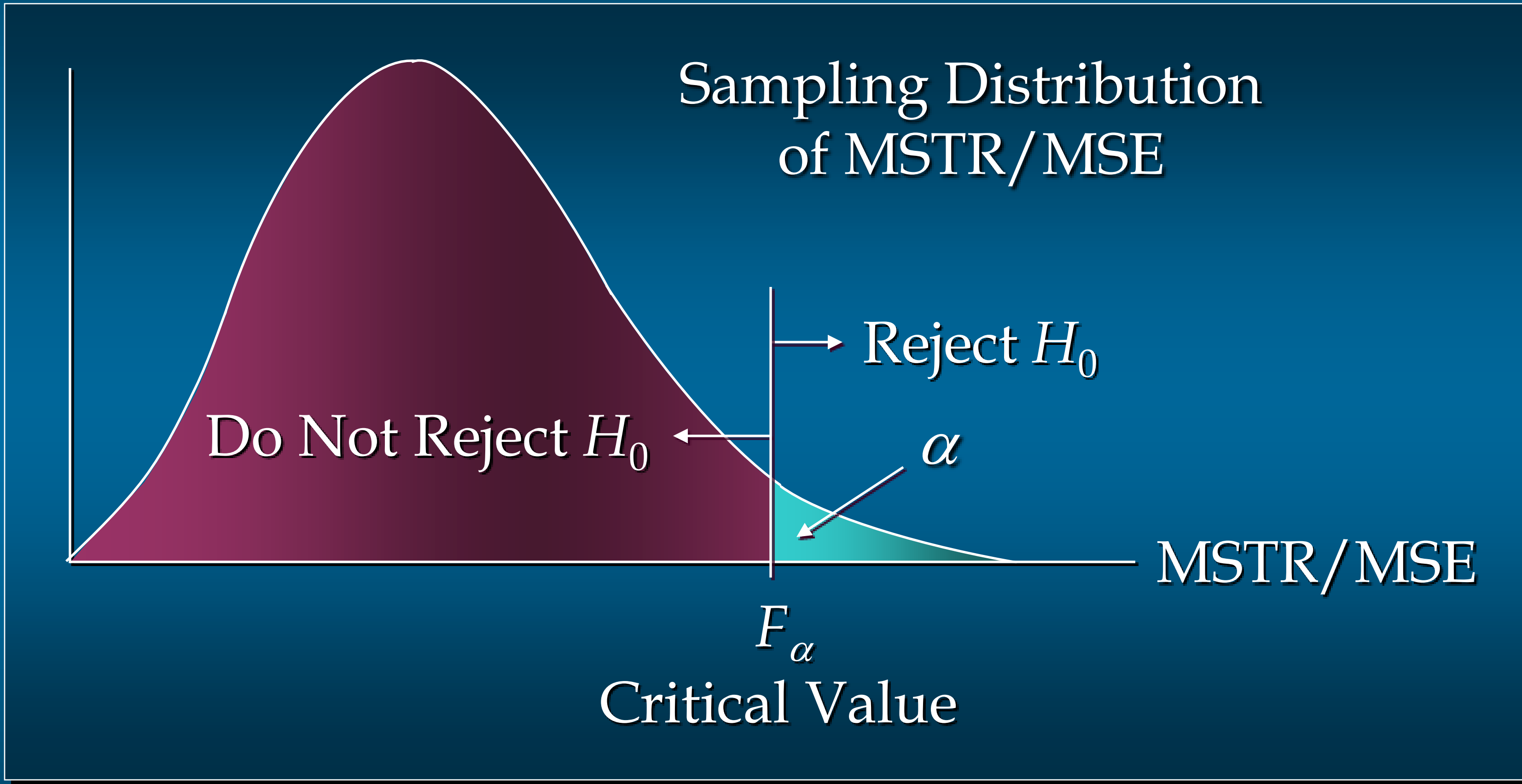
Numerator is called the sum of squares due to error (SSE)

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 d.f. equal to k - 1 and MSE d.f. 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.

Comparing the Variance Estimates: The *F* Test

- Sampling Distribution of MSTR/MSE



ANOVA Table
for a Completely Randomized Design

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	$n_T - 1$			

SST is partitioned into SSTR and SSE.

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

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:

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

ANOVA Table for a Completely Randomized Design

- 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 and the *F* value used to test the hypothesis of equal population means.

Test for the Equality of k Population Means

- Hypotheses

$$H_0\text{: } \mu_1 = \mu_2 = \mu_3 = \cdot \cdot \cdot = \mu_k$$

H_a: Not all population means are equal

- Test Statistic

$$F = \text{MSTR}/\text{MSE}$$

Test for the Equality of k Population Means

- Rejection Rule
  - p-value Approach: **Reject H_0 if 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 d.f. and n_T - k denominator d.f.

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

- Example: AutoShine, Inc.
  - 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.

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

- Example: AutoShine, Inc.
  - 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 I, 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

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

- Hypotheses

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

H_a: Not all the means are equal

where:

μ_1 = mean number of washes using Type 1 wax

μ_2 = mean number of washes using Type 2 wax

μ_3 = 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:

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

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

  $$MSTR = 5.2/(3 - 1) = 2.6$$

- Mean Square Error

  $$SSE = 4(2.5) + 4(3.3) + 4(2.5) = 33.2$$

  $$MSE = 33.2/(15 - 3) = 2.77$$

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

- Rejection Rule

p-Value Approach: Reject H_0 if p-value $\leq .05$

Critical Value Approach: Reject H_0 if $F \geq 3.89$

where $F_{.05} = 3.89$ is based on an F distribution with 2 numerator degrees of freedom and 12 denominator degrees of freedom

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

- Test Statistic

$$F = \text{MSTR}/\text{MSE} = 2.60/2.77 = .939$$

- Conclusion

The p-value is greater than .10, where $F = 2.81$.
(Excel provides a p-value of .42.)
Therefore, we cannot reject H_0.

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

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

- ANOVA Table

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	F	p-Value
Treatments	5.2	2	2.60	.939	.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 $\alpha = .05$.

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

- Example: Reed Manufacturing

  - 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

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

- p -Value and Critical Value Approaches

1. Develop the hypotheses.

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

H_a: Not all the means are equal

where:

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

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

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

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

- p -Value and Critical Value Approaches

  - 2. Specify the level of significance. $\alpha = .05$

  - 3. Compute the value of the test statistic.

    - Mean Square Due to Treatments

      (Sample sizes are all equal.)

      $\bar{\bar{x}} = (55 + 68 + 57)/3 = 60$

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

      MSTR $= 490/(3 - 1) = 245$

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

- p -Value and Critical Value Approaches

  - **3. Compute the value of the test statistic.** (con't.)

    - Mean Square Due to Error

$$\text{SSE} = 4(26.0) + 4(26.5) + 4(24.5) = 308$$

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

$$F = \text{MSTR}/\text{MSE} = 245/25.667 = 9.55$$

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			

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

- p –Value Approach

  - **4. Compute the p –value.**

    With 2 numerator d.f. and 12 denominator d.f., the p-value is .01 for F = 6.93. Therefore, the p-value is less than .01 for F = 9.55.

  - **5. Determine whether to reject H_0.**

    The p-value $\leq$ .05, so we reject H_0.

    We have sufficient evidence to conclude that the mean number of hours worked per week by department managers is not the same at all 3 plant.

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 d.f. and 12 denominator d.f., $F_{.05} = 3.89$.

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 have sufficient evidence to conclude that the mean number of hours worked per week by department managers is not the same at all 3 plant.

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.

Fisher's LSD Procedure

- Hypotheses

$$H_0: \mu_i - \mu_j$$
$$H_a: \mu_i \neq \mu_j$$

- Test Statistic

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

Fisher's LSD Procedure

- Rejection Rule
  - p-value Approach:

Reject H_0 if p-value $\leq \alpha$

  - Critical Value Approach:

Reject H_0 if $t < -t_{a/2}$ or $t > t_{a/2}$

where the value of $t_{a/2}$ 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

$$H_0 : \mu_i - \mu_j$$
$$H_a : \mu_i \neq \mu_j$$

- Test Statistic

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

- Rejection Rule

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

where

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

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.

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

- For $\alpha = .05$ and $n_T - k = 15 - 3 = 12$ degrees of freedom, $t_{.025} = 2.179$

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

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

MSE value was computed earlier

Fisher's LSD Procedure Based on the Test Statistic x_i - x_j

- LSD for Plants 1 and 2
  - Hypotheses (A) $H_0: \mu_1 - \mu_2$
    $H_a: \mu_1 \neq \mu_2$
  - Rejection Rule

    Reject H_0 if $|\bar{x}_1 - \bar{x}_2| > 6.98$
  - Test Statistic

    $|\bar{x}_1 - \bar{x}_2|$ = | 55 – 68 | = 13
  - Conclusion

    The mean number of hours worked at Plant 1 is not equal to the mean number worked at Plant 2.

Fisher's LSD Procedure Based on the Test Statistic x_i - 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

    $$\text{Reject } H_0 \text{ if } \left|\bar{x}_1 - \bar{x}_3\right| > 6.98$$
  - Test Statistic

    $$\left|\bar{x}_1 - \bar{x}_3\right| = |55 - 57| = 2$$
  - 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 $x_i - x_j$

- LSD for Plants 2 and 3
  - Hypotheses (C) $H_0:\ \mu_2 - \mu_3$
    $H_a:\ \mu_2 \neq \mu_3$
  - Rejection Rule

    Reject H_0 if $|\bar{x}_2 - \bar{x}_3| > 6.98$
  - Test Statistic

    $|\bar{x}_2 - \bar{x}_3| = |68 - 57| = 11$
  - Conclusion

    The mean number of hours worked at Plant 2 is not equal to the mean number worked at Plant 3.

Type I Error Rates

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

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

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

End Part A