# Analysis of Variance
# Single Factor

**Dr. Supaporn Erjongmanee**

Department of Computer Engineering
Kasetsart University
fengspe@ku.ac.th



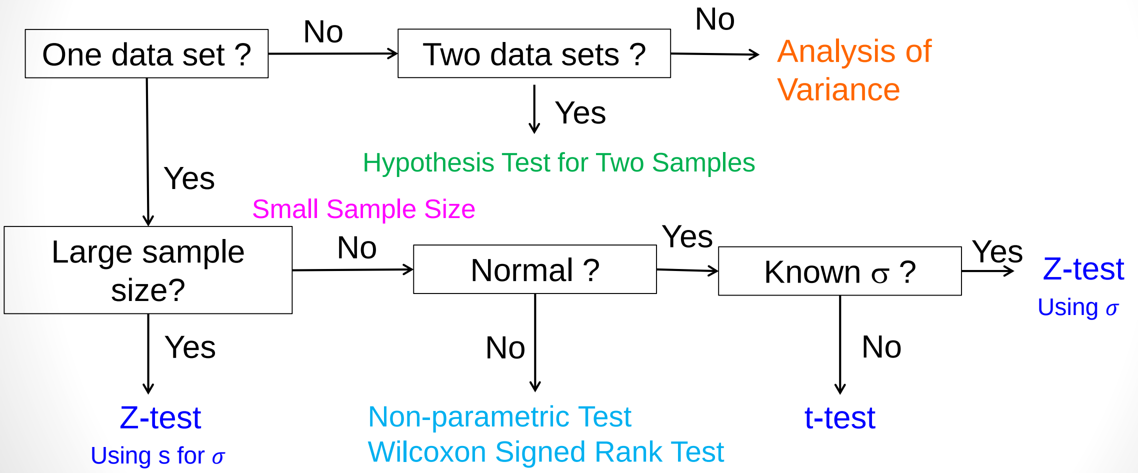


## Outline

- Introduction
- Equal sample size
  - F-test
  - Multiple comparison
- Unequal sample size
  - F-test
  - Multiple comparison
- Multiple comparison discussion
- Random effects model

# Which Test to Choose

- How to choose test to fit data
- <u>For population mean</u>




Supaporn Erjongmanee
fengspe@ku.ac.th

**Statistics in Computer Engineering Applications**
Slide 3

Department of Computer Engineering
Kasetsart University

3

# Introduction : Analysis of Variance

- Analysis of quantitative responses
- In short, ANOVA
- Simplest ANOVA
  - Single-factor or One-way
  - Factorial or Multiple-way
- Examples
  - Study of five brands of gasoline on car efficiency
  - Study of four types of sugar on bacteria growth

Supaporn Erjongmanee
fengspe@ku.ac.th

**Statistics in Computer Engineering Applications**
Slide 4

Department of Computer Engineering
Kasetsart University

4

# Outline

- Introduction
- Equal sample size
  - F-test
  - Multiple comparison
- Unequal sample size
  - F-test
  - Multiple comparison
- Multiple comparison discussion
- Random effects model





# Single-Factor ANOVA : Equal Sample Size

- Compare two or more populations on one factor
- Let
  - $\mu_1$ = mean of treatment (population) 1
  - $\mu_2$ = mean of treatment (population) 2
  - …
  - $\mu_I$ = mean of treatment (population) I

I = number of compared treatments

- Hypothesis
  - $H_0$: $\mu_1 = \mu_2 = \ldots = \mu_I$
  - $H_a$: Not all $\mu_i$'s are equal (at least two of the $\mu_i$'s are different)

## Notation

- Let
  - $X_{ij}$ = Random variable for measurement j of treatment i
  - $x_{ij}$ = Sample value for measurement j of treatment i
  - J = Number of samples in one treatment
  - I = Number of treatments
- (Treatment) sample mean: $\bar{X}_i = \frac{\sum_{j=1}^{J} X_{ij}}{J}$
- Grand mean: $\bar{X} = \frac{\sum_{i=1}^{I} \bar{X}_i}{I} = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} X_{ij}}{IJ}$
- (Treatment) sample variance: $S_i^2 = \frac{\sum_{j=1}^{J}(X_{ij}-\bar{X}_i)^2}{J-1}$

Divided by number of samples in one treatment

Divided by number of samples from treatments



7

## Example

- Experiment testing strength of 4 shipping boxes

| Type | Compression Strength (lb) | | | | | | Sample Mean | Sample SD |
|---|---|---|---|---|---|---|---|---|
| 1 | 655.5 | 788.3 | 734.3 | 721.4 | 679.1 | 699.4 | 713.00 | 46.55 |
| 2 | 789.2 | 772.5 | 786.9 | 686.1 | 732.1 | 774.8 | 756.93 | 40.34 |
| 3 | 737.1 | 639.0 | 696.3 | 671.7 | 717.2 | 727.1 | 698.07 | 37.20 |
| 4 | 535.1 | 628.7 | 542.4 | 559.0 | 586.9 | 520.0 | 562.02 | 39.87 |

Grand Mean = (713.00+756.93+698.07+563.02) / 4
= 682.50



8

# Review: Sample Variance Distribution

- Let $X_1, X_2, \ldots, X_n$ be random sample from a normal distribution with mean value $= \mu$ and standard deviation = $\sigma$.
- Then, sample variance has distribution to be a chi-square distribution with degree of freedom = n-1

$$S^2 \approx \sigma^2 \frac{\chi^2_{n-1}}{(n-1)} \quad \Rightarrow \quad \frac{(n-1)S^2}{\sigma^2} \approx \chi^2_{(n-1)}$$



10

# Basic Assumption

- Assume
  - Distribution of each population is normal with the same variance = $\sigma^2$
- Therefore, each sample $X_{ij}$ comes from normal distribution with
  - $E(X_{ij}) = \mu_i$
  - $V(X_{ij}) = \sigma^2$
  - Hypothesis
    - $H_0$: $\mu_1 = \mu_2 = \ldots = \mu_I$
    - $H_a$: Not all $\mu_i$'s are equal (at least two of the $\mu_i$'s are different)



11

## Sum of Squares

- When $H_0$ is true, all sample means $(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_I)$ should be the same
- Therefore, test statistics will measured from differences of sample means
- Treatment sum of squares ($SST_r$) : Difference between different treatments
  - Sum of difference between each sample mean and grand mean
  - $$SST_r = J(\bar{X}_1 - \bar{X})^2 + J(\bar{X}_2 - \bar{X})^2 + \cdots + J(\bar{X}_I - \bar{X})^2$$
    $$= J \sum_i (\bar{X}_i - \bar{X})^2$$
- Error sum of squares (SSE): Difference within the same treatment
  - Sum of differences between samples and sample mean
  - $$SSE = \sum_i \sum_j (X_{ij} - \bar{X}_i)^2$$



12

## Sum of Squares (cont.)

Difference between means of each treatment and grand mean

- Treatment sum of squares ($SST_r$) :
  - $$SST_r = J \sum_i (\bar{X}_i - \bar{X})^2$$
    $$= J \sum_i (Y_i - \bar{Y})^2$$
    $$= J\,(I-1)\frac{\sum_i (Y_i - \bar{Y})^2}{(I-1)}$$
    $$= (I-1) J\, S_Y^2$$
    $$\frac{SST_r}{\sigma^2} = \frac{(I-1)S_Y^2}{\sigma^2/J}$$
    $$= \frac{(I-1)}{(\sigma^2/J)} \cdot \frac{\sigma_Y^2 \chi_{I-1}^2}{I-1}$$
    $$= \frac{\cancel{(I-1)}\cancel{(\sigma^2/J)}\chi_{I-1}^2}{\cancel{(\sigma^2/J)}\cancel{(I-1)}} \approx \chi_{I-1}^2$$

$$X_i \sim N(\mu, \sigma^2) \rightarrow Y_i = \bar{X}_i \sim N(\mu, \frac{\sigma^2}{J})$$

$$s^2 = \frac{\sum_{i=1}^{n}(x - \bar{x})^2}{n-1}$$

$$S_X^{\ 2} \approx \sigma_X^2 \frac{\chi_{n-1}^2}{(n-1)}$$



13

## Sum of Squares (cont.)

Difference between each sample and means of each treatment

- Error sum of squares (SSE)
  - $SSE = \sum_i \sum_j (X_{ij} - \bar{X}_i)^2$

$= \sum_j (X_{1j} - \bar{X}_1)^2 + \sum_j (X_{2j} - \bar{X}_2)^2 + \ldots + \sum_j (X_{Ij} - \bar{X}_I)^2$

$= (J-1)S_1^2 + (J-1)S_2^2 + \cdots + (J-1)S_I^2$

$= (J-1)[S_1^2 + S_2^2 + \ldots + S_I^2]$

Each $X_i$ has the same variance

$= I(J-1)S^2$

$$\frac{SSE}{\sigma^2} = \frac{I(J-1)S^2}{\sigma^2} \approx \chi^2_{I(J-1)}$$

$$S^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

$$X_i \sim N(\mu, s^2)$$

$$\frac{(n-1)S^2}{\sigma^2} \approx \chi^2_{(n-1)}$$



14

## F distribution

- Let $X_1$ and $X_2$ be <u>independent</u> *chi-squared* random variables with $v_1$ and $v_2$ degrees of freedom

$$F_{v1,v2} = \frac{X_1/v_1}{X_2/v_2}$$

- Generally, sample variance has sampling distribution in term of chi-squared distribution with degree of freedom = n -1

$$S^2 \approx \sigma^2 \frac{\chi^2_{n-1}}{(n-1)}$$

- Let $S_1^2$ and $S_2^2$ be sample variances with chi-squared distribution

$$\frac{(m-1)S_1^2}{\sigma_1^2} \approx \chi^2_{(m-1)} \quad \text{and} \quad \frac{(n-1)S_2^2}{\sigma_2^2} \approx \chi^2_{(n-1)}$$

$$F_{m-1,n-1} = \frac{\frac{(m-1)S_1^2/\sigma_1^2}{m-1}}{\frac{(n-1)S_2^2/\sigma_2^2}{n-1}} = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}$$



15

## F distribution (cont.)

- Let $X_1$ and $X_2$ be independent chi-squared random variables with $v_1$ and $v_2$ degrees of freedom

$$F_{v1,\,v2} = \frac{X_1/v_1}{X_2/v_2}$$

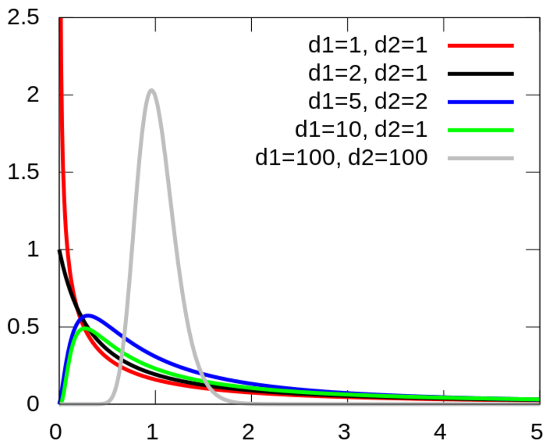



*Image source: http://en.wikipedia.org/wiki/F-distribution*



16

## Sum of Squares (cont.)

- Treatment sum of squares ($SST_r$) :

$$\frac{SST_r}{\sigma^2} = \frac{(I-1)S_Y^2}{\sigma^2/J} = \frac{(I-1)(\sigma^2/J)\chi_{I-1}^2}{(\sigma^2/J)(I-1)} \approx \chi_{I-1}^2$$

- Error sum of squares (SSE)

$$\frac{SSE}{\sigma^2} = \frac{I(J-1)S^2}{\sigma^2} \approx \chi_{I(J-1)}^2$$

Use as test statistic

$$\frac{\frac{SSTr}{\sigma^2}/(I-1)}{\frac{SSE}{\sigma^2}/I(J-1)} = \frac{SSTr/(I-1)}{SSE/I(J-1)} \approx F_{I-1,I(J-1)}$$



17

## To Perform F-Test

Mean of chi-square distribution = degree of freedom

- Test statistic:

$$f = F_{I-1,I(J-1)} = \frac{SSTr/(I-1)}{SSE/(I(J-1))}$$

$$= \frac{MSTr}{MSE}$$

$$Mean = \frac{Sum}{n-1}$$

$$E(\frac{SST_r}{\sigma^2}) = I-1$$

$$E(\frac{SSE}{\sigma^2}) = I(J-1)$$

$$E\left(\frac{SSTr}{I-1}\right) = E(MST_r) = \sigma^2$$

- If $H_o$ is true ($\bar{X}_i's$ are about the same), MSTr and MSE are unbiased estimates of $\sigma^2$. f ~ 1
- If $H_o$ is false ($\bar{X}_i's$ are not the same), E(MSTr) > $\sigma^2$. f is large.

$$E(\frac{SSE}{I(J-1)}) = E(MSE) = \sigma^2$$

- Rejection region for large f?



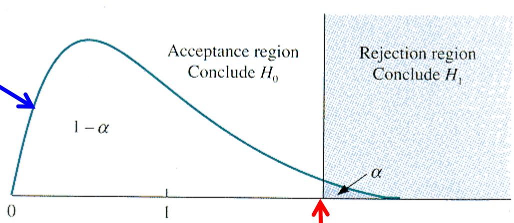
18

## F-Test

- Test statistic:

$$F_{I-1,I(J-1)} = \frac{\frac{SSTr}{\sigma^2}/(I-1)}{\frac{SSE}{\sigma^2}/(I(J-1))} = \frac{SSTr/(I-1)}{SSE/(I(J-1))} = \frac{MSTr}{MSE}$$

F distribution for $v_1$ and $v_2$




$F_{\alpha, v1, v2}$

If test statistic > $F_{\alpha, v1, v2}$, reject $H_0$

*Image source: http://www.unc.edu/~nielsen/soci708/m16/m2009.gif*



19

## Summary : ANOVA (cont.)

- SST = SSTr + SSE

$\sum_i \sum_j (x_{ij} - \bar{x})^2 = \sum_i \sum_j (x_{ij} - \bar{x}_i + \bar{x}_i - \bar{x})^2$

$= \sum_i \sum_j (x_{ij} - \bar{x}_i)^2 + \sum_i \sum_j (\bar{x}_i - \bar{x})^2 - 2 \sum_i \sum_j (x_{ij} - \bar{x}_i)(\bar{x}_i - \bar{x})$

$= \text{SSE} + \text{SSTr} - \cancel{2 \sum_i \sum_j (x_{ij} - \bar{x}_i)(\bar{x}_i - \bar{x})}$ 0



20

## Summary : ANOVA (cont.)

$\sum_i \sum_j (x_{ij} - \bar{x})^2 = \sum_i \sum_j (x_{ij} - \bar{x}_i + \bar{x}_i - \bar{x})^2$

- Test statistic:

$$F_{I-1, I(J-1)} = \frac{\frac{SSTr}{\sigma^2}/(I-1)}{\frac{SSE}{\sigma^2}/(I(J-1))} = \frac{SSTr/(I-1)}{SSE/(I(J-1))} = \frac{MSTr}{MSE}$$

| | df | Sum of Squares (SS) | Mean Square (MS) | f |
|---|---|---|---|---|
| Treatment | I-1 | $J \sum_i (\bar{x}_i - \bar{x})^2$ | SSTr/(I-1) | MSTr / MSE |
| Error | I(J-1) | $\sum_i \sum_j (x_{ij} - \bar{x}_i)^2$ = SST - SSTr | SSE/(I(J-1)) | |
| Total | IJ-1 | $\sum_i \sum_j (x_{ij} - \bar{x})^2$ | | |



21

## Summary : ANOVA (cont.)

- Test statistic:

$$F_{I-1,I(J-1)} = \frac{\frac{SSTr}{\sigma^2}/(I-1)}{\frac{SSE}{\sigma^2}/(I(J-1))} = \frac{SSTr/(I-1)}{SSE/(I(J-1))} = \frac{MSTr}{MSE}$$

Another option: Sample-based computation

| | df | Sum of Squares (SS) | Mean Square (MS) | f |
|---|---|---|---|---|
| Treatment | I-1 | $\frac{1}{J}\sum_{i=1}^{I}(\sum_{j=1}^{J} x_{ij})^2 - \frac{1}{IJ}(\sum_{i=1}^{I}\sum_{j=1}^{J} x_{ij})^2$ | SSTr/(I-1) | MSTr / MSE |
| Error | I(J-1) | SST - SSTr | SSE/(I(J-1)) | |
| Total | IJ-1 | $\sum_{i=1}^{I}\sum_{j=1}^{J} x_{ij}^2 - \frac{1}{IJ}(\sum_{i=1}^{I}\sum_{j=1}^{J} x_{ij})^2$ | | |





## Example 1

- Experiment degree of soiling on 3 mixtures of fabric and polymer
- Prove whether 3 mixture means are the same at $\alpha$ = 0.01

| Mixture | Degree of soiling | | | | |
|---|---|---|---|---|---|
| 1: | 0.56 | 1.12 | 0.90 | 1.07 | 0.94 |
| 2: | 0.72 | 0.69 | 0.87 | 0.78 | 0.91 |
| 3: | 0.62 | 1.08 | 1.07 | 0.99 | 0.93 |

- Let
  - $\mu_1$ = mean of mixture 1
  - $\mu_2$ = mean of mixture 2
  - $\mu_3$ = mean of mixture 3
  - I = 3, J = 5
- Hypothesis
  - $H_0$: $\mu_1 = \mu_2 = \mu_3$
  - $H_a$: Not all $\mu_i$'s are equal (at least two of the $\mu_i$'s are different)

## Example 1 (cont.)

- Experiment degree of soiling on 3 mixtures

| Mixture | Degree of soiling | | | | | $\bar{x}_i$ |
|---|---|---|---|---|---|---|
| 1: | 0.56 | 1.12 | 0.90 | 1.07 | 0.94 | 0.918 |
| 2: | 0.72 | 0.69 | 0.87 | 0.78 | 0.91 | 0.794 |
| 3: | 0.62 | 1.08 | 1.07 | 0.99 | 0.93 | 0.938 |

$\bar{x}$= 0.883

- Fill ANOVA table

| | df | Sum of Squares (SS) | Mean Square (MS) | f |
|---|---|---|---|---|
| Treatment | 2 | 0.0608 | 0.0304 | 0.99 |
| Error | 12 | 0.3701 = 0.4309 – 0.0608 | 0.0308 | |
| Total | 14 | 0.4309 | | |



24

## Example 1 (cont.)

| | df | Sum of Squares (SS) | Mean Square (MS) | f |
|---|---|---|---|---|
| Treatment | 2 | 0.0608 | 0.0304 | 0.99 |
| Error | 12 | 0.3701 = 0.4309 – 0.0608 | 0.0308 | |
| Total | 14 | 0.4309 | | |

- Rejection region: given $\alpha$ = 0.01
  - $F_{0.01, 2, 12} = 6.93$
- $H_0$ is not rejected.
- All mixture means are equals.



25

## Example 2

- Experiment consistency lab measurements from 7 labs

| Lab 1 | Lab 2 | Lab 3 | Lab 4 | Lab 5 | Lab 6 | Lab 7 |
|---|---|---|---|---|---|---|
| 4.13 | 3.86 | 4.00 | 3.88 | 4.02 | 4.02 | 4.00 |
| 4.07 | 3.85 | 4.02 | 3.88 | 3.95 | 3.86 | 4.02 |
| 4.04 | 4.08 | 4.01 | 3.91 | 4.02 | 3.96 | 4.03 |
| 4.07 | 4.11 | 4.01 | 3.95 | 3.89 | 3.97 | 4.04 |
| 4.05 | 4.08 | 4.04 | 3.92 | 3.91 | 4.00 | 4.10 |
| 4.04 | 4.01 | 3.99 | 3.97 | 4.01 | 3.82 | 3.81 |
| 4.02 | 4.02 | 4.03 | 3.92 | 3.89 | 3.98 | 3.91 |
| 4.06 | 4.04 | 3.97 | 3.90 | 3.89 | 3.99 | 3.96 |
| 4.10 | 3.97 | 3.98 | 3.97 | 3.99 | 4.02 | 4.05 |
| 4.04 | 3.95 | 3.98 | 3.90 | 4.00 | 3.93 | 4.06 |





## Example 2

- Experiment consistency lab measurements from 7 labs
- Let
  - $\mu_1$ = mean of measurements from lab 1
  - $\mu_2$ = mean of measurements from lab 2
  - …
  - $\mu_7$ = mean of measurements from lab 7
  - I = 7, J = 10
- Hypothesis
  - $H_0$: $\mu_1 = \mu_2 = \ldots = \mu_7$
  - $H_a$: Not all $\mu_i$'s are equal (at least two of the $\mu_i$'s are different)

# Example 2

- Experiment consistency lab measurements from 7 labs

| Lab 1 | Lab 2 | Lab 3 | Lab 4 | Lab 5 | Lab 6 | Lab 7 |
|---|---|---|---|---|---|---|
| 4.13 | 3.86 | 4.00 | 3.88 | 4.02 | 4.02 | 4.00 |
| 4.07 | 3.85 | 4.02 | 3.88 | 3.95 | 3.86 | 4.02 |
| 4.04 | 4.08 | 4.01 | 3.91 | 4.02 | 3.96 | 4.03 |
| 4.07 | 4.11 | 4.01 | 3.95 | 3.89 | 3.97 | 4.04 |
| 4.05 | 4.08 | 4.04 | 3.92 | 3.91 | 4.00 | 4.10 |
| 4.04 | 4.01 | 3.99 | 3.97 | 4.01 | 3.82 | 3.81 |
| 4.02 | 4.02 | 4.03 | 3.92 | 3.89 | 3.98 | 3.91 |
| 4.06 | 4.04 | 3.97 | 3.90 | 3.89 | 3.99 | 3.96 |
| 4.10 | 3.97 | 3.98 | 3.97 | 3.99 | 4.02 | 4.05 |
| 4.04 | 3.95 | 3.98 | 3.90 | 4.00 | 3.93 | 4.06 |

| $\bar{x}_1$ | $\bar{x}_2$ | $\bar{x}_3$ | $\bar{x}_4$ | $\bar{x}_5$ | $\bar{x}_6$ | $\bar{x}_7$ |
|---|---|---|---|---|---|---|
| 4.062 | 3.997 | 4.003 | 3.92 | 3.957 | 3.955 | 3.998 |

$\bar{x}$=3.985



28

# Example 2 (cont.)

7 labs. Each lab collects 10 samples.

| | df | Sum of Squares (SS) | Mean Square (MS) | f |
|---|---|---|---|---|
| Treatment | 6 | 0.125 | 0.0208 | 5.66 |
| Error | 63 | 0.356-0.125 = 0.231 | 0.0037 | |
| Total | 69 | 0.356 | | |

- There is no $F_{\alpha, 6, 63}$ in the f-table
- Examine closest F critical values: 60 vs. 70
  - $F_{0.1, 6, 60} = 1.8747$, $F_{0.05, 6, 60} = 2.2541$, $F_{0.01, 6, 60} = 3.1187$
  - $F_{0.1, 6, 70} = 1.8600$, $F_{0.05, 6, 70} = 2.2312$, $F_{0.01, 6, 70} = 3.0712$
  - For common values of α, $F_{\alpha, 6, 63} < f$
- $H_0$ is rejected.
- There are differences in 7 means.



29

## Example 2 (cont.)

7 labs. Each lab collects 10 samples.

| | df | Sum of Squares (SS) | Mean Square (MS) | f |
|---|---|---|---|---|
| Treatment | 6 | 0.125 | 0.0208 | 5.66 |
| Error | 63 | 0.356 | | |
| Total | 69 | 0.356 | | |

- Reject $H_0$ for large test statistic -> choose small α
- Do not reject $H_0$ for small test statistic -> choose large α

- There is no $F_{\alpha, 6, 63}$ in the f-table
- Examine closest F critical values: 60 vs. 70
  - $F_{0.1, 6, 60} = 1.8747$, $F_{0.05, 6, 60} = 2.2541$, $F_{0.01, 6, 60} = 3.1187$
  - $F_{0.1, 6, 70} = 1.8600$, $F_{0.05, 6, 70} = 2.2312$, $F_{0.01, 6, 70} = 3.0712$
  - For common values of α, $F_{\alpha, 6, 63} < f$
- $H_0$ is rejected.
- There are differences in 7 means.

How to choose α for rejecting null hypothesis? Small or Large?



30

## Outline

- Introduction
- Equal sample size
  - F-test
  - Multiple comparison
- Unequal sample size
  - F-test
  - Multiple comparison
- Random effects model



31

## Multiple Comparisons

- When $H_0$ for ANOVA is rejected, how many means are different from each other?
- Procedures:
  - Find confidence interval of pairwise difference $\mu_i - \mu_j$
  - If confidence interval for any pairwise difference $\mu_i - \mu_j$ does not include zero, we determine that $\mu_i, \mu_j$ are significantly different from each other





## Studentized Range Distribution

- Let $Z_1, Z_2, \ldots, Z_m$ be *m* independent *standard normal* random variables
- Let W be a chi-squared random variable with degree of freedom = v, and independent of the $Z_i$'s
- Then, Q distribution, called studentized range distribution is

$$Q = \frac{\max |Z_i - Z_j|}{\sqrt{\frac{W}{v}}} \qquad \text{where } W = \frac{SSE}{\sigma^2} = \frac{I(J-1)MSE}{\sigma^2}$$

- This Q distribution has 2 parameters: m and v
  - Hence, it is denoted by $Q_{\alpha,m,v}$

## Studentized Range Distribution (cont.)

- From

$$Z_i = \frac{\bar{X}_i - \mu_i}{\sigma/\sqrt{J}}, \qquad W = \frac{SSE}{\sigma^2} = \frac{I(J-1)MSE}{\sigma^2}, \qquad m = I, \qquad v = I(J-1)$$

$$Q = \frac{\max|Z_i - Z_j|}{\sqrt{\frac{W}{v}}} = \frac{\max\left|\frac{\bar{X}_i - \mu_i}{\sigma/\sqrt{J}} - \frac{\bar{X}_j - \mu_j}{\sigma/\sqrt{J}}\right|}{\sqrt{\frac{\frac{I(J-1)MSE}{\sigma^2}}{I(J-1)}}} = \frac{\max|\bar{X}_i - \bar{X}_j \;- (\mu_i - \mu_j)|}{\sqrt{MSE/J}}$$

$$1 - \alpha = P\left(\frac{\max|\bar{X}_i - \bar{X}_j \;- (\mu_i - \mu_j)|}{\sqrt{MSE/J}} \le Q_{\alpha,I,I(J-1)}\right)$$

$$= P\left(\frac{|\bar{X}_i - \bar{X}_j \;- (\mu_i - \mu_j)|}{\sqrt{MSE/J}} \le Q_{\alpha,I,I(J-1)} \; for\ all\ i,j\right)$$



34

## Studentized Range Distribution (cont.)

$$1 - \alpha = P\left(\frac{|\bar{X}_i - \bar{X}_j \;- (\mu_i - \mu_j)|}{\sqrt{MSE/J}} \le Q_{\alpha,I,I(J-1)} \; for\ all\ i,j\right)$$

$$= P\left(-Q_{\alpha,I,I(J-1)}\sqrt{\frac{MSE}{J}} \le \bar{X}_i - \bar{X}_j \;- (\mu_i - \mu_j) \le Q_{\alpha,I,I(J-1)}\sqrt{\frac{MSE}{J}} \, for\ all\ i,j\right)$$

$$= P\left(\underbrace{\boxed{\bar{X}_i - \bar{X}_j - Q_{\alpha,I,I(J-1)}\sqrt{\frac{MSE}{J}}} \le \mu_i - \mu_j \le \boxed{\bar{X}_i - \bar{X}_j + Q_{\alpha,I,I(J-1)}\sqrt{\frac{MSE}{J}}}}_{\text{Confidence intervals between one pair of } \mu_i - \mu_j} for\ all\ i,j\right)$$

- There are $\binom{I}{2} = \frac{I(I-1)}{2}$ confidence intervals



35

## Studentized Range Distribution (cont.)

$$1-\alpha = P\left(\bar{X}_i - \bar{X}_j - Q_{\alpha,I,I(J-1)}\sqrt{\frac{MSE}{J}} \le \mu_i - \mu_j \le \bar{X}_i - \bar{X}_j + Q_{\alpha,I,I(J-1)}\sqrt{\frac{MSE}{J}}\ for\ all\ i,j\right)$$

Expect difference of two means is not more than this value $Q_{\alpha,I,I(J-1)}\sqrt{\frac{MSE}{J}}$

- The value w = $Q_{\alpha,I,I(J-1)}\sqrt{\frac{MSE}{J}}$ is called
  Tukey's honestly significantly difference (HSD)





## Multiple Comparisons : Equal Sample Size

- When some means are not all equal, how to specify which mean is different from others
- Procedure
  1. Find Tukey's Honestly Significant Difference (HSD)

     $$HSD_\alpha = q_{\alpha,I,I(J-1)}\sqrt{\frac{MSE}{J}}$$

     - $q_{\alpha,\ I,\ I(J-1)}$ = q-value from studentized range distribution with 2 degrees of freedom *I, I(J-1)*
  2. Sort sample means in increasing order
  3. Underline pairs that differ less than $HSD_\alpha$
  4. Any pair without underline are considered as significantly different.

# Example 1

| | df | Sum of Squares (SS) | Mean Square (MS) | f |
|---|---|---|---|---|
| Treatment | 2 | 0.0608 | 0.0304 | 0.99 |
| Error | 12 | 0.3701 = 0.4309 – 0.0608 | 0.0308 | |
| Total | 14 | 0.4309 | | |

- Experiment degree of soiling on 3 mixtures
- Group mixture means at $\alpha$ = 0.01

| Mixture | Degree of soiling | | | | | $\bar{x}_i$ |
|---|---|---|---|---|---|---|
| 1: | 0.56 | 1.12 | 0.90 | 1.07 | 0.94 | 0.918 |
| 2: | 0.72 | 0.69 | 0.87 | 0.78 | 0.91 | 0.794 |
| 3: | 0.62 | 1.08 | 1.07 | 0.99 | 0.93 | 0.938 |

$$HSD_\alpha = q_{\alpha,I,I(J-1)}\sqrt{\frac{MSE}{J}} = q_{0.01,3,12}\sqrt{\frac{0.0308}{5}} = 5.05\sqrt{\frac{0.0308}{5}} = 0.396$$

- Sort sample means: $\underset{\bar{x}_2}{0.794}$, $\underset{\bar{x}_1}{0.918}$, $\underset{\bar{x}_3}{0.938}$

0.124 0.020

- One group of mixture means

Note that $H_0$ is not rejected.



38

# Example 2

- Test on 5 brands of automobile oil filters
  - Use 9 samples for each brands
- $\bar{x}_1$= 14.5, $\bar{x}_2$= 13.8, $\bar{x}_3$= 13.3, $\bar{x}_4$= 14.3, $\bar{x}_5$= 13.1

| | df | Sum of Squares (SS) | Mean Square (MS) | f |
|---|---|---|---|---|
| Treatment | 4 | 13.32 | 3.33 | 37.84 |
| Error | 40 | 3.53 | 0.088 | |
| Total | 44 | 16.85 | | |

- Rejection region:
  - $F_{0.05, 4, 40}$ = 2.61
- $H_0$ is rejected
- Find Tukey's HSD to see mean differences



39

## Example 2 (cont.)

- $\bar{x}_1$= 14.5, $\bar{x}_2$= 13.8, $\bar{x}_3$= 13.3, $\bar{x}_4$= 14.3, $\bar{x}_5$= 13.1

| | df | Sum of Squares (SS) | Mean Square (MS) | f |
|---|---|---|---|---|
| Treatment | 4 | 13.32 | 3.33 | 37.84 |
| Error | 40 | 3.53 | 0.088 | |
| Total | 44 | 16.85 | | |

$$HSD_\alpha = q_{\alpha,I,I(J-1)}\sqrt{\frac{MSE}{J}} = q_{0.05,5,40}\sqrt{\frac{0.088}{9}} = 4.04\sqrt{\frac{0.088}{9}} = 0.399$$

$\bar{x}_5$ $\bar{x}_3$ $\bar{x}_2$ $\bar{x}_4$ $\bar{x}_1$

- Sort sample means: 13.1, 13.3, 13.8, 14.3, 14.5
- 3 groups of means:
  - $\bar{x}_5, \bar{x}_3$ are not significantly different from each other
  - $\bar{x}_4, \bar{x}_1$ are not significantly different from each other
  - $\bar{x}_2$ is significantly different from $\bar{x}_5, \bar{x}_3$ and $\bar{x}_4, \bar{x}_1$



## Example 2 (cont.)

- If use another value for sample mean and same $HSD_\alpha$:
  - $\bar{x}_1$= 14.5, $\bar{x}_2$= 14.15, $\bar{x}_3$= 13.3, $\bar{x}_4$= 14.3, $\bar{x}_5$= 13.1

$$HSD_\alpha = q_{0.05,5,40}\sqrt{\frac{0.088}{9}} = 4.04\sqrt{\frac{0.088}{9}} = 0.399$$

$\bar{x}_5$ $\bar{x}_3$ $\bar{x}_2$ $\bar{x}_4$ $\bar{x}_1$

- Sort sample means: 13.1, 13.3, 14.15, 14.3, 14.5

- 2 groups of means
  - $\bar{x}_5, \bar{x}_3$ are not significantly different from each other
  - $\bar{x}_2, \bar{x}_4, \bar{x}_1$ are not significantly different from each other

## Example 3

| | df | Sum of Squares (SS) | Mean Square (MS) | f |
|---|---|---|---|---|
| Treatment | 3 | 5882.3575 | 1960.7858 | 21.09 |
| Error | 16 | 1487.4000 | 92.9625 | |
| Total | 19 | 7369.7575 | | |

- For another data set:
  - $\bar{x}_1$= 79.28, $\bar{x}_2$= 61.54, $\bar{x}_3$= 47.92, $\bar{x}_4$= 32.76
- I = 4, J = 5

$$HSD_\alpha = q_{\alpha,I,I(J-1)}\sqrt{\frac{MSE}{J}} = q_{0.05,4,16}\sqrt{\frac{92.9625}{5}} = 4.05\sqrt{\frac{92.9625}{5}} = 17.47$$

- Sort sample means: $\underset{\bar{x}_4}{32.76}$, $\underset{\bar{x}_3}{47.92}$, $\underset{\bar{x}_2}{61.54}$, $\underset{\bar{x}_1}{79.28}$

Although $\bar{x}_4, \bar{x}_2$ are different from each other, they are not different from $\bar{x}_3$

- 2 groups of means
  - $\bar{x}_4, \bar{x}_3, \bar{x}_2$ are not significantly different from each other
  - $\bar{x}_1$ is significantly different from each other



42

## Outline

- Introduction
- Equal sample size
  - F-test
  - Multiple comparison
- Unequal sample size
  - F-test
  - Multiple comparison
- Multiple comparison discussion
- Random effects model



43

## Single ANOVA with Unequal Size

- ANOVA table:
  - N = Number of all samples = $\sum_{i=1}^{I} J_i$
  - $\boldsymbol{I}$ = Number of treatments
  - $J_i$ = Number of samples for treatment i, i = 1, 2, …, $\boldsymbol{I}$

$$SST_r = \sum_{i=1}^{I}\sum_{j=1}^{J_i}(\bar{x}_{ij} - \bar{x})^2 = \sum_{i=1}^{I}\frac{1}{J_i}(\sum_{j=1}^{J_i} X_{ij})^2 - \frac{1}{n}(\sum_{i=1}^{I}\sum_{j=1}^{J_i} X_{ij})^2 \quad \boxed{\text{df = I-1}}$$

$$SST = \sum_{i=1}^{I}\sum_{j=1}^{J_i}(x_{ij} - \bar{x})^2 = \sum_{i=1}^{I}\sum_{j=1}^{J_i} X_{ij}^2 - \frac{1}{n}(\sum_{i=1}^{I}\sum_{j=1}^{J_i} X_{ij})^2 \quad \boxed{\text{df = N-1}}$$

$$SSE = \sum_{i=1}^{I}\sum_{j=1}^{J_i}(x_{ij} - \bar{x}_i)^2 = SST - SSTr \quad \boxed{\text{df = N-I}}$$

Supaporn Erjongmanee
fengspe@ku.ac.th

Department of Computer Engineering
Kasetsart University

44

## Single ANOVA with Unequal Size

- ANOVA table:
  - N = Number of all samples = $\sum_{i=1}^{I} J_i$
  - $\boldsymbol{I}$ = Number of treatments
  - $J_i$ = Number of samples for treatment i, i = 1, 2, …, $\boldsymbol{I}$

| | df | Sum of Squares (SS) | Mean Square (MS) | f |
|---|---|---|---|---|
| Treatment | $\boldsymbol{I}$ - 1 | $\sum_{i=1}^{I}\sum_{j=1}^{J_i}(\bar{x}_{ij} - \bar{x})^2$ | SSTr/($\boldsymbol{I}$-1) | MSTr / MSE |
| Error | N - $\boldsymbol{I}$ | SSE = $\sum_{i=1}^{I}\sum_{j=1}^{J_i}(x_{ij} - \bar{x}_i)^2$ = SST - SSTr | SSE/(N - $\boldsymbol{I}$) | |
| Total | N - 1 | $\sum_{i=1}^{I}\sum_{j=1}^{J_i}(x_{ij} - \bar{x})^2$ | | |

…of Computer Engineering
Kasetsart University

45

## Single ANOVA with Unequal Size

- ANOVA table:
  - N = Number of all samples = $\sum_{i=1}^{I} J_i$
  - $\boldsymbol{I}$ = Number of treatments
  - $J_i$ = Number of samples for treatment i, , 2, …, $\boldsymbol{I}$

Another option: Sample-based computation

| | df | Sum of Squares (SS) | Mean Square (MS) | f |
|---|---|---|---|---|
| Treatment | $\boldsymbol{I}$ - 1 | $\sum_{i=1}^{I} \frac{1}{J_i} (\sum_{j=1}^{J_i} X_{ij})^2 - \frac{1}{n} (\sum_{i=1}^{I} \sum_{j=1}^{J_i} X_{ij})^2$ | SSTr/($\boldsymbol{I}$-1) | MSTr / MSE |
| Error | N - $\boldsymbol{I}$ | SSE = $\sum_{i=1}^{I} \sum_{j=1}^{J_i} (X_{ij} - \bar{X}_i)^2$ = SST - SSTr | SSE/(N - $\boldsymbol{I}$) | |
| Total | N - 1 | $\sum_{i=1}^{I} \sum_{j=1}^{J_i} X_{ij}^2 - \frac{1}{n} (\sum_{i=1}^{I} \sum_{j=1}^{J_i} X_{ij})^2$ | | |



46

## Example 1

- Measure strength of Mg-based alloys from 3 processes
- Show that alloys from 3 processes are the same are the same at $\alpha$ = 0.001

| Process | Observations | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Permanent Molding (1) | 45.5 | 45.3 | 45.4 | 44.4 | 44.6 | 43.9 | 44.6 | 44.0 |
| Die casting (2) | 44.2 | 43.9 | 44.7 | 44.2 | 44.0 | 43.8 | 44.6 | 43.1 |
| Plaster molding (3) | 46.0 | 45.9 | 44.8 | 46.2 | 45.1 | 45.5 | | |

- Let
  - $\mu_1$ = mean of alloy 1, $\mu_2$ = mean of alloy 2
  - $\mu_3$ = mean of alloy 3, I = 3, $J_1 = J_2 = 8$, $J_3 = 6$, N = 22
- Hypothesis
  - $H_0$: $\mu_1 = \mu_2 = \mu_3$
  - $H_a$: Not all $\mu_i$'s are equal (at least two of the $\mu_i$'s are different)



47

## Example 1 (cont.)

| Process | Observations | | | | | | | | $\sum_{j=1}^{J_i} x_{ij}$ |
|---|---|---|---|---|---|---|---|---|---|
| Permanent Molding (1) | 45.5 | 45.3 | 45.4 | 44.4 | 44.6 | 43.9 | 44.6 | 44.0 | 357.7 |
| Die casting (2) | 44.2 | 43.9 | 44.7 | 44.2 | 44.0 | 43.8 | 44.6 | 43.1 | 352.5 |
| Plaster molding (3) | 46.0 | 45.9 | 44.8 | 46.2 | 45.1 | 45.5 | | | 273.5 |

- Fill ANOVA table

| | df | Sum of Squares (SS) | Mean Square(MS) | f |
|---|---|---|---|---|
| Treatment | **2** | 7.93 | **3.97** | **12.56** |
| Error | **19** | **13.93 – 7.93 = 6.00** | **0.32** | |
| Total | **21** | **13.93** | | |

Supap
fengspe

puter Engineering

48

## Example 1 (cont.)

| | df | Sum of Squares (SS) | Mean Square(MS) | f |
|---|---|---|---|---|
| Treatment | **2** | 7.93 | **3.97** | **12.56** |
| Error | **19** | **13.93 – 7.93 = 6.00** | **0.32** | |
| Total | **21** | **13.93** | | |

- Rejection region: given $\alpha = 0.001$
  - $F_{0.001,\ 2,\ 19} = 10.16$
- $H_0$ is rejected.
- Alloys from 3 processes are not the same.



49

## Outline

- Introduction
- Equal sample size
  - F-test
  - Multiple comparison
- Unequal sample size
  - F-test
  - Multiple comparison
- Multiple comparison discussion
- Random effects model

Supaporn Erjongmanee
fengspe@ku.ac.th

Department of Computer Engineering
Kasetsart University



## Multiple Comparisons : Unequal Sample Size

- When some means are not all equal, how to specify which mean is different from others
- Procedure
  1. Find Tukey's Honestly Significant Difference (HSD)

     $$HSD_\alpha = q_{\alpha,I,N-I}\sqrt{\frac{MSE}{2}\left(\frac{1}{J_i}+\frac{1}{J_j}\right)}$$

     - $q_{\alpha,\,I,\,N-I}$ = q-value from studentized range distribution with 2 degrees of freedom $I$, $N-I$
  2. Sort sample means in increasing order
  3. Underline pairs that differ less than $HSD_\alpha$
  4. Any pair without underline are considered as significantly different.

Supaporn Erjongmanee
fengspe@ku.ac.th

Department of Computer Engineering
Kasetsart University

## Example 1 : Mg-based Alloy (cont.)

$$HSD_{\alpha,ij} = q_{\alpha,I,N-I}\sqrt{\frac{MSE}{2}\left(\frac{1}{J_i}+\frac{1}{J_j}\right)}$$

- $\bar{x}_1$= 44.71, $\bar{x}_2$= 44.06, $\bar{x}_3$= 45.58

$$w_{12} = HSD_\alpha = q_{\alpha,I,N-I}\sqrt{\frac{MSE}{2}\left(\frac{1}{J_i}+\frac{1}{J_j}\right)} = q_{0.05,3,19}\sqrt{\frac{0.32}{2}\left(\frac{1}{8}+\frac{1}{8}\right)} = 3.59\sqrt{\frac{0.32}{2}\left(\frac{1}{8}+\frac{1}{8}\right)} = 0.718$$

$$w_{13} = w_{23} = HSD_\alpha = q_{\alpha,I,N-I}\sqrt{\frac{MSE}{2}\left(\frac{1}{J_i}+\frac{1}{J_j}\right)} = q_{0.05,3,19}\sqrt{\frac{0.32}{2}\left(\frac{1}{8}+\frac{1}{6}\right)} = 3.59\sqrt{\frac{0.32}{2}\left(\frac{1}{8}+\frac{1}{6}\right)} = 0.775$$

- Sort sample means: $\underset{\bar{x}_2}{44.06}$, $\underset{\bar{x}_1}{44.71}$, $\underset{\bar{x}_3}{45.58}$

  0.65 0.87

- 2 groups of means



52

## Outline

- Introduction
- Equal sample size
  - F-test
  - Multiple comparison
- Unequal sample size
  - F-test
  - Multiple comparison
- Multiple comparison discussion
- Random effects model



53

# More on Multiple Comparison

- Is it possible that $H_{0A}$ (or $H_{0B}$) is not rejected but result to multiple group of means?
- Is it also possible that $H_{0A}$ (or $H_{0B}$) is rejected but have one group of means?

Measured by MSE

ANOVA tests on ALL means whether they are identical

Multiple comparison tests on PAIRWISE means

Not measured by MSE

ANOVA detects variability among all means

ANOVA test is more sensitive than Multiple comparison





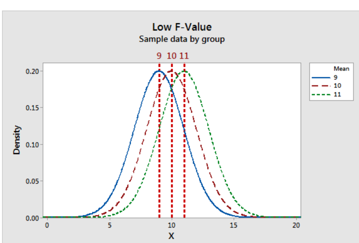
# More on Multiple Comparison (cont.)

ANOVA tests on ALL means whether they are identical

Multiple comparison tests on PAIRWISE means

ANOVA detects lower variability among all means

ANOVA test is more sensitive than Multiple comparison



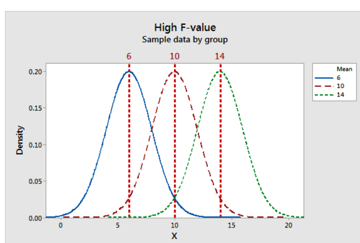




*Image source: http://blog.minitab.com/blog/adventures-in-statistics-2/understanding-analysis-of-variance-anova-and-the-f-test*

# Outline

- Introduction
- Equal sample size
  - F-test
  - Multiple comparison
- Unequal sample size
  - F-test
  - Multiple comparison
- Multiple comparison discussion
- Random effects model



# Random Effects Model

- Single factor can be considered as fixed-effects ANOVA model
- The single-factor fixed-effects model is defined as

$$\boxed{X_{ij} = \mu \quad + \alpha_i + \varepsilon_{ij}} \qquad \boxed{\sum \alpha_i = 0}$$

- $X_{ij}$ = random sample j of treatment i
- $\mu$ = overall mean of all treatment i's
- $\alpha_i$ = effect of treatment i
- $\varepsilon_{ij}$ = random error in sample j of treatment i
  - Assumed to be independent and normally distributed with mean = 0, variance = $\sigma^2$

## Random Effects Model (cont.)

$\boxed{X_{ij} = \mu \quad + \alpha_i + \varepsilon_{ij}}$ $\boxed{\sum \alpha_i = 0}$

- Explanations of the single-factor fixed-effects model:
  - Sample is corrupted by random errors
  - Error in one sample is independent from error of other samples
  - Expected response of treatment i

$$\boxed{E(X_{ij}) = \mu \quad + \ \alpha_i}$$

  - If $\alpha_i = 0$, then all treatment i's have the same response

$$\boxed{E(X_{ij}) = \mu}$$

That's what we try to prove in single-factor ANOVA





## Null Hypothesis: Single-Factor ANOVA

$\boxed{E(X_{ij}) = \mu \quad + \ \alpha_i}$

- Besides the prior null hypothesis in single-factor ANOVA
  - $H_0$: $\mu_1 = \mu_2 = \ldots = \mu_I$
- Sometimes, the following null hypothesis is also used instead:
  - $H_0$: $\alpha_1 = \alpha_2 = \ldots = \alpha_I = 0$

# References


1. J.L. Devore and K.N.Berk, Modern Mathematical Statistics with Applications, Springer, 2012.
2. S. Few, Now You See It: Simple Visualization Techniques for Quantitative Analysis, Analytics Press, 2009