Chapter 10 - Comparison of several means

STAT 251

Lecture 32

Analysis of Variance (ANOVA)
Examples

Dr. Lasantha Premarathna

Chapter 10 - Learning Outcomes

- Identify situations where one-way ANOVA is and is not appropriate.
- State the modeling assumptions underlying ANOVA.
- State the null and alternative hypothesis for the ANOVA test.
- Explain the partitioning of the total sum of squares into the within and between group components.
- Identify the degrees of freedom associated with each sum of squares.
- Interpret an ANOVA table.
- Perform the F test in ANOVA.
- Use the data to estimate the underlying within-group variance.
- Explain the output from a software package for an ANOVA study

Re-cap: SST = SSTr + SSE

In ANOVA, variations are measured by sums of squares (SS)

$$SST = SSTr + SSE$$

Overall variability = Between group SS + Within group SS

$$SST = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..})^2 = \sum_{i=1}^{k} \sum_{j=1}^{n_i} y_{ij}^2 - \frac{1}{n}\, y_{..}^2$$

$$SSTr = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (\bar{y}_{i.} - \bar{y}_{..})^2 = \sum_{i=1}^{k} \frac{1}{n_i}\, y_{i.}^2 - \frac{1}{n}\, y_{..}^2$$

$$SSE = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2 = \sum_{i=1}^{k} \sum_{j=1}^{n_i} y_{ij}^2 - \sum_{i=1}^{k} \frac{y_{i.}^2}{n_i}$$

also $SSE = \sum_{i=1}^{k} (n_i - 1) s_i^2$ where $s_i^2 = \frac{\sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2}{n_i - 1}$

Re-cap: Degrees of Freedom

.

$$df(SST) = n - 1$$

$$df(SSTr) = k - 1$$

$$df(SSE) = n - k$$

$$SST = SSTr + SSE$$

$$\Rightarrow \quad df(SST) = df(SSTr) + df(SSE)$$

Mean Squares

- Mean Squares Treatment $= MSTr = \dfrac{SSTr}{k-1}$
- Mean Squares Error $= MSE = \dfrac{SSE}{n-k}$
  - MSE is a measure of within-samples variability

<u>Note:</u>

- If we increase sample size, SSE will be increased.
- If we consider more groups, $SSTr$ will be increased.
- We want a test statistic that works in general to test hypotheses, no matter the size of your sample or the number of groups we have. Therefore, we consider Mean Squares as above to calculate the test statistic.

ANOVA Test Procedure

- Hypotheses

  $H_0 : \mu_1 = \mu_2 = ... = \mu_k$

  H_a: at least one of the means is different from the others

- Test statistic

$$F_{obs} = \frac{\frac{SSTr}{k-1}}{\frac{SSE}{n-k}} = \frac{MSTr}{MSE} \sim F_{\nu_1,\, \nu_2}$$

Under $H_0, F_{obs} = \dfrac{MSTr}{MSE}$ follows the F**-distribution** with degrees of freedom ν_1 (numerator df) and ν_2 (denominator df)

$$\nu_1 = df(SSTr) = k - 1$$

$$\nu_2 = df(SSE) = n - k$$

ANOVA Test Procedure

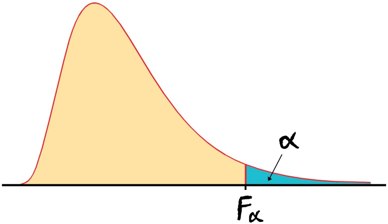
- Hypotheses

  $H_0 : \mu_1 = \mu_2 = ... = \mu_k$

  H_a: at least one of the means is different from the others

- Test statistic
  $$F_{obs} = \frac{MSTr}{MSE} \sim F_{\nu_1,\, \nu_2}$$

- Reject H_0 if $F_{obs} \geq F_\alpha$ or p-value $\leq \alpha$



- F**-distribution** is an asymmetric distribution.

The ANOVA Table

Source of Variation	df	Sum of Squares (SS)	Mean Square (MS)	F-ratio
Treatment (Between group)	$k-1$	$SSTr$	$MSTr = \dfrac{SSTr}{k-1}$	$\dfrac{MSTr}{MSE}$
Error (Within group)	$n-k$	SSE	$MSE = \dfrac{SSE}{n-k}$	
Total	$n-1$	SST		

The ANOVA Model

- The assumptions of one-way ANOVA can be described succinctly by mean of the "model equation"

- Each measurement will be represented as the sum of two terms; unknown constant μ_i and a random variable ϵ_{ij}

$$Y_{ij} = \mu_i + \epsilon_{ij}$$

$$i = 1, 2, \cdots, k \qquad \text{and} \qquad j = 1, 2, \cdots, n_i$$

Where ϵ_{ij} represents a random deviation from the population or true treatment mean μ_i

$\Rightarrow \quad \epsilon_{ij}$'s are iid rvs such that $\epsilon_{ij} \sim N(0, \sigma^2)$

ANOVA using R

- We can run our ANOVA in R using different functions. The most basic and common functions we can use are
  - aov()
  - lm()
- Because ANOVA is a type of linear model, we can use the lm() function.

Multiple Comparisons

- ANOVA can only tell us whether one or more groups are different from the others
  - Cannot tell you which groups are different
- Multiple Comparisons
  - Post-hoc method
    - Performs a set of pairwise comparisons to assess which groups differ
  - Involves performing several tests
    - We must adjust for the increased risk of making a Type I error
- There are many versions of multiple comparisons test (some are more likely to detect siginficant differences and some are less likely to detect significant differences)

Multiple Comparison tests

- We often want to compare pairs of treatment group means to see if they differ
- To do so, we used a t-test with the within group mean square as the pooled sample variance
- If we want to test several such pairs
  - We must adjust for performing several tests to keep the overall risk of Type I error from growing too large
- We adjust the significance level of each pairwise test to:
  - $\dfrac{\alpha}{h}$ where h is the number of pairwise tests.

    This is the Bonferroni Correction.
  - Ensure that the chance of making at least one type I error would be no more than α

Example 1

The article "Compression of Single-Wall Corrugated Shipping Containers Using Fixed and Floating Test Platens" (*J. Testing and Evaluation*, 1992: 318–320) describes an experiment in which several different types of boxes were compared with respect to compression strength (lb).

Example 1

Following table presents the results of a single factor ANOVA experiment involving $k = 4$ types of boxes (the sample means and standard deviations are in good agreement with values given in the article).

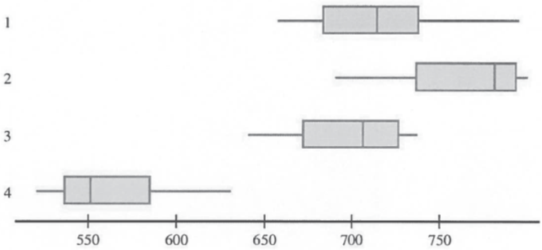
Type of Box	Compression Strength (lb)						Sample Mean	Sample SD
1	655.5	788.3	734.3	721.4	679.1	699.4	713.00	46.55
2	789.2	772.5	786.9	686.1	732.1	774.8	756.93	40.34
3	737.1	639.0	696.3	671.7	717.2	727.1	698.07	37.20
4	535.1	628.7	542.4	559.0	586.9	520.0	562.02	39.87
					Grand mean =		682.50	

Example 1

Assumptions :

1) Random samples
2) Normality assumption
3) Equal population variances

comparative boxplot for the four samples.



Example 1 - Solutions

Let μ_k denote the true average compression atrength for boxes of type $k (k = 1, 2, 3, 4)$

- Hypotheses

  $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$

  H_a: at least one of the means is different from the others

Example 1 - Solutions

Type of Box	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 =		682.50	

Form the data, we can find

$\bar{y}_{1\cdot} = 713.00$ $\quad s_1 = 46.55$

$\bar{y}_{2\cdot} = 756.93$ $\quad s_2 = 40.34$

$\bar{y}_{3\cdot} = 698.07$ $\quad s_3 = 37.20$

$\bar{y}_{2\cdot} = 562.02$ $\quad s_4 = 39.87$

- $\bar{y}_{\cdot\cdot} = 682.50$

$$n = n_1 + n_2 + n_3 + n_4$$
$$n = 6 + 6 + 6 + 6 + 6 = 24$$

Example 1 - Solutions

$$SSTr = \sum_{i=1}^{4}\sum_{j=1}^{n_i}(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2 = \sum_{i=1}^{4} n_i(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2$$

$$= 6(713.00 - 682.50)^2 + 6(756.93 - 682.50)^2 + 6(698.07 - 682.50)^2 + 6(562.02 - 682.50)^2$$

$$SSTr = 127367.58$$

$$SSE = \sum_{i=1}^{4}(n_i - 1)s_i^2 = 5(46.55^2 + 40.34^2 + 37.20^2 + 39.87^2)$$

$$SSE = 33854.5$$

$$\Rightarrow SST = SSTr + SSE$$

$$SST = 127367.58 + 33854.5 = 161222.1$$

Example 1 - Solutions

ANOVA Table

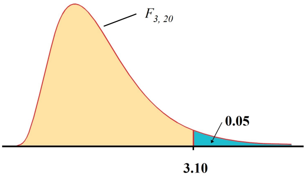
Source of Variation	df	Sum of Squares	Mean Squares	F-ratio
Treatment (Box type)	....	....	....	....
Error	....	....		
Total	....			

Example 1 - Solutions

ANOVA Table

Source of Variation	df	Sum of Squares	Mean Squares	F-ratio
Treatment (Box type)	k - 1 = 4-1 = 3	127367.58	(127367.58)/3 = 42455.9	42455.9/1692.7 = 25.08
Error	n - k = 24 -4 = 20	33854.5	(33854.5)/20 =1692.7	
Total	n - 1 = 23	161222.1		

Example: 1 - Solutions



Critical value approach

consider $\alpha = 0.05$

Observed test statistic is $F_{obs} = 25.08$

critical value is $F_{3,20,0.05} = 3.10 \leftarrow$ get this value using the F-table

Calculated test statistic is in the rejection region.

$$i.e.\ F_{obs} = 25.08 > 3.10 = F_{3,20,0,.05}$$

$$\Rightarrow \quad \text{Reject } H_0 \text{ at } \alpha = 0.05$$

p-value approach

$\Rightarrow$ p-value < 0.001 (according to the F-table, $F_{3,20,0,.001} = 8.10$)

Conclusion:

We have strong evidence to conclude that at least one mean compression strength is different among the four box types.

Example 2

From the article "on the Development of a New Approach for the Determination of Yield strength in Mf-based Alloy" we have following summaries on elastic modulus (GPa) obtained by a new ultrasonic method for specimens of a certain alloy produced using three different casting process. Is there sufficient evidence of a difference in the true mean elastic moduli of three casting process?

Casting process	Sample size	Sample total
Permanent molding	$n_1 = 8$	$x_{1\cdot} = 357.7$
Die casting	$n_2 = 8$	$x_{2\cdot} = 352.5$
Plaster molding	$n_3 = 6$	$x_{3\cdot} = 273.5$

$$\sum_i^3 \sum_j^{n_i} x_{ij} = 983.7$$

$$\sum_i^3 \sum_j^{n_i} x_{ij}^2 = 43{,}998.73$$

1) Identify the population parameter(s)
2) Set up the null and alternative hypotheses
3) What assumptions do you need to make in order to carry out your test?
4) Create ANOVA table to test the hypotheses defined in part 2)
5) What is your conclusion about the difference in the true mean elastic moduli at the 3 casting processes at a 1% significance level? Note the critical value is $F_{0.01,2,19} = 5.93$.

Example 2 - Solutions

1) Population Parameters

    True average elastic moduli for the permanent molding (μ_1)

    True average elastic moduli for the die casting process (μ_2)

    True average elastic moduli for the plaster molding (μ_3)

    common population variance (σ^2)

2) Hypotheses

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

    H_a: at least one of the means is different from the others

3) Assumptions

    different samples are independent of one another

    observations within any particular sample are independent

    treatment distributions are normal with the same variance, σ^2

Example 2 - Solution

4) $n_1 = 8, n_2 = 8, n_3 = 6, \quad n = n_1 + n_2 + n_3 = 22, \quad k = 3, \quad x_{..} = 983.7$

$$SST = \sum_{i=1}^{k=3} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{..})^2$$

$$= \sum_{i=1}^{3} \sum_{j=1}^{n_i} x_{ij}^2 - \frac{1}{n}\, x_{..}^2 = 43998.73 - \frac{1}{22}(983.7)^2 = 13.93$$

$$SSTr = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (\bar{x}_{i.} - \bar{x}_{..})^2 \;= \sum_{i=1}^{k} \frac{1}{n_i}\, x_{i.}^2 - \frac{1}{n}\, x_{..}^2$$

$$= \left[\frac{(357.7)^2}{8} + \frac{(352.5)^2}{8} + \frac{(273.5)^2}{6}\right] - \frac{1}{22}(983.7)^2 = 7.93$$

$$\Rightarrow SST = SSTr + SSE$$

$$\Rightarrow SSE = SST - SSTr \;= 13.93 - 7.93 \;= 6$$

Example 2 - Solutions

4) ANOVA Table

Source of Variation	df	Sum of Squares	Mean Squares	F-ratio
Treatment	$k - 1 = 3 - 1 = 2$	7.93	$7.93/2 = 3.965$	$3.965/0.3158 = 12.55$
Error	$n - k = 22 - 3 = 19$	6	$6/19 = 0.3158$	
Total	$n - 1 = 21$	13.93		

Example: 2 - Solutions

5)

consider $\alpha = 0.01$

critical value is $F_{2,19,0.01} = 5.93 \leftarrow$ this is given in the problem.
otherwise need to get this value using the F-table

$$F_{obs} = 12.55 > 5.93 = F_{2,19,0.01}$$

i.e. Calculated test statistic is in the rejection region.

$$\Rightarrow \quad \text{Reject } H_0 \text{ at } \alpha = 0.01$$

Conclusion:

We conclude that there is evidence that a true average elastic modulus is different for at least one process at 1% significance level (i.e. true average elastic modulus somewhat depends on which casting process is used).

Before the next class ...

Visit the course website at canvas.ubc.ca

- Review Lecture 32 and related sections in the text book
- Topic of next class: **ANOVA activity, Chapter 11 & 2 - Correlation & Simple Linear Regression**