

STAT 515 -- Chapter 10: Analysis of Variance

Designed Experiment – A study in which the researcher controls the levels of one or more variables to determine their effect on the variable of interest (called the **response variable** or **dependent variable**).

Response variable: Main variable of interest (continuous)

Factors: Other variables (typically discrete) which may have an effect on the response.

- Quantitative factors are numerical.
- Qualitative factors are categorical.

The **levels** are the different values (for each factor) used in the experiment.

Example 1:

Response variable: College GPA

Factors: Gender (levels: Male, Female)

of AP courses (levels: 0, 1, 2, 3, 4+)

The **treatments** of an experiment are the different factor level combinations.

Treatments for Example 1:

Experimental Units: the objects on which the factors and response are observed or measured.

Example 1?

Designed experiment: The analyst controls which treatments to use and assigns experimental units to each treatment.

Observational study: The analyst simply observes treatments and responses for a sample of units.

Example 2: Plant growth study:

Experimental Units: A sample of plants

Response: Growth over one month

Factors: Fertilizer Brand (levels: A, B, C)

Environment

(levels: Natural Sunlight, Artificial Lamp)

There are how many treatments?

(Could also have a quantitative factor...)

If 5 plants are assigned to each treatment (5 replicates per treatment), there are how many observations in all?

Completely Randomized Design (CRD)

A Completely Randomized Design is a design in which independent samples of experimental units are selected for each treatment.

Suppose there are p treatments (usually $p \geq 3$).

We want to test for any differences in mean response among the treatments.

Hypothesis Test:

$H_0: \mu_1 = \mu_2 = \dots = \mu_p$

H_a : At least two of the treatment population means differ.

Visually, we could compare all the sample means for the different treatments. (Dot plots, p. 522)

If there are more than two treatments, we cannot just subtract sample mean values.

Instead, we analyze the variance in the data:

Q: Is the variance within each group small compared to the variance between groups (specifically, between group means)?

Top figure?

Bottom figure?

How do we measure the variance within each group and the variance between groups?

The Sum of Squares for Treatments (SST) measures variation between group means.

$$\text{SST} = \sum_{i=1}^p n_i (\bar{X}_i - \bar{X})^2$$

n_i = number of observations in group i

\bar{X}_i = sample mean response for group i

\bar{X} = overall sample mean response

SST measures how much each group sample mean varies from the overall sample mean.

The Sum of Squares for Error (SSE) measures variation within groups.

$$\text{SSE} = \sum_{i=1}^p (n_i - 1) s_i^2$$

s_i^2 = sample variance for group i

SSE is a sum of the variances of each group, weighted by the sample sizes by each group.

To make these measures comparable, we divide by their degrees of freedom and obtain:

$$\text{Mean Square for Treatments (MST)} = \frac{\text{SST}}{p - 1}$$

$$\text{Mean Square for Error (MSE)} = \frac{\text{SSE}}{n - p}$$

The ratio $\frac{MST}{MSE}$ is called the ANOVA F-statistic.

If $F = \frac{MST}{MSE}$ is much bigger than 1, then the variation between groups is much bigger than the variation within groups, and we would reject $H_0: \mu_1 = \mu_2 = \dots = \mu_p$ in favor of H_a .

Example (Table 10.1)

Response: Distance a golf ball travels

4 treatments: Four different brands of ball

$$\bar{X}_1 = 250.8, \bar{X}_2 = 261.1, \bar{X}_3 = 270.0, \bar{X}_4 = 249.3.$$

$$\Rightarrow \bar{X} = 257.8.$$

$$n_1 = 10, n_2 = 10, n_3 = 10, n_4 = 10. \Rightarrow n = 40.$$

Sample variances for each group:

$$s_1^2 = 22.42, s_2^2 = 14.95, s_3^2 = 20.26, s_4^2 = 27.07.$$

SST =

SSE =

MST =

MSE =

F =

This information is summarized in an ANOVA table:

<u>Source</u>	<u>df</u>	<u>SS</u>	<u>MS</u>	<u>F</u>
Treatments	$p - 1$	SST	MST	MST/MSE
<u>Error</u>	<u>$n - p$</u>	<u>SSE</u>	MSE	
Total	$n - 1$	SS(Total)		

**Note that $df(\text{Total}) = df(\text{Trt}) + df(\text{Error})$
and that $SS(\text{Total}) = SST + SSE$.**

For our example, the ANOVA table is:

Note: $\frac{SS(Total)}{df(Total)}$ is simply the sample variance for the entire data set, $\frac{\sum (X - \bar{X})^2}{n - 1}$.

In example, we can see $F = 43.99$ is “clearly” bigger than 1 ... but how much bigger than 1 must it be for us to reject H_0 ?

ANOVA F-test:

If H_0 is true and all the population means are indeed equal, then this F-statistic has an F-distribution with numerator d.f. $p - 1$ and denominator d.f. $n - p$.

We would reject H_0 if our F is unusually large.

Picture:

H_0 : $\mu_1 = \mu_2 = \dots = \mu_p$

H_a : At least two of the treatment population means differ.

Rejection Region: $F > F_\alpha$, where F_α based on $(p - 1, n - p)$ d.f.

Assumptions:

- **We have random samples from the p populations.**
- **All p populations are normal.**
- **All p population variances are equal.**

Example: Perform ANOVA F-test using $\alpha = .10$.

Which treatment means differ? Section 10.3 (Multiple Comparisons of Means) covers this issue.