Introduction to ANOVA

Dr Tom Ilvento
Department of Food and Resource Economics
University of Delaware

Overview

- Our next set of lectures provides an introduction to **ANOVA** and **Experimental Design**
- This is a direct extension of the difference of means test we focused on earlier - ANOVA will do difference of means tests
- ANOVA is heavily used in Designed Experiements
- ANOVA will now allow us to compare multiple groups, not just two
- In order to do this, we have to be able to make some assumptions - equal variances and such
- This lecture will get us started with some terms and what is a designed experiment

ANOVA

- ANOVA provides a strategy to compare **two or more** population means associated with various treatments
- It is used when we have
  - A **Dependent Variable** (called a **Response Variable**) which is measured continuously
  - One or more **Factors** - levels of one or more categorical variables, thought of as **Independent Variables**
    - independent variables are thought to influence the dependent variable
    - E.g. treatment versus control; group membership; different levels of a treatment

ANOVA

- **ANOVA** is used heavily in experimental designs in the biological sciences
  - Treatment versus control groups
  - Levels of treatment of a drug
  - Levels of applications of fertilizers or pesticides
- It is possible to have both continuous and categorical independent variables
- Its origins were in agricultural studies, but it applications can now be found in biological and health related research; business; and economics

Designed Experiment versus an Observational Study

- **Designed Experiment**
  - The researcher manipulates the treatments and randomly assigns subjects to the treatments
  - The specification of the treatments and the way experimental units are assigned to treatments is under the control of the researcher
- **Observational Study**
  - The researcher observes the treatments and the response on a sample of experimental units
  - In essence we sample from a population where the treatments are already present

Elements of a Designed Experiment

- **Response Variable:** the variable of interest to be measured in the experiment. Also known as the dependent variable.
- **Factors:** variables which are thought to influence the response variable
  - Quantitative
  - Qualitative
- **Factor Levels:** the levels of the factor that are experimentally manipulated
- **In a single factor experiment, the factor levels can also be called the treatments**

Elements of a Designed Experiment

- **Treatments:** when two or more factors are utilized, the treatments are the combinations of factor levels used in the experiment.
- Factor 1: fertilizer (low; medium; high)
- Factor 2: water (low; high)
- The Treatments are:
  - Treatment 1: low fertilizer, low water
  - Treatment 2: low fertilizer, high water
  - Treatment 3: medium fertilizer, low water
  - And so forth……
- A special treatment that is used as a benchmark to compare the other treatments is called the **control treatment**.

Elements of a Designed Experiment

- **Experimental Unit:** the physical entity to which each treatment is randomly assigned
- **Measurement Unit:** the physical entity from which a measurement is taken.
- **Replications:** it is best to repeat the treatment on more than one experimental unit to get a better measurement of an effect.
  - If there is only one unit per treatment it is a single replication
  - We generally prefer to randomly assign the treatment to several experimental units.

Completely Randomized Design

- The treatments are **randomly assigned** to the experimental units
- **Or independent random samples** of experimental units are selected from target populations for each treatment
- Many texts refer to both designed and observational studies as being **randomized designs**
- **Most think of the best use of ANOVA for designed experiments**

An example of a designed experiment: shrimp weight gain

- A researcher is studying the conditions under which commercially raised shrimp reach maximum weight gain.
- Three water temperatures (25°, 30°, 35°) and four water salinity levels (10%, 20%, 30%, 40%) are examined
- While there are many other factors which influence shrimp growth, such as density of shrimp in a container, variety, and type of feeding - these two are the focus of this research.
- A specific variety and size of shrimp are selected for the study.
- The a fixed density of shrimp (40 per container) are randomly assigned to the treatments in 24 experimental containers, one shrimp to a container
- After 6 weeks the individual shrimp are weighed

Taken from Statistical Methods and Data Analysis, 5th Edition, Ott and Longnecker, 2001

An example of a designed experiment: shrimp weight gain

- The treatments are (temperature, salinity):
  - (25, 10%) (25, 20%) (25, 30%) (25, 40%)
  - (30, 10%) (30, 20%) (30, 30%) (30, 40%)
  - (35, 10%) (35, 20%) (35, 30%) (35, 40%)
- Replications: 2 for each treatment

An example of a designed experiment: shrimp weight gain

- Here is the experimental design
  - **Response Variable:** Weight of the shrimp in the container
  - **Factors:** 2 factors (temperature and salinity)
  - **Treatments:** 12
  - **Replication:** 2 per treatment
  - **Experimental Unit:** the container
  - **Measurement Unit:** the container
- Total sample size: 24
- This is an efficient design undertaken to estimate the effects of salinity and temperature on shrimp weight gain

What is ANOVA? Hint!!!

It is all about the variance!

It is all about the variance!

It is all about the variance!

It is all about the variance!

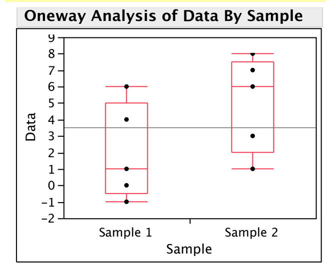
It is... ...all about the variance!

It is all about th
eVARIANCE!

Let's look at a simple problem

- I have a small data set of 5 observations for two variables.
- The difference between sample means is relatively small when compared to the variability within the sample observations
- We would say, **"The effect is small relative to the variability within each sample."**

Obs	Sample 1	Sample 2
1	6	8
2	-1	1
3	0	3
4	4	7
5	1	6
Sum	10.0	25.0
Mean	2.0	5.0
Var	8.5	8.5
Std Dev	2.9	2.9



Look at the result of the Difference of Means test from Excel

- The means are different
- The variances are similar so we pool the variance
- But we have a hard time rejecting a Null Hypothesis that the two means are equal to each other
- The t* = 1.627
- The p-value for a two-tailed test is .142
- We cannot reject Ho: $\mu_2 - \mu_1 = 0$

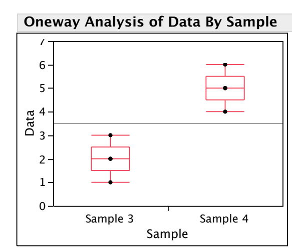
t-Test: Two-Sample Assuming Equal Variances

	Sample 2	*Sample 1*
Mean	5.00	2.00
Variance	8.50	8.50
Observations	5	5
Pooled Variance	8.5	
Hypothesized Mean Diff	0	
df	8	
t Stat	1.627	
P(T<=t) one-tail	0.071	
t Critical one-tail	1.860	
P(T<=t) two-tail	0.142	
t Critical two-tail	2.306	

Compare those results to this data

- I have the same mean differences
- But the spread of the data are much different - there is no overlapping of values
- The difference between sample means is relatively large when compared to the variability within the sample observations

Obs	Sample 3	Sample 4
1	2	5
2	3	5
3	2	5
4	2	4
5	1	6
Sum	10.0	25.0
Mean	2.0	5.0
Var	0.5	0.5
Std Dev	0.7	0.7



And when I do the difference of Means test in Excel

- Now it is easy to reject a Null Hypothesis that the two means are equal to each other
- The t* = 6.708
- The p-value for a two-tailed test is .000 or p < .001
- We can reject Ho: $\mu_2 - \mu_1 = 0$
- **When the effect is large relative to the variances within each group, I will be better able to reject the Null Hypothesis**

t-Test: Two-Sample Assuming Equal Variances

	Sample 4	*Sample 3*
Mean	5.00	2.00
Variance	0.50	0.50
Observations	5	5
Pooled Variance	0.5	
Hypothesized Mean Difference	0	
df	8	
t Stat	6.708	
P(T<=t) one-tail	0.000	
t Critical one-tail	1.860	
P(T<=t) two-tail	0.000	
t Critical two-tail	2.306	

Example of Contrasts of Two Samples

- One way to determine whether a difference exists between the population means is to examine the difference between the sample means and compare it to a measure of variability within the samples.
- This is what we do in a Difference of Means Test - the variability is accounted for in the standard error
  - The difference of the two means is only part of the story.
  - The other part is the variability and separation of the two samples
- ANOVA uses this strategy to compare and test the difference between two or more means
  - In essence it compares the variability across group means
  - To the variability within each group

ANOVA Strategy

- In ANOVA, we will decompose the variance of our dependent variable
  - Part due to the treatments or independent variables – this part is thought to be “explained” by our model.
  - Part that is “unexplained” or random “error”
- I will adjust these variances from different sources by dividing by degrees of freedom to get an average deviation
- We will decompose the Total Sum of Squares =
  - Sum of Squares for Treatment
  - Sum of Squares for Error

$$SS(Total) = \sum_{i=1}^{n} (y_i - \overline{Y})^2$$

ANOVA Model

- This is one way to look at the ANOVA model with i observations and j factor levels
- Each observation is a function of:
  - A Grand Mean
  - The deviations of each Factor Level Mean from the Grand Mean
  - Some random error

$$Y_{ij} = \mu + \tau_{\bullet j} + \varepsilon_{ij}$$

Another viewpoint

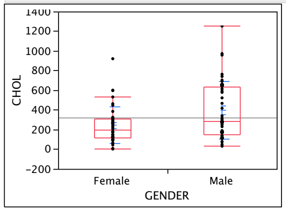
- The difference of an individual value from the Grand Mean is a function of: $Y_{ij} - \overline{Y} =$
- **The difference of the Factor Level Mean from the Grand Mean** — **Variability between Factor Levels**
- **The difference between the value and its Factor Level Mean** — **Variability within Factor Levels**

We want the variability between Factor Levels to be large relative to the variability within Factor Levels

Another Data Example

- This is some data comparing the cholesterol level of 40 males to 40 females
- We can see that the level for males is higher
- Since this a sample, I need a statistical tests to see if there is a difference.
- We can do a difference of means test from JMP
  - $t^* = 2.8169$
  - $p < .01$ for a 2-tailed test
  - We can conclude there is a difference in cholesterol levels between males and females

Oneway Analysis of CHOL By GENDER



Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
Female	40	240.875	185.982	29.406	181.39	300.36
Male	40	395.225	292.412	46.234	301.71	488.74

I can do the same analysis using ANOVA

- ANOVA - randomized design for an observational study
- **Response Variable:** cholesterol level
- **Factor:** gender
  - Females
  - Males
- **Experimental Units:** people
- We conduct an F-test with ANOVA

Excel: ANOVA, Single Factor

Mean, variance, and sample size for each group

Anova: Single Factor

SUMMARY

Groups	Count	Sum	Average	Variance
Females	40	9635.000	240.875	34589.446
Males	40	15809.000	395.225	85504.640

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	476478.450	1	476478.450	7.935	0.006	3.963
Within Groups	4683669.350	78	60047.043			
Total	5160147.8	79				

ANOVA Table with Sums of Squares breakdown

ANOVA

- ANOVA Table with Sums of Squares breakdown
  - SS Treatment **1,393.425** Between Groups
  - SS Error **41,846.879** Within Groups
  - SS Total **43,240.304**
- F-test

Anova: Single Factor

SUMMARY

Groups	Count	Sum	Average	Variance
Females	40	9635.000	240.875	34589.446
Males	40	15809.000	395.225	85504.640

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	476478.450	1	476478.450	7.935	0.006	3.963
Within Groups	4683669.350	78	60047.043			
Total	5160147.8	79				

Summary

- This is just an introduction to ANOVA. We will spend several lectures on the approach and some more complex models
- ANOVA comes with its own terms and jargon!
- There is a connection between the difference of means test and ANOVA - ANOVA will generate the same conclusion.
- But ANOVA will allow us to extend this to many means, not just two
- The next steps are looking at the F-distribution and the logic of the hypothesis test in ANOVA

Extra Problem

- An experiment is designed to examine protective coatings on frying pans.
- Four different coatings are being examined.
- Five frying pans are randomly assigned to each of the four coatings.
- A measure of abrasive resistance of the coatings is taken at three locations on each of the 20 pans.
- **Identify the following for this study:**
  - **Response variable**
  - **Treatments**
  - **Replications**
  - **Experimental Unit**
  - **Measurement Unit**
  - **Total sample size**