ANalysis Of VAriance III

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

Overview

- Next we will discuss two variations of the ANOVA model introducing a third variable
  - **Two-way ANOVA**
    - Now we have two factors that influence the response variable
    - Plus an interaction term
  - **Randomized Block Design**
    - We take into account another source of variability
    - This source is referred to as a block
    - We want to account for the blocks in the analysis in order to better understand the factor

Two-Way ANOVA

- When we deal with more than one Factor it is called a **Factorial Design**
- A **Complete Factorial Design** is an experiment that includes **all possible treatments**.
- The treatments are the combinations of the factor levels
- We also want replications to better estimate
- The data should be **balanced** – equal number of observations for each treatment
- Excel will do a two-way ANOVA if the data are balanced
- JMP will run a two-way ANOVA for unbalanced designs

Interaction within a Two-Way ANOVA

- A new concept will be an **interaction term**
- We always test for the interaction first before looking at the main effects of the factors
- An interaction means that the effect of one factor depends upon the level of the other factor – and vice versa
- In other words, we can't say something about the effect of one factor without knowing something about the other factor
- We didn't make this explicit, but the alternative to the Null Hypothesis in the Chi-square test for contingency tables is really saying there is an interaction between the two variables
- The ANOVA model will build in a test for the interaction terms

Terms in the Two-Way ANOVA

- In the Two-way ANOVA with r replications we will have Sum and Mean Squares for each term in the model
- Which means we will have separate F-tests for Factor A, Factor B, and the Interaction Term
- r stands for replicants

Source	d.f.	Excel Term	JMP Term
Factor A with a levels	a-1 d.f.	Sample	Variable name
Factor B with b levels	b-1 d.f.	Column	Variable name
AB Interaction	(a-1)(b-1) d.f.	Interaction	NameA*NameB
Error	a*b(r-1)	Within	Error

An Example: Gaskets

- A company stamps gaskets out of three materials - rubber, plastic and cork – on two machines.
- They want to compare the mean number of gaskets produced per hour in order to see if:
  - One machine is better than they other
  - If that difference in machines is the same for all three materials
- They conduct an experiment where each machine is operated for three one-hour time periods for each material.
- The time periods for each machine/material combination is assigned randomly.
- The response variable is the number of gaskets per hour (1,000)

Design: 3 x 2 Factorial Design

- **Response:**
  - the number of gaskets per hour (1,000)
- **Factor A**
  - Machine (2) - M1, M2
- **Factor B**
  - Type of gasket (3) – cork, rubber, plastic
- **Treatments**
  - M1 & Cork; M1 & Rubber; M1 & Plastic;
  - M2 & Cork; M2 & Rubber; M2 & Plastic;
- **Replications**
  - 3 per type per machine

18 Total Observations

Terms in the Gasket Two-Way ANOVA

- In this two-way ANOVA with 3 replications we will have Sum and Mean Squares for
- We will have separate F-tests for Machine, Material, and the Interaction Term

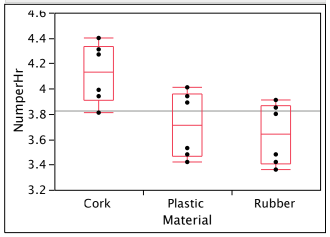
Source	d.f.	Excel Term	JMP Term
Machine with 2 levels	2-1 = 1 d.f.	Sample	Machine
Material with 3 levels	3-1 = 2 d.f.	Column	Material
Machine/Material Interaction	(2-1)(3-1) = 2 d.f.	Interaction	Machine*Material
Error	2*3*(3-1) = 12 d.f.	Within	Error

The data in Excel (ANOVA Gasket.xls) and JMP (Gasket.jmp)

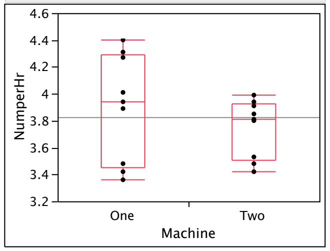
	Cork	Rubber	Plastic
Machine 1	4.31	3.36	4.01
	4.27	3.42	3.94
	4.40	3.48	3.89
Machine 2	3.94	3.91	3.48
	3.81	3.80	3.53
	3.99	3.85	3.42

Gasket Descriptive Statistics

Boxplot of NumperHr By Material



BoxPlot of NumperHr By Machine



	Cork	*Rubber*	*Plastic*	*Machine 1*	*Machine 2*	*Total*
Mean	**4.120**	**3.637**	**3.712**	**3.898**	**3.748**	**3.823**
Standard Error	0.097	0.099	0.107	0.133	0.071	0.075
Median	4.130	3.640	3.710	3.940	3.810	3.870
Mode	#N/A	#N/A	#N/A	#N/A	#N/A	3.420
Standard Deviation	0.238	0.243	0.263	0.398	0.214	0.319
Sample Variance	**0.056**	**0.059**	**0.069**	**0.158**	**0.046**	**0.102**
Kurtosis	-2.185	-2.824	-2.883	-1.603	-1.420	-0.846
Skewness	-0.126	-0.015	0.015	-0.224	-0.620	0.157
Range	0.590	0.550	0.590	1.040	0.570	1.040
Minimum	3.810	3.360	3.420	3.360	3.420	3.360
Maximum	4.400	3.910	4.010	4.400	3.990	4.400
Sum	24.720	21.820	22.270	35.080	33.730	68.810
Count	6	6	6	9	9	18

Let's Look at a 1-Way ANOVA with Material

- F* is significant at the .01 level (p = .0088)
- R^2 = .46826: 46.8% of the variability in units per hour is due to the material type
- If we looked at the differences in means using LSD we would find:
  - LSD = .30510, α = .05
  - Cork is significantly different from Plastic and Rubber
  - Plastic and Rubber are not different from each other

Oneway Anova

Summary of Fit

Rsquare	0.46826
Adj Rsquare	0.397361
Root Mean Square Error	0.247927
Mean of Response	3.822778
Observations (or Sum Wgts)	18

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Material	2	0.8119444	0.405972	6.6046	0.0088*
Error	15	0.9220167	0.061468		
C. Total	17	1.7339611			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
Cork	6	4.12000	0.10122	3.9043	4.3357
Plastic	6	3.71167	0.10122	3.4959	3.9274
Rubber	6	3.63667	0.10122	3.4209	3.8524

Std Error uses a pooled estimate of error variance

Means Comparisons

Comparisons for each pair using Student's t

t	Alpha
2.13145	0.05

Abs(Dif)-LSD

	Cork	Plastic	Rubber
Cork	-0.30510	0.10324	0.17824
Plastic	0.10324	-0.30510	-0.23010
Rubber	0.17824	-0.23010	-0.30510

Positive values show pairs of means that are significantly different.

Level			Mean
Cork	A		4.1200000
Plastic		B	3.7116667
Rubber		B	3.6366667

Levels not connected by same letter are significantly different.

2-Way ANOVA

- Now I will take into account Machine as another factor
- **If I use Excel**
  - Tools
  - Data Analysis
  - ANOVA: Two-Factor With Replication
- I grab the entire data
- Include the row labels
- Note there are 3 rows per sample
- Set alpha and where the output should go
- **If I use JMP**
  - Fit Model
  - NumperHr is Response
  - Factors are Material and Machine
  - I also need to cross Materials and Machine by grabbing both and clicking **Cross**

Excel Output

- Excel gives the means and variances by each Machine
- Notice the order (highest to lowest) of the means by material:
  - Machine 1: Cork, Plastic, Rubber
  - Machine 2: Cork Rubber, Plastic
- Note the Excel ANOVA terms
  - Sample, Columns, Interaction and within
  - Notice the SS for Material (Columns) is the same as the 1-Way ANOVA
  - $R^2 = 1 -.0527/1.7340 = .9696$

Anova: Two-Factor With Replication

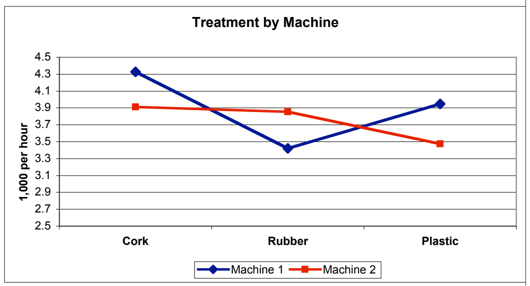
SUMMARY	Cork	Rubber	Plastic	Total
Machine 1				
Count	3	3	3	9
Sum	12.98	10.26	11.84	35.08
Average	4.3267	3.4200	3.9467	3.8978
Variance	0.0044	0.0036	0.0036	0.1584
Machine 2				
Count	3	3	3	9
Sum	11.74	11.56	10.43	33.73
Average	3.9133	3.8533	3.4767	3.7478
Variance	0.0086	0.0030	0.0030	0.0457
Total				
Count	6	6	6	
Sum	24.72	21.82	22.27	
Average	4.1200	3.6367	3.7117	
Variance	0.0565	0.0590	0.0689	

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Sample	0.1013	1	0.1013	23.0405	0.0004	4.7472
Columns	0.8119	2	0.4060	92.3831	0.0000	3.8853
Interaction	0.7680	2	0.3840	87.3869	0.0000	3.8853
Within	0.0527	12	0.0044			
Total	1.7340	17				

What's an Interaction Look Like?

- Instead of seeing the same pattern for each Material by Machine
- We see a different pattern
- The effect of the Material on output depends upon which Machine we are using
- **In a 2-Way NOVA**
  - You should **first test to see if there is an interaction**
  - If YES, then that is the story
  - IF NO, then you can examine the individual factors



Partial JMP Output

- JMP gives a more complete analysis
- **Summary of Fit**
- Analysis of Variance
  - Model and Error SS
  - Overall F-test
- **Effects Tests**
  - The individual SS for each element in the model
- We can see **the effect of Material and Machine is significant**
- But most important, **the Interaction between Machine and Material is very significant**
- In fact, about 46% of the Model SS (1.6812) is due to the Interaction term in the model

Summary of Fit

RSquare	0.969588
RSquare Adj	0.956916
Root Mean Square Error	0.066291
Mean of Response	3.822778
Observations (or Sum Wgts)	18

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	1.6812278	0.336246	76.5161
Error	12	0.0527333	0.004394	Prob > F
C. Total	17	1.7339611		<.0001*

Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Material	2	2	0.81194444	92.3831	<.0001*
Machine	1	1	0.10125000	23.0405	0.0004*
Material*Machine	2	2	0.76803333	87.3869	<.0001*

Once we establish a significant Interaction Term, it is difficult to talk about the other Factors in the model.

Hypothesis Test Strategy for a Two-way ANOVA

1. **Look at the interaction term first**
   - If it is significant, it means that you cannot assess the effects of the Machines without looking at which gasket is being produced
   - You should not look at the individual factors
   - `Do not perform the individual F-tests for the factors!
2. **If the Interaction term is not significant, then you can focus on the individual factor tests and make separate conclusions for each factor**
   - Including testing differences of means
   - Decomposing R^2 to see which Factor is more important

ANOVA Hypothesis Test for Interaction Term

- **Ho:** **Ho: No Interaction**
- **Ha:** **Ha: Interaction Present**
- **Assumptions** **Equal variances, normal distribution**
- **Test Statistic** **$F^* = 87.39$ $p < .001$**
- **Rejection Region** **$F_{.05,\,2,\,12} = 3.89$**
- **Conclusion:** **$F^* > F_{.05,\,2,\,12}$**

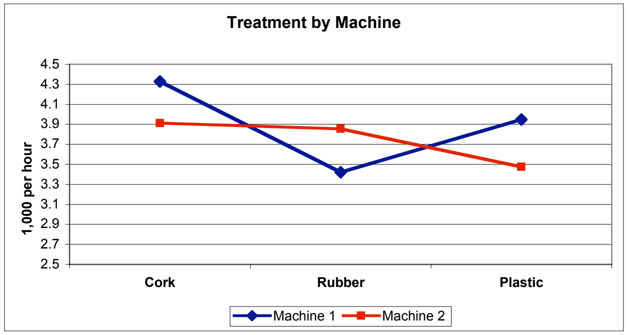
  or $p < .001$

  Reject Ho: No Interaction

There is a significant interaction: output by gasket type depends upon the machine

Interactions in ANOVA

- This is as clear a visual as you will see.
- Sometimes it is more subtle - different slopes or the effect is stronger/weaker based on levels of the third variable



Randomized Block Designs in ANOVA

- In a **Randomized Block Design**, we use matched sets of experimental units to each treatment
- The matched sets of experimental units are called BLOCKS
  - The blocks should consist of experimental units that are as similar as possible.
  - May be the same experimental unit
- The strategy is to reduce the MSE by accounting for another source of variability in the data (other than the factor)
- Two Step Process in the Design
  1. Matched sets of experimental units, called Blocks, are formed, each block consisting of k experimental units (where k is the number of treatments).
  2. One experimental unit from each block is randomly assigned to each treatment, resulting in n=b*k responses.

Randomized Block Design

- One book gives a good example relating to golfers hitting four different golf club brands
  - They prefer testing under “real” conditions with live golfers
  - Rather than 40 different golfers randomly assigned to four different brands
  - A Randomized Block Design would have 10 golfers each hitting a ball using each brand of club
  - This allows us to account for the effect of the golfers (blocks)
- The blocks are not the main focus of the research, but they can help us better see the effect of a Factor by removing some of the variability in the data through the design itself.
- If done correctly, the SS for the main treatment is unaffected by accounting for the Blocks, but the SSE is reduced by blocking.

Randomized Block Design

- Now we add a new source of variation - due to the Blocks
- The Blocks account for some of the variation in the response variable
- **Sum of Squares for Blocks (SSB)**
- And the SSE is a residual, derived as:
- SSE = SS_{Total} - SST - SSB
  - This shows that blocking, when successful, will reduce the SSE
  - This makes the F-test for the Factor larger, since MSE is the denominator for the test

$$SSB = \sum_{i=1}^{b} k(y_{B_i} - \overline{Y})^2$$

Terms in the Randomized Block Design

- In the Randomized Block Design we will have Sum and Mean Squares for
- n= b*k
- We will have separate F-tests for Treatments and the Blocks

Source	d.f.	Excel Term	JMP Term
SST	k-1 d.f.	Columns	Var Name
SSB	b-1 = 2 d.f.	Row	Var Name
SSE	n-b-k+1 or, (b-1)(k-1)	Error	Error

Randomized Block Example

- A bank out-sources home appraisals to three companies
- They want to conduct a formal test to see if the companies tend to differ on average – some being higher or lower
- One approach would be to randomly assign different homes to different companies
- However, even a random approach could lead to one company receiving larger or higher quality homes relative to the others
- The Blocks in this case would need to
  - match similar homes by size, quality, and location
  - or, assign the same house to all three companies

The data

- Three companies (Factor Levels)
- Five properties (Blocks)
- Response variable is the appraised value in $1,000s

Appraisal Data

	Appriasal Companies		
Property	Allen	Heist	AL
1	78.0	82.0	79.0
2	102.0	102.0	99.0
3	68.0	74.0	70.0
4	83.0	88.0	86.0
5	95.0	99.0	92.0

Single Factor Approach: Treatment is Appraisal Company

- F* = .162 and p = .852
- We will not be able to reject the Null Hypothesis
- Pay attention to the SST

Anova: Single Factor

SUMMARY

Groups	Count	Sum	Average	Variance
Allen	5	426	85.2	182.7
Heist	5	445	89.0	136.0
AL	5	426	85.2	126.7

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	48.133	2	24.067	0.162	0.852	3.885
Within Groups	1781.600	12	148.467			
Total	1829.733	14				

Block Models in Excel

- Use Tools, Data Analysis
- ANOVA: Two-Factor Without Replication
  - The Rows reflect the Blocks
  - The Columns the Factor Levels

Property	Allen	Heist	AL
1	78.0	82.0	79.0
2	102.0	102.0	99.0
3	68.0	74.0	70.0
4	83.0	88.0	86.0
5	95.0	99.0	92.0

Excel Output

- The top part gives means and variances
- The second part gives the ANOVA
- Notice the Rows SS reflect the Blocks - SS = 1759.07
- While SS Columns (the Treatments) are unchanged, the **F* = 8.54 with a p-value = .01**
- R^2 = 1 - 22.53/1829.73 = .9877 - **Most of R^2 due to Blocks**

Anova: Two-Factor Without Replication

SUMMARY	Count	Sum	Average	Variance
1	3	239	79.67	4.33
2	3	303	101.00	3.00
3	3	212	70.67	9.33
4	3	257	85.67	6.33
5	3	286	95.33	12.33
Allen	5	426	85.20	182.70
Heist	5	445	89.00	136.00
Al	5	426	85.20	126.70

Notes on the Block Model

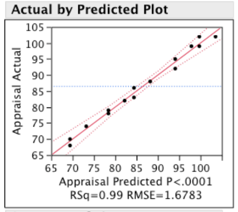
- The Sum of Squares due to the Factor (i.e., the companies) was the same in the one-way and block approaches
  - 48.13 for SST
  - 24.07 for the MST
- Accounting for the Blocks (properties) reduces the SSE
- Which allows us to make a better estimate for the effect of the factor
- And R-square increased dramatically (.9877)

JMP Output

- Use Fit Model
  - Response is the Appraisal
  - Factors are Block and Company
- As you might Expect, JMP gives more information
- And we can test to see which means are different using LSD
  - Hesit is significantly higher than the other two companies

Whole Model

Actual by Predicted Plot



Summary of Fit

RSquare	0.987685
RSquare Adj	0.978449
Root Mean Square Error	1.678293
Mean of Response	86.46667
Observations (or Sum Wgts)	15

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	6	1807.2000	301.200	106.9349
Error	8	22.5333	2.817	Prob > F
C. Total	14	1829.7333		<.0001*

Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Block	4	4	1759.0667	156.1302	<.0001*
Company	2	2	48.1333	8.5444	0.0103*

ANOVA Hypothesis Test for Main Factor Term

- **Ho:** Ho: $\mu_1 = \mu_2 = \mu_3$
- **Ha:** Ha: at least two means are different
- **Assumptions** Equal variances, normal distribution
- **Test Statistic** $F^* = 8.54$ $p = .01$
- **Rejection Region** $F_{.05, 2, 8} = 4.46$
- **Conclusion:** $F^* > F_{.05, 2, 8}$
  or $p = .01$
  Reject Ho: $\mu_1 = \mu_2 = \mu_3$

Now we can say that the appraisal price differs by company. And, based on the LSD test, the appraisal for Heist is higher than the other two.

ANOVA Hypothesis Test for Block Factor Term

- **Ho:** Ho: $\mu_{B1} = \mu_{B2} = \mu_{B3} = \mu_{B4} = \mu_{B5}$
- **Ha:** Ha: at least 2 block means are different
- **Assumptions** Equal variances, normal distribution
- **Test Statistic** $F^* = 156.13$ $p < .001$
- **Rejection Region** $F_{.05, 4, 8} = 3.84$
- **Conclusion:** $F^* > F_{.05, 4, 8}$
  or $p < .001$
  Reject Ho: $\mu_{B1} = \mu_{B2} = \mu_{B3} = \mu_{B4} = \mu_{B5}$

The Blocks make a difference in the model - it is useful to control for the Blocks.

Summary for ANOVA

- ANOVA provides us additional tools to examine the results of experimental designs.
- The analysis is tied to the design.
- ANOVA allows us to examine differences of many means in a single factor, or in multiple factors.
- In two-way ANOVA, with two factors, we also must examine the possibility of an interaction between the factors.
- We also introduced a Block design, which seeks to remove a source of variability via the design, so that we can better examine the effect of a Factor.
- This class gave a brief introduction to these approaches – there is so much more!