ANalysis Of VAriance II

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

Overview

- Let's continue our discussion of the ANOVA Model
- We will solve for the sum of squares for a basic model with two means
- See how software displays the results
- We will look at the Basic Test for ANOVA
  - F-test
  - Based on the F-distribution
- Give an example of more that two means

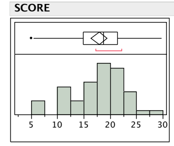
Let's get started with a problem

- A study was done to examine at how **Intrinsic** versus **Extrinsic Motivation influences Creativity**
- **Art for art's sake; money for God's sake!**
- Students who were experienced in creative writing were randomly assigned to two groups
  - Intrinsic motivation - motivation to be creative from within
  - Extrinsic motivation - motivation based on external reasons
- Motivation was supplied by a survey mechanism - they were asked to rank motivational questions in one of two areas
- Later they were asked to write a poem and it was judged by a panel of 12 judges on a 40 point scale
- **Does the type of motivation influence creativity?**

Let's look at the data

- **What do you see?**
- Somewhat normal distribution
- A few outliers, but not too bad
- The mean score was 17.855; the median is 18.70
- The CV is 29.335; a moderate amount of variability

SCORE



Quantiles

100.0%	maximum	29.700
99.5%		29.700
97.5%		29.100
90.0%		24.000
75.0%	quartile	21.300
50.0%	median	18.700
25.0%	quartile	14.800
10.0%		11.620
2.5%		5.080
0.5%		5.000
0.0%	minimum	5.000

Moments

Mean	17.855
Std Dev	5.238
Std Err Mean	0.764
upper 95% Mean	19.393
lower 95% Mean	16.317
N	47.000
Sum Wgt	47.000
Sum	839.200
Variance	27.435
Skewness	-0.548
Kurtosis	0.515
CV	29.335
N Missing	0.000

Stem and Leaf

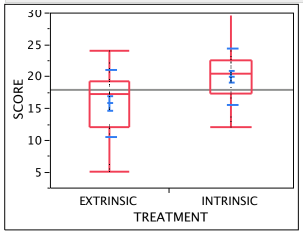
Stem	Leaf	Count
3	0	1
2		
2	7	1
2	444	3
2	222233	6
2	00011111	8
1	888999999	9
1	777777	6
1	455	3
1	222223	6
1	1	1
0		
0	6	1
0	55	2

0|5 represents 5

Let's compare the means

- We have everything we need to conduct a test of the two means
  - Ho: $\mu_i - \mu_e = 0$
  - Ha: $\mu_i - \mu_e \neq 0$

Oneway Analysis of SCORE By TREATMENT



Means and Std Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
EXTRINSIC	23	15.7391	5.25260	1.0952	13.468	18.011
INTRINSIC	24	19.8833	4.43951	0.9062	18.009	21.758

Difference of Means Hypothesis Test

- **Ho:** **$\mu1-\mu2=0$**
- **Ha:** **$\mu1-\mu2\neq0$**
- **Test Statistic** **2.93**
- **Rejection Region** **2.01**
- **p-value** **.01**
- **Conclusion:** **Reject**

t-Test: Two-Sample Assuming Equal Variances

	Intrinsic	Extrinsic
Mean	19.88	15.74
Variance	19.71	27.59
Observations	24	23
Pooled Variance	23.56	
Hypothesized Mean Difference	0	
df	45	
t Stat	2.93	
P(T<=t) one-tail	0.00	
t Critical one-tail	1.68	
P(T<=t) two-tail	0.01	
t Critical two-tail	2.01	

ANOVA Output for the same problem

- ANOVA output would typically give the means and variances of each group
- And then the ANOVA table, which is the breakdown of the **Sums of Squares**, and the **Mean Squares** (SS/df)

t-Test: Two-Sample Assuming Equal Variances

	Intrinsic	Extrinsic
Mean	19.88	15.74
Variance	19.71	27.59
Observations	24	23
Pooled Variance	23.56	
Hypothesized Mean Difference	0	
df	45	
t Stat	2.93	
P(T<=t) one-tail	0.00	
t Critical one-tail	1.68	
P(T<=t) two-tail	0.01	
t Critical two-tail	2.01	

Anova: Single Factor

SUMMARY

Groups	Count	Sum	Average	Variance
Extrinsic	23	362.000	15.739	27.590
Intrinsic	24	477.200	19.883	19.709

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	201.708	1	201.708	8.561	0.005	4.057
Within Groups	1060.288	45	23.562			
Total	1261.996	46				

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
TREATMENT	1	201.7080	201.708	8.5608	0.0054*
Error	45	1060.2882	23.562		
C. Total	46	1261.9962			

How does ANOVA work?

- **We decompose the Total Sum of Squares**
  - **Part due to Treatment or Between Factor Levels**
  - **Part due to Error or Within each Factor Level**
- We compute the variability of the treatment means from the Grand Mean (the mean from the whole sample, i.e., all groups)
  - **Sum of Squares for Treatments (SST)**
  - **Which measures Between Factor Level variation**
- And the variability within the treatment levels
  - **Sum of Squares for Error (SSE)**
  - **Which measures within each Factor Level variation**

How does ANOVA work?

- We adjust SST and SSE to reflect a mean sum of squares – divide by degrees of freedom
- Then we compare to see if the Mean Sum of Squares for Treatments (MST) is larger relative to the Mean Sum of Squares for Error (MSE)
- We do this by taking a ratio of the two sources of sums of squares

$$F^* = \frac{MST}{MSE}$$

How does ANOVA work?

- We ask, **"Is there more variability across the means for the factor levels (or treatments), than collectively within each of the Factor Levels?"**
- We look to see if there is more variability across factor levels (treatments),
  - with respect to a probability framework
  - using an F-distribution with specified degrees of freedom
- If yes, we will conclude that the factors influence the response variable – there are differences in the means between Factor Levels
- In the case of the single Factor, 2 levels, the result is identical to a difference of means test

Decomposing the Total Sums of Squares for a one-Factor Model

Let k = # treatments	k
The Grand Mean is	$\overline{Y}$
Each group mean is	$\overline{y}_k$
The sample size for each factor level	n_k
Degrees of Freedom	
Total Sum of Squares	n-1
Sum of Squares for Treatment	k-1
Sum of Squares for Error	n-k

The degrees of freedome should add up

- **SS_{Total} = SST + SSE**
- Degrees of freedom
- n-1 = (k-1) + (n-k)
- e.g. n = 100 k=3
  - 100-1 = (3-1) + (100-3)
  - 99 = 2 + 97

Formulas for the Sums of Squares

- **Total Sum of Squares (n-1 d.f.)**
- If I divided by the degrees of freedom, I would have the sample variance

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

- Sum of Squares for Treatments (k-1 d.f.)
- Dividing by the degrees of freedom gives the Mean Square for Treatment, MST = SST/(k-1)

$$SST = \sum_{i=1}^{k} n_i (\bar{y}_i - \bar{Y})^2$$

- Sum of Squares for Error (n-k d.f.)

$$SSE = \sum_{j=1}^{n_1} (y_{1j} - \bar{y}_1)^2 + \sum_{j=1}^{n_2} (y_{2j} - \bar{y}_1)^2 + \ldots \sum_{j=1}^{n_k} (y_{kj} - \bar{y}_k)^2$$

- Another way to calculate this is:

$$SSE = (n_1 - 1)s_1^{\,2} + (n_2 - 1)s_2^{\,2} + \ldots (n_k - 1)s_k^{\,2}$$

- Dividing by degrees of freedom gives the Mean Square Error, MSE = SSE/(n-k)

Let's Work through the computations for the Creativity Example

$$SST = \sum_{i=1}^{k} n_i (\bar{y}_i - \bar{Y})^2$$

SST = 24(19.883-17.855)² + 23(15.739-17.855)²

SST = 24(1.854)² + 23(-2.116)²

SST = 82.495 + 102.981 = 201.688

MST = 201.688/1 = 201.688

$$SSE = (n_1 - 1)s_1^{\,2} + (n_2 - 1)s_2^{\,2}$$

SSE = 23(19.709) + 22(27.590)

SSE = 453.307 + 606.980 = 1060.287

MSE = 1060.287/(47-2) = 23.562

	Intrinsic	Extrinsic
Sample size	n_1=24	n_2=23
Mean	19.883	15.739
Variance	19.709	27.590
Deviation from Grand Mean	1.854	-2.116
Grand Mean	17.855	
SS_{total}	1262.010	
Treatment Effect	19.883-15.739 = 4.144	

SS_{total} = SST + SSE = 201.688 + 1060.287 = 1261.975

JMP Output

- Our values are similar except for rounding error
- SST = 201.688
- SSE = 1060.287

Oneway Analysis of SCORE By TREATMENT

Oneway Anova

Summary of Fit

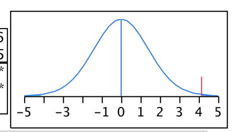
Rsquare	0.159833
Adj Rsquare	0.141162
Root Mean Square Error	4.854066
Mean of Response	17.85532
Observations (or Sum Wgts)	47

t Test

INTRINSIC-EXTRINSIC

Assuming equal variances

Difference	4.14420	t Ratio	2.925876
Std Err Dif	1.41640	DF	45
Upper CL Dif	6.99697	Prob > |t|	0.0054*
Lower CL Dif	1.29143	Prob > t	0.0027*
Confidence	0.95	Prob < t	0.9973



Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
TREATMENT	1	201.7080	201.708	8.5608	0.0054*
Error	45	1060.2882	23.562		
C. Total	46	1261.9962			

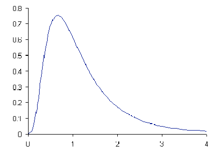
F-Test

- An F* is the ratio of two variances
  - Between treatments
  - Within treatments
- An F* = 1 mean equality of variances and that is the Null Hypothesis for our test

$$F^* = \frac{MST}{MSE}$$

$$F^* = \frac{201.688}{23.562} = 8.556$$

F Distribution



- The F-distribution is a positively skewed distribution with two degrees of freedom – the numerator d.f. and the denominator d.f.
- We compare our F* to a F distribution at a specified α, with v_1 and v_2 degrees of freedom
- v_1 represents the degrees of freedom for the numerator (k-1)
- v_2 represents the degrees of freedom for the denominator (n-k)
- The table will give the critical value to compare with α
- F-tests are always one-tailed tests

F-table of critical values for α = .05

- For each level of (.10, .05, .01) there is a different F-table with
  - Degrees of freedom for the numerator and denominator
  - The numerator degrees of freedom have less options (k-1)
  - What would be the F critical value value for 2 and 18 d.f.?

F-Distribution Tables

This table is best for regression & ANOVA tests were d.f.numerator tend to be small

Alpha is: 0.05

Degrees of Freedom Numerator

	d.f.	1	2	3	4	5	6	7	8	9	10	11	12
	1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	242.98	243.91
	2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.40	19.41
	3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.76	8.74
D	4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.94	5.91
e	5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.70	4.68
g	6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00
r	7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.60	3.57
e	8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.31	3.28
e	9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.10	3.07
s	10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91
	11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.82	2.79
o	12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.72	2.69
f	13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.63	2.60
	14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.57	2.53
F	15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.51	2.48
r	16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.46	2.42
e	17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.41	2.38
e	18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34
d	19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.34	2.31
o	20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.31	2.28

ANOVA Hypothesis Test for Creativity Data

- **Ho:** **Ho: $\mu_1 = \mu_2$**
- **Ha:** **Ha: At least two means are different**
- **Assumptions** **Equal variances, normal distribution**
- **Test Statistic** **F* = 8.5608 p =.0054**
- **Rejection Region** **$F_{.05,\,1,\,45} = 4.057$**
- **Conclusion:** **$F^* > F_{.05,\,1,\,45}$**

or p < .01

Reject Ho: $\mu_1 = \mu_2$

Let's revisit the Output

- Excel's output includes
  - Means and Variance
  - ANOVA Table
- JMP gives more

Anova: Single Factor

SUMMARY

Groups	Count	Sum	Average	Variance
Extrinsic	23	362.000	15.739	27.590
Intrinsic	24	477.200	19.883	19.709

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	201.708	1	201.708	8.561	0.005	4.057
Within Groups	1060.288	45	23.562			
Total	1261.996	46				

A note about a Difference of Means Test and the ANOVA Test

- t* from difference of means test = 2.926
- F* = 8.561
- In case of comparing just two means, if you square t*, it roughly equals F*
- The two tests (t-test and F-test for ANOVA) result in the same conclusion
- The advantage of ANOVA is that you can compare and make inferences on more than two means

More than two means: Coupon Example

- A greeting company wanted to use a coupon offer to increase sales
- They developed four different coupon designs, and used each design with a number of customers
- They took a sample of 8 customers for each design and noted their purchase amount as a result of the coupon
- Did the coupons have different effects on sales?

Customer	Design1	Design 2	Design 3	Design 4
1	$4.10	$6.90	$4.60	$12.50
2	$5.90	$9.10	$11.40	$7.50
3	$10.45	$13.00	$6.15	$6.25
4	$11.55	$7.90	$7.85	$8.75
5	$5.25	$9.10	$4.30	$11.15
6	$7.75	$13.40	$8.70	$10.25
7	$4.78	$7.60	$10.20	$6.40
8	$6.22	$5.00	$10.80	$9.20
Mean	$7.00	$9.00	$8.00	$9.00
Variance	$7.34	$8.42	$7.63	$5.02
GRAND MEAN		$8.25		

Experimental Design

1 Factor: Coupon Design

4 Treatments

8 replications per treatment

Experimental Unit: Person

Measurement Unit: Person

Total Sample Size: 32

Calculations

- The Variances are very similar - no problem with the assumptions
- **SST = 8(7.00-8.25)^2 + 8(9.00-8.25)^2 + 8(8.00-8.25)^2 + 8(9.00-8.25)^2**
  - **SST= 22.00**
  - **MST = 22.00/(4-1) = 7.33**
- **SSE = (8-1)7.34 + (8-1)8.42 + (8-1)7.63 + (8-1)5.02**
  - **SSE=198.87**
  - **MSE = 198.87/(32-4) = 7.10**
- **F* = 7.33/7.10 = 1.03**

But we can use the output from Excel orJMP

- Excel gives the means and Variance
  - F* = 1.032
  - Critical Value = 2.947
  - p-value = .3933
- JMP gives more
  - Box plot shows us the differences are small relative to the spread
  - We cannot reject Ho:

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
Design 1	8	56	7.000	7.341
Design 2	8	72	9.000	8.423
Design 3	8	64	8.000	7.632
Design 4	8	72	9.000	5.016

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	22	3	7.333	1.032	0.393	2.947
Within Groups	198.8818	28	7.103			
Total	220.8818	31				

ANOVA Hypothesis Test for Greeting Card Data

- **Ho:** Ho: $\mu_1 = \mu_2 = \mu_3 = \mu_4$
- **Ha:** Ha: At least two means are different
- **Assumptions** Equal variances, normal distribution
- **Test Statistic** $F^* = 1.032$ $p = .393$
- **Rejection Region** $F_{.05, 3, 28} = 2.947$
- **Conclusion:** $F^* > F_{.05, 3, 28}$

or $p = .393$

Cannot Reject Ho: $\mu_1 = \mu_2 = \mu_3 = \mu_4$

Using Excel to do an ANOVA

- Arrange data in columns – each factor level is a column
- It is a good idea to label the columns
- Use: Tools, Data Analysis, ANOVA: Single Factor
- Identify
  - the input range of columns
  - Alpha for the test
  - Whether labels are present
  - Output range
- I would also suggest using Data Descriptive on each column to get Factor Level statistics
- On the total sample to get the Grand Mean

Summary

- We now know the components of the ANOVA table
- And how to do an F-test for the difference of the means for the treatment levels
- There still is more terms to learn - R-square
- And the strategy of what to do next should we find significant differences in the means
- And we will explore a few other models – two factors and a block design