WWW.KRISHANPANDEY.COM

ANALYSIS OF VARIANCE

Dr. Krishan K. Pandey

2009

OP JINDAL GLOBAL UNIVERSITY

ANOVA: AN INTRODUCTION WITH EXAMPLES

When we wish to look at differences among three or more sample means, we use a statistical test called analysis of variance or ANOVA. Analysis of variance yields a statistic, F, which indicates if there is a significant difference among three or more sample means. When conducting an analysis of variance, we divide the variance of the scores into two components.

1. The variance between groups, that is the variability among the three or more group means.
2. The variance within the groups, or how the individual scores within each group vary around the mean of the group.

We measure these variances by calculating $\mathbf{SS_B}$, the sum of squares between groups, and $\mathbf{SS_W}$, the sum of squares within groups.

Each of these sum of squares is divided by its degrees of freedom, ($\mathbf{df_B}$, degrees of freedom between, and $\mathbf{df_W}$, degrees of freedom within) to calculate the mean square between groups, $\mathbf{MS_B}$, and the mean square within groups, $\mathbf{MS_W}$.

Finally we calculate **F**, the F-ratio, which is the ratio of the mean square between groups to the mean square within groups. We then test the significance of F to complete our analysis of variance.

Example for One-way Analysis of Variance

Three groups of students, 5 in each group, were receiving therapy for severe test anxiety. Group 1 received 5 hours of therapy, group 2 - 10 hours and group 3 - 15 hours. At the end of therapy each subject completed an evaluation of test anxiety (the dependent variable in the study). Did the amount of therapy have an effect on the level of test anxiety?

The three groups of students received the following scores on the Test Anxiety Index (TAI) at the end of treatment.

TAI Scores for Three Groups of Students		
Group 1 - 5 hours	**Group 2 - 10 hours**	**Group 3 - 15 hours**
48	55	51
50	52	52
53	53	50
52	55	53
50	53	50

The following table contains the quantities we need to calculate the means for the three groups, the sum of squares, and the degrees of freedom:

Worksheet for Test Anxiety Study					
Group 1 - 5 hours		**Group 2 - 10 hours**		**Group 3 - 15 hours**	
X_1	$(X_1)^2$	X_2	$(X_2)^2$	X_3	$(X_3)^2$
48	2304	55	3025	51	2601
50	2500	52	2704	52	2704
53	2809	53	2809	50	2500
52	2704	55	3025	53	2809
50	2500	53	2809	50	2500
----------	----------	----------	----------	----------	----------
253	12817	268	14372	256	13114

The mean for group 1 is 253/5 = 50.6, the mean for group 2 is 268/5 = 53.6, and the mean for group 3 is 256/5 = 51.2

Are the differences between these three means significant? We can use analysis of variance to answer that question. Since we only have one independent variable, amount of therapy, we will use one-way analysis of variance. If we were concerned with the effect of two independent variables on the dependent variable, then we would use two-way analysis of variance.

First we will calculate SS_B, the sum of squares between groups, where X_1 is a score from Group 1, X_2 is a score from Group 2, X_3 is a score from Group 3, n_1 is the number of subjects in group

1, n_2 is the number of subjects in group 2, n_3 is the number of subjects in group 3, X_T is a score from any subject in the total group of subjects, and N_T is the total number of subjects in all groups.

$$SS_B = \frac{(\sum X_1)^2}{n_1} + \frac{(\sum X_2)^2}{n_2} + \frac{(\sum X_3)^2}{n_3} - \frac{(\sum X_T)^2}{n_T}$$

$$= \frac{(253)^2}{5} + \frac{(268)^2}{5} + \frac{(256)^2}{5} - \frac{(777)^2}{15}$$

$$= \frac{64009}{5} + \frac{71824}{5} + \frac{65536}{5} - \frac{603729}{15}$$

$$= 12801.8 + 14364.8 + 13107.2 - 40248.6 = 25.2$$

The degrees of freedom between groups are: $df_B = K - 1 = 3 - 1 = 2$,Where K is the number of groups. Next we calculate SS_W, the sum of squares within groups.

$$SS_W = \left(\sum X_1^2 - \frac{(\sum X_1)^2}{n_1}\right) + \left(\sum X_2^2 - \frac{(\sum X_2)^2}{n_2}\right) + \left(\sum X_3^2 - \frac{(\sum X_3)^2}{n_3}\right)$$

$$= \left(12817 - \frac{(253)^2}{5}\right) + \left(14372 - \frac{(268)^2}{5}\right) + \left(13114 - \frac{(256)^2}{5}\right)$$

$$= (12817 - 12801.8) + (14372 - 14364.8) + (13114 - 13107.2)$$

$$= 15.2 + 7.2 + 6.8 = 29.2$$

The degrees of freedom within groups are: $df_W = N_T - K = 15 - 3 = 12$, Where N_T is the total number of subjects. Finally, we will calculate SS_T, the total sum of squares.

$$SS_T = \sum X_T^2 - \frac{(\sum X_T)^2}{N_T} = 40303 - \frac{(777)^2}{15} =$$

$$40303 - \frac{603729}{15} = 40303 - 40248.6 = 54.4$$

As a check $SS_T = SS_B + SS_W$

54.4 = 25.2 + 29.2

We can now calculate MS_B, the mean square between groups, MS_W, the mean square within groups, and F, the F ratio.

$$MS_B = \frac{SS_B}{df_B} = \frac{25.2}{2} = 12.60$$

$$MS_W = \frac{SS_W}{df_W} = \frac{29.2}{12} = 2.43$$

$$F = \frac{MS_B}{MS_W} = \frac{12.60}{2.4333} = 5.178$$

To test the significance of the F value we obtained, we need to compare it with the critical F value with an alpha level of .05, 2 degrees of freedom between groups (or degrees of freedom in the numerator of the F ratio), and 12 degrees of freedom within groups (or degrees of freedom in the denominator of the F ratio). We can look up the critical value of F in Appendix Table D of the text book (The 5 percent (Lightface Type) and 1 percent (Boldface Type) points for the Distribution of F), pages 319-326. Look in the table under column 2 (2 degrees of freedom for the numerator) and row 12 (12 degrees of freedom for the denominator) and read the non-boldfaced entry (for .05 level) of 3.88 - this is the critical value for F.

One way of indicating this critical value of F at the .05 level, with 2 degrees of freedom between groups and 12 degrees of freedom within groups is $F_{.05}(2,12) = 3.88$.

When using analysis of variance, it is a common practice to present the results of the analysis in an analysis of variance table. This table which shows the source of variation, the sum of squares, the degrees of freedom, the mean squares, and the probability is sometimes presented in a research article. The analysis of variance table for our problem would appear as follows:

Analysis of Variance Table

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio	p
Between Groups	25.20	2	12.60	5.178	<.05
Within Groups	29.20	12	2.43		
Total	54.40	14			

We now have the information we need to complete the six step process for testing statistical hypotheses for our research problem. We will also be adding another analysis of the individual means.

1. **State the null hypothesis and the alternative hypothesis based on your research question**.

   $$H_0: \mu_1 = \mu_2 = \mu_3$$
   $$H_1: not H_0$$

   Note: Our null hypothesis, for the F test, states that there are no differences among the three means. The alternate hypothesis states that there are significant differences among some or all of the individual means. An unequivocal way of stating this is **not H_0**.

2. **Set the alpha level**. $\alpha = .05$

   Note: As usual we will set our alpha level at 0.05; we have 5 chances in 100 of making a type I error.

3. **Calculate the value of the appropriate statistic. Also indicate the degrees of freedom for the statistical test if necessary. F(2, 12) = 5.178;**

   Note: We have indicated the value of F from our analysis of variance table. We have also indicated by (2, 12) that there are 2 degrees of freedom between groups, and 12 degrees of freedom within groups.

4. **Write the decision rule for rejecting the null hypothesis.** Reject H_0 if F is >= 3.88

   Note: To write the decision rule we had to know the critical value for F, with an alpha level of .05, 2 degrees of freedom in the numerator (df between groups) and 12 degrees of freedom in the denominator (df within groups). We can do this by looking at Appendix Table D and noting the tabled value for the .05 level in the column for 2 df and the row for 12 df.

5. **Write a summary statement based on the decision.** Reject H_0, $p < 0.05$

   Note: Since our calculated value of F (5.178) is greater than 3.88, we reject the null hypothesis and accept the alternative hypothesis.

6. **Write a statement of results in Standard English.**

   There is a significant difference among the scores the three groups of students received on the Test Anxiety Index.

In the problem above, we rejected the null hypothesis and found that there is indeed a significant difference among the three cell means. We know that Group 1 had the lowest mean (50.6), while group 3 had a higher mean (51.2) while group 2 had the highest mean of all (53.6). We would like to know which of these differences in means are significant. We can analyze the significance of the difference between pairs of means in analysis of variance by the use of post hoc (after the fact) comparisons. We only do these post hoc comparisons when there is a significant F ratio. It would make no sense to look for differences with a post hoc test if no differences exist.

Problem 1: A research study was conducted to examine the clinical efficacy of a new antidepressant. Depressed patients were randomly assigned to one of three groups: a placebo group, a group that received a low dose of the drug, and a group that received a moderate dose of the drug. After four weeks of treatment, the patients completed the Beck Depression Inventory. The higher the score, the more depressed the patient. The data are presented below. Compute the appropriate test.

Placebo	Low Dose	Moderate Dose
38	22	14
47	19	26
39	8	11
25	23	18
42	31	5

Source	SS	df	MS	F
Between	1484.9333333	2	742.4666666	11.26
Within	790.8	12	65.9	
Total	2275.733333	14		

1. What is your computed answer? **Answer:** $F = 11.26\ (2,12)\ p < .01$
2. What would be the null hypothesis in this study? **Answer:** There will be no difference in depression levels between the three groups. The groups taking the drug will not be different than the groups taking the placebo.
3. What would be the alternate hypothesis? **Answer:** There will be a difference somewhere in depression levels between the three levels of drug groups.

4. What probability level did you choose and why? **Answer:** $p = .01$. There is a risk involved with a Type I error. I do not want to erroneously say the drug works and then later find out that it doesn't.
5. What is your F_{crit}? **Answer:** $F_{crit} = 6.93$
6. Is there a significant difference between the groups? **Answer:** Yes - a significant difference exists somewhere between the three groups.
7. If there is a significant difference, where specifically are the differences? **Answer:** There is a significant difference between the placebo group and the low dose group ($F_{comp} = 11.75$ and $q_{obs} = 4.84$, $p < .05$). There is a significant difference between the placebo group and the moderate dose group ($F_{comp} = 20.77$ and $q_{obs} = 6.44$, $p < .01$). There is no significant difference between the low dose and the moderate dose groups ($F_{comp} = 1.27$ and $q_{obs} = 1.59$, n.s.).
8. Interpret your answer. **Answer:** The drug appears to help alleviate depression. However, as there is no significant difference between taking a low or moderate dose, a low dose would be recommended.

Problem 2: A researcher is concerned about the level of knowledge possessed by university students regarding United States history. Students completed a high school senior level standardized U.S. history exam. Major for students was also recorded. Data in terms of percent correct is recorded below for 32 students. Compute the appropriate test for the data provided below.

Education	Business/Management	Behavioral/Social Science	Fine Arts
62	72	42	80
81	49	52	57
75	63	31	87
58	68	80	64
67	39	22	28
48	79	71	29
26	40	68	62
36	15	76	45

Source	SS	df	MS	F
Between	63.25	3	21.0833333333	.04
Within	12298.25	28	439.2232143	
Total	12361.5	31		

1. What is your computed answer? **Answer:** F = .04 (3,28), not significant
2. What would be the null hypothesis in this study? **Answer:** There will be no difference in history test scores between students with different academic major.
3. What would be the alternate hypothesis? **Answer:** There will be a difference somewhere in history scores between the four groups with different academic major.
4. What probability level did you choose and why? **Answer:** p = .05 There is little risk involved if either a Type I or a Type II major is made.
5. What were your degrees of freedom? **Answer:** 3, 28
6. Is there a significant difference between the four testing conditions? **Answer:** No significant differences were found between the four groups in terms of performance on a U.S. history exam.
7. Interpret your answer. **Answer:** Students regardless of academic major performed equally (in this case poorly) on a high school senior standardized U.S. history exam.
8. If you have made an error, would it be a Type I or a Type II error? Explain your answer. **Answer:** If I have made an error, it would be a Type II error. There really is a difference in history knowledge between academic major but somehow I failed to demonstrate that with this study.

Problem 3: Neuroscience researchers examined the impact of environment on rat development. Rats were randomly assigned to be raised in one of the four following test conditions: Impoverished (wire mesh cage - housed alone), standard (cage with other rats), enriched (cage with other rats and toys), super enriched (cage with rats and toys changes on a periodic basis). After two months, the rats were tested on a variety of learning measures (including the number of trials to learn a maze to a three perfect trial criteria), and several neurological measure (overall cortical weight, degree of dendritic branching, etc.). The data for the maze task is below. Compute the appropriate test for the data provided below.

Impoverished	Standard	Enriched	Super Enriched
22	17	12	8
19	21	14	7
15	15	11	10
24	12	9	9
18	19	15	12

Source	SS	df	MS	F
Between	323.35	3	107.7833	12.71
Within	135.6	16	8.475	
Total	458.95	19		

1. What is your computed answer? **Answer:** F = 12.71 (3,16) $p < .01$
2. What would be the null hypothesis in this study? **Answer:** Environment will have no impact on learning ability as operationalized by maze performance in rats.
3. What would be the alternate hypothesis? **Answer:** Environment will have an impact on learning ability as operationalized by maze performance in rats.
4. What is your F_{crit}? **Answer:** $F_{crit} = 5.29$
5. Are there any significant differences between the four testing conditions? **Answer:** Yes - There is no significant difference between the impoverished group and the standard group ($F_{comp} = 2.32$ and $q_{obs} = 2.15$, n.s.). There is a significant difference between the impoverished group and both the enriched and super enriched group ($F_{comp} = 16.15$ and $q_{obs} = 5.68$, $p < .01$) and $F_{comp} = 31.90$ and $q_{obs} = 7.98$, $p < .01$), respectively). There is no significant difference between the standard group and the enriched group ($F_{comp} = 6.24$ and $q_{obs} = 3.53$, n.s.). There is a significant difference between the standard group and the

super enriched group (F_{comp} = 17.03 and q_{obs}= 5.83, $p < .05$). There is no significant difference between the enriched group and the super enriched group (F_{comp} = 2.65 and q_{obs}= 2.30, $p < .05$)).

6. Interpret your answer. **Answer:** Environment may have an impact on ability to learn. Differences were found between groups when each group is compared to a group at least two levels above the one under study. Thus for example, there is a difference between the impoverished and the enriched and super enriched but not between the impoverished and the standard groups.

Problem 4: A research study was conducted to examine the impact of eating a high protein breakfast on adolescents' performance during a physical education physical fitness test. Half of the subjects received a high protein breakfast and half were given a low protein breakfast. All of the adolescents, both male and female, were given a fitness test with high scores representing better performance. Test scores are recorded below.

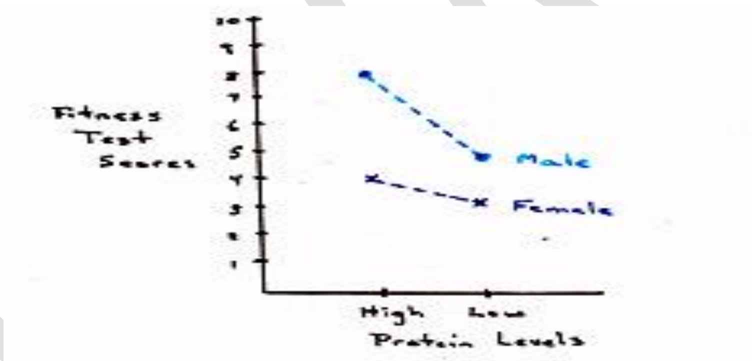
Group	**High Protein**	**Low Protein**
Males	10 7 9 6 8 Mean=8.0	5 4 7 4 5 Mean=5.0
Females	5 4 6 3 2 Mean=4.0	3 4 5 1 2 Mean=3.0

1. Complete the following ANOVA table.

Source	**SS**	**df**	**MS**	**F**
Protein Level	20	1	20.00	8.89

Gender	45	1	45.00	20.00
Protein Level x Gender	5	1	5.00	2.22
Within	36	16	2.25	
Total	106	19		

2. Graph your data.



3. Are there any significant main effects or an interaction effect.

   Answer: There appears to be significant main effects for both protein level (F=8.89 (1, 16), p<.01) and gender (F=20.00 (1,16), p<.01). There was not a significant interaction effect (F=2.22 (1, 16), not significant).

4. Interpret your answer.

   Answer: Based on this data, it appears that a high protein diet results in a better fitness test score. Additionally, young men seem to have a significantly higher fitness test score than women.

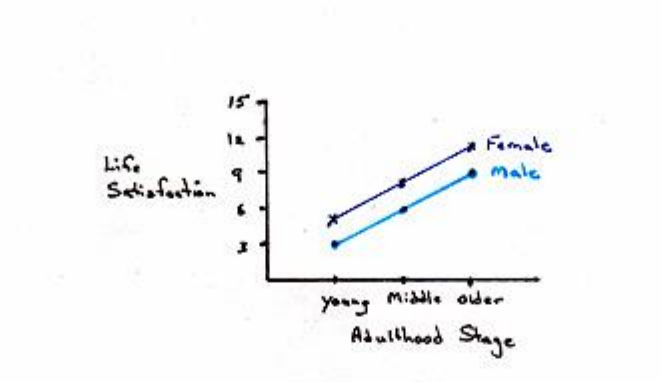
Problem 4: A study examining differences in life satisfaction between young adult, middle adult, and older adult men and women was conducted. Each individual who participated in the study completed a life satisfaction questionnaire. A high score on the test indicates a higher level of life satisfaction. Test scores are recorded below.

Group	**Young Adult**	**Middle Adult**	**Older Adult**
Male	4 2 3 4 2 Mean=3.0	7 5 7 5 6 Mean=6.0	10 7 9 8 11 Mean=9.0
Female	7 4 3 6 5 Mean=5.0	8 10 7 7 8 Mean=8.0	10 9 12 11 13 Mean=11.0

1. Complete the following ANOVA table.

Source	**SS**	**df**	**MS**	**F**
Age	180	2	90.00	49.09
Gender	30	1	30.00	16.36
Age x Gender	0	2	0.00	0.00
Within	44	24	1.83	
Total	254	29		

2. Graph your data.



3. Are there any significant main effects or an interaction effect.

   Without post hoc tests it is difficult to know where differences lie specifically. There are significant main effects for age (F=49.09 (2,24), p<.01) and gender (F=16.36 (1, 24), p<.01). There is no interaction effect (F=0.00 (2,24), not significant).

4. Interpret your answer.

   It appears from the data that older adults have the highest life satisfaction and younger adults have the lowest life satisfaction. Women also have significantly higher life satisfaction than men.