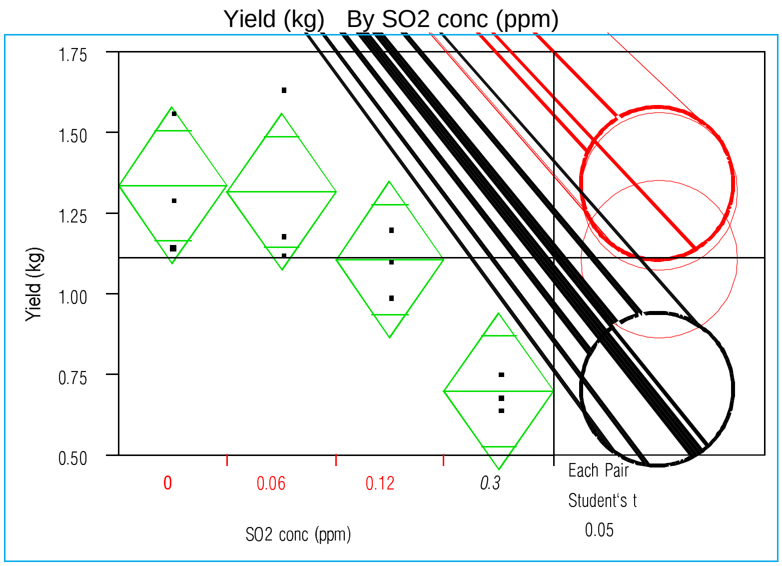
ONE−WAY ANOVA EXAMPLE

In this example, we reanalyze the soybean yield data, but treat the explanatory variable as a categorical (ordinal) variable with 4 levels and use one−way anova to test for a significant SO2 effect. We can think of the data as samples from four different populations, corresponding to the SO2 concentration the chambers received, with possibly different means. In each treatment group we have 3 observations.

The null hypothesis is that SO2 does not effect yields, so that the 4 means under the different levels of SO2 are all the same. This is the same model as in the null hypothesis from the regression analysis, that the data all have a common mean. However, the alternative hypothesis is very different. Here, the alternative hypothesis is that the SO2 treatment leads to at least one of the means being different from the others. This doesn't specify which one; they could all be different or only some of them are different. This is less specific than in the case of the regression model, which implied that the mean yield increased (decreased) linearly with SO2 concentration.

Here is the JMP output. The top plot shows the data plotted by SO2 levels. The vertical points of the diamonds represent 95% confidence intervals, with the estimated means represented by the horizontal line in the center of the diamond. The circles on the right are used to graphically view which treatments are significantly different from each other. In JMP, clicking on the circles will show which treatments are different by using a different color.



The ANOVA analysis assumes that all observations with the same treatment come from a normal population with the same mean and standard deviation. Futhermore, the SD's are the same over all treatment groups. The confidence intervals above use a "pooled" estimate of the standard deviation,

$$\text{MSE = Mean Square Error} = \frac{\sum_{ij} (y_{ij} - \bar{y}_j)^2}{(n-4)} = 0.03428$$

SD = sqrt(MSE) = sqrt(0.03428) = 0.1851

SE = SD/sqrt(n_i) = 0.1851/ $\sqrt{3}$ = 0.1069

SE(diff i and j) = sqrt|(SEi^2 + SEj^2) = 0.1512

Like in regression, we have an ANOVA table that shows how the variation in yields is broken down. Note that the df and Sum of Squares for the "Total" are the same as before. However, now we have 3 df for the Model, while the regression model had just one. Under the null hypothesis there is 1 df for the overall mean. Under the alternative hypothesis we have 4 means that we must estimate; the model df represent the additional parameters that we must estimate beyond the number under the null, so df = 4 − 1 = 3. The error df are n − 4; the sample size − the total number of parameters we must estimate. The same formula from before holds for the Sum of Squares, but now the "fitted" values are just the sample means in each group. The F ratio is used to test the null hypothesis that there is no SO2 effect. If the means are significantly different from each other, and hence the overall mean, then the numerator of the F ratio is large. The corresponding p−value (the probability of getting an F ratio larger than the 7.7278 that we observed) is 0.0095. Since this is less than 0.05, we can conclude that at least one of the means is significantly different from the others, and that there is a statistically significant effect of SO2 on soy bean yields.

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	3	0.7948000	0.264933	7.7278
Error	8	0.2742667	0.034283	Prob>F
C Total	11	1.0690667	0.097188	0.0095

OK, so which means are the different! Multiple comparisons of means are often used to show which treatments lead to difference. The simplest approach is to do all pairwise tests to compare means.

Means for Oneway Anova

Level	Number	Mean	Std Error
0	3	1.34000	0.10690
0.06	3	1.32000	0.10690
0.12	3	1.10667	0.10690
0.3	3	0.70000	0.10690

Std Error uses a pooled estimate of error variance

Means Comparisons

Dif=Mean[i]−Mean[j]	0	0.06	0.12	0.3
0	0.000000	0.020000	0.233333	0.640000
0.06	−0.02	0.000000	0.213333	0.620000
0.12	−0.23333	−0.21333	0.000000	0.406667
0.3	−0.64	−0.62	−0.40667	0.000000

Comparisons for each pair using Student's t alpha=0.05; t=2.30603

Least Significant Difference = LSD = t*SE(diff)

Abs(Dif)−LSD	0	0.06	0.12	0.3
0	−0.34863	−0.32863	−0.11529	0.291374
0.06	−0.32863	−0.34863	−0.13529	0.271374
0.12	−0.11529	−0.13529	−0.34863	0.058040
0.3	0.291374	0.271374	0.058040	−0.34863

Positive values show pairs of means that are significantly different.

This has a problem in that the overall Type I error rate for the multiple comparisons is much higher than the rate for a single test. If for each single test you fix the Type I error rate at 0.05, then what is the probability of at least one Type I error in 6 comparisons? Other procedures adjust the t value so that the overall Type I error rate is smaller. ie Tukey−Kramer. Does using another procedure change the conclusions? Check it out in lab!