The ANOVA for 2 Independent Groups -- Analysis of 2-Between-Group Data with a Quantitative Response Variable

Application: To compare means of a quantitative response variable obtained from 2 independent groups.

About the terms used: Descriptions of ANOVA are often in terms of the IV and the DV. However, this can be confusing when working with non-experimental data. In addition, different terminology is often used for between groups ANOVA than for within-groups ANOVA, adding to the confusion. The terms *Condition* and *Error* will be used below for two reasons. First, it will allow you to see that the between groups ANOVA and within-groups ANOVA do basically the same thing -- comparing size of the mean differences (variability associated with which condition the score came from) to the variability around those means (an index of error). Second, it avoids the use of "causal words".

H0: The populations represented by the conditions of the qualitative variable have the same mean on the quantitative response variable.

To reject H0: is to say that the populations represented by the conditions of the qualitative variable have different means on the quantitative response variable.

The data: This analysis involves the grouping variable **reptdept** (1 = not separate reptile department, 2 = separate department), and the quantitative response variable **reptgood** (rating of reptile quality - 1-10 scale) . Below are the data for the 12 stores (reptdept, reptgood).

     1,2 2,8 2,9 2,7 1,4 1,7 2,4 1,4 1,5 2,9 2,7 1,2

Research Hypothesis for this analysis: The researcher hypothesized that stores with separate reptile departments would have reptiles of better overall quality than stores which did not have separate reptile departments.

H0: for this analysis: Pet shops that do not have separate reptile departments have the same mean reptile quality ratings as shops that do have separate reptile departments.

Note: **k** = the # of conditions   **n** = # of participants in a condition   **N** = total # of participants

Assemble the data for analysis. Rearrange the data so that the subjects in each condition are in the same column. There should be as many columns as there are conditions. Label the quantitative variable as X and each condition of the grouping variable as a different value of k, to simplify the presentation and use of the formulas below.

not separate dept. (k1)	separate dept. (k2)
X	X
2	8
4	9
7	7
4	4
5	9
2	7

Compute the square of each score and place it in an adjacent column.

not separate dept. (k1)		separate dept. (k2)	
X	X²	X	X²
2	4	8	64
4	16	9	81
7	49	7	49
4	16	4	16
5	25	9	81
2	4	7	49

Compute $\sum X$, $\sum X^2$ for each group and determine the **sample size** for each group.

<pre>
        not separate dept. (k1)   separate dept. (k2)
</pre>

$$\sum X_{k1} \quad = \quad 24 \qquad\qquad\qquad\qquad \sum X_{k2} \quad = \quad 44$$
$$\sum X_{k1}^2 \quad = \quad 114 \qquad\qquad\qquad\qquad \sum X_{k2}^2 \quad = \quad 340$$
$$n_{k1} \quad = \quad 6 \qquad\qquad\qquad\qquad n_{k2} \quad = \quad 6$$

Compute the mean and std for each condition.

Computational examples using data from not separate department condition:

$$\text{Mean} = \frac{\sum X_{k1}}{n_{k1}} = \frac{24}{6} = 4.0$$

$$\text{Standard deviation} = \sqrt{\left[\frac{\sum X_{k1}^2 - [(\sum X_{k1})^2/n_{k1}]}{n_{k1}-1}\right]} = \sqrt{\left[\frac{114 - [(24)^2/6]}{6-1}\right]} = 1.90$$

Compute SS_{Total}, a measure of the total variability among all the scores (ignoring the scores come from different groups).

$$SS_{Total} = (\sum X_{k1}^2 + \sum X_{k2}^2) - \frac{(\sum X_{k1} + \sum X_{k2})^2}{N}$$

$$= (114 + 340) - \frac{(24 + 44)^2}{12}$$

$$= 454 - \frac{(68)^2}{12} = 454 - 385.33 = 68.67$$

Compute SS_{Cond} (sometimes referred to as SS_{BG} for between groups of SS_{IV} if the data are from an experiment), a   measure of the variability (difference) among the means from the different groups. (Notice the "right side" of this step is the same as for Step 4)

$$SS_{Cond} = \left[\frac{(\sum X_{k1})^2}{n_{k1}} + \frac{(\sum X_{k2})^2}{n_{k2}}\right] - \frac{(\sum X_{k1} + \sum X_{k2})^2}{N}$$

$$= \left[\frac{24^2}{6} + \frac{44^2}{6}\right] - 358.33 = 418.67 - 358.33 = 33.34$$

Compute SS_{Error} (also called SS_{WG} or SS_{Within}), a measure of the variability among the scores within each condition. (Notice that the "right side" of this equation is the same as the "left side" of Step 5, also the "left side" of this step is the   same as for Step 4-- saves some computation).

$$SS_{Error} = (\sum X_{k1}^2 + \sum X_{k2}^2) - \left[\frac{(\sum X_{k1})^2}{n_{k1}} + \frac{(\sum X_{k2})^2}{n_{k2}}\right]$$

$$= 454 - 418.67 = 35.33$$

Perform a computational check to be sure that $SS_{Total} = SS_{Cond} + SS_{Error}$.

Compute df$_{Cond}$ (also called df$_{BG}$)

$$df_{Cond} = k - 1 = 2 - 1 = 1$$

Compute MS$_{Cond}$

$$MS_{Cond} = \frac{SS_{Cond}}{df_{cond}} = \frac{33.34}{1} = 33.34$$

Compute df$_{Error}$ (also called df$_{WG}$ or df$_{Within}$)

$$df_{Error} = N - k = 12 - 2 = 10$$

Compute MS$_{Error}$ (Also called the MSE for mean squared error)

$$MS_{Error} = \frac{SS_{Error}}{df_{Error}} = \frac{35.33}{10} = 3.53$$

Compute the F (also called the F-obtained, F-computed or F-omnibus)

$$F = \frac{MS_{Cond}}{MS_{Error}} = \frac{33.34}{3.53} = 9.44$$

Prepare a Summary Table (also called a Source Table)

Occasionally you will be asked to prepare a summary of the ANOVA computations, here is a common way to organize that summary.

Note: the Total df is calculated as the sum of df(Cond) and df(WG) or as N -1.

Summary Table Showing Values from the Example

Source	df	SS	MS	F	p
Cond	1	33.34	33.34	9.44	<.05
Error	10	35.33	3.53		
Total	11	68.67			

Look up the F-critical for α = .05 and the appropriate degrees of freedom using the F-table.
Use numerator degrees of freedom = df$_{IV}$ = 1 and denominator degrees of freedom = df$_{Error}$ = 10.

$$F(1, 10, \alpha = .05) = 4.96$$

Determine whether to retain or reject H0:

- if the obtained F is less than the critical F, then retain the null hypothesis -- conclude that the populations represented by the different conditions o the qualitative grouping variable have the same mean score on the quantitative variable
- if the obtained F is greater than the critical F, then reject the null hypothesis -- conclude that the populations represented by the different conditions of the qualitative grouping variable have different mean scores on the quantitative variable

For the example data, we would decide to reject the null hypothesis, because the obtained F value of 9.44 is larger than the critical F value of 4.84.

Determine whether the pattern of the means supports or does not support the research hypothesis.

- Usually the researcher hypothesizes that there is a difference between the conditions. If so, then to support the research hypothesis will require:
  - Reject H0: that there is no mean difference
  - The mean difference must be in the same direction as that specified in the research hypothesis

- Sometimes, however, the research hypothesis is that there is **no** difference between the conditions. If so, the research hypothesis and H0: are the same!
  - When this is the case, retaining H0: provides support for the research hypothesis, whereas rejecting H0: provides evidence that research hypothesis is incorrect.

- **Please note:** When you have decided to retain H0: (because F < F-critical), then don't talk about one group having a mean that is larger or smaller than the other -- retaining H0: is saying that the means are the same (and any apparent difference is probably due to sampling variation, chance, etc.)

The researcher hypothesized that stores with a separate reptile department would have the highest reptile quality ratings. This "translates" into an expected difference between the store types that rept. separate > rept. not-separate.

Based on rejecting the H0: and an inspection of the means, we would conclude there is support for the research hypothesis. Stores with separate reptile departments did have a higher overall mean reptile quality than did store without a separate reptile department.

By the Way: Usually the researcher hypothesizes that there is a difference between the conditions of the qualitative grouping variable. Sometimes, however, the research hypothesis is that there is NO mean difference between the conditions. If so, the research hypothesis and H0: are the same! When this is the case, retaining H0: provides support for the research hypothesis, whereas rejecting H0: provides evidence that research hypothesis is incorrect.

Describe the results of the Independent Groups Analysis of Variance -- be sure to include the following

- Name and tell the conditions of the qualitative or grouping variable
- Name the quantitative variable and tell its mean and standard deviation for each condition of the qualitative variable
- Tell the he F-value, df (in parentheses) and p-value ($p < .05$ or $p > .05$).
- If you reject H0:, tell which condition has the larger mean
  - If you retain H0:, tell that the condition means aren't significantly different
- Tell whether or not the results support the research hypothesis

Please note: Reporting ANOVA results is not a form of "creative writing". The idea is to be succinct, clear, and follow the prescribed format -- it is really a lot like completing a fill-in-the-blanks sentence. After you write and read enough of these you'll develop some "style", but for now just follow the format.

Here are two write-ups of these results that say the same thing. The 1st reports the univariates and then the significance test. The 2nd combines them into a single sentence -- either is fine.

Those stores without separate reptile departments displayed reptiles with a mean quality rating of 4.0 ($\underline{S} = 1.90$), whereas those that did have separate departments had a mean rating of 7.33 ($\underline{S} = 1.86$). As hypothesized, pet stores with separate reptile departments had significantly higher mean ratings than those without separate departments, $\underline{F}(1,10) = 9.44$, $\underline{p} < .05$, $\underline{Mse} = 3.53$.

Consistent with the research hypothesis, stores with a separate reptile department had higher reptile quality ratings ($\underline{M} = 7.33$, $\underline{S} = 1.86$) than stores without a separate reptile department ($\underline{M} = 4.0$, $\underline{S} = 1.90$), $\underline{F}(1,10) = 9.44$, $\underline{p} < .05$, $\underline{MSE} = 3.53$.

ANOVA Write-up Examples for Other Occasions

Here's are examples (not all using the same data as the computational example above) of what we would write for different combinations of RH:, condition means & significance test results. **Note**: You can follow these for WG ANOVA too.

For…
- RH: those stores with separate reptile departments would have higher means than those without separate departments
- retained H0:

Stores without separate reptile departments displayed reptiles with a mean quality rating of 6.75 (S = 1.90) and those that did have separate departments had a mean rating of 7.33 (S = 1.86). Contrary to the research hypothesis pet stores with and without separate reptile depart- ments did not have significantly different mean ratings, $F(1,10)$ = 2.09, p > .05, Mse = 3.53.

Contrary to the research hypothesis, stores with a separate reptile department did not have significantly different mean reptile quality ratings (M = 7.33, S = 1.86) than stores without a separate reptile department (M = 6.75, S = 1.90) $F(1,10)$ = 2.09, p > .05, MSE = 3.53.

For…
- RH: there would be no mean difference between the conditions (RH: = H0:)
- retained H0:

Stores without separate reptile departments displayed reptiles with a mean quality rating of 6.75 (S = 1.90) and those that did have separate departments had a mean rating of 7.33 (S = 1.86). As hypothesized here was no significant difference between the mean reptile quality of these two types of stores, $F(1,10)$ = 2.09, p > .05, MSE = 3.53.

As hypothesized, stores with a separate reptile department had about the same reptile quality, ratings (M = 7.33, S = 1.86) as stores without a separate reptile department (M = 6.75, S = 1.90) $F(1,10)$ = 2.09, p > .05, MSE = 3.53.

For…
- RH: those with separate reptile departments would have higher means
- rejected H0:
- but found the mean difference was "backwards" (these are not the same data as the write-up above)

Stores with a separate reptile department had mean reptile quality ratings of 3.33 (S = 1.86), whereas stores without a separate reptile department had a mean rating of 6.75 (S = 1.90). Contrary to the research hypothesis, stores with a separate reptile department had lower reptile quality ratings than stores without a separate reptile department, $F(1,10)$ = 9.09, p < .05, MSE = 3.53.

Contrary to the research hypothesis, stores with a separate reptile department had lower reptile quality ratings (M = 3.33, S = 1.86) than stores without a separate reptile department (M = 6.75, S = 1.90), $F(1,10)$ = 9.09, p < .05, MSE = 3.53.

For…
- RH: there would be no mean difference between the conditions (RH: = H0:)
- Rejected H0:

Stores with a separate reptile department had a mean reptile quality rating of 3.33 (S = 1.86), whereas stores without a separate reptile department had a mean rating of 6.75 (S = 1.90). Contrary to the research hypothesis that the two types of stores would have the same mean quality ratings, stores with a separate reptile department had lower reptile quality ratings than stores without a separate reptile department, $F(1,10)$ = 9.09, p < .05, MSE = 3.53.

Contrary to the research hypothesis that the two types of stores would have the same mean quality ratings, stores with a separate reptile department had lower reptile quality ratings (M = 3.33, S = 1.86) than stores without a separate reptile department (M = 6.75, S = 1.90), $F(1,10)$ = 9.09, p < .05, MSE = 3.53.

ANOVA Write-up Example Using a Table or Figure

Here is another write-up of the example analysis using a Table to present the univariate statistics. Tables reduce the about of parenthetical information that can clutter a write-up.

The quality of reptiles at each type of store is summarized in Table 1. Consistent with the research hypothesis, stores with a separate reptile department had higher reptile quality ratings than stores without a separate reptile department, $\underline{F}(1,10) = 9.44$, $\underline{MSE} = 3.53$, $\underline{p} < .05$.

Table 1
Quality of reptiles at each Store Type

	Store Type	
Quality of Reptiles	Separate Reptile Department	No Separate Reptile Department
$\underline{M}$	7.33	4.0
$\underline{SD}$	1.86	1.90
$\underline{n}$	6	6

Finally, here is what a Figure representing these data would look like. The height of the bars represents the mean of each condition and the "whiskers" show one std above and below the mean of each condition.
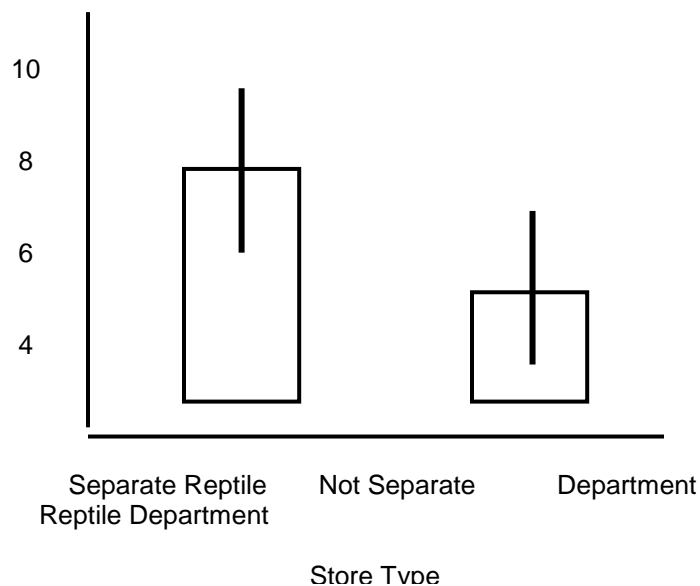

Figure 1.
Mean quality for each type of animal (+/- 1 std shown)