

BIOL 310

Two Factor Analysis of Variance

In the previous discussions of analysis of variance (ANOVA), only one “factor” was involved. For example, in Chapter 7 the variable of interest in the sample problem was cholesterol concentration and the statistical test was to determine if sex affected the concentration. In that problem sex is termed the factor. In Chapter 9, the variable of interest in the sample problem was body weight and the test was to determine if the feed used (the factor) affected body weight. In that problem, four feeds were used and each type of feed is said to be a *level* of the factor.

In many biological experiments, there are more than one factor involved. This lesson deals with a two factor analysis, which is dealt with by running a two-way ANOVA, which can then be followed by a post-test if necessary.

One of the strengths of a two-way ANOVA is that it allows you to test three null hypotheses at one time, all of which may be biologically important.

1. Ho: there is no effect of factor 1 on the **response variable** (variable of interest)
2. Ho: there is no effect of factor 2 on the response variable
3. Ho: there is no interaction between factor 1 and factor 2 on the response variable

The first two of these can be tested using the one factor ANOVA (one-way ANOVA) but the third one, the interaction, can be obtained only through the two-way ANOVA approach. Besides it is much more efficient to test all three null hypotheses in one single test (two-way ANOVA) rather than using a series of tests (one-way ANOVA).

If it is found that there is a significant interaction between the two factors, then neither factor can be considered independently. An **interaction** means that the effect of one factor depends on the level of another factor. If the interaction is significant it could be due to *synergism* or *interference*. When two factors act synergistically, the combined effect is greater than the sum of the separate effects. In contrast, when two factors in combination inhibit each other's effects it is termed interference. Biologically, interactions are very important and can be extremely interesting; however, they can sometimes be a challenge to interpret.

One way to think about all of the possible Hos being tested by a two-way ANOVA is to consider them as a series of linear equations. For instance, in a one-way ANOVA we are interested in the effect of a single factor (X) on the response variable (Y) then you could write this as a word equation in the form of $Y = X$. Using the example from Chapter 7, the word equation describing the analysis would be: cholesterol = sex. A two-way ANOVA then investigates three scenarios:

$$\text{Response} = \text{Factor1}$$

$$\text{Response} = \text{Factor2}$$

$$\text{Response} = \text{Factor1} * \text{Factor2}$$

with the later being the potential interaction between the two factors.

Calculating a Two-way ANOVA Using MINITAB

Although only slightly more complex mathematically, it is usually easiest to use a statistical package and a **general linear model** to conduct a two-way ANOVA. We will be using MINITAB in this class, but the data entry and output from the analysis are fairly standard. A standard MINITAB screen consists of two windows, a “worksheet” similar to EXCEL where data is entered and a “session” window where the results of the analysis are shown. We will first consider how to enter data for a two-way ANOVA.

In MINITAB it is necessary to enter each factor as a separate column and the response variable in another. For example, for Problem 15.1 the worksheet would have three columns (Figure 15.1): 1) hormone or no hormone, 2) sex, and 3) calcium levels. The first two are factors and the last, calcium levels, is the response variable. The next step is to choose “ANOVA” from the “Stat” menu and then “General Linear Model.”

The general linear model window is fairly self explanatory (Figure 15.2, below). The columns with appropriate data are listed in the left-hand box, after selecting the appropriate one you chose “Select” and it will be transferred into the place selected. For instance, “Ca level” was entered into the “Response:” box by placing the cursor in the

box and then choosing “Ca level” from the options at the left. In fact, “Ca level” is the only option for the response variable since the other two columns are filled with categorical data. The “Model” box has each factor tested independently as well as the interaction (in this case, “Hormone*Sex”). Once the model is set up as shown above the next step is to simply choose “OK.” The results will then appear in the “Session” window as shown in Figure 15.3.

↓	C1-T	C2-T	C3
	Treatment	Sex	Ca Level
1	NoHormone	Female	16.5
2	NoHormone	Female	18.4
3	NoHormone	Female	12.7
4	NoHormone	Female	14.0
5	NoHormone	Female	12.8
6	Hormone	Female	39.1
7	Hormone	Female	26.2
8	Hormone	Female	21.3
9	Hormone	Female	35.8
10	Hormone	Female	40.2
11	NoHormone	Male	14.5
12	NoHormone	Male	11.0
13	NoHormone	Male	10.8

Figure 15.1. Sample MINITAB worksheet from Problem 15.1.

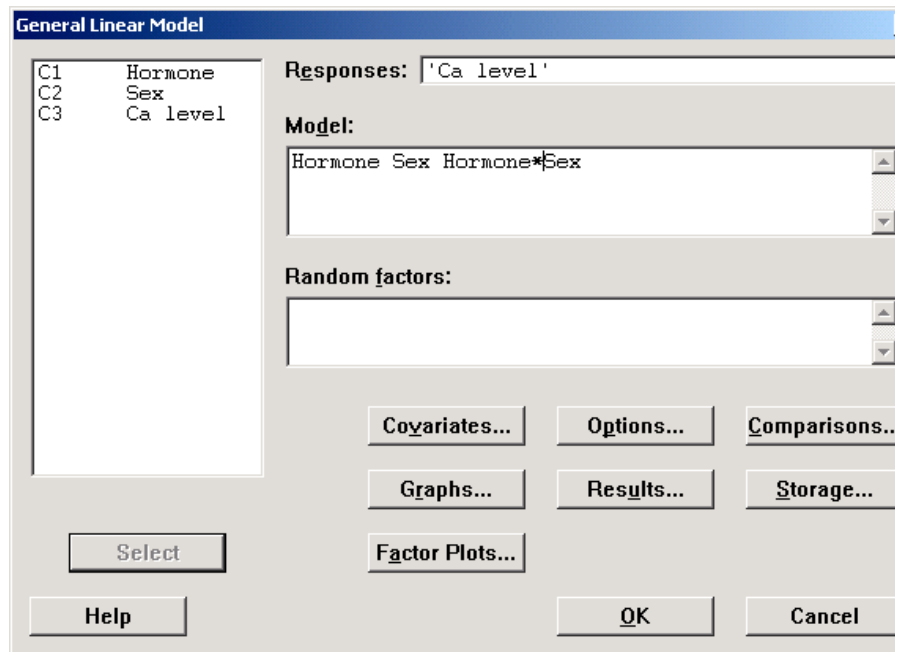


Figure 15.2. Example of general linear model window

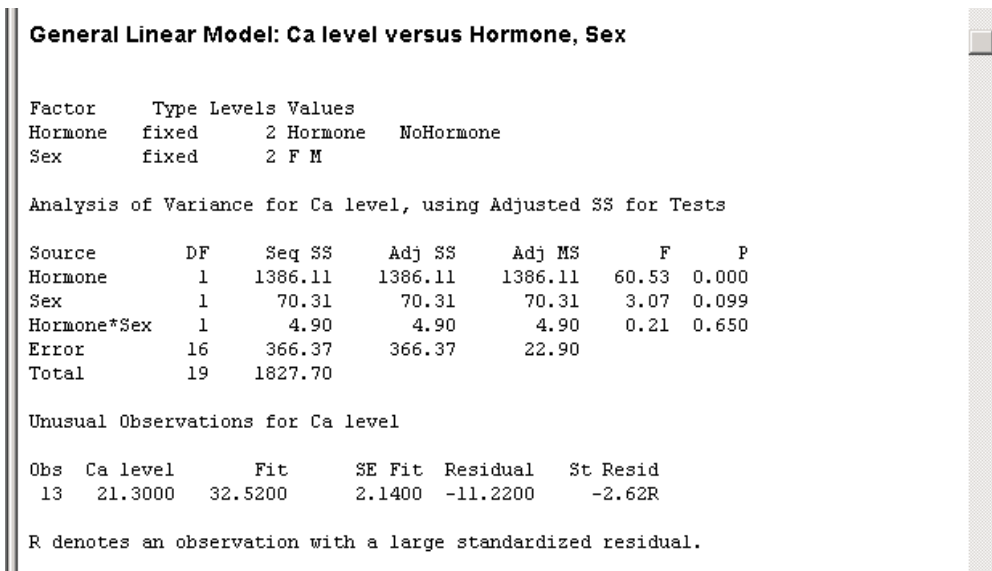


Figure 15.3. MINITAB statistical output for a two-way ANOVA run as a general linear model using data from Problem 15.1.

Example Problem (Problem 15.1)

Investigators wished to know if there was a difference in plasma calcium levels of male and female birds given a hormone treatment. They obtained the data presented in Problem 16.1. The two factors are sex and hormone treatment. Both factors have two levels since there are male and female categories for sex and with and without hormone for hormone treatment. In this 2 x 2 analysis (two levels of each factor) there are five replicates (five birds) for each combination of the two factors; in other words, there are equal sample sizes for each of the experimental conditions.

I. STATISTICAL STEPS

A. Statement of Ho

- Ho: No effect of hormone treatment on the plasma Ca concentration
- Ho: No difference in plasma Ca concentration between male and female birds
- Ho: No interaction between sex and hormone treatment on the plasma Ca concentration

B. Statistical test

Since two factors are involved we need to use a two-way ANOVA, with possibly a multiple comparison post-test if necessary.

C. Computation of test statistics

The Minitab analysis results are shown above, in Figure 15.3.

Ho: No effect of hormone treatment on the plasma Ca concentration

$$F = \frac{MS \text{ Hormone}}{MS \text{ Error}} = \frac{1386.1}{22.9} = 60.5, 1/16 \text{ df}, P < 0.001$$

Ho: No difference in plasma Ca concentration between male and female birds

$$F = \frac{MS_{Sex}}{MS_{Error}} = \frac{70.3}{22.9} = 3.07, 1/16 \text{ df}, P = 0.099$$

Ho: No interaction between sex and hormone treatment on the plasma Ca concentration

$$F = \frac{MS_{Interaction}}{MS_{Error}} = \frac{4.9}{22.9} = 0.21, 1/16 \text{ df}, P = 0.650$$

D. Determine the P of the test statistic

Ho: No effect of hormone treatment on the plasma Ca concentration, $P < 0.05$

Ho: No difference in plasma Ca concentration between male and female birds, $P > 0.05$

Ho: No interaction between sex and hormone treatment on the plasma Ca concentration, $P > 0.05$

E. Statistical Inference

1. The hormone treatment did have an effect on plasma Ca concentration.
2. There was no difference in plasma Ca concentrations of male and female birds

Note of caution – the only reason we can examine the effects of each of the factors is because the interaction term was *not* significant. Thus, it is often best to look at the analysis of the interaction first.

II. BIOLOGICAL INTERPRETATION

Results

The hormone treatment significantly increased plasma calcium concentrations in birds, regardless of sex (Table 1). Calcium levels were on average 55% higher in birds that received the hormone treatment than those that did not (Figure 1).

Table 1. Analysis of variance table for plasma calcium levels in birds of both sexes, half of which received a hormone treatment. Use of an asterisk denotes a significant effect at the 5% significance level.

Source	df	SS	MS	F	P
Hormone Treatment	1	1386.11	1386.11	60.53	< 0.001*
Sex	1	70.31	70.31	3.07	0.099
Interaction	1	4.90	4.90	0.21	0.650
Error	16	366.37	22.90		
Total	19	1827.70			

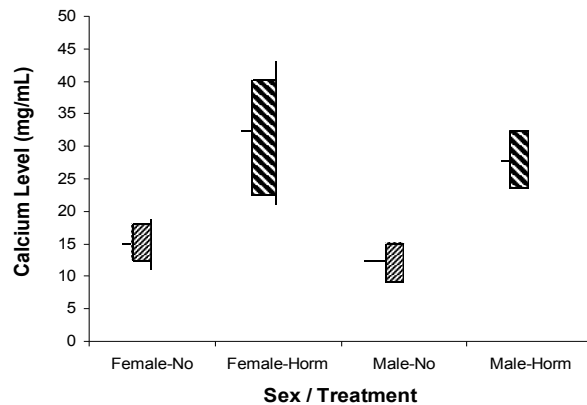


Figure 1. Plasma calcium levels (mg/mL) of male and female birds both with (closed symbols) and without (open symbols) a hormone treatment. Horizontal lines are the means and the boxes are the 95% confidence limits about the mean.

Special Case of Random Factors

In Problem 15.1, as with many biological experiments, the factors are assigned by the investigator and we refer to them as being *fixed*. However there are some situations in which one or both the factors are *random* (e.g. geographic location). In order to determine if a factor is random or fixed, biostatisticians have suggested that one should ask the question, “If the observations were to be made several times, would the biology of the problem force the same choice of levels (e.g. sex) each time, or would any set of levels work?” If the choice of levels is fixed by the problem, the factor is a fixed one; if the choice of levels is arbitrary, it is a random factor.

It is important to determine if the factors are fixed or random because the general linear model two-way ANOVA procedure in MINITAB assumes that they are and calculates the F values based that case (Table 15.1). If the factors are not fixed, but instead are random, then a different procedure is appropriate: a “Balanced ANOVA.” The balanced ANOVA procedure works the same way as the general linear model procedure except that you also specify which of the factors are random, so that the F statistic is calculated appropriately. The model is the same as would be entered for a regular two-way ANOVA. Finally, if sample sizes are unequal, use “General Linear Model,” which is used in the same manner as Balanced ANOVA.

Table 15.1. Computation of the F statistic for tests of significance in a two-factor ANOVA.

Hypothesized Effect	A & B Fixed	A Fixed, B Random	A & B Random
Factor A	<u>Factor A MS</u>	<u>Factor A MS</u>	<u>Factor A MS</u>
	Error MS	A x B MS	A x B MS
Factor B	<u>Factor B MS</u>	<u>Factor B MS</u>	<u>Factor B MS</u>
	Error MS	Error MS	A x B MS
Interaction	<u>A x B MS</u>	<u>A x B MS</u>	<u>A x B MS</u>
	Error MS	Error MS	Error MS

Problem Set - Two-way ANOVA

- 15.1) Investigators wished to know if there was a difference in plasma calcium levels (mg Ca mL⁻¹) of male and female birds given a hormone treatment.

No Hormone		Hormone	
<u>Female</u>	<u>Male</u>	<u>Female</u>	<u>Male</u>
16.5	14.5	39.1	32.0
18.4	11.0	26.2	23.8
12.7	10.8	21.3	28.8
14.0	14.3	35.8	25.0
12.8	10.0	40.2	29.3

- 15.2) Fish of each sex were given one of three different hormone treatments, after which the blood calcium was measured (mg Ca per 100 mL blood). The data are as follows:

Males			Females		
<u>Treatment1</u>	<u>Treatment2</u>	<u>Treatment3</u>	<u>Treatment1</u>	<u>Treatment2</u>	<u>Treatment3</u>
16.87	19.07	32.45	15.86	17.20	30.54
16.18	18.77	28.71	14.92	17.64	32.41
17.12	17.63	34.65	15.63	17.89	28.97
16.83	16.99	28.79	15.24	16.78	28.46
17.19	18.04	24.46	14.8	16.92	29.65

- 15.3) Investigators wished to know if the oxygen consumption rates ($\mu\text{L O}_2 \text{ mg dry weight}^{-1} \text{ min}^{-1}$) varied between two species of limpets, *Acmaea scabra* and *A. digitalis*, at three concentrations of seawater.

100%				75%				50%			
<u><i>A. scabra</i></u>		<u><i>A. digitalis</i></u>		<u><i>A. scabra</i></u>		<u><i>A. digitalis</i></u>		<u><i>A. scabra</i></u>		<u><i>A. digitalis</i></u>	
7.16	8.26	6.14	6.14	5.20	13.20	4.47	4.95	11.11	10.50	9.63	14.50
6.78	14.00	3.86	10.00	5.20	8.39	9.90	6.49	9.74	14.60	6.38	10.20
13.60	16.10	10.40	11.60	7.17	10.40	5.75	5.44	18.80	11.10	13.40	17.70
8.93	9.66	5.49	5.80	6.37	7.18	11.80	9.90	9.74	11.80	14.50	12.30

- 15.4) Investigators wished to know if the number of insects varied between two creeks and at various times of year. They collected the following data:

<u>Month</u>	<u>Shope</u>			<u>Ball</u>		
December	7	9	18	25	10	9
	19	1	15	16	14	28
June	124	51	63	20	44	26
	106	81	83	127	38	52
March	29	24	37	35	18	45
	114	64	49	22	27	29
September	100	87	72	40	45	263
	68	67	9	100	129	115