

Data Set 6: Effects of Host Species and Size on Parasitoid Size

Statistical Setting

This handout illustrates two-factor ANOVA with both factors being fixed effects. Two versions of essentially the same data set are used:

(A) a completely balanced data set, *i.e.* one in which there is equal replication of each treatment combination [this is a sub-set of the actual data set]; and

(B) An unbalanced data set, with unequal replication of treatments but all treatment combinations present [this data set was artificially constructed but mimics the actual treatment effects].

The simpler balanced situation is used to illustrate the analysis of the three overall effects and comparisons among levels of factors; assessment of the validity of the ANOVA model also is briefly discussed for this data set. The second version illustrates the complications produced by unbalanced designs, and how these are dealt with.

At the end of this handout are a detailed description of how to conduct a two-way fixed-effects ANOVA in Minitab, and a sketch of how to do so in SAS.

Background

The data are from my study of larval dynamics of the parasitic wasp *Bracon hebetor* (see Data Set 5), and describe the effect of host type on the size of surviving wasps. The two explanatory variables (factors) are the species and size class of the host. There are two species ("levels" of the species variable), "A" and "B" (actually *Anagasta kuhniella* and *Plodia interpunctella*). Size classes are large, medium and small ("L", "M" and "S"); these are intervals of live weights, and are defined differently for the two species. The response variable is the mean head width (in mm) of all wasps surviving to adulthood on a given host. The initial density was 20 eggs per host.

Part A: Balanced Design

The Data

There are seven replicate observations of each of the six treatment combinations:

A	L	0.5903	0.5833	0.5903	0.5685	0.6354	0.5694	0.5990
	M	0.4514	0.5744	0.5685	0.5972	0.6250	0.5764	0.5714
	S	0.5685	0.5972	0.5208	0.5000	0.5833	0.5451	0.6125
B	L	0.5595	0.5729	0.5792	0.5104	0.6042	0.5655	0.5799
	M	0.5026	0.5833	0.5694	0.5179	0.5500	0.5451	0.5937
	S	0.4757	0.5234	0.5156	0.5792	0.4688	0.4583	0.5156

Preliminary Data Exploration

Descriptive statistics

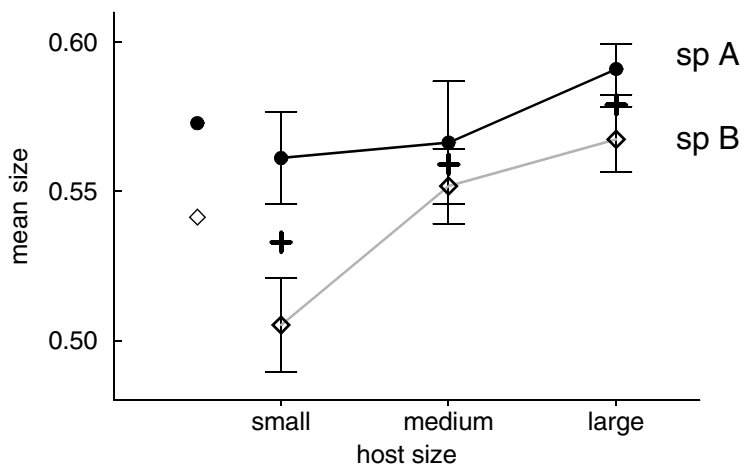
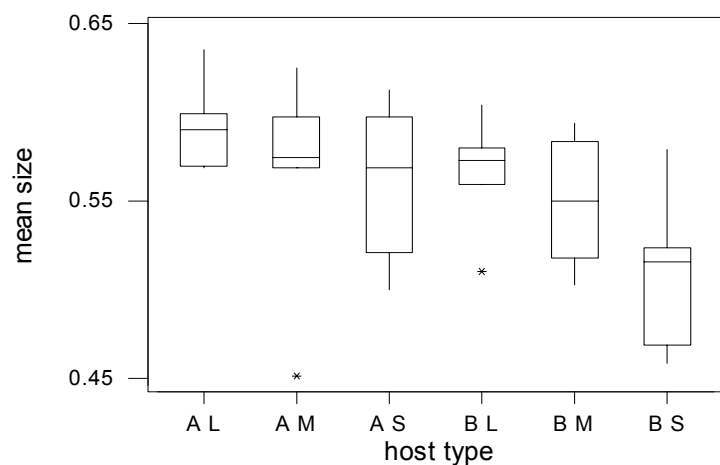
	N	MEAN	MEDIAN	STDEV	MIN	MAX	Q1	Q3	
A	L	7	0.5909	0.5903	0.0226	0.5685	0.6354	0.5694	0.5990
	M	7	0.5663	0.5744	0.0545	0.4514	0.6250	0.5685	0.5972
	S	7	0.5611	0.5685	0.0410	0.5000	0.6125	0.5208	0.5972
B	L	7	0.5674	0.5729	0.0289	0.5104	0.6042	0.5595	0.5799
	M	7	0.5517	0.5500	0.0334	0.5026	0.5937	0.5179	0.5833
	S	7	0.5052	0.5156	0.0417	0.4583	0.5792	0.4688	0.5234

marginal means:	species A	0.5728	size L	0.579
	B	0.5414	M	0.559
			S	0.533

The means and medians show that the sizes of the parasitoids decreased from the large to the small classes of each host species, and also was smaller on host species **B**. The means and medians are quite similar for all but one of the host types, indicating that the distributions are symmetrical. The exception is medium hosts of species **A**, on which the mean is pulled down (relative to the median) by one exceptionally low value, represented by the small minimum for that host type. This outlier also caused the standard deviation to be substantially larger for this host type than for the others.

Plots

These same patterns are seen more easily in the set of boxplots and the interaction plot of means below. Plotting treatment means against host size, by host species, shows the effects quite clearly: size is greater on species **A**, increases with host size, and does so in relatively the same way on the two species:.



The points to the left are the means for the two species, averaging over sizes, while the pluses are the means for the three sizes, averaging over the species.

Analysis

ANOVA Results

Source	DF	SS	MS	F	P
Model	5	0.028472	0.005694	3.86	0.0066
Error	36	0.053064	0.001474		
Total	41	0.081536			
h_spec	1	0.010299	0.010299	6.99	0.012
h_size	2	0.014881	0.007440	5.05	0.012
h_spec*h_size	2	0.003292	0.001646	1.12	0.338
R-Square			Root MSE		
		0.349196			0.0383928

These results provide strong evidence that wasp size is affected by host type, and more specifically that both host species and host size class affect wasp size. These species and size effects appear to be additive: there is little statistical evidence that they interact.

Analysis of Effects

Since we can conclude that there is no interaction between the host species and size factors, it is appropriate to estimate and compare means for the various levels of the two factors.

Host Species Effect

Since there were only two species, little further analysis is needed of their effects: the marginal means were presented earlier, and the ANOVA results above already gave us a test rejecting the hypothesis that the two means are the same. To complete the analysis we can calculate confidence intervals for each mean and for the difference between the means. For the former CIs, with individual $\alpha = 0.05$ (or simultaneous Bonferroni $\alpha = 0.10$), we use

$$s\{\bar{Y}_{i..}\} = \sqrt{\frac{MSE}{bn}} = \sqrt{\frac{0.001474}{3 \cdot 7}} = 0.008378 \quad \text{and} \quad t_{0.975;36}^* = 2.028.$$

The resulting CIs then are:

$$\text{A: } 0.5728 \pm 2.028 \cdot 0.008378 = 0.5728 \pm 0.01699 = (0.5558, 0.5898)$$

$$\text{B: } 0.5414 \pm 0.01699 = (0.5244, 0.5584)$$

The CI for the difference is calculated in the same way, except that the standard error is multiplied by $\sqrt{2}$. The $\alpha = 0.05$ CI for $\mu_{\mathbf{A}} - \mu_{\mathbf{B}}$ thus is

$$(0.5728 - 0.5414) \pm 0.0240 = (0.0073, 0.0553)$$

Host Size Effect

The ANOVA results above indicate that wasp sizes on the three size categories are not all the same, but do not indicate where the differences are. There were no *a priori* hypotheses that would suggest preplanned contrasts more specific than all pairwise comparisons. It was expected that wasps would be smaller on the smaller hosts, so it could be argued that one-sided tests would be legitimate, but because it is possible that the effect could be opposite of that expected, and I would want to detect this, I used two-sided comparisons. To control the overall significance level either Tukey or Bonferroni comparisons can be used; with only three comparisons their results will be very similar:

Tukey's Studentized Range (HSD) Test for variable: MEANSIZE

H_SIZE Comparison		Simultaneous Lower Confidence Limit	Difference Between Means	Simultaneous Upper Confidence Limit	
L	- M	-0.0154	0.0201	0.0556	
L	- S	0.0105	0.0460	0.0815	***
M	- S	-0.0096	0.0259	0.0613	

Comparisons significant at the 0.05 level are indicated by '***'.

These results indicate that only the difference between **Large** and **Small** size classes is significant at a combined $\alpha = 0.05$.

Assessment of "Aptness" of ANOVA Model

One problem applying ANOVA to these data has already been mentioned: the standard deviations differ somewhat more than is desirable. This is largely due to the one outlier in the **A-M** category. For diagnostic purposes I repeated the analysis without that observation. The results were similar to those above, but all *P*-values were smaller because the *MSE* was considerably smaller.

The other important issue to assess is normality. A normal quantile-quantile plot (not shown) of all 42 residuals (deviations from group means) showed the data to be quite close to a normal distribution, except for the one outlier.

I therefore see no reason not to accept the ANOVA analysis as valid.

Conclusions

There is good evidence of additive effects of host species and size on wasp size. Wasps are larger on hosts of species **A** than species **B**. Within a host species, **Large** hosts produce significantly larger wasps than do **Small** hosts. Wasps from **Medium**-sized hosts are intermediate in size, though not statistically different (at overall $\alpha = 0.05$) from those on **Large** or **Small** hosts.

Part B: Unbalanced Design, No Empty Cells

The Data

A	L	(n=3)	0.5990	0.5833	0.5917			
	M	(n=6)	0.5764	0.5536	0.6125	0.5714	0.6250	0.5685
	S	(n=9)	0.5833	0.5451	0.5972	0.5208	0.6125	
			0.5799	0.5000	0.5451	0.5685		
B	L	(n=8)	0.5792	0.5104	0.5595	0.5799	0.5729	
			0.5500	0.6042	0.5655			
	M	(n=5)	0.5833	0.5500	0.5937	0.5179	0.5694	
	S	(n=3)	0.5156	0.4757	0.5156			

Preliminary Data Exploration

Descriptive statistics

	N	MEAN	MEDIAN	STDEV	MIN	MAX	Q1	Q3
A L	3	0.5913	0.5917	0.0079	0.5833	0.5990	0.5833	0.5990
M	6	0.5846	0.5739	0.0278	0.5536	0.6250	0.5648	0.6156
S	9	0.5614	0.5685	0.0366	0.5000	0.6125	0.5329	0.5903
B L	8	0.5652	0.5692	0.0274	0.5104	0.6042	0.5524	0.5797
M	5	0.5629	0.5694	0.0300	0.5179	0.5937	0.5339	0.5885
S	3	0.5023	0.5156	0.0230	0.4757	0.5156	0.4757	0.5156

Treatment means (sample sizes):

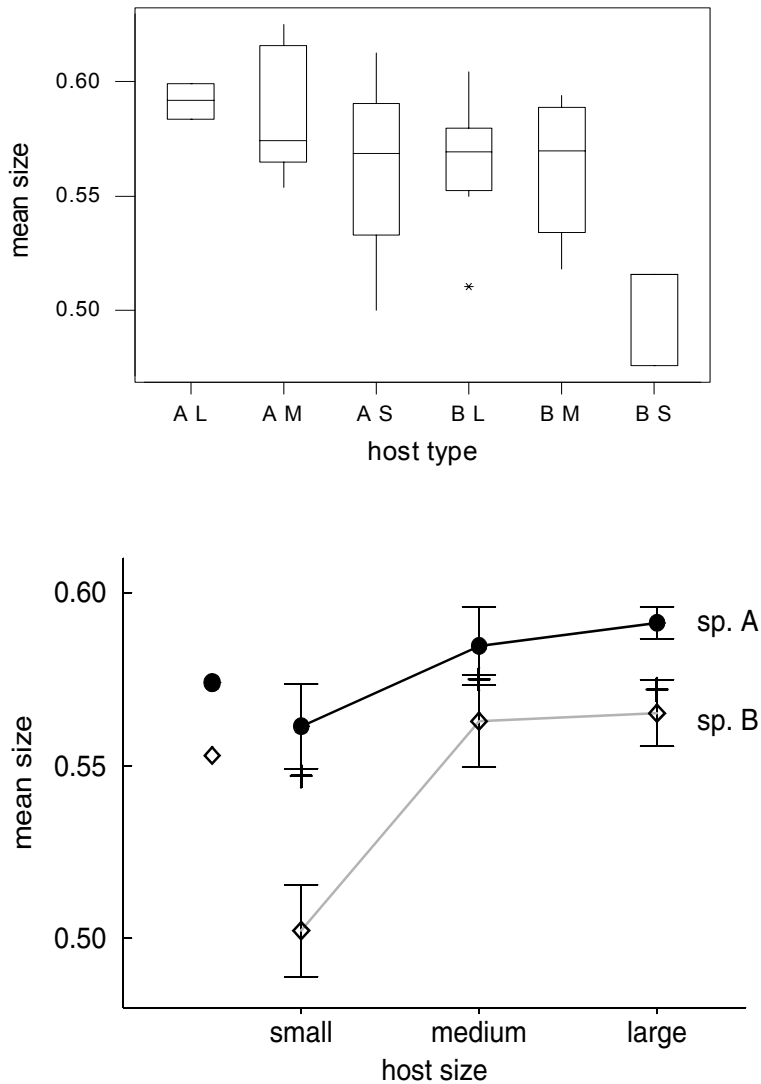
	size:	L	M	S	marg. mean
species: A		0.591 (3)	0.585 (6)	0.561 (9)	0.574 (18)
B		0.565 (8)	0.563 (5)	0.502 (3)	0.553 (16)
marg. mean		0.572 (11)	0.575 (11)	0.547 (12)	

These data (intentionally) show the same basic effects as in the balanced data set analyzed above: wasp size is greater on species-**A** hosts, and decreases with decreasing host size within a host species. (One minor difference is that in these data wasp size on **Medium** hosts of a given species was nearly as large as on **Large** hosts of the same species, while in the balanced data set there was a greater difference between **Large** and **Medium** hosts.)

Because of the unequal replication, however, the simple marginal means do not accurately represent the effects of host type. For instance, because a higher proportion of hosts in the **Large** category than in the **Medium** category are species **B**, which tends to produce smaller wasps, the simple mean for **Large** hosts actually is smaller than for **Medium** hosts. Similarly, since species-**A** hosts are predominantly in the **Small** category while species-**B** hosts are primarily in the **Large** category, the simple means for the species under-represent the effect of host species as seen within particular size categories.

Plots

Both boxplots and a plot of the treatment means shows that the trends are quite similar to those in the balanced data set. The plot of means, however, also shows how the simple means for host species or size categories do not properly represent the effects shown by the treatment means.



The points to the left are the means for the two species, averaging over sizes, while the pluses are the means for the three sizes, averaging over the species.

Analysis

ANOVA Results

Overall model (with interaction):

Source	DF	Sum of Squares	Mean Square	F	P
Model	5	0.01627982	0.00325596	3.70	0.0107
Error	28	0.02462935	0.00087962		
Corr. Total	33	0.04090917			

R-Square	C.V.	Root MSE	MEANSIZE	Mean
0.397950	5.258414	0.0296584		0.5640176

This result shows that there is strong evidence that at least one of the factors (host species or host size-class) affects the mean size of the wasps.

Tests for the three effects:

Because the design is unbalanced, there is some association between the factors, *i.e.* some confounding. Results for a given factor therefore will differ depending on whether they are “adjusted” for other factors. In other words, there are various types of Sums of Squares, adjusted in various ways (or not) for other effects in the model.

Minitab’s SS Types:

“Seq SS” are added-in-order SS and so are generally of little use.

“Adj SS” are added-last SS). The F test and P value use the “Adj SS”, so the tests for the main effects are adjusted for the interaction.

Source	DF	Seq SS	Adj SS	Adj MS	F	P
sp	1	0.0038883	0.0090050	0.0090050	10.24	0.003
size	2	0.0104340	0.0119932	0.0059966	6.82	0.004
sp*size	2	0.0019576	0.0019576	0.0009788	1.11	0.343
Error	28	0.0246294	0.0246294	0.0008796		
Total	33	0.0409092				

To get tests for main effects not adjusted for the (non-significant) interaction, equivalent to SAS’ Type II SS, it is necessary to run the analysis again without the interaction term:

Source	DF	Seq SS	Adj SS	Adj MS	F	P
sp	1	0.0038883	0.0086704	0.0086704	9.78	0.004
size	2	0.0104340	0.0104340	0.0052170	5.89	0.007
Error	30	0.0265870	0.0265870	0.0008862		
Total	33	0.0409092				

SAS’ SS Types:

“Type I SS” are added-in-order tests and therefore of little general use.

“Type II SS” for the main effects each are adjusted for the other main effect but not for the interaction. They therefore are essentially “added-last” tests in a model without the

interaction term (except that the MSE used is that from the full model with the interaction.) They can be used for testing the main effects when the interaction is not significant (as is the case here).

The Type II SS for the interaction is adjusted for both the main effects. It is therefore an added-last test in this model. It can be used to determine the significance of the interaction.

Source	DF	Type II SS	MS	F	P
H_SPEC	1	0.00867042	0.00867042	9.86	0.0040
H_SIZE	2	0.01043396	0.00521698	5.93	0.0071
H_SPEC*H_SIZE	2	0.00195760	0.00097880	1.11	0.3428

“Type III SS” are added-last tests: every term is adjusted for all other terms in the model. When there are no empty cells, “Type IV SS” are identical to Type III.

Source	DF	Type III SS	MS	F	P
H_SPEC	1	0.00900505	0.00900505	10.24	0.0034
H_SIZE	2	0.01199316	0.00599658	6.82	0.0039
H_SPEC*H_SIZE	2	0.00195760	0.00097880	1.11	0.3428

Summary of results:

These analyses indicate that there is quite significant evidence of both host-species and host-size “main effects” and little evidence of an interaction between them.

The confounding of the factors (and the interaction term) is shown by the differences between sequential and adjusted (= Type II–IV) SSES for the main effects, and between the SSES adjusted for the interaction and the other main effect (Type III, IV) or only for the other main effect (Type II). In this case the main effects would be “significant” at $\alpha=0.05$ using any of the SSES, but this will often not be true.

Analysis of Effects

With the interaction not being significant, the two significant main effects would be analyzed similarly to the case for the balanced data set, *e.g.* by multiple comparisons among the levels. Because of the unbalanced design, however, these analyses would not use the simple marginal means given above, but instead would use “least squares means”: the means estimated by the least-squares–fit model.

For the model with the interaction term, the least-squares means for the levels of the main effects are the means of the constituent cell means. These means, with Tukey-Kramer multiple comparisons among the sizes, are as follows:

species	Mean	StDev	Tukey-Kramer <i>P</i> -value for difference between means	
A	0.5791	0.007728		
B	0.5435	0.008021		
size			vs L	vs S
L	0.5783	0.010039		
M	0.5737	0.008980	M	0.9391
S	0.5318	0.009886	S	0.0073

For the model without the interaction, the least-squares means (“fits”) for the cells are constrained to have additive effects of the two factors. The least-squares means for the levels of the main effects then again are means of these cell least-squares means. Since the interaction estimated in the preceding model is slight, the least-squares means for the no-interaction model are not very different from those given above for the interaction model:

species	Mean	StDev	Tukey-Kramer <i>P</i> -value for difference between means	
A	0.5812	0.007369		
B	0.5464	0.007742		
size			vs L	vs S
L	0.5802	0.009325		
M	0.5731	0.008990	M 0.8496	0.0241
S	0.5379	0.009033	S 0.0105	

The Tukey-Kramer multiple comparisons (for either model) indicate that while mean wasp sizes on the **Large** and **Medium** hosts are not significantly different, they are quite significantly larger than those of wasps on the **Small** hosts.

Assessment of “Aptness” of ANOVA Models

The principal concern about the aptness of the ANOVA models for these data is that the standard deviations differ somewhat more than is desirable: the standard deviation for **Large-A** hosts is considerably smaller than the others. It helps, though, that this also is the smallest sample; unequal variances are most problematic when the small samples also are most variable. Given this, and the lack of good alternatives to the ANOVA, I feel this moderate violation of the assumption probably can be accepted.

A normal quantile-quantile plot of all the residuals showed the data were acceptably close to normally distributed. The experimental procedure guaranteed independence of the observations.

I therefore feel the results of the analyses are valid.

Conclusions

As with the balanced data set (**Part A**), this unbalanced data set gives good evidence of additive effects of host species and size on wasp size. Wasps are larger on hosts of species **A** than species **B**. Within a host species, both **Large** and **Medium**-sized hosts produce significantly larger wasps than do **Small** hosts, while wasps from **Large** and **Medium**-sized hosts are not statistically different (at overall $\alpha = 0.05$).

Appendix — Computing Procedures

Minitab

For a balanced two-way ANOVA, three procedures are available: **Two-way...**, **Balanced ANOVA...**, and **General Linear Model...** All three of these procedures will produce the ANOVA table for models either with or without the interaction term, and will list the means for the levels of each factor. Only the **General Linear Model (GLM)** procedure, though, can do multiple comparisons among these means

For unbalanced data sets only the **GLM** procedure can be used.

General Linear Model

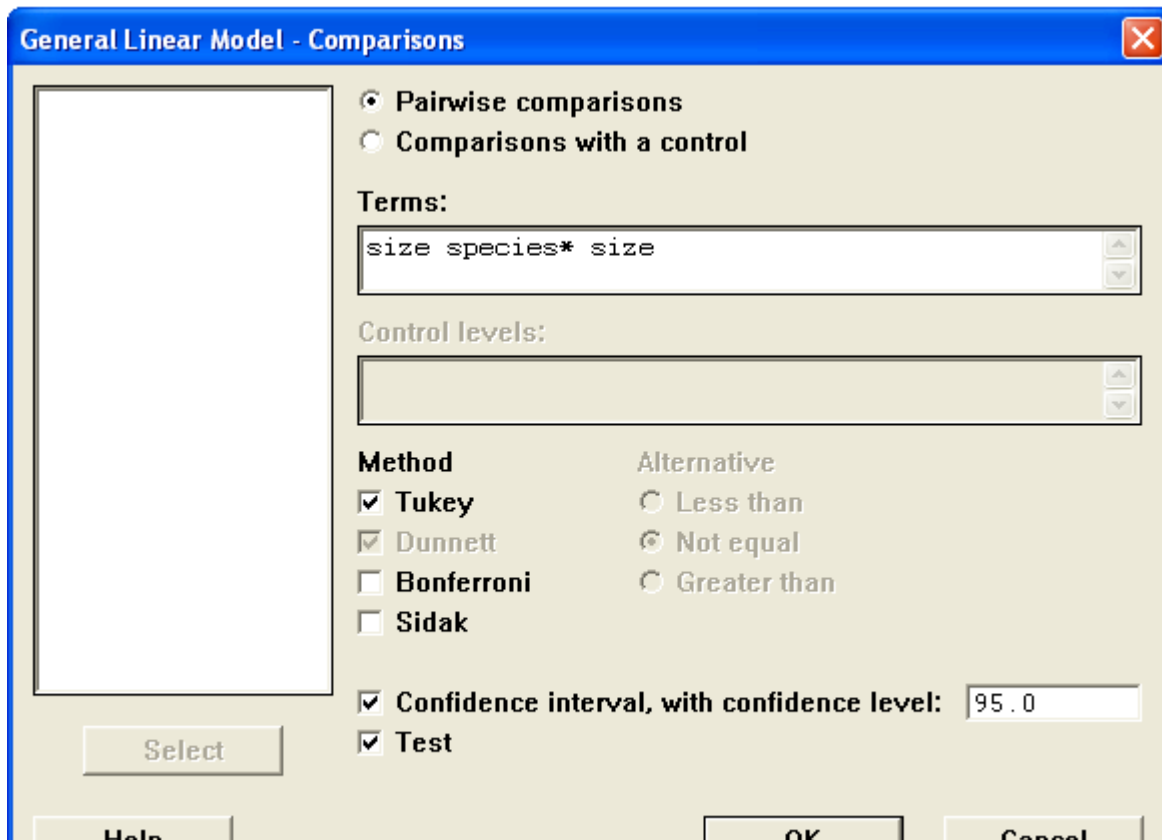
Stat → **ANOVA** → **Linear Models...**

In the main **General Linear Model** window, the response variable goes in the **Responses:** box and the factors and interactions go in the **Model:** box. (Interactions do not need to be pre-calculated, but can be specified as “**X1 * X2.**”) The **Random factors:** box does not need to be used for these analyses.

The **Covariates...** button is not needed (all variables are categorical).

The **Options...** allows you to request weighted least squares and/or Sequential SS, neither of which typically is needed.

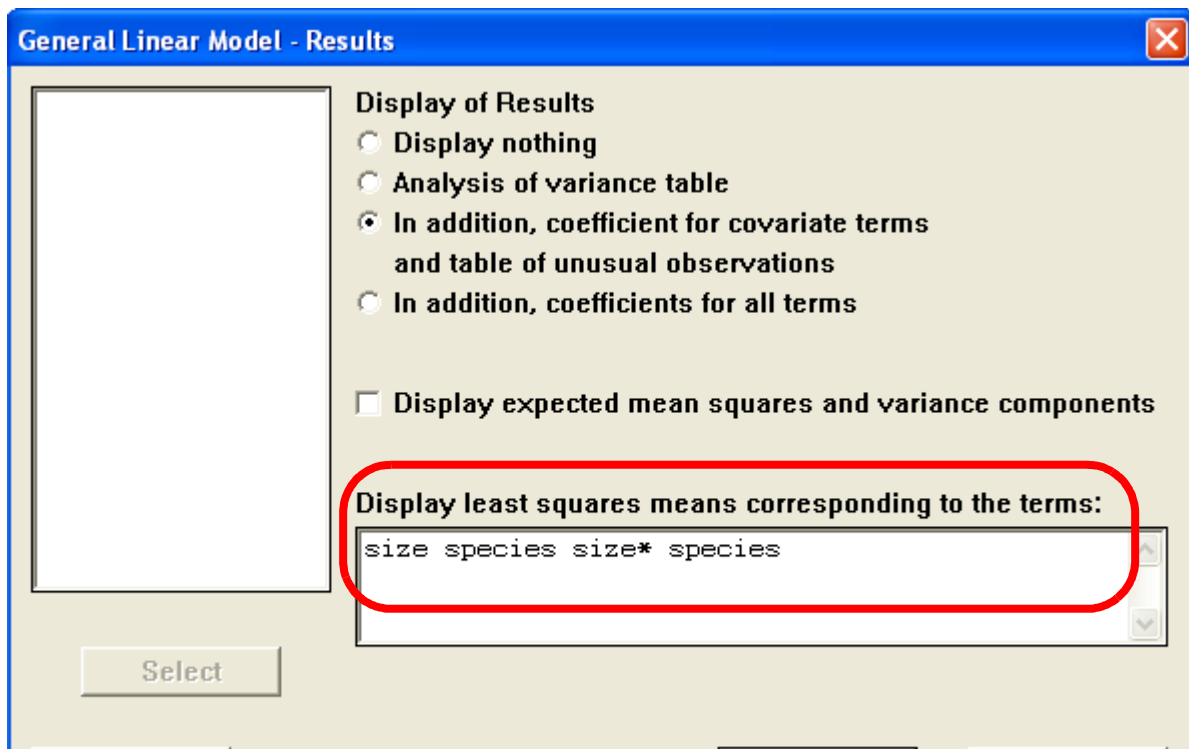
The **Comparisons...** button opens a window allowing you to request multiple comparisons among factor levels or combinations. These comparisons use the least



squares (adjusted) means. The default is to do all **Pairwise comparisons**, but **Comparisons with a control** (Dunnett's comparisons) can be requested instead. The factor(s) and/or interaction(s) for which comparisons are desired are entered in the **Terms:** box. For example, to get comparisons among the three host-size categories in this example data set, **size** would be put in this box; to get comparisons among all six cells (combinations of species and size class), **size * species** would be put in this box. Check boxes allow you to select which type(s) of comparisons are done (default is Tukey), whether to produce CIs and/or *P*-values, and (for comparisons with a control), whether to do one- or two-sided tests.

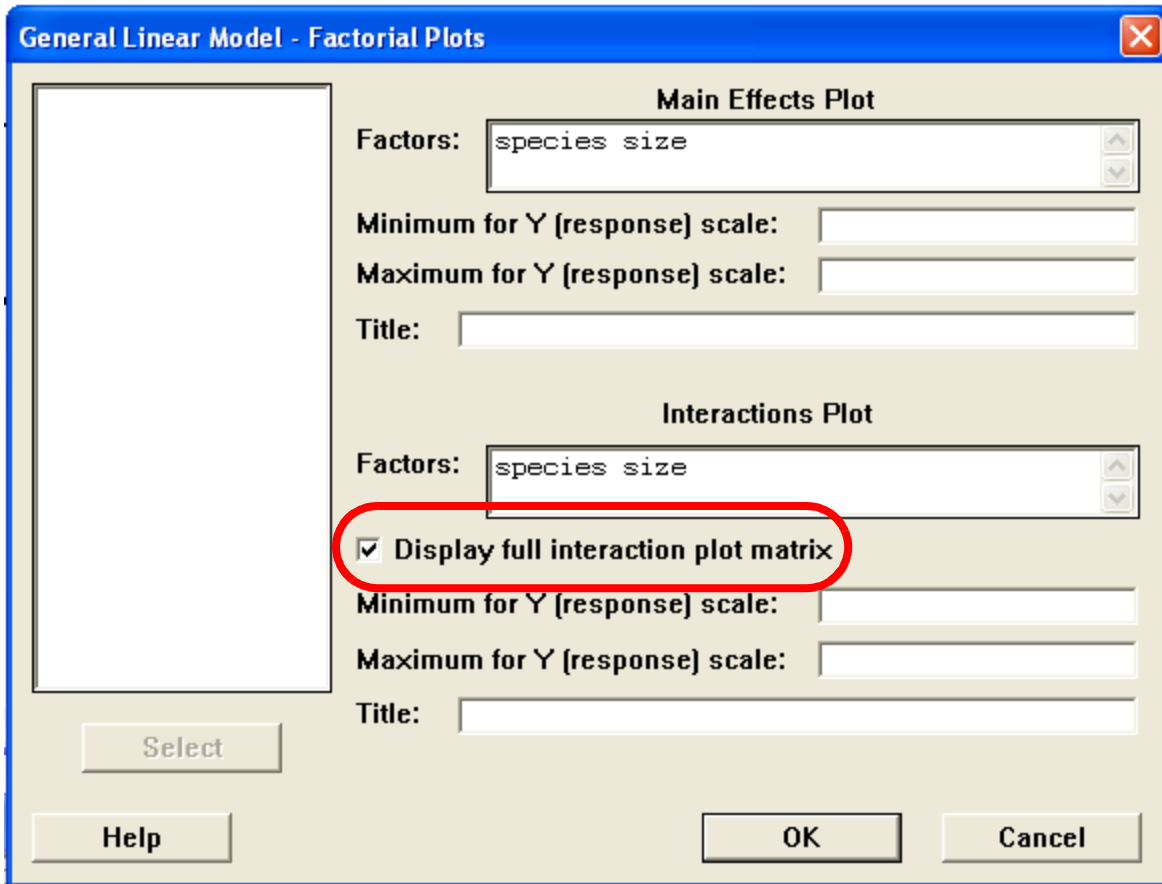
The **Graphs...** button provides the usual residual plots.

The **Results...** button allows you to request that the least-squares (*i.e.* adjusted) means for one or more factors or interactions be included in the output; this is done by entering the appropriate terms in the lower box (**Display least squares means corresponding to the terms:**). As usual, the **Results** window also allows you to choose how much of the standard output is produced.



The **Storage...** button provides the usual diagnostic statistics, fits, *etc.*

The **Factor Plots...** button opens a window (next page) in which you can request two kinds of plots of estimated means. The upper box, **Main Effects Plot**, allows you to request plots of the means of the response variable for the different levels of one or more of the factors in the model. In the example below, plots for both factors are requested. The box in the middle, under the heading **Interactions Plot**, allows you to request plots of the means of the response variable for all the combinations of levels of two (or more) factors. If **Display full interaction plot matrix** is checked, an interaction plot has two panels; in one, the second factor is on the horizontal axis and the first factor defines groups (different



symbols and lines), and in the other panel the factors' roles are reversed (see output below). If **Display full interaction plot matrix** is not checked, only the first of these panels is produced. The options for controlling the ranges of the vertical axes, and assigning titles to the plots, are obvious.

Output:

The first part of the output is the usual ANOVA table, giving both sequential (Type

General Linear Model: mean size versus species, size

Factor	Type	Levels	Values
species	fixed	2	A, B
size	fixed	3	L, M, S

Analysis of Variance for mean size, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
species	1	0.0038883	0.0090051	0.0090051	10.24	0.003
size	2	0.0104340	0.0119932	0.0059966	6.82	0.004
species*size	2	0.0019576	0.0019576	0.0009788	1.11	0.343
Error	28	0.0246293	0.0246293	0.0008796		
Total	33	0.0409092				

S = 0.0296584 R-Sq = 39.80% R-Sq(adj) = 29.04%

I) and adjusted (Type III) SSES for each of the terms in the model. The *F* tests use the adjusted SSES.

If least squares means were requested on the **Results** sub-window, they follow the ANOVA table (see example to right). Their standard errors also are printed.

Least Squares Means for mean size

size	Mean	SE Mean
L	0.5783	0.010039
M	0.5737	0.008980
S	0.5318	0.009886
species		
A	0.5791	0.007728
B	0.5435	0.008021
species*size		
A L	0.5913	0.017123
A M	0.5846	0.012108
A S	0.5614	0.009886
B L	0.5652	0.010486
B M	0.5629	0.013264
B S	0.5023	0.017123

If multiple comparisons were requested (in the **Comparisons** subwindow), they come next in the output. The output for comparisons for main effects (single factors) and for interactions is essentially the same. The partial output below shows the format of the comparisons when confidence intervals are requested.

Tukey 95.0% Simultaneous Confidence Intervals
 Response Variable mean size
 All Pairwise Comparisons among Levels of size
 size = L subtracted from:

size	Lower	Center	Upper	
M	-0.03789	-0.00455	0.02878	-----+-----+-----+----- (-----*-----)
S	-0.08130	-0.04643	-0.01156	(-----*-----) -----+-----+-----+-----
				-0.060 -0.030 0.000

This output gives, for each pair of levels, the difference between the least squares means (in the *Center* column) and the bounds of a 95% simultaneous confidence interval for the difference in population means (in the *Lower* and *Upper* columns). These also are shown graphically to the right. Each line is a comparison. In the example above, the first line is estimating the difference $\mu_M - \mu_L$ and the second line the difference $\mu_S - \mu_L$. A similar table would give results for the comparison of **M** and **S** size classes. Comparisons among

levels of interactions are shown similarly, as below; the first line estimates $\mu_{AM} - \mu_{AL}$, the second lines estimates $\mu_{AS} - \mu_{AL}$, and subsequent lines give results for the remaining 13 comparisons among the six cell means.

All Pairwise Comparisons among Levels of species*size
species = A
size = L subtracted from:

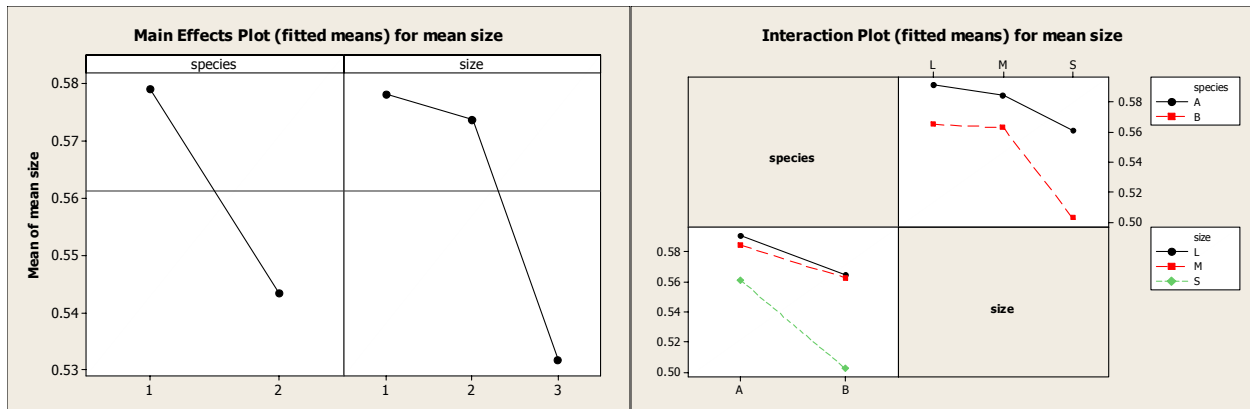
species	size	Lower	Center	Upper
A	M	-0.0708	-0.00677	0.05730
A	S	-0.0904	-0.02996	0.03044

If tests are requested for the comparisons, the output is similar in format except that standard errors, *t* statistics, and adjusted *P*-values are given:

Tukey Simultaneous Tests
Response Variable mean size
All Pairwise Comparisons among Levels of size
size = L subtracted from:

size	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
M	-0.00455	0.01347	-0.338	0.9391
S	-0.04643	0.01409	-3.295	0.0073

Main effects plots (requested in the **Factor Plots** subwindow) plot least square means for levels of the factors, as the left graph below. Interaction plots (right graph) similarly show least square means for the combinations of levels of the variables.



SAS

“Interactive Data Analysis”:

The “Interactive Data Analysis” (“Proc Insight”) can fit the appropriate ANOVA models and provide any of the forms of SS.

“Analyst”:

The “Analyst” can fit the appropriate ANOVA models, provide any of the forms of SS, and in addition can list factor-level means and least-square means, and provide multiple comparisons among them. Either of the following can be used (and are used very similarly): **Factorial ANOVA...** and **Linear Models...**

The first panel (in either **Factorial ANOVA** or **Linear Models**) only specifies the response variable and the factors. The desired model (with or without interaction) is specified after clicking the **Model** button. By default only Type III SS are reported; to get other forms, click the **Statistics** button and specify the desired forms. Factor-level means and multiple comparisons are gotten by clicking the **Means** button; the **Comparisons** tab allows simple (unadjusted) level means to be requested, with a large set of possible methods of controlling the overall significance level, while the **LS Means** tab does the same for the (adjusted) least-squares means.

Program code:

```
proc glm data=advbiom.bracon;
    class species hostsize;
    model waspsize = species hostsize species * hostsize;
    lsmeans species hostsize species*hostsize / pdiff adjust=tukey;
    means species hostsize / tukey;
run;
```

The preceding line invokes the procedure and specifies the dataset to use.

This line specifies which variables are categorical (any variables in the model that are not in the class statement are assumed to be quantitative).

This line specifies the model.

The first part of this line requests the least squares means be output, for the levels of each main effect and for the combinations of levels of both. The second part, after the slash /, specifies options for the **lsmeans** statement, in this case requesting *P*-values for each pairwise difference, adjusted using the Tukey method.

This line requests that simple means for the levels of the factors be output, with Tukey pairwise comparisons.