

Data Set 6-II: Mixed-Model Analysis of the Effects of Host Species and Size on Parasitoid Size

Statistical Setting

This Handout illustrates analyses of the data from my study of host effects on parasitoid larvae (Data Set 6) when the host species is treated as a random factor. Only the ANOVA results and analyses of effects are shown for this mixed model; preliminary data exploration and model diagnostics would be the same as presented in the previous handout.

Analyses of both the balanced and unbalanced versions of the data set are shown, using both the restricted and unrestricted ANOVA models; results of maximum likelihood and restricted maximum likelihood analyses of the unbalanced data also are shown.

Balanced Data

Basic ANOVA Results

From the previous handout (Data Set 6), we have the *SSes*, *MSEs* and *dfs* needed for all the following analyses:

Source	DF	SS	MS
species	1	0.010299	0.010299
size-class	2	0.014881	0.007440
species * size-class	2	0.003292	0.001646
error	36	0.053064	0.001474
Total	41	0.081536	

Restricted Model

Expected mean squares:

species:	$\sigma^2 + 21 \sigma_{sp}^2$
size class	$\sigma^2 + 7 \sigma_{sp \times size}^2 + 14 \sum \alpha_i^2 / 2$
species x size class	$\sigma^2 + 7 \sigma_{sp \times size}^2$
error	σ^2

From these we see that the MSE is the proper denominator for the *F* statistics for both the species effect and the interaction, but that the size-class MS should be divided by the interaction MS.

ANOVA:

Source	DF	SS	MS	F	P
species	1	0.010299	0.010299	6.99	0.012
size-class	2	0.014881	0.007440	4.52	0.181
species * size-class	2	0.003292	0.001646	1.12	0.338
Error	36	0.053064	0.001474		

The results for host species and the species x size-class interaction are identical to those in the fixed-effects ANOVA (original handout): the species effect is significant ($P = 0.012$) and the interaction is nonsignificant ($P = 0.338$).

The result for host size-class, however, is very different from that in the fixed-effects model. Although the F statistic is not very different (since the interaction MS is similar to MSE), it has far fewer denominator df . This causes the test to be nonsignificant ($P = 0.181$). The use of the interaction df in this test is one reason it is preferable to use a large number of levels of the random factor (which will give large df for both that factor and the interaction).

Analysis of Effects:

The estimate of the variance component due to host species is obtained as

$$s_{sp}^2 = \frac{MS \text{ species} - MSE}{nb} = \frac{0.010299 - 0.001474}{7 \cdot 3} = 0.00042.$$

The estimate for the interaction variance component is

$$s_{sp \times size}^2 = \frac{MS \text{ species} \times \text{size} - MSE}{n} = \frac{0.001646 - 0.001474}{7} = 0.00002.$$

Clearly the random variation among replicates — with variance component σ^2 estimated by the MSE of 0.001474 — contributes much more to the total variance in the data set than does either the variance among host-species effects or that among the species—size-class combinations, despite the fairly high statistical significance of the species effect.

Since the host size-class effect is not significant in this analysis, further analysis of effects is probably not appropriate. If it were, multiple comparisons or contrasts among host size-classes could be performed using marginal means (*i.e.* averaging across the two host species). The standard error for any such analysis would use the interaction MS in place of *MSE*.

Unrestricted Model

Expected mean squares:

species:	$\sigma^2 + 7 \sigma_{sp \times size}^2 + 21 \sigma_{sp}^2$
size class	$\sigma^2 + 7 \sigma_{sp \times size}^2 + 14 \Sigma \alpha_i^2 / 2$
species x size class	$\sigma^2 + 7 \sigma_{sp \times size}^2$
error	σ^2

From these we see that the proper denominator for the F statistics for both main effects now is the interaction MS, with the MSE being used only in the denominator of the F for the interaction.

ANOVA:

Source	DF	SS	MS	F	P
species	1	0.010299	0.010299	6.26	0.130
size-class	2	0.014881	0.007440	4.52	0.181
species * size-class	2	0.003292	0.001646	1.12	0.338
Error	36	0.053064	0.001474		

These results for the size-class effect and interaction are identical to those in the restricted model above. The test for the species effect, however, is very different: the F happens to be similar but due to the much smaller df , the P -value is much larger and the null hypothesis would not be rejected.

The estimated variance component for the interaction is the same as in the restricted model above, but that for the species effect is (slightly) different:

$$s_{sp}^2 = \frac{MS \text{ species} - MS \text{ species} \times \text{size}}{nb} = \frac{0.010299 - 0.001646}{7 \cdot 3} = 0.00041$$

Again the variation among replicates (MSE) dwarfs the variance components from the species and species \times size-class terms.

Unbalanced Data

Basic ANOVA Results

From the previous handout (Data Set 6), we have the SS es, MS es and df s needed for all the following analyses. Note that the Type III SS es will be used.

Source	DF	SS	MS
species	1	0.009005	0.009005
size-class	2	0.011993	0.005997
species * size-class	2	0.001958	0.000979
Error	28	0.024629	0.000880

Restricted Model

Since neither Minitab or SAS will perform a restricted-model analysis with unbalanced data, the results shown here for the random term, species, were calculated by hand using the MS s above, and the EMS s below.

Expected mean squares:

species:	$\sigma^2 + 14.179 \sigma_{sp}^2$
size class	$\sigma^2 + 4.7481 \sigma_{sp \times \text{size}}^2 + Q\{\alpha_i\}$
species \times size class	$\sigma^2 + 4.7481 \sigma_{sp \times \text{size}}^2$
error	σ^2

(In the EMS for the fixed-effect factor size class, “ $Q\{\alpha_i\}$ ” represents a “quadratic form,” i.e. the sum of squared α_i s multiplied by a constant determined by the numbers of replicates of the treatment combinations and the number of levels of the other factor.)

The basic structure of these EMS es is similar to that in the balanced case: each EMS contains σ^2 as well as its own effect — $Q\{\alpha_i\}$ or a variance component — and the EMS for the fixed factor (size class) also includes the interaction variance component. The effect of the unequal replication is on the constants associated with each variance component, which no longer are integers.

Since the constant (4.7481) for $\sigma_{\text{sp} \times \text{size}}^2$ is the same in the EMSs for size-class and the species x size-class interaction, the F -statistic for the size-class effect is constructed in the same way as in the balanced case, using the interaction MS as the denominator. The species main effect and the interaction both are tested using MSE as the denominator in the F statistics.

ANOVA:

Source	DF	SS	MS	denom df	F	P
species	1	0.009005	0.009005	28	10.24	0.0034
size-class	2	0.011993	0.005997	2	6.13	0.14
species * size-class	2	0.001958	0.000979	28	1.11	0.34
Error	28	0.024629	0.000880			

Because the interaction and error MS es are so similar for these data, the F -statistics are not greatly different from those in the fixed-effects analyses presented in the earlier handout. The denominator df for the test for the size-class main effect, however, is much smaller than in the fixed-effects model; this causes the test now to be not significant.

Variance components:

Estimation of the variance components is done just as in the balanced case, using the differences between a given MS and MSE :

$$s_{\text{sp}}^2 = \frac{MS_{\text{Sp}} - MSE}{k} = \frac{0.00900505 - 0.0008796}{14.179} = 0.000573$$

$$s_{\text{sp} \times \text{size}}^2 = \frac{MS_{\text{Sp} \times \text{Size}} - MSE}{k} = \frac{0.0009788 - 0.0008796}{4.7481} = 0.00002$$

$$s^2 = MSE = 0.0008796$$

The variance among the replicates — the “error” variance — is considerably larger than the variance component due to host species (though this difference is much less than for the balanced data set), and the variance component due to the species x size-class interaction is much smaller than either of the other variance components.

Unrestricted Model

Expected mean squares:

species:	$\sigma^2 + 4.7265 \sigma_{\text{sp} \times \text{size}}^2 + 14.179 \sigma_{\text{sp}}^2$
size class	$\sigma^2 + 4.7481 \sigma_{\text{sp} \times \text{size}}^2 + Q\{\alpha_i\}$
species x size class	$\sigma^2 + 4.7481 \sigma_{\text{sp} \times \text{size}}^2$
error	σ^2

In this case, as usual with the unrestricted model, the interaction variance component is part of the EMS for the species effect. Note, though, that **the constant for the interaction variance is not the same in the EMS for host species (4.7265) as in the EMS for the size-class x species interaction (4.7481)**. The test for the species effect therefore must use a composite of the interaction and error MS es to construct a “quasi- F ” statistic: the denominator will be

$$0.9955 \text{ MS Sp x Size} + 0.0045 \text{ MSE} = 0.000978351 \quad \text{with } 2.02 \text{ df}$$

As in the restricted model, the size-class main effect can be tested using simply the interaction MS as the denominator of F statistic, and the interaction is tested using the MSE.

ANOVA:

Source	DF	SS	MS	denom df	F	P
species	1	0.009005	0.009005	2.02	9.20	0.093
size-class	2	0.011993	0.005997	2	6.13	0.140
species * size-class	2	0.001958	0.000979	28	1.11	0.343
Error	28	0.024629	0.000880			

Because the interaction and error $MSEs$ are so similar for these data, the F -statistics are not greatly different from those in the fixed-effects analyses presented in the earlier handout. The denominator df for the tests for the two main effects, however, are much smaller than in the fixed-effects model; this causes the results now to be not significant.

Variance components:

Estimation of the variance components for the interaction and error is done just as in the balanced case:

$$s_{\text{sp x size}}^2 = \frac{\text{MS Sp x Size} - \text{MSE}}{k} = \frac{0.0009788 - 0.0008796}{4.7481} = 0.00002$$

$$s^2 = \text{MSE} = 0.0008796$$

The variance component for the host species effect is gotten by subtracting off the denominator used in the “quasi- F ” test and dividing by the appropriate constant:

$$s_{\text{sp}}^2 = \frac{\text{MS Sp} - [0.9955 \cdot \text{MS Sp x Size} + 0.0045 \cdot \text{MSE}]}{k}$$

$$= \frac{0.009005 - 0.0009784}{14.179} = 0.000566$$

As with the restricted model, the variance among the replicates — the “error” variance — is considerably larger than the variance component due to host species and both are much larger than the variance component due to the species x size-class interaction.

Maximum Likelihood and REML Analyses

The following results are from Proc Mixed in SAS; only the parts of the output corresponding to the results of the standard ANOVA analyses above are shown.

Fixed factor:

The fixed factor, host size class, is tested by an F statistic (though it is not obtained as a ratio of mean squares).

		df	error df	F	P
size-class	ML	2	32.8	5.64	0.0079
	REML	2	1	5.23	0.295

Although the F statistics produced by the maximum likelihood (ML) and restricted maximum likelihood (REML) methods are similar, the denominator dfs are very different, resulting in very different P values. The ML method's denominator df are close to those that would be used in a fixed-effects ANOVA, while the REML method's df are similar to — and indeed even smaller than — those of the mixed-model ANOVAs shown above.

The REML method is considered preferable for these sorts of analyses, and it is reassuring that its results are close to those of the (approximate) ANOVA F tests shown above.

Random factors:

The output includes estimates of the variance components for the random factors, as well as their standard errors, and Z tests conducted by dividing each estimate by its standard error and comparing the result to the standard normal distribution.

		var	SE	Z	P
species	ML	0.000235	0.000294	0.80	0.21
	REML	0.000541	0.000861	0.63	0.26
species x size-class	ML	0	.	.	
	REML	0.000014	0.000208	0.07	0.47

These results, especially those from the REML method, are consistent with those of the ANOVAs presented above: both the variance components and the P values are roughly similar in magnitude (though the REML P values are consistently larger). The ML estimates are smaller — that for the interaction apparently actually was negative, and thus set equal to 0 — but the (non-) significance of the tests again is similar.