1(a) (5 points each) true. Each squirrel belongs to only one hibernation group. There are more squirrels than groups, so model 2 is more complex than model 1.

(b) false. $R^2$ is necessarily larger for model 2.

(c) true. 5 time categories will make for 4 adjustments to the baseline time category.

(d) true. Here $\log(n) = \log(40) > 2$ so BIC penalizes complexity more than AIC (this is true whenever $n \geq 8$) and AIC prefers models that are as complex as or more complex than those BIC favors. If BIC already favors the most complex model (model 2), then AIC will favor the same (most complex) model.

(e) false. RSS can be made artificially small, just like $R^2 = 1 - \frac{RSS}{SS\text{total}}$ can be artificially large. For nested models, the RSS of the complex model is necessarily smaller than the RSS of the simpler model, so if we were to select the model with the smallest RSS, we would necessarily choose the most complex model, regardless of the data... which is not appropriate.

(f) false: the interaction terms are the $4 - 1 = 3$ adjustments to the $Y \sim X_1$ slope, one for each of the $X_2$ categories other than the reference category.

2(a) (7 pts) The LRT statistic is $X^2 = 212.2 - 210.0 = 2.2$. The p-value is obtained by comparing 2.2 to the chi-square distribution with 1 df (because $X_3$ is numerical and takes only 1 coefficient).

(b) (5 pts) We accept the null hypothesis that model 1 is correct, meaning that $X_3$ does not influence the response $Y$ (assuming that $X_1$ and $X_2$ may influence $Y$).

(c) (7 pts) For model 1, $AIC = 212.2 + 2 \times 4 = 220.2$. For model 2, $AIC = 210.0 + 2 \times 5 = 220.0$. Therefore, model 2 is favored by AIC because it has a lower AIC value. Note: we may ignore the intercept and the variance parameters throughout all models to compare their AIC values. We would then compare $212.2 + 2 \times 2 = 216.2$ for model 1 to $210.0 + 2 \times 3 = 216.0$ for model 2, and still select model 2. we would find the same AIC difference (0.2) between the two models.

(d) (6 pts) The one or 2 points to the right appear to have a very strong leverage and might have a undue influence on the analysis based on the untransformed $x$ values. Based on the transformed $x$ values, these 2 points no longer seem to have extreme leverage, so their influence should be more acceptable. Note that the residuals (on the $y$ axis) are almost unchanged between the two analyses. Variances seem to be equally homogeneous in both graphs. The one lonely point to the right (with largest leverage) is not enough evidence of a decreased variability at larger predicted values.

3(a) (7 pts) For more information on the study, see Mauck, De Moraes & Mescher (2010) PNAS.

(b) (10 pts) $\exp(4.939 + 2.494 + 0.257 - 0.064) = 2051$ aphids (double-check with the graph!).

(c) (7 pts) 4.9391 is the mean log-number of aphids at 7 days on CMV-infected plants, corresponding to $\exp(4.9391) = 139$ aphids at 7 days on infected plants. (again, double-check with the graph!)

(d) (7 pts) 0.2252 is the difference in mean log-number of aphids between untouched and CMV-infected plants. It means that untouched plants have about 22% more aphids than infected plants. The exact predicted increase is by 25% (rather than 22%) because $\exp(0.2252) = 1.25$

(e) (7 pts) Model A2 assumes that the effect of day on the log-number of aphids is the same across all treatments (i.e. no interaction). Or, equivalently, model A2 assumes that the effects of plant treatments are the same on day 7 and on day 14.

(f) (7 pts) No evidence of interaction: we apply an f-test. The test statistic is $f = 0.05$, which is compared to an F distribution on df = 2 and 66. The associated p-value is 0.95, indicating a lack of interaction.

1www.pnas.org/content/early/2010/02/02/0907191107
Summary of grades: