A0. In this paper, the authors attempt to measure whether raising two sets of flies on different feed for several generations results in the two groups being less likely to mate with each other. To test this, the researchers performed two experiments. For both experiments, the distinct populations were raised on the respective feed for a number of generations. Then, a new generation from each population were both raised on one feed. Finally, a number of different wells were filled with one male and one female fly from the final generation of each population. The mating in each well was observed, and the number of pairings was tallied by male and female population. In the first experiment, these counts were summed across all the wells for one generation. In the second experiment, this process was repeated after different numbers of generations had been raised on the two feed, and the results were summed across wells and generations.

To analyze this statistics, the researchers assumed that a mating represented an independent Bernoulli random draw of whether the mating occurred between flies of the same population. Thus, for n_h, the number of matings with both members from the same population, and p, the probability of homogamic mating given mating, $n_h \sim Bin(n, p)$. They transformed this into a statistic

$$SII = \frac{n_h - (n - n_h)}{n} = \frac{2n_h}{n} - 1$$

Then, they calculated its standard error based on the binomial assumption:

$$
\begin{aligned}
\mathrm{SD}(SII) &= \mathrm{SD}(\frac{2n_h}{n}) \\
&= \sqrt{\frac{2}{n}\frac{n_h}{n}\left(1 - \frac{n_h}{n}\right)} \\
&= \sqrt{\frac{1}{n}\left(1 - \left(\frac{n_h - (n - n_h)}{n}\right)^2\right)} \\
&= \sqrt{\frac{1}{n}\left(1 - SII^2\right)}
\end{aligned}
$$

For the first experiment, with $n_h = 29, n = 38$ they calculated an SII of .53 with a standard error of .14. Also, they attempted to calculate a p-value as $P(n_h \geq 29 \mid p = .5)$, which should have been .00083, but which they reported as .0012. In the second experiment, $n = 900, n_h = 571$, yielding a SII of .27, a SE of .02, and a p-value less than .0001. They accordingly concluded that the different diets reduced inter-mating.

The biggest issue with this analysis was the binomial assumption. The fact that there was only two males and two females in each well prevented independence; while one pair of flies is mating, the other two can only possibly be mating with each other. This means that two homogamic matings or two heterogamic matings would be more common than one of each in a well (indeed, depending on how they counted mating, which was not clear, it may be impossible). This means that their confidence intervals and p-values are invalid.

Other set of potential issues stem from the pooling of data across generation in the second experiment. Its questionable whether SII is constant across the generations. When I fit a quadratic regression by least squares, our p-value from an F test is .1534, but as you see below, it does appear to be a better fit than just a constant.

SII by Generation With Quadratic Fit
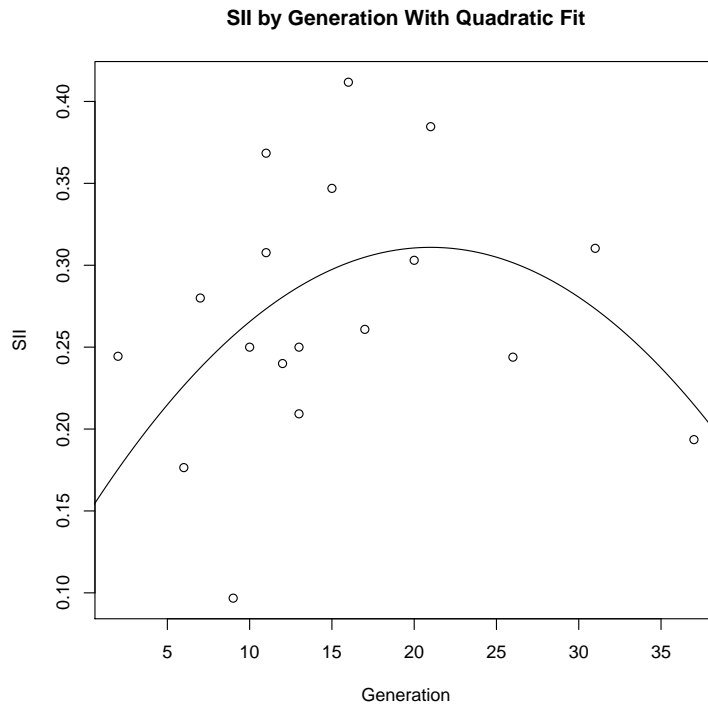
Thankfully, since flies which were placed in wells to mate were permanently removed from the population, there is little concern of dependence among mating between generation.

The one remaining issue then is whether the last three sets of tests in the second experiment, which were from a parallel pair of populations, is comparable to the rest of the data set. If we look at the number of matings per well, the answer seems to be no. As one would expect, the total number of matings grows linearly with wells. However, the last three test groups do not seem to fit the pattern from the first 15:

Matings vs. Wells
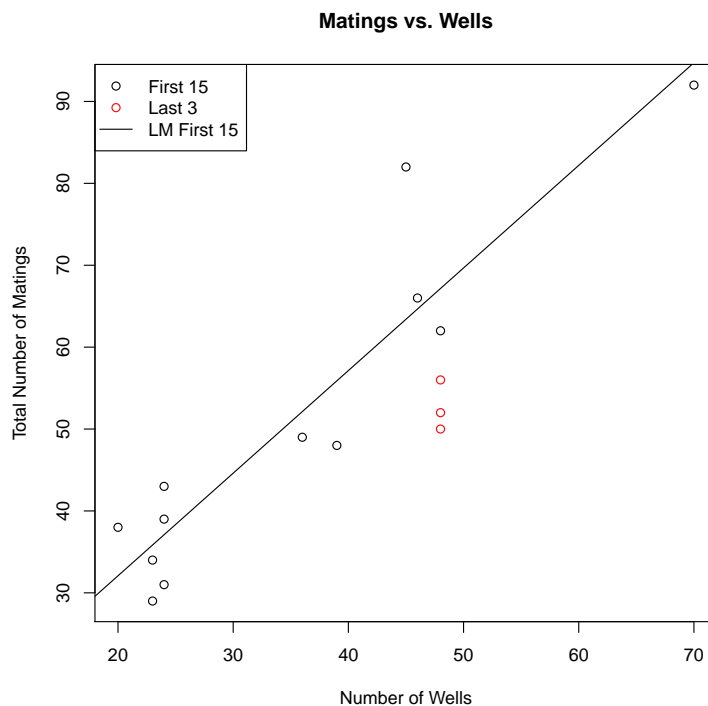
When we test for the probability of seeing the mean of the three points fall so far away from the prediction under the linear model, we get a p-value of .013, which is sufficient to reject that they follow the same

2

distribution. (I used the usual formula for prediction distribution of a linear model taking into account the smaller standard error for the mean of three observations).

A3. Fitting a logistic model to the data, we get this model:

	Estimate	Std. Error	z value	Pr($>$|z|)
(Intercept)	-2.5411	0.7719	-3.29	0.0010
dlog2	1.0121	0.2626	3.85	0.0001

Superimposed over the data, it seems to be a very good fit:

Mortality by Log Dosage
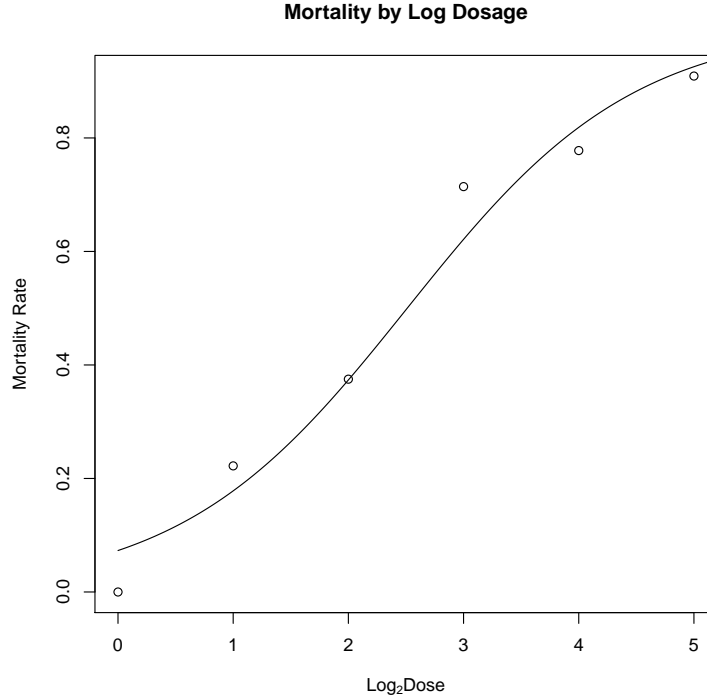
We can obtain an estimate of $\log_2 LD_{50}$ by solving this equation:

$$.5 = (1 + e^{-1.01x + 2.54})^{-1}$$
$$0 = -1.01x + 2.54$$
$$x = 2.51$$

Where $\hat{\alpha}$ is in the intercept and $\hat{\beta}$ is the coefficient on x, we want a confidence interval on $-\frac{\hat{\alpha}}{\hat{\beta}}$, which will set the linear predictor to 0 and yield $\log_2 LD_{50}$. Thus, we can apply Fieller's method.

$$1.96^2 > \frac{(\hat{\alpha} - x\hat{\beta})^2}{\hat{\sigma}_\alpha^2 - 2x\hat{\sigma}_{\alpha,\beta} + x^2\sigma_\beta^2}$$

$$U, L = \frac{(\hat{\alpha}\hat{\beta} - 1.96^2\hat{\sigma}_{\alpha,\beta}) \pm \sqrt{(\hat{\alpha}\hat{\beta} - 1.96^2\hat{\sigma}_{\alpha,\beta})^2 - (\hat{\beta}^2 - 1.96^2\sigma_{\hat{\beta}})(\hat{\alpha}^2 - 1.96^2\sigma_{\hat{\alpha}})}}{\hat{\beta}^2 - 1.96^2\sigma_{\hat{\alpha}}}$$

$$U = -1.66, L = -3.31$$

Since this is the confidence interval on $\frac{\hat{\alpha}}{\hat{\beta}}$ while we want $-\frac{\hat{\alpha}}{\hat{\beta}}$, our confidence interval for $\log_2 LD_{50}$ is $(1.66, 3.31)$. Its worth noting that this ignores the uncertainty in the covariance matrix.

Fitting a model under the constraint that $\log_2 LD_{50} = 4$ means that the linear predictor must be 0 at x. Thus, we fit the model $E(y) = logit(\beta(x - 4))$. This yields the model

	Estimate	Std. Error	z value	Pr($>$|z|)
I(dlog2 - 4)	0.5412	0.1776	3.05	0.0023

Calculating the likelihood ratio statistic for this model as compared to the original model, we get a value of 11.59, which corresponds to a p-value of .00066 assuming a χ_1^2 distribution. Extending this concept, we might calculate a confidence interval set based on the .95 percentile of the χ_1^2 distribution, including only values of $\log_2 LD_{50}$ whose likelihood ratio statistic would fall below it. Indeed, this is plotted below, with the dotted line representing the cutoff:
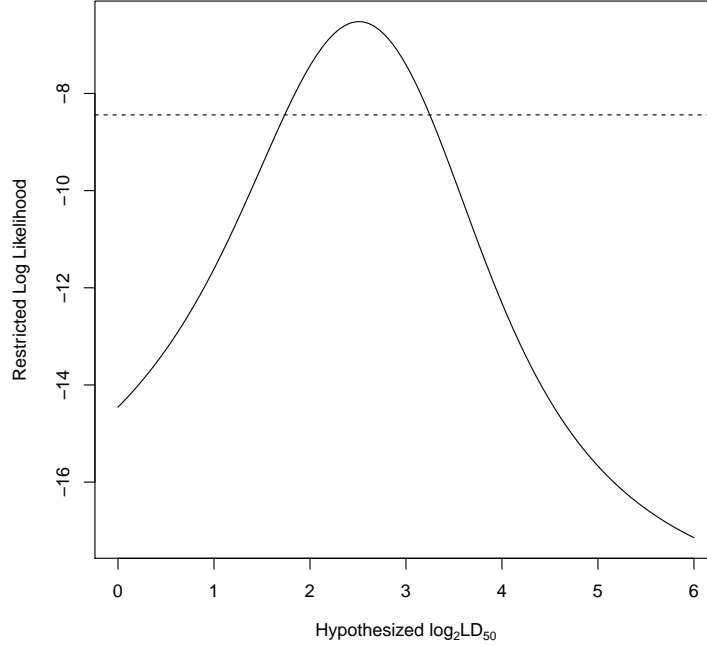
The region which has log likelihoods above the cut off corresponds to $\log_2 LD_{50} \in (1.75, 3.21)$. This is pretty close to the confidence interval we calculated earlier.