Homework1

Conceptual Exercises

1. Suppose we plan to use data to estimate one parameter, p_B.

- When using a likelihood to obtain an estimate for the parameter, which is preferred a large or a small likelihood value? Why?
  - A relatively large likelihood value, because under the fit model our probability of seeing the data we did in fact see will be larger.
- The height of a likelihood curve is the probability of the data for the given parameter. The horizontal axis represents different possible parameter values. Does the area under the likelihood curve for an interval from .25 to .75 equal the probability that the true probability of a boy is between 0.25 and 0.75?
  - .

2. Suppose the families with an "only child" were excluded for the Sex Conditional Model. How might the estimates for the three parameters be affected? Would it still be possible to perform a Likelihood Ratio Test to compare the Sex Unconditional and Sex Conditional Models? Why or why not?

Guided Exercises

2. In Case 1 we used hypothetical data with 30 boys and 20 girls. Case 2 was a much larger study with 600 boys and 400 girls. Consider Case 3, a hypothetical data set with 6000 boys and 4000 girls.

- Use the methods for Case 1 and Case 2 and determine the MLE for p_B for the independence model. Compare your result to the MLEs for Cases 1 and 2.
  - a
- Describe how the graph of the log-likelihood for Case 3 would compare to the log-likelihood graphs for Cases 1 and 2.
  - .
- Compute the log-likelihood for Case 3. Why is it incorrect to perform an LRT comparing Cases 1, 2, and 3?
  - .

3. Write out an expression for the likelihood of seeing our data (5,416 boys and 5,256 girls) if the true probability of a boy is:

a. $p_B = 0.5$

   answer

b. $p_B = 0.45$

   answer2

c. $p_B = 0.55$

   answer3

d. $p_B = 0.5075$

   answer4