Exercise 3: Sieve Analysis- If and how does VE_S depend on characteristics of HIV?

Due Tuesday February 21

The data are available at http://faculty.washington.edu/peterg/Vaccine2006.html

Consider variable 24. in the mock efficacy trial dataset, the HIV genetic distance 1. Most subjects who became HIV infected during the trial have a value for this variable-the genetic distance (in the neutralizing face core region of gp120) of their infecting virus to the GNE8 virus represented in the tested vaccine.

1. Come up with a reasonable way to discretize infecting viruses into K ordered categories, based on the HIV genetic distance 1. The number of categories K should be at least 3 and not so large that there are too few data points in categories to make stable inferences.

2. Based on the Lecture 7 Sieve Analysis slides, do some analyses to evaluate if and how vaccine efficacy to prevent infection varies with the K genotype categories. Several approaches are described in Lecture 7; select a small subset of these approaches to apply to the data. Address the following questions:

   - Is there evidence for differential vaccine efficacy against the different HIV genotypes? [Overall test]

   - Based on point estimates, 95% confidence intervals, and a plot, how does the vaccine efficacy seem to vary across the genotype category?