

HW 3 for Stat 7249 - Spring 2009

Due February 17

Reading in text for this assignment

- Chapter 7

1. Show that the first four cumulants of $Z = \frac{Y - m\pi}{\sqrt{m\pi(1-\pi)}}$, where $Y \sim \text{Bin}(m, \pi)$ are 0, 1, $O(m^{-1/2})$, and $O(m^{-1})$, respectively. This implies that for fixed π , as $m \rightarrow \infty$, the cumulants of Z approach those of a standard normal random variable (so converge in distribution to a standard normal). Also, derive the first four cumulants of a standard normal random variable to verify they are 0, 1, 0, and 0, respectively.
2. Variance stabilizing transformations: Suppose that $Y \sim \text{Bin}(m, \pi)$ with m large. By expanding in a Taylor series, show that the random variable, $Z = \arcsin\{(Y/m)^{1/2}\}$ has approximate first two moments

$$E[Z] \approx \arcsin(\pi^{1/2}) - \frac{1-2\pi}{8\sqrt{m(1-\pi)}} \quad (1)$$

$$\text{Var}[Z] \approx (4m)^{-1}. \quad (2)$$

Comment on the form of both moments.

3. Derive the log likelihood and score equations for a beta-binomial regression using the notation from class.
4. The National Institute of Diabetes and Digestive and Kidney Diseases conducted a study on 768 adult female Pima Indians living near Phoenix. The purpose of the study was to investigate factors related to diabetes. The data can be found in the class directory, pima.txt. The response variable is 'test' (whether tested positive for diabetes or not). Find a good model for predicting diabetes. Which covariates are included? Interpret your results. Does your final model fit well? Give some evidence for or against. Compute a confidence interval for the predicted probability of a positive test in your final model for an individual with the average value of the included covariates.