International Conference on Advances in Interdisciplinary Statistics and Combinatorics, October 12–14, 2007, University of North Carolina at Greensboro, Greensboro, North Carolina, USA

Coauthors: Elizabeth Slate, Debajyoti Sinha and Dipak Dey

# BAYESIAN ANALYSIS OF ZERO-INFLATED COUNT DATA WITH APPLICATIONS TO DENTAL CARIES

DIPANKAR BANDYOPADHYAY

The Gullah-speaking inhabitants of the Sea Islands of South Carolina are a unique population because of their minimal Caucasian genetic admixture and high propensity for diabetes. A clinical study was conducted to determine their dental health status of Type-2 diabetic Gullah African Americans. Dental caries was assessed using the total number of decayed, missing and filled surfaces, an index known as DMFS in the dental literature. Data resulted from examining 4 (for canines and incisors) or 5 (for premolars and molars) surfaces per tooth, for all (up to 32) teeth, for over 260 individuals. Also recorded were covariates including age, gender, smoking and brushing/flossing habits, etc., which may influence caries development. We model the tooth-level contributions to DMFS, which range from 0 to 5, and evaluate associations with covariates. Histograms suggest a zero-inflated binomial model for the tooth-level counts. As in a Hurdle Model, the process determining a healthy tooth (with a count of 0) is treated as a structural zero and hence separated from the remaining counts (1 to 5), which are modeled using a zero-truncated binomial distribution. We develop a multivariate model where covariates enter through a random effects logistic regression on the logit of the probability of a carious surface. To preserve marginal logit structure for interpretability, we use a bridge density for the subject-specific random effects. The tooth-specific zero-inflation probability is modeled as arising from a beta distribution whose shape/scale parameters are linked to the odds of a healthy tooth. We compare our model with alternatives to assess improvements in fit, prediction and interpretability.

Dept. of Biostatistics, Bioinformatics and Epidemiology, Medical University of South Carolina, Charleston, SC, 29425

*E-mail address*: bandyopd@musc.edu