Computational Statistics

The attached data consist of 14 variables on 390 subjects who were interviewed in a study to understand the prevalence of obesity, diabetes, and other cardiovascular risk factors in central Virginia for African Americans. One of the objectives of the study of these data was to investigate the relationship between the glycosolated hemoglobin and the associated set of the observed risk factors. Notice that i) glycosolated hemoglobin (outcome) > 7.0 is usually taken as a positive diagnosis of diabetes, ii) gamma or other positive variable distribution may be appropriate for the regression model, iii) if there is nothing in a cell below, that means NA (not available) observation.

Variable name	Description
ID	Subject identification
CHOL	Total cholesterol
SGLU	Stabilized glucose
HDL	High density lipoprotein
GHB	Glycosolated hemoglobin
LOCATION	Buckingham, Louisa
AGE	(years)
GENDER	(male, female)
HHT	Height (inches)
WHT	Weight (pounds)
FRAME	(large, medium, small)
SBP	First systolic blood pressure
DSP	First diastolic blood pressure
W	Waist (inches)
H	Hip (inches)

Perform a thorough statistical analysis of the data, in particular using MCMC methods to select a suitable gamma model and interpret the associated results concerning the objectives of the underlying study.

References: Schorling JB, Roach J, Siegel M, Baturka N, Hunt DE, Guterbock TM, Stewart HL (1997). A trial of church-based smoking cessation interventions for rural African Americans. Preventive Medicine, 26, 92-101. Willems JP, Saunders JT, DE Hunt, JB Schorling (1997). Prevalence of coronary heart disease risk factors among rural blacks: A community-based study. Southern Medical Journal, 90, 814-820.