

Lab assignment 12

The attached data set contains information on 35 young automobile drivers. Recorded for each individual is the response, y_i , which is 1 if the individual has been involved in 1 or more automobile accidents in the last 6 months, 0 otherwise; the driver's age in years; whether or not the driver has taken an extensive driver's education course (1 for yes); and how the driver responded when asked if s/he enjoys the discipline of statistics (1 for yes).

- a) Using multivariate normal prior with mean β_0 and covariance Σ_0 (report your choice), perform Bayesian probit regression on these data. Report posterior means and 95% intervals of parameters and discuss about the effect of each predictor on the response.
- b) Based on this data set, whom would you choose to park your new Mercedes-Benz: a 17-year-old who has done a driver's education course and hates statistics, or an 18-year old who has done a driving course and loves statistics? Use the probit output. Assume that the probability of crashing while valet parking is proportional to the probability of having crashed in the past 6 months.
- c) Generate simulation data; $s_{1i} \sim N(0, 1)$, $s_{2i} \sim N(0, 1.5)$, $s_{3i} \sim N(0, 2)$ independently for $i = 1, \dots, 35$. Let $x_i = (1, \text{age}_i, \text{course}_i, \text{likeStats}_i, s_{1i}, s_{2i}, s_{3i})$. Assuming you don't know s_{1i} , s_{2i} and s_{3i} have no effect on y_i , perform Bayesian probit regression with $\{y_i, x_i\}$ using stochastic search variable selection (SSVS) approach (show your prior setting). Report posterior means, 95% intervals of parameters and inclusion probabilities of all predictors and discuss the result.
- d) (optional:+1point) Derive \hat{p}_j in the slide 58 in lecture 18.