Problem Set 5

Fitting of multiple logistic regression

Question: You are provided with data on birth weight status of new born children. A child is classified as low birth weight (lbw) if she weighs less than 2400 gms. Here lbw=1 implies a low birth weight child and 0 otherwise. Along with the birth weight status, you are also provided with information on mothers race, smoking status of mother and age of mother at child birth.

(i) Using the information on mother as covariates fit a multiple logistic regression to the given data.

(ii) Plot the fitted (predicted) probabilities against distinct age values.

(iii) Classify the predicted probabilities into two groups according to the smoking status of the mother. Compute the mean and dispersion of each group and comment.

(iv) Obtain the mean of the predicted probabilities. Using the mean value as the cut point obtain the predicted lbw status of a child. Prepare a confusion matrix and obtain the probability of misclassification.

(v) Obtain the median of the predicted probabilities. Using the median value as the cut point obtain the predicted lbw status of a child. Prepare a confusion matrix and obtain the probability of misclassification.

(vi) Compare the results in (iv) and (v) and comment.