# Homework 1:
## Probability Review and Priors

4. (15) You have started an online business selling books that are of interest to your customers. A publisher has just given you a large book with photos from famous $20^{th}$ century photographers. You think this book will appeal to people who have bought art books, history books and coffee table books. In an initial offering of the new book you collect data on purchases of the new book and combine these data with data from the past purchases (see ArtHistBooks.csv).

   Use Bayesian analysis to give the posterior probabilities for purchases of art books, history books and coffee table books, as well as, the separate probabilities for purchases of the new book given each possible combination of prior purchases of art books, history books and coffee table books. Do this by first using beta priors with values of the hyperparameters that represent lack of prior information. Then compute these probabilities again with beta priors that show strong weighting for low likelihood of a book purchase. Compare your results.

5. (15) The data set CHDdata.csv contains cases of coronary heart disease (CHD) and variables associated with the patient's condition: systolic blood pressure, yearly tobacco use (in kg), low density lipoprotein (ldl), adiposity, family history (0 or 1), type A personality score (typea), obesity (body mass index), alcohol use, age, and the diagnosis of CHD (0 or 1).

   Perform a Bayesian analysis of these data that finds the posterior marginal probability distributions for the means for the data of patients with and without CHD. You should first standard scale (subtract the mean and divide by the standard deviation) all the numeric variables (remove family history and do not scale CHD). Then separate the data into two sets, one for patients with CIID and one for patients without CIID.

   Your priors for both groups should assume means of 0 for all variables and a correlation of 0 between all pairs of variables. You should assume all variances for the variables are 1. Use a prior alpha equal to one plus the number of predictor variables. Compute and compare the Bayesian estimates for the posterior means for each group.

   For 5 extra credit points, compute the probability of observing a point at least as extreme as the posterior mean of patients without coronary heart disease under the posterior distribution for the patients with coronary heart disease. Then compute the probability of observing a point at least as extreme as the posterior mean of patients with coronary heart disease under the posterior distribution for the patients without coronary heart disease.