Economics 696F, Homework 5 Solutions

1. The likelihood function in the probit model can be written as

$$f(y|x,\alpha,\beta) = \prod_{i=1}^{n} \Phi(\alpha+\beta x_i)^{y_i}(1-\Phi(\alpha+\beta x_i)^{1-y_i}.$$

   Given the discretization

$$\alpha \in \{-10.0, -9.9, \ldots, 9.9, 10.0\},$$

$$\beta \in \{0, .01, .02, \ldots, .49, .50\},$$

   we can write the joint posterior density for α and β as:

$$Pr(\alpha = a, \beta = b|y,x) = p(a,b|y,x) = \frac{f(y|x,a,b)p(a,b)}{\sum_{\alpha\in\mathcal{A}}\sum_{\beta\in\mathcal{B}} f(y|x,\alpha,\beta)p(\alpha,\beta)}.$$

   Then the marginal posterior density for β can be obtained by summing over the possible values of α:

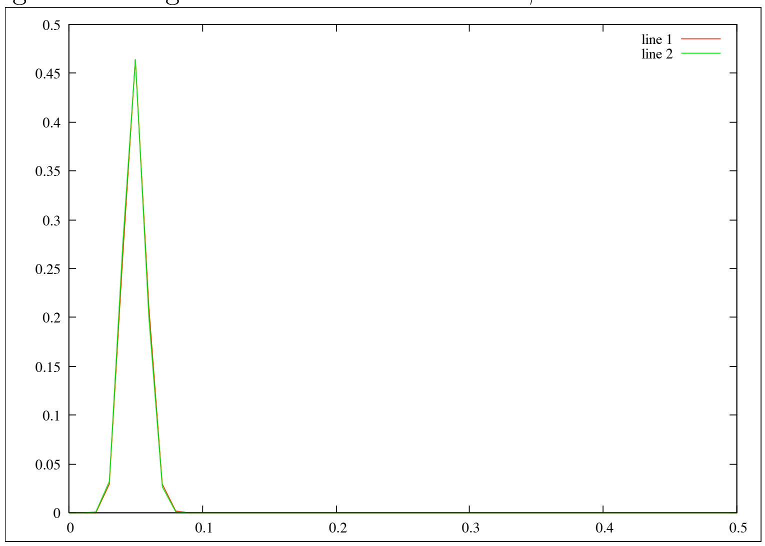
$$Pr(\beta = b|y,x) = \sum_{\alpha\in\mathcal{A}} p(\alpha,b|y,x).$$

   The program `hw5a.m` implements these calculations in Matlab, and Figure 1 plots the resulting marginal densities. The program calls another small program, `stdn_cdf.m` to calculate the standard normal cdf $\Phi(\cdot)$. The two prior densities for β correspond to normal distributions centered at 0 with variances of ∞ and 1/100. We see from the plot that the priors lead to nearly indistinguishable results. This seems to indicate that the sample size is large enough that the contribution of the likelihood dominates the informative prior.

2. The program `hw5b.m` implements the data augmentation algorithm. It iterates between drawing for the latent y_i^* given β, y, and X, and drawing for β given y^*, y, and X. Some notes on the procedure:

   Draws for y_i^*: recall from LN10 and if $y_i = 1$, the distribution of y_i^* is a normal distribution with mean $x_i'\beta$ and variance 1, truncated so that y_i^* is strictly positive. Likewise, if $y_i = 0$ then the distribution is truncated so that y_i^* is negative. We implement this in a very crude, inefficient way, by simply drawing a candidate $y_i^* \sim N(x_i'\beta, 1)$, and checking if satisfies the truncation requirement. If not, we keep redrawing until we get a draw that satisfies the truncation requirement. This can be very slow; a better way to draw for univariate truncated normals is given in the Hajivassiliou and McFadden paper.

Figure 1: Marginal Posterior Densities for β via Discretization



Draws for β: from LN10, the draw for β is:

$$\beta | y^*, X \sim \mathcal{N}(b^*, (X'X)^{-1}),$$

To implement this, we generate a 2×1 vector of independent standard normal draws z, and use the result that

$$b + chol((X'X)^{-1}) \cdot z \sim N(b, (X'X)^{-1}).$$

Here *chol* indicates the Cholesky factor of a matrix, which is a function available in Matlab.

Results:

The posterior mean for β was approximately .05, and the posterior standard deviation was approximately .0084, implying a posterior variance of approximately 7×10^{-5}. Figure 2 shows the time series plots of the posterior draws for α and β, while Figure 3 plots the histogram of the last 400 draws for β. It's a little difficult to read Figure 2 because the two variables have rather different variances, but it looks like 100 iterations is probably enough for convergence.

Figure 2: Draws for α and β

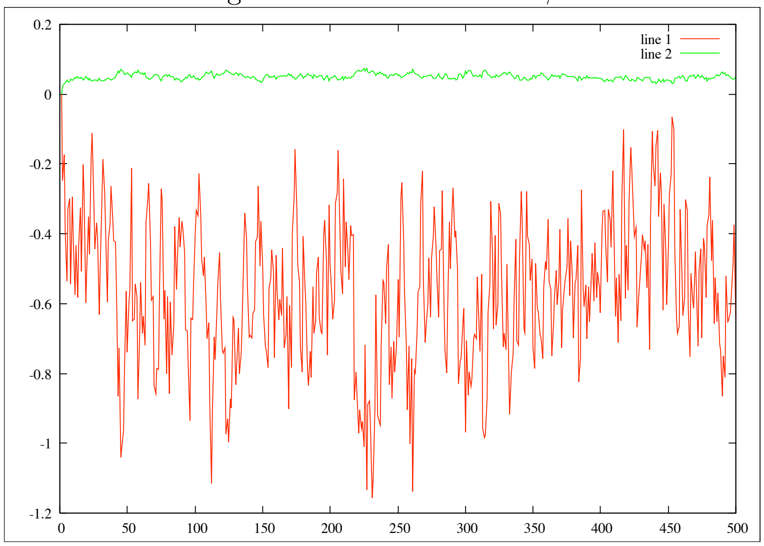


Figure 3: Histogram of Posterior Draws for β