4. In class we discussed the latent variable representation for the probit regression model, and the EM algorithm for finding maximum likelihood estimates by taking advantage of the form of the 'complete-data' likelihood.

Implement the EM algorithm for fitting a probit regression model. How does this algorithm seem to compare with the usual IRWLS algorithm? Obvious comparisons would involve whether the algorithms converge from various starting values, and how many iterations each takes.

Note: if you look at the R documentation for glm(), it is possible for the user to specify starting values.

Thanks to Paul Gustafson at UBC