Name: _____

MA 578 — Bayesian Statistics
Fall 2015

Midterm Exam

Thursday 10/15/15, 12:30–2:00 PM

1. There are roughly two types of parameters in generalized linear models: fitting parameters β, whose quality is given by a function of the data y and β called "deviance" and denoted by $D(y, \beta)$, and a dispersion parameter ϕ that scales the variance of the data. Moreover, there is a subclass of models called "extended quasi-likelihood models" in which the likelihood is given by

$$\mathbb{P}(y \,|\, \beta, \phi) \propto \phi^{-\frac{n}{2}} \exp\left\{ -\frac{D(y, \beta)}{2\phi} \right\},$$

where n is the number of data observations.

   (a) What is a conjugate prior for ϕ given β?

   (b) Derive the conditional posterior distribution for ϕ given β. How do you update the parameters in the prior?

   (c) If $\widehat{\beta}$ is the MLE for β, give a point estimate for ϕ conditional on $\beta = \widehat{\beta}$. How does your point estimate behave as $n \to \infty$? Does it converge to the MLE for ϕ?

Name: _____

2. Given your good work on a previous project we discussed in class, the MBTA manager wants your (Bayesian) opinion on a dataset he has collected on a new T station. He observed the number of trains that have stopped during one hour, in *five* different days (assume that the data on each day are independent). The total number of trains he has observed across all days is 8. The manager is mostly interested in the *rate* θ of trains that stop at the new station per hour. He expects, based on his experience with similar T stations, an average rate of one train every two hours.

   (a) Assume an exponential prior for θ, $\theta \sim \mathsf{Exp}(\beta)$. What is β based on the prior information given by the manager?

   (b) What is the posterior on θ? Write the parameters of your posterior as a function of β and the data.

   (c) Based on your previous answer, what are the sufficient statistics? Do you need the number of observations on each observation day to derive the posterior? Why?

   (d) Give an approximate 95% posterior interval for θ. Use an asymptotic approximation.

   (e) The manager plans to go to the new station and wait for three hours. What is the (posterior) probability that he sees no trains stopping?