Stat238: Problem Set 1
Due Wednesday Sep. 14

August 30, 2016

Comments:

- Please note my comments in the syllabus about academic integrity. You can work together on the problems but your final writeup must be your own and you should wrestle with the problems on your own first. A couple of the problems are from lab; overlap with your groupmates on those problems is fine.
- It's due at the start of class on paper. The syllabus discusses the penalty for turning it in late.

Problems

1. Problem 1 from Lab 1 (DNA testing).
2. In class we discussed the posterior variance for the binomial probability of success (θ) under a beta prior distribution.
   (a) Find an example of a prior and data where the posterior variance is larger than the prior variance.
   (b) Explain in a sentence or two why it makes intuitive sense in this situation that your information about the parameter has decreased based on seeing the data.
   (c) For your example, plot the prior, posterior, and likelihood on the same plot.
3. Conjugacy and predictive distributions (Problem 2 from Lab 2):
   (a) Show that if $y|\theta$ is exponentially distributed with rate θ, then the gamma prior distribution is conjugate for θ given independent and identically distributed y values.
   (b) Derive the predictive distribution $p(y)$, including its normalizing constant.
   (c) Show that if we parameterize $y|\phi$ as exponential with mean ϕ (note that $\phi = 1/\theta$), the conjugate prior specification for ϕ is inverse-gamma.
4. Predictive distributions:
   (a) In class we covered the predictive distribution for a scalar under the normal-normal setting, with $\theta \sim \mathcal{N}(\theta_0, \sigma_0^2)$. Let $y_i \stackrel{iid}{\sim} \mathcal{N}(\theta, \sigma^2),\ i = 1, \ldots, n$, where σ^2 is known and "iid" means independent and identically distributed. In the Bayesian context in which θ is random, this means that the $y_i s$ are conditionally independent given θ. Let $\tilde{y}$ be a **vector** of m new iid observations. What is the predictive distribution of $\tilde{y}$? You may make use of our result for $p(\theta|y)$ from the

book/class.
Hint: since the joint for $\tilde{y}, \theta | y$ is quadratic in the exponent, it must be normal and therefore the marginal for $\tilde{y}|y$ (i.e., marginal with respect to θ) is also normal. Given this you shouldn't need to work with any integrals to determine the result.

(b) What is the covariance of $(\tilde{y}_i, \tilde{y}_j | y)$ for $i \neq j$. Are they correlated? Explain why your answer makes sense even though the data are iid.

5. Consider the usual regression model, $Y \sim \mathcal{N}(X\beta, \sigma^2 I)$ where Y is an n-vector of IID observations, and β a k-vector of regression coefficients. Derive Jeffreys' prior for (β, σ^2). To ease the algebra (and to give you practice working in matrix-vector format), work with the likelihood for n observations in matrix-vector notation:

$$\frac{1}{(\sigma^2)^{n/2}} \exp\left(-\frac{(Y - X\beta)^T (Y - X\beta)}{2\sigma^2}\right)$$

where X is an $n \times k$ matrix. When you differentiate with respect to β, do the derivatives with respect to the entire vector. Note that apart from the multivariate notation, this is similar to the Jeffreys' prior we derived in class. Is Jeffreys' prior proper or improper?

6. (Extra credit) Consider the absolute error loss function, $L(\theta, a) = |\theta - a|$. Prove that the Bayes rule under this loss function is the posterior median.
Hints:

- You should be able do some algebraic manipulations to avoid having to take a derivative inside an integral (and thereby avoid getting into technical issues concerning interchanging differentiation and integration).
- You may take as given the following variation on Leibnitz' rule:

$$\frac{d}{dc} \int_{-\infty}^{c} h(x)dx = h(c)$$

.