

# Stat238: Lab 3 DRAFT

September 9

- This lab explores the influence of priors on the posterior in more detail.

## Problems

1. I call this problem “Brian’s paradox” as it arose from a student question when I taught a previous version of the class. Consider  $y_i, i = 1, \dots, n$  conditionally independent with  $y_i|\theta \sim N(\theta, \sigma^2)$  where for the purpose of what we want to illustrate, let  $\sigma^2 = n$ . Consider the prior  $\theta \sim N(0, \sigma_0^2)$  where  $\sigma_0^2 = 1$ . Now suppose we observe the  $y$ s and find  $\bar{y} = 100$ .
  - (a) What is the posterior? Plot the prior, likelihood, and posterior on a single plot. What is the ‘paradox’? Why does it happen?
  - (b) Now try using a  $t$  distribution as the prior for  $\theta$ . We don’t have a closed form solution, but you can still plot the prior, likelihood, and (up to a normalizing constant) the posterior. Does that give a more sensible answer?
2. Let  $y \sim N(\theta, \sigma^2)$ . Suppose that there is a prior probability,  $q$ , that  $\theta = 0$ . With probability  $1 - q$ , the prior is  $\theta \sim N(0, \omega^2)$ . I.e., the prior is a mixture of a point mass at zero and a normal distribution centered at 0 with variance  $\omega^2$ . Let’s consider the posterior probability that  $\theta = 0$ .
  - (a) Define  $A$  to be the event  $\theta = 0$ , namely that  $\theta$  comes from the point mass part of the prior and  $A^c$  that  $\theta$  comes from the continuous part of the prior. Consider the posterior probability based on Bayes theorem involving a probability and a density:

$$P(A|y) = \frac{p(y|A)P(A)}{p(y)}$$

Write out  $p(y)$ ; you should see that it involves an integral that you should be able to do without any explicit integration.

- (b) Let  $\omega^2 \rightarrow \infty$ . What does that posterior probability converge to as  $\omega^2 \rightarrow \infty$ ? Notice anything strange? This is basically a model selection problem in disguise (choosing between  $\theta = 0$  and  $\theta \neq 0$ ), and we’ll see later in the semester that standard Bayesian model selection does not work with non-informative priors.