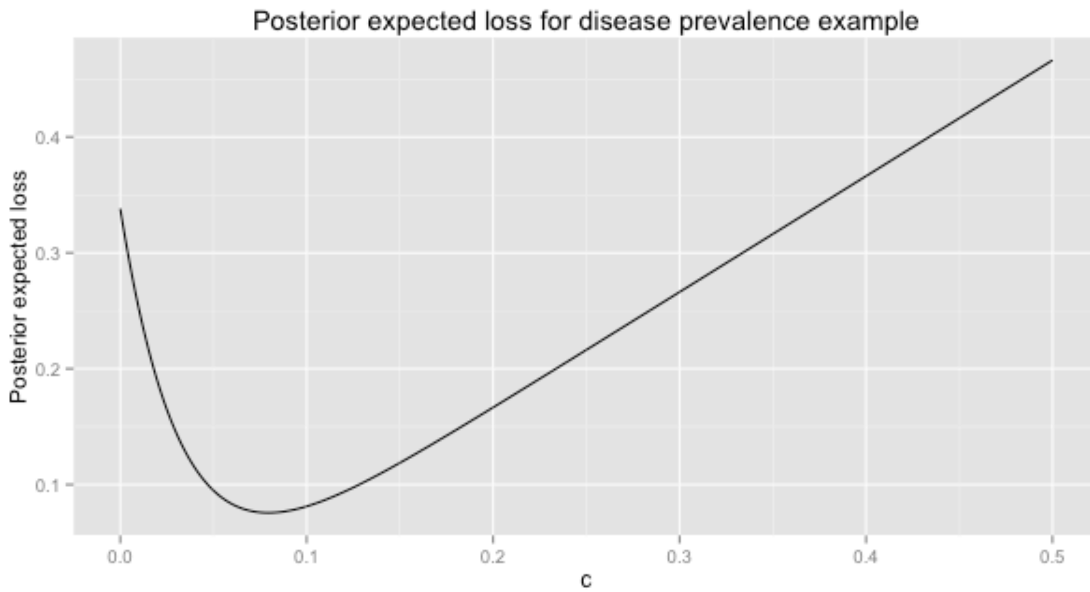
Kelvin Niu

Homework 2

Question 8



The above plot is obtained from the equation:

$$\rho(c,x) = \int_0^c \frac{|\theta - c|\theta^{0.05}(1-\theta)^{29}}{B(1.05,30)}d\theta + \int_c^1 \frac{10|\theta - c|\theta^{0.05}(1-\theta)^{29}}{B(1.05,30)}d\theta$$

using an Riemann approximation with N = 1000

Question 9

A 0 – 1 loss would not be appropriate in a binary scenario in which the impact of false negatives and false positives are not the same. For example, in the case of detecting cancer, it is much better to have a false positive than a false negative. In the case of a false positive, the patient will have an emotional scare (say, a loss of 10). However, in the case of a false negative, the patient will go untreated and die (a much greater loss, say 1000). Accurately predicting cancer would be good for the patient so he/she can start treatment (say, a loss of -10). Accurately predicting that someone does not have cancer is pretty neutral (say, a loss of 0). Thus a more appropriate loss function would be:

Diagnosis

Reality

	1 (positive)	**0**
1 (cancer)	-10	1000
0	10	0

Using this loss function, a test should probably be developed to be more sensitive.

In-class question

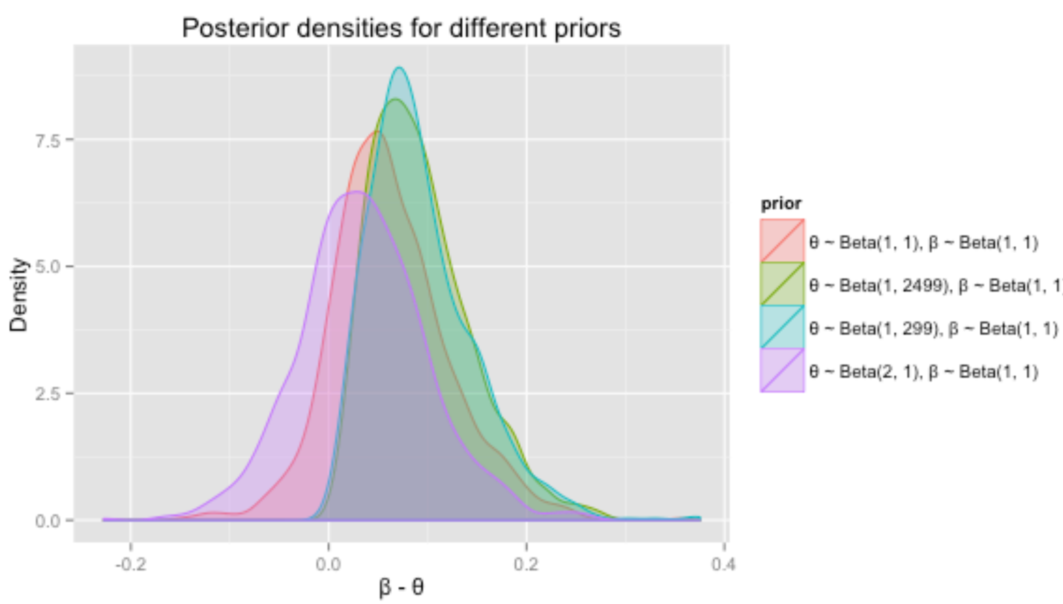
I compared the following 4 choices of priors (where θ is the mean of the control group and β is the mean of the exposed group):

- $\theta \sim$ Beta(2, 1) and $\beta \sim$ Beta(1, 1) – This is a bad prior since the data suggests evidence in the opposite direction of the prior
- $\theta \sim$ Beta(1, 1) and $\beta \sim$ Beta(1, 1) – This is an uninformative prior that does not assume any prior knowledge
- $\theta \sim$ Beta(1, 299) and $\beta \sim$ Beta(1, 1) – This is the informative prior that was mentioned in class with a relatively strong belief that the chances of a rat developing a tumor in the control group is very small
- $\theta \sim$ Beta(1, 2499) and $\beta \sim$ Beta(1, 1) – This is a informative prior with a stronger prior belief than the previous one

I used a Monte Carlo approach with 1000 simulations to estimate the distribution, probability that $\beta > \theta$ given y and x, and 95% credible intervals for $\beta - \theta$. The results are summarized in the table and plot below.

Table: Summary of results for different choices of priors for θ and β

	$\theta \sim$ Beta(2, 1) $\beta \sim$ Beta(1, 1)	$\theta \sim$ Beta(1, 1) $\beta \sim$ Beta(1, 1)	$\theta \sim$ Beta(1, 299) $\beta \sim$ Beta(1, 1)	$\theta \sim$ Beta(1, 2499) $\beta \sim$ Beta(1, 1)
Description	Bad prior, contradicts direction of data	Uninformative prior	Informative prior (from class)	More informative prior
$Pr(\beta > \theta \mid y, x)$	0.714	0.882	0.998	1
95% CI for $\beta - \theta$	[-0.0956, 0.1708]	[-0.0472, 0.1940]	[0.0167, 0.2175]	[0.0205, 0.2142]



As the priors become more informative in direction that the data supports, the probability that $\beta > \theta$ given y and x increases and becomes closer to 1. In addition, the length of the credible intervals and the spread of the posterior tend to become smaller. This makes sense because if our prior beliefs are more adamant and turn out to be supported by the data, we have more reason to be sure of them.