STAT 491: Midterm Exam
Name:

You have been hired as a data scientist by the Yellowstone Club. Your favorite activity is celebrity spotting and you have decided to estimate the probability of spotting a celebrity during a day in the Big Sky area.

1. Priors (11 points)

a. (4 points)

State and defend a prior for this situation. Make sure to use notation that clearly defines parameters and values.

b. (3 points)

Discuss your philosophy for choosing prior distributions.

c. (4 points)

Define and give an example of conjugate prior.

2. Sampling Model (5 points)

a. (2 points)

What is a sampling model and why is it necessary in Bayesian statistics?

b. (3 points)

Write out the sampling model for this problem. You can assume you work N days and observe celebrities on z days.

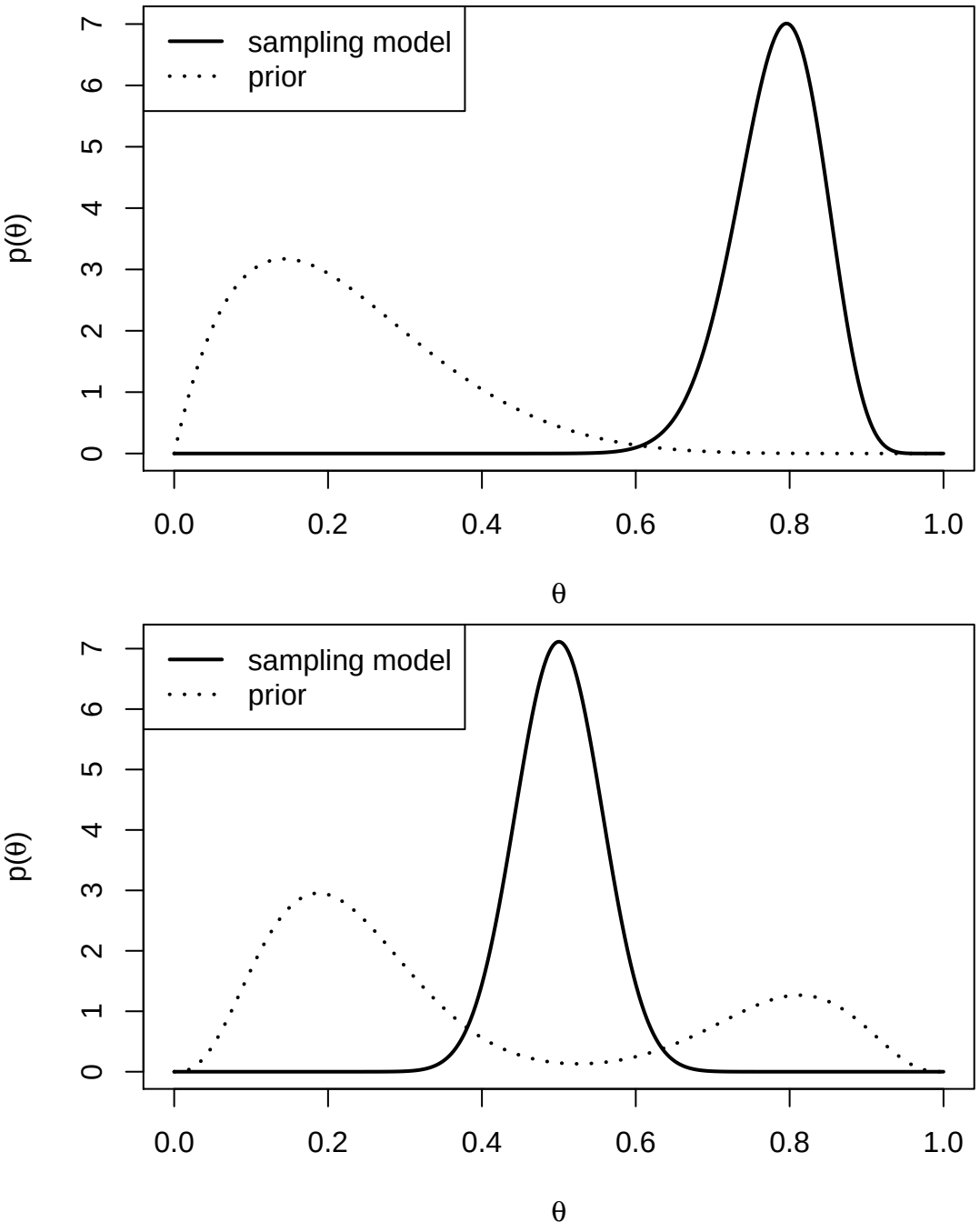
3. Posterior (19 points)

a. (4 points)

It turns out that your office overlooks the golf course at the Yellowstone Club. After working for 21 days, you have spotted celebrities (including Justin Timberlake and Tom Brady) on 14 of those days. What is the posterior distribution for seeing celebrities?

b. (5 points)

Consider the following two sketches that contain a sampling model and a prior. Draw the curve where you think the posterior distribution will be.



Finally defend your answers.

c. (2 points)

Given this posterior distribution in part 3a, what is the posterior mean for the probability of seeing a celebrity? You do not need to simplify this expression.

d. (4 points)

After running JAGS and using the `coda()` package in R, your posterior HDI is (.53,.75). Assume your aunt is big fan of Justin Timberlake (and all things celebrity), describe what the posterior HDI means to her.

e. (4 points)

After hearing your explanation, your aunt remembers her Intro to Stat course and asks "So this posterior HDI is just a confidence interval, right?" Respond to your aunt's question with a yes or no and an explanation of your answer.