STAT314-2016S2 ASSIGNMENT 1: Bayes' Theorem and the Basics of Posterior Inference. Beta-Binomial Model.

SET: Tue 25 July

To be submitted by: Monday 7 Aug, 5pm.

Please type everything in either Word or LaTex, and submit it as a ***PDF file*** *via Learn. No handwritten submissions. Please name your file* ***Surname_Name_STAT314A1.pdf****, and please indicate your name in the file as well.*

Problem 1: (3pt)

Suppose there are two species of panda bear. Both are equally common in the wild and live in the same places. They look exactly alike and eat the same food, and there is as yet no genetic assay capable of telling them apart. They differ however in their family sizes. Species A gives birth to twins 10% of the time, otherwise birthing a single infant. Species B births twins 30% of the time, otherwise birthing singleton infants. Assume these numbers are known with certainty, from many years of field research.

Now suppose you are managing a captive panda breeding program. You have a new female panda of unknown species and she has just given birth to twins.

- **(a)** Compute the probability that the panda we have is from species A, given this information (i.e., the birth of twins). **(1pt)**
- **(b)** What is the probability that her next birth will also be twins? **(1pt)**
- **(c)** Suppose that the same panda mother has a second birth and that it is not twins, but a singleton infant. Compute the posterior probability that this panda is species A. **(1pt)**

Problem 2: (2pt)

You have two coins in your pocket. One is a standard fair coin, so that P(heads)=0.50. The other one is a two-sided coin, so that P(heads)=1. You draw one of the two coins at random and toss it twice. Both tosses result in heads.

- **(a)** Use Bayes' formula to obtain the probability that you have selected the fair coin, given your observations. **(1pt)**
- **(b)** Find the classical p-value for the null hypothesis: the coin is fair. **(1pt)**

Problem 3: (3pt)

The geometric distribution is a probability distribution of the number **X** of Bernoulli trials needed to get one success. For example, how many attempts does a basketball player need to get a goal. Given the probability of success in a single trial is *p*, the probability that the *x*th trial is the first success is:

$$\Pr(X = x|p) = (1-p)^{x-1}p$$

for *x=1,2,3,….*

Suppose, you observe *n* basketball players trying to score and record the number of attempts required to get the first goal: $x_1, x_2, \dots, x_n$ for each of them.

- **(a)** Write down your likelihood and find the MLE of *p.* **(1pt)**
- **(b)** Using Bayesian approach, assign a beta prior to *p* and obtain the posterior distribution for p. **(1pt)**
- **(c)** Compare the posterior mean of *p* to the MLE. What can you say about sensitivity of posterior inference to prior assumptions? What happens to the difference between the posterior mean and the MLE of *p* as the sample size *n* increases? **(1pt)**

Problem 4: (4p)

A sample of 100 people were asked how many days in the last week did they exercise. The resulting frequency table of responses is shown below

Days/week	0	1	2	3	4	5
Frequency	35	19	23	17	5	1

- **(a)** Use beta-binomial model to estimate the average number of 'physically active' days a week for a **uniform prior**. Simulate a random sample from the analytically evaluated **posterior distribution**. Use it to summarize the posterior distribution using posterior means and 95% CIs and interpret the statistics. **(1pt)**
- **(b)** Simulate a random sample **from the posterior predictive distribution.** Use it to estimate the probability of people exercising at least three days a week. **(1pt)**
- **(c)** Suppose, another similar study has been conducted recently, and their posterior distribution for the probability of drinking on any day of the week was found to be **B(302,458)**. How would you change your analysis in (a) and (b) to accommodate that result? Report your responses to (a) and (b) in light of this new information.**(1pt)**
- **(d)** Is beta-binomial a good model for these data? Explain. **(1pt)**