Stat 474/574 Homework 2 - Due Wed., March 1

Each part is worth 10 points, for a total of 60 points.

Problem 1
Kobe Bryant was one of the greatest basketball players of all time. He played 21 regular seasons with the Los Angeles Lakers (LAL on score boards), and retired after the 2015-2016 season. He tragically died in a helicopter accident in early 2020.
How good a shooter was Kobe Bryant? And did his shooting average and shooting reliability increase over time?
To answer these questions, we first download Kobe Bryant's statistics from Yahoo!Sports at `http://sports.yahoo.com/nba/players/3118/`. The data are saved in the file entitled *kobe-regular-seasons.csv*. For each of the regular seasons, we focus on field goals made (FGM) and field goals attempted (FGA) per game and realize that FGA is the number of trials of a binomial random variable with number of successes equal to FGM and unknown probability of success θ.
We could compute the posterior distribution of θ in each of the 21 regular seasons. Instead, we accumulate information over seasons to compare Bryant's performance after 1, 5, 10, 15 and 20 seasons. Was Bryant's field goal percentage still improving at the time of his retirement?
Even though you can compute everything explicitly, **you are expected to use simulation** to draw values of θ from its appropriate posterior distribution.

1a Using a non-informative prior, obtain the posterior distribution of Bryant's probability of successful shots using the season 1 data. Summarize the posterior distribution using the posterior mean and the 95% credible set for the probability of success.

1b Repeat the analysis four more times accumulating the data from seasons 2-5, 6-10, 11-15 and 16-20. Each time, use the posterior distribution you had estimated earlier as the current prior.

1c Which was Bryant's best season in terms of shooting success?

1d Was Bryant still improving at the time he retired or was his shooting performance decreasing?

Problem 2 In this exercise, we write Metropolis-Hastings routines to sample θ from its posterior distribution, even though the distributions are of standard form and therefore we do not really need to use simulation.

2a In a clinical trial for a Covid-19 vaccine, 100 participants were allocated to the placebo (no treatment) group. Seventeen out of those 100 had contracted Covid-19 by the end of the trial. We let Y denote the number of infected participants in the placebo group and let n denote the size of the placebo group. We assume that the observations are distributed as binomial, with parameters $n = 100$ and unknown probability θ. Researchers wish to estimate θ, and they choose a conjugate beta prior with parameters $a = 7.5$ and $b = 42.5$. (Note that this means that they expected that about 15 in 100 placebo participants would be infected and given that $a + b = 50$, were rather confident about that guess.) With this information, use the Metropolis algorithm to draw 1000 values of θ from its posterior distribution. Using the draws, compute **and interpret** the 95% credible set for θ.

2b Let $x \sim p(x)$, where

$$p(x) = \exp\{-x\}, \quad x \geq 0.$$

Using the Metropolis algorithm, draw 1000 values of x from $p(x)$. With the draws, construct a histogram that shows the shape of the distribution of x.