MATH 341 / 650.3 Spring 2018 Homework #3

Professor Adam Kapelner

Due Monday 5PM, March 19, 2018 under the door of KY604

(this document last updated Tuesday 13$^{\text{th}}$ March, 2018 at 5:09pm)

Instructions and Philosophy

The path to success in this class is to do many problems. Unlike other courses, exclusively doing reading(s) will not help. Coming to lecture is akin to watching workout videos; thinking about and solving problems on your own is the actual "working out." Feel free to "work out" with others; **I want you to work on this in groups.**

Reading is still *required*. For this homework set, review the beta-binomial conjugate model, Bayesian Hypothesis Testing, Bayes Factors, Empirical Bayes, the beta-binomial distribution (and also, posterior predictive distributions), uninformative priors and read the relevant sections of McGrayne.

The problems below are color coded: green problems are considered *easy* and marked "[easy]"; yellow problems are considered *intermediate* and marked "[harder]", red problems are considered *difficult* and marked "[difficult]" and purple problems are extra credit. The *easy* problems are intended to be "giveaways" if you went to class. Do as much as you can of the others; I expect you to at least attempt the *difficult* problems.

Problems marked "[MA]" are for the masters students only (those enrolled in the 650.3 course). For those in 341, doing these questions will count as extra credit.

This homework is worth 100 points but the point distribution will not be determined until after the due date. See syllabus for the policy on late homework.

Up to 10 points are given as a bonus if the homework is typed using LaTeX. Links to instaling LaTeX and program for compiling LaTeX is found on the syllabus. You are encouraged to use `overleaf.com`. If you are handing in homework this way, read the comments in the code; there are two lines to comment out and you should replace my name with yours and write your section. The easiest way to use overleaf is to copy the raw text from hwxx.tex and preamble.tex into two new overleaf tex files with the same name. If you are asked to make drawings, you can take a picture of your handwritten drawing and insert them as figures or leave space using the "\vspace" command and draw them in after printing or attach them stapled.

The document is available with spaces for you to write your answers. If not using LaTeX, print this document and write in your answers. I do not accept homeworks which are *not* on this printout. Keep this first page printed for your records.

NAME: _____

Problem 1

These are questions about McGrayne's book, chapters 8-10.

(a) [easy] When was experimentation introduced to medical science and who introduced it? Are you surprised that it was this recent?

(b) [easy] Sir Ronald A. Fisher, the founder of modern experiments, did not believe cigarettes caused lung cancer. What were his two hypotheses for the cause of lung cancer?

(c) [easy] Who invented, and what are Bayes Factors? (p116)

(d) [easy] Trick question: who convinced Cornfield to stop smoking?

(e) [easy] Why were frequentists at a loss to estimate the probability of a nuclear bomb being detonated by accident?

(f) [easy] What is Cromwell's Rule? And, when applying this principle to a Bayesian model what would it imply? (See the Wikipedia link and p123).

(g) [easy] Did Bayesian Statistics prevent nuclear accidents? Discuss.

(h) [easy] What is the main reason why there are so many variations of Bayesian interpretation? (p129)

(i) [easy] What is a large practical drawback of Bayesian inference? (See mid-end of chapter 8).
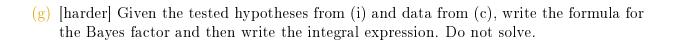
We will again be looking at the beta-prior, binomial-likelihood Bayesian model and practicing hypothesis testing.
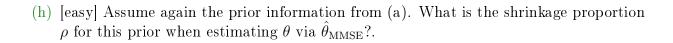
(a) [harder] [Also on HW#2] Design a prior where you believe $\mathbb{E}[\theta] = 0.5$ and you feel as if your belief represents information contained in five coin flips.

(b) [harder] You flip the same coin 100 times and you observe 39 heads. Test the hypothesis that this coin is fair given prior information from (a). Use the credible region method. Make sure you say whether you retain or reject the null and justify why.

(c) [harder] Test the hypothesis that this coin has a bias towards Heads (not tails) given prior information from (a) and the data from (c). Calculate the Bayesian p-val for the test to determine if you should retain or reject H_0.

(d) [easy] Let's say you wanted to test whether the coin is fair but you are indifferent to any θ which is different from 0.5 by a margin of 0.1. Write out the hypotheses for this test.

(e) [harder] Test the hypotheses from (h) given prior information from (a) and the data from (c). Make sure you say whether you retain or reject the null and justify why.

(f) [harder] Calculate the Bayesian p-val for the test in (i).

(g) [harder] Given the tested hypotheses from (i) and data from (c), write the formula for the Bayes factor and then write the integral expression. Do not solve.

(h) [easy] Assume again the prior information from (a). What is the shrinkage proportion ρ for this prior when estimating θ via $\hat{\theta}_{\mathrm{MMSE}}$?.

(i) [difficult] Prove that $\hat{\theta}_{\mathrm{MMSE}}$ is a biased estimator (i.e. its expectation is *not* θ).

(j) [easy] Prove that $\lim\limits_{n \to \infty} \rho = 0$ and therefore this bias $\to 0$ as your dataset gets larger.

(k) [difficult] [MA] Why on Earth should anyone use shrinkage estimators if they're biased? Google it. Discuss.

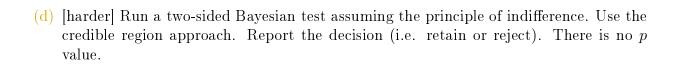This question is about a famous extra sensory perception (ESP) experiment.

According to quantum mechanics, an event should happen with exactly probability 50%. So if someone comes with a claim that they have ESP and can affect this event with their mind, then the event no longer has probability 50% (this is what we are trying to prove).

An experiment was run with someone claiming they have ESP. They tried to affect the event with their mind. The data is as follows: of $n = 104,490,000$ observations, the number of events was $x = 52,263,970$. We will now test to see if the person has ESP a number of ways and try to reconcile the differences.

(a) [easy] What is the MLE for the probability of the event?

(b) [easy] Run a two-sided frequentist test at $\alpha = 5\%$ and report the decision (i.e. retain or reject) and the p value.

(c) [harder] Run a two-sided Bayesian test assuming the principle of indifference. Use the equivalence region approach with $\delta = 0.01$. Report the decision (i.e. retain or reject) and the p value.

(d) [harder] Run a two-sided Bayesian test assuming the principle of indifference. Use the credible region approach. Report the decision (i.e. retain or reject). There is no p value.

(e) [difficult] Calculate the Bayes Factor B where the alternative hypothesis of $\theta \neq 0.5$ can be summed up with $\theta \sim U(0, 1)$.

(f) [difficult] Try to reconcile parts (b), (c), (d), (e). Why are there different answers? What is going on??

Some quick question on mixture distributions.

(a) [easy] If X is independent to W and X is independent to Z and X is independent to U, can you write $\mathbb{P}\left(X \mid U, V, W, Y, Z\right)$ more compactly? Do so below.

(b) [easy] Let X be $\mathcal{N}\left(0,\ 1^2\right)$ $1/3$ of the time and $\mathrm{Exp}\left(3\right)$ $2/3$ of the time. What is its pdf?

(c) [difficult] Let's say $X \mid \beta \sim \mathrm{Beta}\left(1,\ \beta\right)$ where $\beta \mid \lambda \sim \mathrm{Exp}\left(\lambda\right)$. Write an integral expression which when solved, finds the compound / marginal density of X. DO NOT solve.

(d) [difficult] [MA] Let's say $X \mid \theta,\ \sigma^2 \sim \mathcal{N}\left(\theta,\ \sigma^2\right)$ where $\theta \mid \mu_0,\ \tau^2 \sim \mathcal{N}\left(\mu_0,\ \tau^2\right)$. Write an integral expression which when solved, finds the compound / marginal density of X. DO NOT solve.

Problem 5

These are questions about other vague priors: improper priors and Jeffreys priors.

(a) [easy] What is an improper prior?

(b) [easy] Is $\theta \sim \text{Beta}\,(100,\,0)$ improper? Yes / no and provide a proof.

(c) [easy] When are improper priors "legal"?

(d) [easy] When are improper priors "illegal"?

(e) [difficult] What does $I(\theta)$ tell you about the random variable with respect to its parameter θ?

(f) [harder] If I compute a posterior on the θ scale and then measure the parameter on another scale, will I (generally) get the same posterior probability? Yes/no explain.

(g) [easy] What is the Jeffrey's prior for θ under the binomial likelihood? Your answer must be a distribution.

(h) [difficult] What is the Jeffrey's prior for $\theta = t^{-1}(r) = \frac{e^r}{1+e^r}$ (i.e. the log-odds reparameterization) under the binomial likelihood?

(i) [difficult] Explain the advantage of Jeffrey's prior.

(j) [difficult] [MA] Prove Jeffrey's invariance principle i.e. prove that the Jeffrey's prior makes your prior probability immune to transformations. Use the second proof from class.

Problem 6

This question is about "batting averages" in baseball.

Every hitter's *sample* batting average (BA) is defined as:

$$BA := \frac{\text{sample \# of hits}}{\text{sample \# of at bats}}$$

In this problem we care about estimating a hitter's *true* batting average which we call θ. Each player has a different θ but we focus in this problem on one specific player. In order to estimate the player's true batting average, we make use of the sample batting average as defined above (with Bayesian modifications, of course).

We assume that each at bat (for any player) are conditionally $\overset{iid}{\sim}$ based on the players' true batting average, θ. So if a player has n at bats, then each successful hit in each at bat can be modeled via $X_1 \mid \theta, \; X_2 \mid \theta, \ldots, \; X_n \mid \theta \overset{iid}{\sim} \text{Bernoulli}\,(\theta)$ i.e. the standard assumption.

(a) [easy] Looking at the entire dataset for 6,061 batters who had 100 or more at bats, I fit the beta distribution PDF to the sample batting averages using the maximum likelihood approach and I'm estimating $\alpha = 42.3$ and $\beta = 127.7$ (which we will call the "empirical Bayes" estimates in class). Consider building a prior from this estimate as $\theta \sim \text{Beta}\,(42.3, \; 127.7)$ Would a prior based on these hyperparameter estimates be "objective"? Yes / No. Why?

(b) [easy] Using prior data to build the prior is called...

(c) [easy] Is the prior from (a) considered a "conjugate prior"? Yes / No.

(d) [easy] Using the prior from (a), find the $\hat{\theta}_{\text{MMSE}}$.

(e) [difficult] We now observe four at bats for a new player and there were no hits. Write an exact expression for the batter getting 14 or more hits on the next 20 at bats. You can leave your answer in terms of the beta function. Do not compute explicitly.