

MATH 390 / 650.3 Spring 2022 Homework #3

Professor Adam Kapelner

Due by email, Sunday 11:59PM, March 24, 2022

(this document last updated Tuesday 15th March, 2022 at 7:20pm)

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, read about credible regions, high-density regions, Bayesian one-, two-sided hypothesis tests, improper priors, Haldane’s prior, mixture/compound distributions, the beta-binomial model, the prior predictive distribution, the posterior predictive distribution and read chapters 8-10 in 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 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 L^AT_EX. Links to installing L^AT_EX and program for compiling L^AT_EX is found on the syllabus. You are encouraged to use [overleaf.com](https://www.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 L^AT_EX, 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).

Problem 2

We will continue considering the beta-binomial model: $\mathcal{F} : \text{Binomial}(n, \theta)$ and $\mathbb{P}(\theta) = \text{Beta}(\alpha, \beta)$. Assume the dataset from the previous homework: $x = \langle 0, 1, 1, 1, 1, 1 \rangle$, $n = 6$. Assume Laplace's flat prior.

- (a) [easy] Write out an expression for the 95% Bayesian credible region (CR) for θ . Then write out the answer using the `qbeta` function from the `R` language.

- (b) [easy] Compute the 95% Bayesian credible region (CR) for θ . Use `R` on your computer (or `rdr.io` online) and its `qbeta` function.

- (c) [easy] Compute a 95% frequentist confidence interval (CI) for θ . Is there a problem with it? If so what is wrong and why is it a problem?

- (d) [difficult] [MA] Let $\mu : \mathbb{R} \rightarrow \mathbb{R}^+$ be the Lebesgue measure which measures the length of a subset of \mathbb{R} . Why is $\mu(\text{CR}) < \mu(\text{CI})$? That is, why is the Bayesian CR in (b) tighter than the Frequentist CI in (c)?

- (e) [easy] Now assume a different dataset. You flip the same coin 100 times and you observe 39 heads. Calculate a 95% credible region (CR) for θ , the parameter representing the coin's inherent propensity to flip heads. Round to the nearest 3 decimal points. Use R on your computer (or rdrr.io online) and its `qbeta` function.
- (f) [harder] Test the theory that the coin is unfairly weighted towards tails at significance level $\alpha_0 = 5\%$. Write out null and alternative hypothesis as mathematical statements (i.e. of the form θ belongs to some subset of Θ). Perform the calculation and leave in terms of the computer notation we discussed in class. Calculate the Bayesian p-value. Round to the nearest 3 decimal points. Indicate your decision (rejection of the null or retainment of the null) and write a sentence in English about the decision in this context. Use R on your computer (or rdrr.io online) and its `qbeta` function.
- (g) [easy] If you were to test the theory that the coin is unfair, write out null and alternative hypothesis as mathematical statements (i.e. of the form θ belongs to some subset of Θ). This is different from the previous question! Here you will need to declare another constant. Declare that constant. And justify why you chose its value.

- (h) [harder] Test the theory that the coin is unfair at significance level $\alpha_0 = 5\%$ using the hypotheses from the previous question. Perform the calculation and leave in terms of the computer notation we discussed in class. Calculate the Bayesian p-value. Round to the nearest 3 decimal points. Indicate your decision (rejection of the null or retainment of the null) and write a sentence in English about the decision in this context. Use R on your computer (or rdrv.io online) and its `qbeta` function.
- (i) [easy] Explain the disadvantages of the Bayesian highest density region (HDR) method for computing confidence sets.
- (j) [difficult] Assume the dataset $x = < 1, 1, 1, 1, 1, 1 >$, $n = 6$ and Laplace's flat prior. Find the 95% Bayesian credible region (HDR) for θ . Hint: plot the posterior and you should be able to see the answer

- (k) [easy] What is Haldane's prior $\mathbb{P}(\theta)$ in the beta-binomial Bayesian model? What was he trying to accomplish with this prior?
- (l) [easy] What is the definition of "a proper prior"? Is Haldane's prior $\mathbb{P}(\theta)$ a proper prior? Yes/no and why.
- (m) [easy] How many pseudotrials n_0 , pseudosuccesses x_0 and pseudofailures $n_0 - x_0$ are contributed by Haldane's prior? Is it the same as for Laplace's prior?
- (n) [easy] What is $\hat{\theta}_{\text{MMSE}}$ under Haldane's prior? Is it the same as $\hat{\theta}_{\text{MLE}}$?
- (o) [easy] For what datasets is the posterior distribution $\mathbb{P}(\theta \mid X)$ proper under Haldane's prior?
- (p) [harder] Assume the dataset $x = \langle 1, 1, 1, 1, 1, 1 \rangle$, $n = 6$ and Haldane's prior. Is it possible to compute a point estimate? Is it possible to compute CR's? Is it possible to compute Bayesian p-val's when testing hypotheses?

Problem 3

Assume \mathcal{F} = binomial with n fixed and $\mathbb{P}(\theta) = \text{Beta}(2.5, 2.5)$, $n = 100$ and $x = 39$.

- (a) [easy] Find the posterior predictive distribution, $X_* | X$ where X_* denotes the random variable that counts the number of successes in n_* future trials. MA students — do this yourself. Other students — use my notes and justify each step.

- (b) [difficult] Show that for the case of predicting only $n_* = 1$ future trials, the posterior predictive distribution is $X_* | X \sim \text{Bernoulli}(\hat{\theta}_{\text{MMSE}})$.

(c) [easy] If $n_* = 17$, what is the expectation and variance of $X_* \mid X$?

(d) [harder] Plot the PMF of $X_* \mid X$ as best as you can. Mark critical points and label the axes.

(e) [easy] What is the probability of $x_* = 10$ given your data and prior? Write your answer in terms of the numerical function(s) we've been using in class.

(f) [harder] Answer the previous problem exactly and then round to two decimal places using the <https://rdr.io/snippets/> website if you don't have access to **R** on your computer.

Problem 4

These are some questions on mixture / compound distributions in general.

- (a) [easy] Let X be $\mathcal{N}(0, 1^2)$ 1/3 of the time and $\text{Exp}(3)$ 2/3 of the time. What is its pdf?
- (b) [difficult] Let's say $X \mid \beta \sim \text{Beta}(1, \beta)$ where $\beta \mid \lambda \sim \text{Exp}(\lambda)$. Write an integral expression which when solved, finds the compound / marginal density of X . DO NOT solve.
- (c) [difficult] [MA] Let's say $X \mid \theta, \sigma^2 \sim \mathcal{N}(\theta, \sigma^2)$ where $\theta \mid \mu_0, \tau^2 \sim \mathcal{N}(\mu_0, \tau^2)$. Write an integral expression which when solved, finds the compound / marginal density of X . DO NOT solve.

Problem 5

We will now have lots of examples finding kernels from common distributions. Some of these questions are silly, but they will force you to think hard about what the kernel is under different situations. And... they're fun! Probabilistic pyromania!

(a) [easy] What is the kernel of $X \mid \theta, n \sim \text{Binomial}(n, \theta)$?

(b) [difficult] What is the kernel of $X, n \mid \theta \sim \text{Binomial}(n, \theta)$? Be careful...

(c) [easy] What is the kernel of $X \mid \alpha, \beta \sim \text{BetaBinomial}(n, \alpha, \beta)$?

(d) [easy] What is the kernel of $X \mid \theta \sim \text{Poisson}(\theta) := \frac{e^{-\theta}\theta^x}{x!}$?

(e) [difficult] What is the kernel of $\theta \mid X \sim \text{Poisson}(\theta)$? Be careful...

(f) [easy] What is the kernel of $X \mid \alpha, \beta \sim \text{Beta}(\alpha, \beta)$?

(g) [easy] What is the kernel of $X \mid \theta \sim \text{Exp}(\theta) := \theta e^{-\theta x}$? This is the exponential distribution.

(h) [easy] What is the kernel of $X \mid \theta, \sigma^2 \sim \mathcal{N}(\theta, \sigma^2)$?

(i) [difficult] What is the kernel of $\theta, \sigma^2 \mid X \sim \mathcal{N}(\theta, \sigma^2)$? Be careful...

(j) [easy] What is the kernel of $X \mid k \sim \chi_k^2 := \frac{1}{2^{k/2}\Gamma(k/2)} x^{k/2-1} e^{-x/2}$, the chi-squared distribution with k degrees of freedom?

(k) [harder] What is the kernel of

$$X \mid N, \theta, n \sim \text{Hypergeometric}(N, \theta, n) := \frac{\binom{\theta}{x} \binom{N-\theta}{n-x}}{\binom{N}{n}}$$

where N is the number of total balls in the bag, θ is the number of success balls in the bag and n is the number drawn out of the bag?

(l) [harder] [MA] If $X \sim F_{k_1, k_2}$, Snecedor's F-distribution, what is its kernel?

(m) [difficult] [MA] If $X \mid \theta, \sigma^2 \sim \mathcal{N}(\theta, \sigma^2)$ and $\theta \mid \mu_0, \tau^2 \sim \mathcal{N}(\mu_0, \tau^2)$, what is the kernel of $\theta \mid X, \sigma^2, \mu_0, \tau^2$?