MATH 390.03-02 / 650 Fall 2015 Homework #5

Professor Adam Kapelner

Due 4PM in my mail slot, Friday, March 11, 2016

(this document last updated Thursday 3$^{\text{rd}}$ March, 2016 at 9:58am)

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 mixture models, kernels of PMFs / PDFs, Jeffrey's priors, improper priors, Bayesian model checking. No readings in McGrayne this week; use the time to study for the exam.

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 course). For those in 390, 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

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 $\text{Exp}\left(3\right)$ 2/3 of the time. What is its pdf?

(c) [difficult] Let's say $X \mid \beta \sim \text{Beta}\left(1, \beta\right)$ where $\beta \mid \lambda \sim \text{Exp}\left(\lambda\right)$. 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.

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!

(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 $X \mid \theta \sim \text{Poisson}(\theta)$?

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

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

(f) [easy] What is the kernel of $X \mid \theta \sim \text{Exp}(\theta)$?

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

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

(i) [harder] [MA] 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?

(j) [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 $\theta \mid X, \sigma^2, \mu_0, \tau^2$?

Problem 3

Some mostly definition questions.

(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's the name of the prior that I should use to always guarantee I get the same posterior probability regardless of the scale I measure θ on?

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

(i) [easy] State the definition of $\text{logit}(\theta)$.

(j) [difficult] What is the Jeffrey's prior for $\text{logit}(\theta)$ under the binomial likelihood?

(k) [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.

(l) [easy] State the Bayesian CLT.

(m) [difficult] In the standard binomial data / beta prior model, use the Bayesian CLT to approximate the distribution of the posterior.

(n) [easy] State the Law of Iterated Expectation for two r.v.'s X and Y.

(o) [harder] Draw the picture from class and explain why the picture illustrates the Law of Iterated Expectation for two r.v.'s X and Y.

(p) [easy] State the Law of Total Variance for two r.v.'s X and Y.

(q) [harder] Draw the picture from class and explain why the picture illustrates the Law of Total Variance for two r.v.'s X and Y.

(r) [easy] State and explain the relationship between prior and posterior variance.

(s) [easy] If your prior variance is less than your posterior variance, generally speaking, what happened?

We will now perform model checking by looking at pictures.

(a) [easy] If $\theta \sim \text{Beta}\,(\alpha,\,\beta)$, what is the prior predictive distribution for n observations?

(b) [easy] If $\alpha = 100$ and $\beta = 3$ and $n = 50$. Run the following code to look at the prior predictive distribution. No need to print out the illustration.

If you now get 10 successes on your 50 observations (the red line), does this indicate a problem? Yes/No

```
#begin copying this line
options(repos = "http://lib.stat.cmu.edu/R/CRAN/")
install.packages("VGAM")
library(VGAM)
n = 50
alpha = 100
beta = 3
barplot(dbetabinom.ab(1 : n, size = n, shape1 = alpha, shape2 = beta),
names = 1 : n, xlab = "x", ylab = "prob(x)")
abline(v = 10, col = "red", lwd = 5)
#copy this line last and paste it in R and press "enter"
```

(c) [easy] If you use the Jeffrey's prior $\alpha = 0.5$ and $\beta = 0.5$, run the following code to look at the prior predictive distribution. No need to print out the illustration.

If you now get 10 successes on your 50 observations, does this indicate a problem? Yes/No

```
#begin copying this line
alpha = 0.5
beta = 0.5
barplot(dbetabinom.ab(1 : n, size = n, shape1 = alpha, shape2 = beta),
names = 1 : n, xlab = "x", ylab = "prob(x)")
abline(v = 10, col = "red", lwd = 5)
#copy this line last and paste it in R and press "enter"
```

(d) [easy] Why should the data look plausible under the prior predictive distribution?

(e) [easy] If you use $\alpha = 100$ and $\beta = 3$ prior and you got 10 successes, we now look at the posterior predictive distribution by running the following code. No need to print out the illustration.

Does this indicate a problem? Yes/No

```
#begin copying this line
alpha = 100 + 10
beta = 3 + 40
barplot(dbetabinom.ab(1 : n, size = n, shape1 = alpha, shape2 = beta),
names = 1 : n, xlab = "x*", ylab = "prob(x*|x)")
abline(v = 10, col = "red", lwd = 5)
#copy this line last and paste it in R and press "enter"
```

(f) [easy] If you use the Jeffrey's prior with the same data, we now look at the posterior predictive distribution by running the following code. No need to print out the illustration.

Does this indicate a problem? Yes/No

```
#begin copying this line
alpha = 0.5 + 10
beta = 0.5 + 40
barplot(dbetabinom.ab(1 : n, size = n, shape1 = alpha, shape2 = beta),
names = 1 : n, xlab = "x*", ylab = "prob(x*|x)")
abline(v = 10, col = "red", lwd = 5)
#copy this line last and paste it in R and press "enter"
```

(g) [easy] Why should the data look plausible under the posterior predictive distribution?

(h) [easy] Why are looking at these illustrations called "graphical checks"?

(i) [easy] Why are looking at these illustrations called "self-consistency checks"?

(j) [difficult] [MA] George Box said, "all models are wrong but some are useful". What does this mean?

(k) [difficult] [MA] Andrew Gelman refines George Box's quotation by saying, we know all models are wrong "but do the model's deficiencies have a noticable impact on inference"? What does he mean by asking this question? Keep in mind Gelman loves these graphical checks a la (b), (c), (e) and (f).