MATH 390.03-02 / 650 Fall 2015 Homework #7

Professor Adam Kapelner

Due *in class*, Monday, April 11, 2016

(this document last updated Wednesday 6^{th} April, 2016 at 11:43am)

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 the normal-normal conjugate and semi-conjugate model. Also read ch13 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 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

These are questions about McGrayne's book, chapters 13 and 14.

(a) [easy] Write a one paragraph biography of John Tukey.

(b) [easy] Why did Alfred Kinsey's wife want to poison John Tukey?

(c) [easy] Tukey helped NBC with polling predictions for the presidential campaign. What was NBC's polling algorithm based on?

(d) [easy] Why is "objectivity an heirloom ... and ... a fallacy?"

(e) [easy] Why do you think Tukey called Bayes Rule by the name "borrowing strength?"

(f) [easy] Why is it that we don't know a lot of Bayes Rule's modern history?

(g) [easy] Generally speaking, how does Nate Silver predict elections?

(h) [easy] How many Bayesians of import were there in 1979?

(i) [easy] What advice did Chernoff give to Susan Holmes? (Note: Susan Holmes was my undergraduate advisor).

(j) [easy] How did Rasmussen's team estimate the probability of a nuclear plant core meltdown?

(k) [easy] How did the Three Mile Island accident vindicate Rasmussen's committee report?

Problem 2

We will review classical *frequentist* concepts from "Math 241/242". Much of this can be drawn from lecture 14 first page.

(a) [easy] If $X_1, \ldots, X_n \overset{iid}{\sim} \mathcal{N}(\theta, \sigma^2)$ and $\bar{X} := \frac{1}{n}\sum_{i=1}^n X_i$, what is the distribution of the following:

$$\frac{\bar{X} - \theta}{\frac{\sigma}{\sqrt{n}}} \sim$$

(b) [easy] If $X_1, \ldots, X_n \overset{iid}{\sim} \mathcal{N}(\theta, \sigma^2)$ and $\bar{X} := \frac{1}{n}\sum_{i=1}^n X_i$, what is the distribution of $\bar{X}$ assuming σ^2 is known? This can be derived from (a) or found in your Math 241 notes.

(c) [easy] Write the definition of S^2, the r.v. which is the sample variance *estimator*. Hint: use capital letters.

(d) [easy] Write the definition of S, the sample standard deviation *estimator* (or standard error estimator — both terms are synonymous). Hint: use capital letters.

(e) [easy] Write the definition of s^2, the r.v. which is the sample variance *estimate*. Hint: use lowercase letters.

(f) [E.C.] [MA] This is for Alina: Prove that $\mathbb{E}[S^2] = \sigma^2$, i.e. S^2 is an unbiased estimator of what it seeks to estiamte, the true variance of the r.v. i.e. σ^2. Note that here you should not use the normality of the underlying r.v.'s (but for the other subparts of this question this is necessary). This vindicates the weird $n-1$ term in the bottom (which is part of Bessel's correction). The answer is eminently Googlable... but don't spoil the fun. Do this on a separate piece of paper as there is not enough room below.

(g) [easy] Write the density function of the r.v. $\frac{n-1}{\sigma^2}S^2$. I gave this in the notes in the beginning of lecture 14. You can find the density online here. Note that the proof of this is really difficult and involves linear algebra.

(h) [harder] Show that the χ^2_{n-1} density from above is actually a gamma density and find its parameters. All you need to do is match the density function with α and β on a gamma PDF.

(i) [harder] What is the distribution of S^2? Hint: use the fact that if $X \sim \chi^2_\nu$ and $c > 0$ then $cX \sim \text{Gamma}\left(\frac{\nu}{2},\, 2c\right)$.

(j) [harder] What is the distribution of $\left(\frac{n-1}{\sigma^2}S^2\right)^{-1} = \frac{\sigma^2}{(n-1)S^2}$? Hint: use the fact that if $X \sim \text{Gamma}\,(\alpha,\, \beta)$ then $\frac{1}{X} \sim \text{InvGamma}\left(\alpha,\, \frac{1}{\beta}\right)$.

(k) [easy] This answer is in the notes. If $X_1, \ldots, X_n \overset{iid}{\sim} \mathcal{N}\left(\theta,\, \sigma^2\right)$ and $\bar{X} := \frac{1}{n}\sum_{i=1}^n X_i$, what is the distribution of the following where S is defined as in (d):

$$\frac{\bar{X} - \theta}{\frac{S}{\sqrt{n}}} \sim$$

(l) [easy] Write the PDF of the general (also called noncentral) T distribution below. You need to use the notation given in class. You can look up the answer here or in the notes (lecture 14, bottom of page 4).

(m) [easy] What is the kernel of the non-central T?

(n) [harder] What is the distribution of $\bar{X}$ assuming σ^2 is *unknown*? This will differ from (b). Use the answer from part (k) above and the fact that $aT_\nu + c \sim T_\nu(c, a)$ which means that if you shift and scale a T with ν degrees of freedom, you get a noncentral T_ν with the new center and scaling as parameters.

Problem 3

Now we will move to the Bayesian normal-normal model for estimating σ^2.

(a) [easy] If $X_1, \ldots, X_n \overset{exch}{\sim} \mathcal{N}(\theta, \sigma^2)$ and X represents all $X_1, \ldots, X_n$, in HW6 6(b) you found the kernel of $\sigma^2 \mid X,\ \theta$. Show that this is the kernel of an inverse gamma. Use the $\hat{\sigma}^2$ substituion we did in class.

(b) [harder] Is $\hat{\sigma}^2$ equivalent to s^2? Why or why not?

(c) [harder] Why is using $\hat{\sigma}^2$ permitted in the setup in (a) but doesn't make sense in the frequentist setup?

(d) [easy] In class we looked at $\sigma^2 \sim \text{InvGamma}(\alpha,\, \beta)$ but we used a different parameterization. Write the different parameterization below and explain why this was done i.e. interpret the meaning of the two new parameters.

(e) [harder] Show that $\sigma^2 \mid X,\; \theta$ is distributed as an inverse gamma with the prior from (d) and find its parameters.

(f) [easy] What is the Jeffrey's prior for σ^2 (look in the notes and write it down — no need to prove it). Is it proper?

(g) [easy] Show that the Jeffrey's prior for σ^2 is an improper inverse gamma distribution and find its parameters. Note these parameters are not in the parameter space of a proper inverse gamma distribution.

(h) [easy] Under the Jeffrey's prior for σ^2, what is the posterior?

(i) [harder] You are in a milk manufacturing plant producing 1 quart cartons of whole milk. You are willing to assume that the nozzle emits 1 qt on average. In your previous job, you remember inspecting 3 cartons of which you saw 1.02, 0.97, 1.03 quarts of milk inside. Create a prior based on what you've seen in your previous job. This forces you to understand (d).

(j) [difficult] The company wishes to test if there's too much variability i.e. that there is more than $\sigma = 0.1$ variability. You take a sample of 10 and see 1.153, 1.045, 1.268, 1.333, 0.799, 1.075, 1.27, 1.07, 1.192 and 1.079 quarts. Find the p value. You can write the answer below as a function of `rinvgamma`, `qinvgamma` or `pinvgamma`. E.C. for computing it and testing this at $\alpha = 5\%$. You may want to use the `actuar` package (see here).

Problem 4

Now we will move to the Bayesian normal-normal model for estimating both the mean and variance and demonstrate similarities with the classical results.

(a) [easy] If $X_1, \ldots, X_n \overset{exch}{\sim} \mathcal{N}(\theta,\ \sigma^2)$ and X represents all $X_1, \ldots, X_n$, in HW6 6(c) you found the kernel of $\theta,\ \sigma^2 \mid X = x$. Do so again below but this time use the substitution that we made in class:

$$\sum_{i=1}^{n} (x_i - \theta)^2 = (n-1)s^2 + n(\bar{x} - \theta)^2$$

where s^2 is your answer from 2(e). We do this here because this substitution is important for what comes next.

(b) [harder] If $\mathbb{P}(\theta,\ \sigma^2) \propto \frac{1}{\sigma^2}$, show that this is a conjugate prior for the posterior of both the mean and variance, $\mathbb{P}(\theta,\ \sigma^2 \mid X)$. We called this two-dimensional distribution the "normal-inverse-gamma" distribution but we did not go into details about it.

(c) [easy] Using Bayes Rule, break up $\mathbb{P}(\theta,\ \sigma^2 \mid X)$ into two pieces.

(d) [easy] Using your answer from (c), explain how you can create samples $[\theta_s, \sigma_s^2]$ from the distribution $\mathbb{P}(\theta,\ \sigma^2 \mid X)$.

(e) [harder] Using these samples, how would you estimate $\mathbb{E}[\theta \mid X]$ and $\mathbb{E}[\sigma^2 \mid X]$?

(f) [difficult] [MA] Using these samples, how would you estimate $\mathbb{C}\text{orr}[\theta \mid X,\ \sigma^2 \mid X]$ i.e. the correlation between the posterior distributions of the two parameters?

(g) [easy] If $X_1, \ldots, X_n \overset{exch}{\sim} \mathcal{N}(\theta,\ \sigma^2)$ and $\theta \sim \mathcal{N}(\mu_0,\ \tau^2)$ write the distribution of $\theta \mid X, \sigma^2$. Hint: it's in the notes and it was HW6 6(d). Note this problem is independent of the other problems.

(h) [easy] Find $\mathbb{P}(\theta \mid X,\ \sigma^2)$ which is the first part of (c) above. Assume the prior from (b) and not from (g). This is in the notes. Note how this prior is equivalent to saying $\mathbb{P}(\theta) \propto 1$.

(i) [harder] Show that $\mathbb{P}(\sigma^2 \mid \theta,\ X)$ which is the second part of (c) above is an inverse gamma and find its parameters.

(j) [difficult] Show that $\mathbb{P}(\theta \mid X)$ is a non-central T distribution. You can reference 2(m) here. The answer is in the notes, but try to do it yourself. How does this compare to 2(n)?

(k) [difficult] Show that $\mathbb{P}(\sigma^2 \mid X)$ is an inverse gamma distribution and find its parameters.

(l) [harder] How does this compare to 2(j)? Note that $X \sim \text{InvGamma}(\alpha, \beta)$ then $cX \sim \text{InvGamma}\left(\alpha, \frac{\beta}{c}\right)$.

(m) [easy] Write down the distribution of $X^* \mid X$ which is in the notes (lec 14, page 6). Note that the answer I wrote down is for the non-informative prior only.

(n) [E.C.] [MA] Prove (m).

(o) [easy] Now consider the informative conjugate prior of

$$\mathbb{P}\left(\theta,\ \sigma^2\right) = \mathbb{P}\left(\theta \mid \sigma^2\right)\mathbb{P}\left(\sigma^2\right) = \mathcal{N}\left(\mu_0,\ \frac{\sigma^2}{m}\right)\text{InvGamma}\left(\frac{\nu_0}{2},\ \frac{\nu_0\sigma_0^2}{2}\right).$$

What is its kernel? Collect common terms and be neat.

(p) [harder] What is the distribution of $\theta \mid X,\ \sigma^2$? The new θ_p here has a nicer form than before.

(q) [easy] What is the distribution of $\theta \mid X$? No need to derive it; just copy it from the notes.

(r) [difficult] [MA] Prove (q) above. Or prove my notes are wrong.

(s) [easy] What is the distribution of $\sigma^2 \mid X$? No need to derive it; just copy it from the notes.

(t) [difficult] [MA] Prove (s) above. Or prove my notes are wrong.

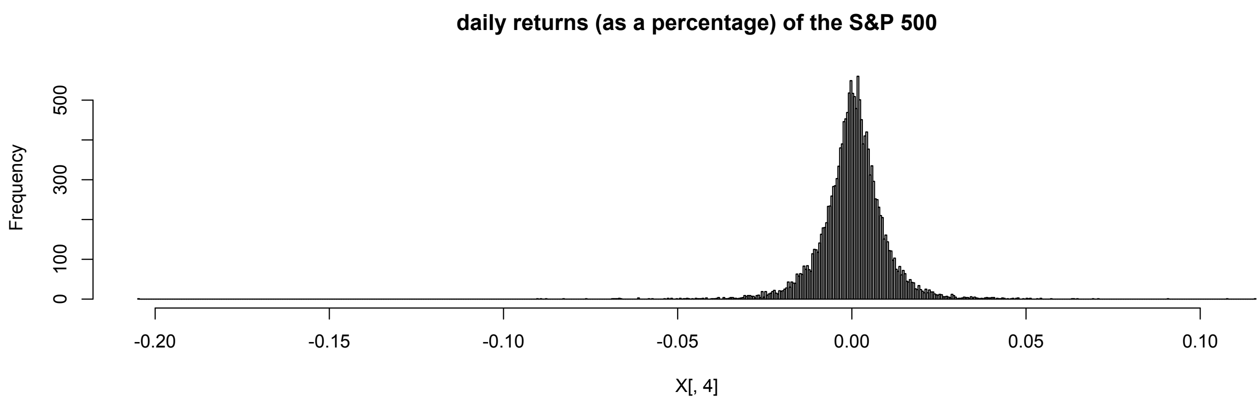
(u) [easy] Explain how to sample from the distribution of $X^* \mid X$. Hint: write it as a double integrel of two conditional distributions and a marginal distribution (all conditional on X).

(v) [easy] Guess the form of the distribution of $X^* \mid X$ given (m), (q) and (s).

Problem 5

We model the returns of S&P 500 here.

(a) [easy] Below are the 16,428 daily returns (as a percentage) of the S&P 500 dating back to January 4, 1950 and the code used to generate it. Does the data look normal? Yes/no



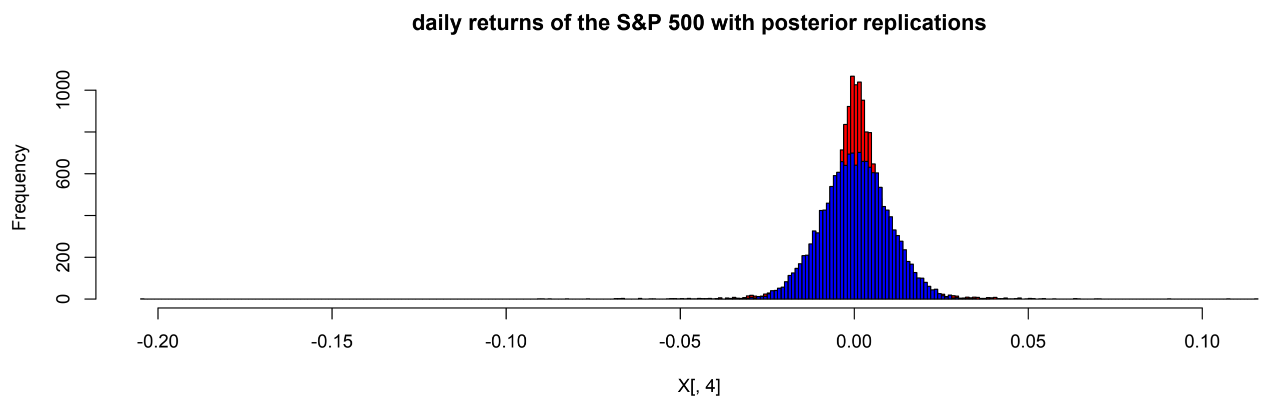

```
X = read.csv('sp_tot_ret_price_1950.csv')
n = nrow(X)
n
hist(X[,4], br = 1000,
  main = 'daily returns (as a percentage) of the S&P 500')
```

(b) [harder] Regardless of normality, do you think the data is exchangeable?

(c) [harder] Assume exchangeability and normality regardless of what you wrote in (a) and (b). The sample average is $\bar{x}$ =0.0003415 and the sample standard deviation is s =0.0096. Under a relatively objective prior, give a 95% credible region for the true mean daily return.

(d) [harder] Give a 95% credible region for *tomorrow's* return.

(e) [harder] Below is a sample of the same n from the posterior predictive distribution in blue graphed atop the actual data in red (and a zoomed view of the tail). Is our model a good fit? Yes/no and explain.




```
hist(X[,4], br = 450,
  main = 'daily returns of the S&P 500 with posterior replications',
  col='red', xlim = c(-.1, -.02), ylim = c(0, 150))
hist(rnorm(16428, 0.0003415, 0.0096), br = 100, add = TRUE,
  col = 'blue', xlim = c(-.1, -.02), ylim = c(0, 150))
```