MATH 341/641 Fall 2024 Homework #7

Professor Adam Kapelner

Due by email 11:59PM December 14

(this document last updated Sunday 1st December, 2024 at 3:40pm)

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 MATH 340 concepts: mixture distributions, poisson, gamma, extended negative binomial, normal, inverse gamma, students T.

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. "[MA]" are for those registered for 621 and extra credit otherwise.

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 5 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: ____

Distribution of r.v.	Quantile Function	PMF / PDF function	CDF function	Sampling Function
beta	`qbeta`(p, α, β)	`d-`(x, α, β)	`p-`(x, α, β)	`r-`(α, β)
betabinomial	`qbetabinom`(p, n, α, β)	`d-`(x, n, α, β)	`p-`(x, n, α, β)	`r-`(n, α, β)
gamma	`qgamma`(p, α, β)	`d-`(x, α, β)	`p-`(x, α, β)	`r-`(α, β)
ext negative binomial	`qnbinom`(p, r, θ)	`d-`(x, r, θ)	`p-`(x, r, θ)	`r-`(r, θ)
normal	`qnorm`(p, θ, σ)	`d-`(x, θ, σ)	`p-`(x, θ, σ)	`r-`(θ, σ)
inversegamma	`qinvgamma`(p, α, β)	`d-`(x, α, β)	`p-`(x, α, β)	`r-`(α, β)
Student's T	`qt.scaled`(p, k, μ, σ)	`d-`(x, k, μ, σ)	`p-`(x, k, μ, σ)	`r-`(k, μ, σ)

Table 1: Functions from `R` (in alphabetical order) that can be used on this assignment and exams. The hyphen in colums 3, 4 and 5 is shorthand notation for the full text of the r.v. which can be found in column 2.

Problem 1

We now discuss the theory of the normal-normal conjugate model. Assume the DGP: $X_1, \ldots, X_n \mid \theta, \sigma^2 \overset{iid}{\sim} \mathcal{N}(\theta,\, \sigma^2)$ and σ^2 known.

(a) [in the notes] Assume $f(\theta \,|\, \sigma^2) = \mathcal{N}(\mu_0,\, \sigma^2/n_0)$. Show that posterior distribution is normal and find its parameters.

(b) [in the notes] Provide pseudocount interpretations of μ_0 and n_0.

(c) [in the notes] Find the Bayesian point estimates as function of the data and prior's hyperparameters (i.e. $\hat{\hat{\theta}}^{\text{MMSE}}$, $\hat{\hat{\theta}}^{\text{MMAE}}$ and $\hat{\hat{\theta}}^{\text{MAP}}$).

(d) [in the notes] Show that $\hat{\theta}^{\text{MMSE}}$ is a shrinkage estimator and find ρ.

(e) [in the notes] What is the posterior distribution under Laplace's prior of indifference?

(f) [easy] Assuming $\sigma^2 = 1.3$, Laplace's prior and a dataset of $n = 10$ with values 0.48 0.39 1.29 1.02 1.55 -0.22 0.01 -0.52 -1.50 0.71, provide a Bayesian point estimate.

(g) [easy] Assuming the prior, σ^2 and the dataset from (f), provide a 95% CR for θ.

(h) [easy] Assuming the prior, σ^2 and the dataset from (f), provide notation that calculates the p value for a test of $H_a : \theta < 1$.

(i) [in the notes] Using the pseudocount interpretations of μ_0 and n_0, what is Haldane's prior of ignorance? Is it proper?

(j) [difficult] Derive the posterior predictive distribution $f(X_* \mid \boldsymbol{X}, \sigma^2)$ when $n_* = 1$. Try to do it yourself. Use kernels. If you get stuck, look in the notes from MATH 340.

Problem 2

We now discuss the theory of the normal-inverse-gamma conjugate model. Assume the DGP: $X_1, \ldots, X_n \mid \theta, \sigma^2 \stackrel{iid}{\sim} \mathcal{N}(\theta,\, \sigma^2)$ and θ known.

(a) [in the notes] Assume the Laplace prior, $f(\sigma^2 \,|\, \theta) \propto \mathbb{1}_{\sigma^2>0}$. Show that posterior is inverse gamma and find the posterior parameters.

(b) [in the notes] Show that posterior is inverse gamma (and find the posterior parameters) if $\mathbb{P}(\sigma^2) = \text{InvGamma}\left(\frac{n_0}{2},\, \frac{n_0\sigma_0^2}{2}\right)$.

(c) [in the notes] What is the pseudodata interpretation of the hyperparameters n_0 and σ_0^2?

(d) [in the notes] Based on the pseudodata interpretation of the hyperparameters n_0 and σ_0^2, what would Haldane's prior be and why?

(e) [in the notes] In the Laplace prior, what are the hyperparameters?

(f) [difficult] Why is the Laplace prior a bad idea to use in this modeling setting?

(g) [in the notes] Provide all three Bayesian point estimates for σ^2 given θ.

(h) [in the notes] Show that the $\hat{\theta}^{\text{MMSE}}$ is a linear shrinkage estimator. Is it valid for every inverse gamma prior?

(i) [harder] Show that the $\hat{\hat{\theta}}^{\text{MAP}}$ is a linear shrinkage estimator (i.e. a linear combination of the MLE and the prior mode). Is it valid for every inverse gamma prior?

(j) [in the notes] What is the predictive distribution $f(X^* \mid \boldsymbol{X}, \theta)$ if $n^* = 1$ and $\theta \sim \text{InvGamma}\left(\frac{n_0}{2}, \frac{n_0\sigma_0^2}{2}\right)$?

(k) [harder] Find a 95% posterior predictive interval (PI) for the next observation.

Problem 3

This question is about building models for the prices of cars sold at dealerships.

The 2016 Honda Accord sells at many different dealerships in New York City but sell it for more and some for less. We'll assume that the final negotiated price is distributed normally because it's most likely the sum of many different negotiation factors.

Our goal here is to determine the mean price at a certain car dealership in Astoria that people have been saying is "too cheap" and if it's too cheap, Honda corporate may wish to investigate.

(a) [harder] Assume that each Accord's price at the Astoria dealership is normal and $\overset{iid}{\sim}$ given the parameters. The nationwide variance for a Honda Accord selling price we're going to assume is $\sigma^2 = \$1000^2$. You and your colleague go down to the Astoria dealership undercover and ask to buy a Honda. After much negotiation, they will sell it to you for \$19,000 and they will sell it to your colleague for \$18,200 but they sense something suspicious so you hesitate to send another one of your guys down there to do another faux negotiation. Unfortunately, we're going to have to estimate the mean with just $x_1 = 19000$ and $x_2 = 18200$. What is your best guess of the mean price of Honda Accords sold here? Assume your prior from (a).

(b) [harder] Based on this data, we wish to test if this dealership is selling Honda Accords below the manufacturer sugested retail price (MSRP) of \$22,205 — if so, they would be subject to a fine. Calculate a p-value for this test below by using notation from Table 1 but do not solve numerically.

(c) [difficult] What is the probability I get a really good deal — that I can buy a car from these Astoria people for under \$17,000? Use the notation from Table 1 but do not solve numerically.

Problem 4

This question is about building a model to understand the accuracy of this beverage-filling machine pictured below:

This machine fills 12oz plastic bottles. There is no doubt the mean amount of liquid filled per bottle is 12oz as been determined by the final weights of pallets of filled bottles. But we are uncertain about the variance. We decide to do an experiment and select $n = 21$ bottles at random and measure the amount of liquid in each bottle. Here are the measurements:

```
12.00 12.05 11.98 11.66 12.05 11.92 12.03
12.23 12.36 11.57 12.04 12.10 11.99 12.47
12.57 11.83 12.20 12.48 12.14 12.14 12.74
```

Assume an $\overset{iid}{\sim}$ normal model.

(a) [easy] Find the MLE for σ^2.

(b) [harder] Under the Jeffrey's prior for σ^2, find the posterior of σ^2 by solving for the parameter values.

(c) [harder] Write an expression for the 95% left-sided credible region for σ^2. This is a one sided CR which will give the upper bound for the machine's variance (since the lower bound for σ^2 is zero).

(d) [harder] The bottles are actually 13.5oz. This means that you wish to test if $\sigma^2 > 0.352$ for if so, about $1/100{,}000$ of the bottles will be overfull and that's the tolerance of the factory. Write an expression for the Bayesian p-value of this test.

(e) [harder] Write an expression for the probability the next bottle has more than 13oz of liquid.

Problem 5

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) [in the notes] If $X_1, \ldots, X_n \overset{iid}{\sim} \mathcal{N}(\theta,\, \sigma^2)$ with θ, σ^2 both known, find the kernel of the posterior if $f(\theta,\, \sigma^2) \propto \frac{1}{\sigma^2}$. 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 $s^2 := \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$.

(b) [in the notes] Using Bayes Rule, break up the posterior into two pieces as we did in class. How are those two pieces distributed?

(c) [difficult] Using Bayes Rule, partition the posterior into two pieces differently than the previous question. How are those two pieces distributed?

(d) [harder] Using your answer from (b), explain in English how you can create S samples from the posterior distribution that look like $\{[\theta_1, \sigma_1^2], [\theta_2, \sigma_2^2], \ldots, [\theta_S, \sigma_S^2]\}$.

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

(f) [difficult] Using these samples, how would you estimate a 95% CR for θ?

(g) [difficult] Using these samples, how would you obtain a p-val for testing if $\sigma^2 > 1.364$?

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

(i) [difficult] Explain how to sample from the posterior predictive distribution of for the next observation.

(j) [in the notes] [MA] Show that $\mathbb{P}(\theta \mid X)$ is a non-standard T distribution and find its parameters. Assume the prior $\mathbb{P}(\theta,\ \sigma^2) \propto \frac{1}{\sigma^2}$.

(k) [in the notes] [MA] Show that $\mathbb{P}(\sigma^2 \mid X)$ is an inverse gamma and find its parameters. Assume the prior $\mathbb{P}(\theta,\ \sigma^2) \propto \frac{1}{\sigma^2}$.

(l) [easy] How is $X^* \mid \boldsymbol{X}$ distributed assuming the prior $f(\theta, \sigma^2) \propto \frac{1}{\sigma^2}$?

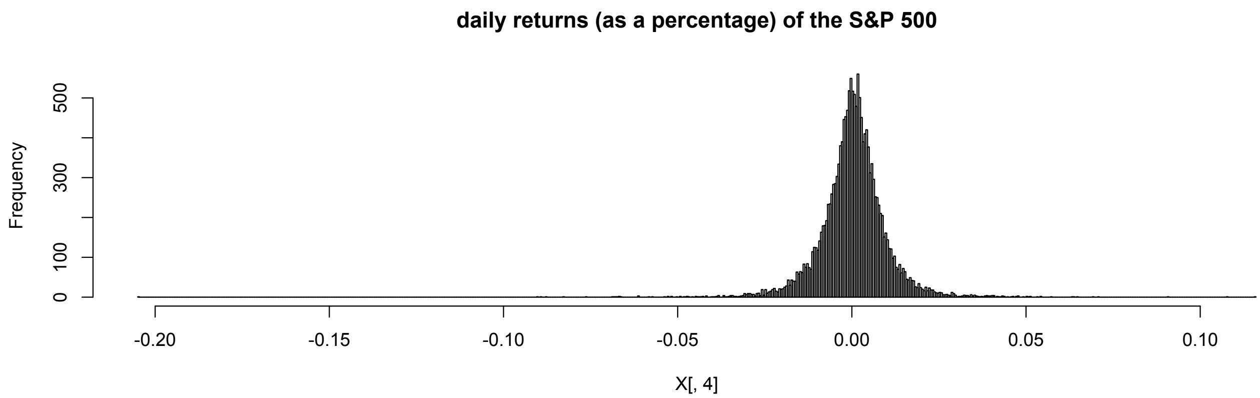
(m) [harder] [MA] Now consider the informative conjugate prior of $\theta \,|\, \sigma^2 \sim \mathcal{N}\left(\mu_0,\, \frac{\sigma^2}{m}\right)$ and $\sigma^2 \sim \text{InvGamma}\left(\frac{n_0}{2},\, \frac{n_0\sigma_0^2}{2}\right)$. Find the posterior and demonstrate it that the normal-inverse gamma is conjugate for the normal likelihood with both mean and variance unknown. This is what I did *not* do in class.

(n) [in the notes] Now consider the informative conjugate prior of $\theta \sim \mathcal{N}\left(\mu_0,\, \tau^2\right)$ independent of $\sigma^2 \sim \text{InvGamma}\left(\frac{n_0}{2},\, \frac{n_0\sigma_0^2}{2}\right)$. Find the kernel of the posterior and demonstrate that the $\mathbb{P}\left(\theta \mid \boldsymbol{X}, \sigma^2\right)$ is normal but the $\mathbb{P}\left(\sigma^2 \mid \boldsymbol{X}\right)$ piece is not any known distribution.

Problem 6

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



(b) [harder] Do you think the data is $\stackrel{iid}{\sim}$? Explain.

(c) [harder] Assume $\stackrel{iid}{\sim}$ normal data 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 an objective prior, give a 95% credible region for the true mean daily return.

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