MATH 341 / 650.3 Spring 2021 Homework #5

Professor Adam Kapelner

Due by email, Sunday 11:59PM, April 18, 2021

(this document last updated Friday 9th April, 2021 at 1:57pm)

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 Poisson-Gamma conjugate model, the extended negative binomial, the normal-normal conjugate model and read chapter 15 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 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, chapter 15.

(a) [easy] During the H-Bomb search in Spain and its coastal regions, RAdm. William Guest was busy sending ships here, there and everywhere even if the ships couldn't see the bottom of the ocean. How did Richardson use those useless searches?

(b) [harder] When the Navy was looking for the *Scorpion* submarine, they used Monte Carlo methods (which we will see in class soon). How does the description of these methods by Richardson (p199) remind you of the "sampling" techniques to approximate integrals we did in class?

(c) [harder] What is a Kalman filter? Read about it online and write a few descriptive sentences.

(d) [harder] Where do frequentist methods practically break down? (end of chapter 15)

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)
binomial	`qbinom`(p, n, θ)	`d-`(x, n, θ)	`p-`(x, n, θ)	`r-`(n, θ)
exponential	`qexp`(p, θ)	`d-`(x, θ)	`p-`(x, θ)	`r-`(θ)
gamma	`qgamma`(p, α, β)	`d-`(x, α, β)	`p-`(x, α, β)	`r-`(α, β)
inversegamma	`qinvgamma`(p, α, β)	`d-`(x, α, β)	`p-`(x, α, β)	`r-`(α, β)
negative-binomial	`qnbinom`(p, r, θ)	`d-`(x, r, θ)	`p-`(x, r, θ)	`r-`(r, θ)
normal (univariate)	`qnorm`(p, θ, σ)	`d-`(x, θ, σ)	`p-`(x, θ, σ)	`r-`(θ, σ)
poisson	`qpois`(p, θ)	`d-`(x, θ)	`p-`(x, θ)	`r-`(θ)
T (standard)	`qt`(p, ν)	`d-`(x, ν)	`p-`(x, ν)	`r-`(ν)
T (nonstandard)	`qt.scaled`(p, ν, μ, σ)	`d-`(x, ν, μ, σ)	`p-`(x, ν, μ, σ)	`r-`(ν, μ, σ)
uniform	`qunif`(p, a, b)	`d-`(x, a, b)	`p-`(x, a, b)	`r-`(a, b)

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 2

We will ask some basic problems on the Gamma-Poisson conjugate model. Please review HW4 #5 as this is a continuation of that problem. Feel free to answer using functions from Table 1.

(a) [harder] Prove that the Poisson likelihood for n observations, i.e. $X_1, \ldots, X_n; \theta \overset{iid}{\sim} \text{Poisson}(\theta)$, with a gamma prior yields a gamma posterior and find its parameters.

(b) [easy] Now that you see the posterior, provide a pseudodata interpretation for both hyperparameters.

(c) [harder] Find the Bayesian point estimates as function of the data and prior's hyper-parameters (i.e. $\hat{\theta}_{\text{MMSE}}$, $\hat{\theta}_{\text{MMAE}}$ and $\hat{\theta}_{\text{MAP}}$).

(d) [harder] If $X_1, \ldots, X_n; \theta \overset{iid}{\sim} \text{Poisson}\left(\theta\right)$, find $\hat{\theta}_{\text{MLE}}$.

(e) [harder] Demonstrate that $\hat{\theta}_{\text{MMSE}}$ is a shrinkage estimator and find ρ.

(f) [harder] Demonstrate that $\mathbb{P}\left(\theta\right) \propto 1$ is improper.

(g) [easy] [MA] Demonstrate that $\mathbb{P}(\theta) \propto 1$ can be created by using an improper Gamma distribution (i.e. a Gamma distribution with parameters that are not technically in its parameter space and thereby does not admit a distribution function).

(h) [harder] Find Jeffrey's prior for the Poisson likelihood model. Try to do it yourself.

(i) [easy] What is the equivalent of the Haldane prior in the Binomial likelihood model for the Poisson likelihood model? Use an interpretation of pseudocounts to explain.

(j) [difficult] If $\mathbb{P}(\theta) = \text{Gamma}(\alpha, \beta)$ where $\alpha \in \mathbb{N}$, prove that prior predictive distribution is $\mathbb{P}(X) = \text{NegBin}(r, p) := \binom{x+r-1}{r-1}(1-p)^{x-r}p^r$ where $p = \beta/(\beta+1)$ and $r = \alpha$. This is a little bit different than that posterior predictive distribution derivation we did in class but mostly the same.

(k) [harder] Why is the extended negative binomial r.v. also known as the gamma-Poisson mixture distribution? Why is it also called the "overdispersed Poisson"?

(l) [harder] If you observe $0, 3, 2, 4, 2, 6, 1, 0, 5$, give a 90% CR for θ. Pick an principled objective (uninformative) prior.

(m) [harder] Using the data and the prior from (l), test if $\theta < 2$.

(n) [difficult] Using the data and the prior from (l), find the probability the next observation will be a 7. Leave in exact form using Table 1's notation.

(o) [easy] Use the R calculator (if you don't have it on your computer, go to `rdrr.io`) to compute it to the nearest two significat digits.

(p) [difficult] [MA] We talked about that the negative binomial is an "overdispersed" Poisson. Show that this negative binomial converges to a Poisson as $n \to \infty$ by showing PMF convergence.

(q) [E.C.] [MA] Find the joint posterior predictive distribution for m future observations. I couldn't find the answer to this myself nor compute the integral.

Problem 3

We now discuss the theory of the normal-normal conjugate model. Assume

$$\mathcal{F} : X_1, \ldots, X_n \mid \theta, \sigma^2 \overset{iid}{\sim} \mathcal{N}\left(\theta,\, \sigma^2\right)$$

and "X", is the usual shorthand for all $X_1, \ldots, X_n$ samples.

(a) [easy] Show that the kernel of the normal distribution is, $\mathbb{P}\left(X_1 \mid \theta,\ \sigma^2\right) \propto k(X_1 \mid \theta,\, \sigma^2) = e^{ax}e^{-bx^2}$ and solve for the values of a and b as functions of θ and σ^2.

(b) [difficult] Assume $\mathbb{P}\left(\theta\right) = \mathcal{N}\left(\mu_0,\, \tau^2\right)$. Using the result from (a), show that posterior distribution is normal. Try to do it yourself and only copy from the notes if you must.

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

(d) [harder] On a previous homework we showed that if $\mathcal{F} : X_1, \ldots, X_n \mid \theta, \sigma^2 \overset{iid}{\sim} \mathcal{N}(\theta, \sigma^2)$ then $\hat{\theta}_{\text{MLE}} = \bar{x}$. Assuming this MLE, show that $\hat{\theta}_{\text{MMSE}}$ is a shrinkage estimator and find ρ.

(e) [harder] Rederive the posterior distribution under Laplace's prior of indifference. This is easier than the exercise in (b) as there are less terms.

(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 $\theta > 1$.

(i) [harder] Rederive Jeffrey's prior. Is it proper?

(j) [harder] Now let $\tau^2 := \sigma^2/n_0$ which is a reparameterization from τ^2 to n_0. Substitute this change into the posterior distribution from (b) to derive the posterior distribution under this reparemterization.

(k) [easy] Provide pseudocount interpretations of μ_0 and n_0.

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

(m) [easy] If the prior is $\mathbb{P}(\theta) = \mathcal{N}(\mu_0, \sigma^2/n_0)$, write the integral that will compute the posterior predictive distribution $\mathbb{P}(X_* \mid X, \sigma^2)$ when $n_* = 1$.

(n) [difficult] Compute the posterior predictive distribution $\mathbb{P}\left(X_* \mid X, \sigma^2\right)$ when $n_* = 1$ using the integral from (m). Try to do it yourself and if you get stuck, look in the notes. Note the change of parameterization from the notes!