MATH 341 / 650.3 Spring 2020 Homework #5

Professor Adam Kapelner

Due by email, Sunday 11:59PM, March 29, 2020

(this document last updated Friday 20$^{\text{th}}$ March, 2020 at 6:01pm)

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 Geometric-Poisson conjugate model with negative binomial posterior predictive distribution. Also ch11-14 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, chapters 11–14.

(a) [easy] Did Savage like Shlaifer? Yes / No and why?

(b) [easy] How did Neyman-Pearson approach statistical decision theory? What is the weakness to this approach? (p145)

(c) [easy] Who popularized "probability trees" (and "tree flipping") similar to exercises we did in Math 241?

(d) [easy] Where are Bayesian methods taught more widely than any other discipline in academia?

(e) [easy] Despite the popularity of his Bayesian textbook on business decision theory, why didn't Schlaifer's Bayesianism catch on in the real world of business executives making decisions?

(f) [easy] Why did the pollsters fail (big time) to predict Harry Truman's victory in the 1948 presidential election?

(g) [easy] When does the diference between Bayesianism and Frequentism grow "immense"?

(h) [easy] How did Mosteller demonstrate that Madison wrote the 12 Federalist papers of unknown authorship?

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

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

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

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

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

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

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

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

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

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

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

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.

(a) [easy] Write the PDF of $\theta \sim \text{Gamma}(\alpha, \beta)$ which is the gamma distribution with the standard parameterization and notated with the hyperparameters we used in class.

(b) [easy] What is the support and parameter space?

(c) [easy] What is the expectation and standard error and mode?

(d) [easy] Draw four different pictures of different hyperparameter combinations to demonstrate this model's flexibility

(e) [harder] Prove that the Poisson likelihood for $n = 1$ with a gamma prior yields a gamma posterior and find its parameters.

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

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

(h) [harder] 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 $\hat{\theta}_{\text{MAP}}$).

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

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

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

(l) [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).

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

(n) [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.

(o) [harder] Prove that posterior predictive distribution for the next Poisson realization (i.e. $n^* = 1$) given n observed Poisson realizations is negative binomially distributed and show its parameters are $p = \beta/(\beta + 1)$ and $r = \alpha$ for $\alpha \in \mathbb{N}$.

(p) [harder] If $\alpha \notin \mathbb{N}$, create an "extended negative binomial" r.v. and find its PMF. You can copy from Wikipedia.

(q) [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"?

(r) [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.

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

(t) [harder] Using the data and the prior from (s), find the probability the next observation will be a 7. Leave in exact form then use a calculator to compute it to the nearest two significat digits.

(u) [difficult] [MA] We talked about that the negative binomial is an "overdispersed" Poisson. Show that the negative binomial converges to a Poisson.

(v) [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

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

but where you see only "X", this is shorthand for all n samples.

(a) [easy] What is the kernel of $\theta \mid X,\ \sigma^2$?

(b) [difficult] Show that posterior of $\theta \mid X,\ \sigma^2$ is normal if $\theta \sim \mathcal{N}(\mu_0,\, \tau^2)$. Try to do it yourself and only copy from the notes if you have to.

(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 $X_1, \ldots, X_n \mid \theta, \sigma^2 \overset{iid}{\sim} \mathcal{N}(\theta, \sigma^2)$ then $\hat{\theta}_{\text{MLE}} = \bar{x}$. Show that $\hat{\theta}_{\text{MMSE}}$ is a shrinkage estimator and find ρ.

(e) [harder] Setup the integral to find $\mathbb{P}(X_* \mid X)$ where $n_* = 1$ but don't solve.