# MATH 341 / 650.3 Spring 2018 Homework #4

## Professor Adam Kapelner

## Due Friday 11:59PM, April 13, 2018 under the door of KY604

(this document last updated Thursday 12$^{\text{th}}$ April, 2018 at 1:43pm)

**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 the Gamma-Poisson and Normal-Normal models and read the relevant sections of 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 650.3 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 L&#x41;T&#x45;X. 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$(p, \alpha, \beta)$ | d-$(x, \alpha, \beta)$ | p-$(x, \alpha, \beta)$ | r-$(\alpha, \beta)$ |
| betabinomial | qbetabinom$(p, n, \alpha, \beta)$ | d-$(x, n, \alpha, \beta)$ | p-$(x, n, \alpha, \beta)$ | r-$(n, \alpha, \beta)$ |
| betanegativebinomial | qbeta_nbinom$(p, r, \alpha, \beta)$ | d-$(x, r, \alpha, \beta)$ | p-$(x, r, \alpha, \beta)$ | r-$(r, \alpha, \beta)$ |
| binomial | qbinom$(p, n, \theta)$ | d-$(x, n, \theta)$ | p-$(x, n, \theta)$ | r-$(n, \theta)$ |
| exponential | qexp$(p, \theta)$ | d-$(x, \theta)$ | p-$(x, \theta)$ | r-$(\theta)$ |
| gamma | qgamma$(p, \alpha, \beta)$ | d-$(x, \alpha, \beta)$ | p-$(x, \alpha, \beta)$ | r-$(\alpha, \beta)$ |
| geometric | qgeom$(p, \theta)$ | d-$(x, \theta)$ | p-$(x, \theta)$ | r-$(\theta)$ |
| inversegamma | qinvgamma$(p, \alpha, \beta)$ | d-$(x, \alpha, \beta)$ | p-$(x, \alpha, \beta)$ | r-$(\alpha, \beta)$ |
| negative-binomial | qnbinom$(p, r, \theta)$ | d-$(x, r, \theta)$ | p-$(x, r, \theta)$ | r-$(r, \theta)$ |
| normal (univariate) | qnorm$(p, \theta, \sigma)$ | d-$(x, \theta, \sigma)$ | p-$(x, \theta, \sigma)$ | r-$(\theta, \sigma)$ |
| poisson | qpois$(p, \theta)$ | d-$(x, \theta)$ | p-$(x, \theta)$ | r-$(\theta)$ |
| T (standard) | qt$(p, \nu)$ | d-$(x, \nu)$ | p-$(x, \nu)$ | r-$(\nu)$ |
| T (nonstandard) | qt.scaled$(p, \nu, \mu, \sigma)$ | d-$(x, \nu, \mu, \sigma)$ | p-$(x, \nu, \mu, \sigma)$ | r-$(\nu, \mu, \sigma)$ |
| 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.

We will ask some basic problems on the Gamma-Poisson conjugate model.

(a) [easy] If $X \sim \text{Poisson}(\theta)$ and prior $\mathbb{P}(\theta)$, what is the kernel of the posterior? Leave the prior in as "$\mathbb{P}(\theta)$".

(b) [easy] Write the PDF of $\theta$ which is the gamma distribution with the standard hyperparameters we used in class.

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

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

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

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

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

(h) [harder] For the Poisson likelihood for $n$ observations with a gamma prior find $\hat{\theta}_{\mathrm{MMSE}}$, $\hat{\theta}_{\mathrm{MAE}}$ and $\hat{\theta}_{\mathrm{MAP}}$.

(i) [harder] [MA] Demonstrate that $\hat{\theta}_{\mathrm{MMSE}}$ is a shrinkage estimator and find $\rho$.

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

(k) [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).

(l) [easy] What is the Jeffrey's prior for the Poisson likelihood model? Do not rederive. Just copy.

(m) [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.

(n) [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 $\alpha \in \mathbb{N}$.

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

(p) [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"?

(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.

(r) [harder] If you observe $0, 3, 2, 4, 2, 6, 1, 0, 5$, give a 95% CR for $\theta$. Pick an objective (uninformative) prior.

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

(t) [E.C.] We talked about that the negative binomial is an "overdispersed" Poisson. Show that the negative binomial converges to a Poisson.

## Problem 3

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

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

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

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

(d) [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.

(e) [difficult] Show that predictive distribution of $X^* \mid X$, $\sigma^2$ is normal if $\theta \sim \mathcal{N}\left(\mu_0,\, \sigma^2/n_0\right)$ by solving the integral and not using the convolution.

## Problem 4

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) [easy] Assume that each Accord's price at the Astoria dealership is normal and $\overset{iid}{\sim}$ given the parameters. Is this a good model? Why or why not? There is no "correct" answer here but I expect you to defend whatever answer you write using the concepts we discussed in class.

(b) [easy] Despite what you wrote in (b), assume $\overset{iid}{\sim}$ for the rest of the problem. The nationwide variance for a Honda Accord selling price we're going to assume is $\sigma^2 = \$1000^2$, an assumption we will relax later, in the next homework. Given a sample with average $\bar{x}$ and sample size $n$, what is the distribution of the mean price of a car from this shady Astoria dealership? Assume an uninformative prior of your choice but ensure to explicitly state it.

(c) [easy] 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). Compute explicitly as a number rounded to two decimals.

(d) [easy] What is the shrinkage value (which we have been denoting $\rho$) for this estimate? Compute explicitly as a number rounded to two decimals.

(e) [easy] 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.

(f) [easy] 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.