

MATH 341 / 650.3 Spring 2020 Homework #5

Frank Palma Gomez

Sunday 29th March, 2020

Problem 1

These are questions about McGrayne's book, chapters 11–14.

- (a) [easy] Did Savage like Schlaifer? 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 difference 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	<code>qbeta(p, α, β)</code>	<code>d-(x, α, β)</code>	<code>p-(x, α, β)</code>	<code>r-(α, β)</code>
betabinomial	<code>qbetabinom(p, n, α, β)</code>	<code>d-(x, n, α, β)</code>	<code>p-(x, n, α, β)</code>	<code>r-(n, α, β)</code>
binomial	<code>qbinom(p, n, θ)</code>	<code>d-(x, n, θ)</code>	<code>p-(x, n, θ)</code>	<code>r-(n, θ)</code>
exponential	<code>qexp(p, θ)</code>	<code>d-(x, θ)</code>	<code>p-(x, θ)</code>	<code>r-(θ)</code>
gamma	<code>qgamma(p, α, β)</code>	<code>d-(x, α, β)</code>	<code>p-(x, α, β)</code>	<code>r-(α, β)</code>
inversegamma	<code>qinvgamma(p, α, β)</code>	<code>d-(x, α, β)</code>	<code>p-(x, α, β)</code>	<code>r-(α, β)</code>
negative-binomial	<code>qnbinom(p, r, θ)</code>	<code>d-(x, r, θ)</code>	<code>p-(x, r, θ)</code>	<code>r-(r, θ)</code>
normal (univariate)	<code>qnorm(p, θ, σ)</code>	<code>d-(x, θ, σ)</code>	<code>p-(x, θ, σ)</code>	<code>r-(θ, σ)</code>
poisson	<code>qpois(p, θ)</code>	<code>d-(x, θ)</code>	<code>p-(x, θ)</code>	<code>r-(θ)</code>
T (standard)	<code>qt(p, ν)</code>	<code>d-(x, ν)</code>	<code>p-(x, ν)</code>	<code>r-(ν)</code>
T (nonstandard)	<code>qt.scaled(p, ν, μ, σ)</code>	<code>d-(x, ν, μ, σ)</code>	<code>p-(x, ν, μ, σ)</code>	<code>r-(ν, μ, σ)</code>
uniform	<code>qunif(p, a, b)</code>	<code>d-(x, a, b)</code>	<code>p-(x, a, b)</code>	<code>r-(a, b)</code>

Table 1: Functions from R (in alphabetical order) that can be used on this assignment and exams. The hyphen in columns 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.

$$\mathbb{P}(\theta) =$$

- (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. $X_1, \dots, X_n; \theta \stackrel{iid}{\sim} \text{Poisson}(\theta)$, 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 $X_1, \dots, X_n; \theta \stackrel{iid}{\sim} \text{Poisson}(\theta)$, 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 significant 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

$$X_1, \dots, X_n \mid \theta, \sigma^2 \stackrel{iid}{\sim} \mathcal{N}(\theta, \sigma^2)$$

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 $X_1, \dots, X_n \mid \theta, \sigma^2 \stackrel{iid}{\sim} \mathcal{N}(\theta, \sigma^2)$ 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.