MATH 341 / 650.3 Spring 2020 Homework #6

Frank Palma Gomez

Monday 20th April, 2020

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?

William Guest used a grid based on location to compute the probability that a bomb is in any location of the grid.

(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?

Using Bayes, he assigned probabilities to generate new probabilities and continued repeating the process for 10,000 times.

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

Kalman filter is an algorithm that produces estimates of unknown variables over time. It does so by estimating the joint probability for a variable for a period of time. As the algorithm continues to predict, it updates to better its prediction.

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

Due to its objectivity, alternative methods cannot compare.

Problem 2

We continue with our discussion of the theory of the normal-normal conjugate model. Let $X_1, \ldots, X_n \mid \theta, \sigma^2 \overset{iid}{\sim} \mathcal{N}(\theta, \sigma^2)$ where "X", this is shorthand for all n samples.

(a) [difficult] If the prior is $\theta \sim \mathcal{N}(\mu_0, \sigma^2/n_0)$, find $\mathbb{P}(X_* \mid X, \sigma^2)$ where $n_* = 1$. Try to do it yourself and if you get stuck, look in the notes. Note the change of parameterization from the notes!

$$\begin{aligned}
\mathbb{P}\left(X_* \mid X, \sigma^2\right) &= \int_\Theta \mathbb{P}\left(X_* \mid \theta, \sigma^2\right)\mathbb{P}\left(\theta \mid \sigma^2\right) d\theta \\
&= \int_\mathbb{R} \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}(X_* - \theta)^2\right)\frac{1}{\sqrt{2\pi\sigma_p^2}}\exp\left(-\frac{1}{2\sigma_p^2}(\theta - \theta_p)^2\right) d\theta \\
&= \int_\mathbb{R} e^{(\frac{X_*}{\sigma^2} + \frac{\theta_p}{\sigma_p^2})\theta - (\frac{1}{2\sigma^2} + \frac{1}{2\sigma_p^2})\theta^2} d\theta \\
&\propto \mathcal{N}\left(\frac{a}{2b}, \frac{1}{2b}\right)
\end{aligned}$$

Where $a = \frac{X_*}{\sigma^2} + \frac{\theta_p}{\sigma_p^2}$ and $b = \frac{1}{2\sigma^2} + \frac{1}{2\sigma_p^2}$

$$\begin{aligned}
\mathcal{N}\left(\frac{a}{2b}, \frac{1}{2b}\right) &= \frac{1}{2\pi\frac{1}{2b}} e^{\frac{1}{2\frac{1}{2b}}(\theta - \frac{a}{2b})^2} \\
&= \frac{1}{\sqrt{\frac{\pi}{b}}} e^{-b\theta^2 + a\theta + \frac{a^2}{4b}} \\
&= \frac{1}{\sqrt{\frac{\pi}{b}}} e^{\frac{a^2}{4b}} e^{a\theta - b\theta^2}
\end{aligned}$$

Therefore,

$$\mathbb{P}\left(X_* \mid X, \sigma^2\right) = e^{\frac{X_*^2}{2\sigma^2}} \sqrt{\frac{\pi}{\frac{1}{2\sigma^2} + \frac{1}{2\sigma_p^2}}} e^{\frac{-(\frac{X_*}{\sigma^2} + \frac{\theta_p}{\sigma_p^2})^2}{\frac{2}{\sigma^2} + \frac{2}{2\sigma_p^2}}} \int_\mathbb{R} e^{(\frac{X_*}{\sigma^2} + \frac{\theta_p}{\sigma_p^2})\theta - (\frac{1}{2\sigma^2} + \frac{1}{2\sigma_p^2})\theta^2} d\theta$$

(b) [E.C.] Find the predictive distribution of $X^* \mid X,\ \sigma^2\ n^* > 1$ and $\theta \sim \mathcal{N}(\mu_0,\ \sigma^2/n_0)$.

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

$$\text{InvGamma}\left(\frac{n_0}{2}, \frac{n_0\sigma_0^2}{2}\right)$$

(d) [easy] What is the parameter space of the inverse gamma r.v.? What is its mean? Mode? Is there a formula for the median in closed form?

Where $\alpha, \beta > 0$

- $\mathbb{E}[Y] = \frac{\beta}{\alpha - 1}$ if $\alpha > 1$
- $\text{Mode}[Y] = \frac{\beta}{\alpha+1}$
- $\text{median}[Y] = \texttt{qinvgamma}(0.5, \alpha, 1)$

(e) [difficult] Show that posterior of $\sigma^2 \mid X,\ \theta$ is inverse gamma (and find the posterior parameters) if $\mathbb{P}(\theta) \propto 1$. Try to do it yourself and only copy from the notes if you have to.

(f) [difficult] Show that posterior of $\sigma^2 \mid X,\ \theta$ is inverse gamma (and find the posterior parameters) if $\theta \sim \text{InvGamma}\left(\frac{n_0}{2},\ \frac{n_0\sigma_0^2}{2}\right)$. Try to do it yourself and only copy from the notes if you have to.

$$\begin{aligned}
\mathbb{P}\left(\sigma^2 \mid X, \theta\right) &\propto (\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\Sigma(x_i-\theta)^2} \mathbb{P}\left(\sigma^2 \mid \theta\right) \\
&= (\sigma^2)^{-\frac{n}{2}} e^{\frac{n\sigma^2_{MLE}/2}{\sigma^2}} \left((\sigma^2)^{-\frac{n_0}{2}-1} e^{\frac{n_0\sigma_0^2/2}{\sigma^2}}\right) \\
&= (\sigma^2)^{-\frac{n+n_0}{2}-1} e^{-\left(\frac{(n\sigma^2_{MLE}+n_0\sigma_0^2)/2}{\sigma^2}\right)} \\
&\propto \text{InvGamma}\left(\frac{n+n_0}{2},\ \frac{n\sigma^2_{MLE}+n_0\sigma_0^2}{2}\right)
\end{aligned}$$

(g) [easy] What is the pseudodata interpretation of the hyperparameters n_0 and σ_0^2?

- n_0: number of pseudo-observations
- σ_0^2: variance of the pseudo-observations

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

If n_0 then its prior is,

$$\mathbb{P}\left(\sigma^2 \mid \theta\right) = \text{InvGamma}\left(0,\ 0\right)$$

then its posterior is,

$$\mathbb{P}\left(\sigma^2 \mid \theta, x\right) = \text{InvGamma}\left(\frac{n}{2},\ \frac{n\sigma^2_{MLE}}{2}\right)$$

only proper when $n \geq 1$

(i) [difficult] In the Laplace prior, what are the hyperparameters? Does this make sense?

$$\mathbb{P}\left(\sigma^2 \mid \theta, x\right) \propto \mathbb{P}\left(x \mid \sigma^2, \theta\right) \propto (\sigma^2)^{-(n/2-1)-1} e^{\frac{n\hat{\sigma^2}_{MLE}}{\sigma^2}} \propto \text{InvGamma}\left(\frac{n-2}{2}, \frac{n\hat{\sigma^2}_{MLE}}{2}\right)$$

only proper when $n \geq 3$

(j) [harder] [MA] Find the Jeffrey's Prior for inference on σ^2 if θ is known.

$$\begin{aligned}
\ell(\sigma^2; \theta, x) &= -\frac{n}{2} ln(2\pi) - \frac{n}{2} ln(\sigma^2) - \frac{1}{2\sigma^2} \Sigma(x_i - \theta)^2 \\
\ell'(\sigma^2; \theta, x) &= -\frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2(\sigma^2)^2} \Sigma(x_i - \theta)^2 \\
-\ell'(\sigma^2; \theta, x) &= -\frac{n}{2}\frac{1}{(\sigma^2)^2} + \frac{1}{(\sigma^2)^3} \Sigma(x_i - \theta)^2 \\
\boldsymbol{I}(\sigma^2, \theta) &= \mathbb{E}\left[\ell''(\sigma^2; \theta, x)\right] = -\frac{n}{2}\frac{1}{(\sigma^2)^2} + \frac{1}{(\sigma^2)^3} \Sigma \mathbb{E}_x[x_i - \theta]^2 \\
\boldsymbol{I}(\sigma^2, \theta) &= -\frac{n}{2}\frac{1}{(\sigma^2)^2} + \frac{1}{(\sigma^2)^3} n\sigma^2 = n\Big(-\frac{1}{2(\sigma^2)^2} + \frac{1}{(\sigma^2)^2}\Big) = \frac{1}{2(\sigma^2)^2} = \frac{n}{2}(\sigma^2)^{-2} \\
P_J(\sigma^2 \mid \theta) &\propto \sqrt{\boldsymbol{I}(\sigma^2; \theta)} = \sqrt{\frac{n}{2}(\sigma^2)^{-2}} \propto (\sigma^2)^{-2} \propto \text{InvGamma}\,(0,\, 0)
\end{aligned}$$

(k) [easy] Provide all three Bayesian estimates for σ^2.

$$\begin{aligned}
\hat{\sigma^2}_{MMSE} &= \frac{n\hat{\sigma^2}_{MLE} + n_0\sigma_0^2}{n + n_0 - 2} \\
\hat{\sigma^2}_{MAP} &= \frac{n\hat{\sigma^2}_{MLE} + n_0\sigma_0^2}{n + n_0 + 2} \\
\hat{\sigma^2}_{MMAE} &= qinvgamma\big(0.5, \frac{n + n_0}{2}, \frac{n\hat{\sigma^2}_{MLE} + n_0\sigma_0^2}{2}\big)
\end{aligned}$$

(l) [harder] Show that the $\hat{\theta}_{\text{MMSE}}$ is a linear shrinkage estimator.

$$\hat{\sigma^2_{MMSE}} = \frac{n\hat{\sigma^2}_{MLE}}{n + n_0 - 2} + \frac{n_0\sigma_0^2}{n + n_0 - 2}\Big(\frac{n_0 - 2}{n_0 - 2}\Big)$$

$$
\begin{aligned}
&= \left(\frac{n}{n+n_0-2}\right)\hat{\sigma^2}_{MLE} + \frac{n_0-2}{n+n_0-2}\mathbb{E}[\sigma^2] \\
&= (1-\rho)\hat{\sigma^2}_{MLE} + \rho\mathbb{E}[\sigma^2]
\end{aligned}
$$

(m) [difficult] [MA] Show that predictive distribution of $X^* \mid X,\ \theta$ is non-standard Student's t distribution if $n^* = 1$ and $\theta \sim \text{InvGamma}\left(\frac{n_0}{2},\ \frac{n_0\sigma_0^2}{2}\right)$ by solving the integral. Try to do it yourself and only copy from the notes if you have to!

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-($n,\ \alpha,\ \beta$)
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-(θ)
gamma	qgamma($p,\ \alpha,\ \beta$)	d-($x,\ \alpha,\ \beta$)	p-($x,\ \alpha,\ \beta$)	r-($\alpha,\ \beta$)
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-(θ)
T (standard)	qt($p,\ \nu$)	d-($x,\ \nu$)	p-($x,\ \nu$)	r-(ν)
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.

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) [easy] Assume that each Accord's price at the Astoria dealership is normal and $\stackrel{iid}{\sim}$ 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.

It's a good model because the support of the normal model is all real numbers so the models appropriately fits our problem.

(b) [easy] Despite what you wrote in (b), assume $\stackrel{iid}{\sim}$ for the rest of the problem. The nationwide variance for a Honda Accord selling price we're going to assume is $\sigma^2 = \$1000^2$. 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.

Let $\mathbb{P}(\theta) \propto 1$ be the Laplace Prior then,

$$\mathcal{N}\left(\bar{x},\, \frac{\sigma^2}{n}\right) = \mathcal{N}\left(\bar{x},\, \frac{(1000)^2}{n}\right)$$

(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.

$$\hat{\theta}_{\text{MMSE}} = \bar{x} = \frac{19000 + 18200}{2} = 18600$$

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

$$\rho = \frac{1}{1 + \frac{n\tau_a}{\sigma^2}} =$$

(e) [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.

$$\mathbb{P}\left(X < \$22,205\right) = \texttt{pnorm}(22205, \bar{x}, \frac{\sigma^2}{n}) = \texttt{pnorm}(22205, 18600, \frac{(1000)^2}{2})$$

(f) [harder] 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.

$$\mathbb{P}\left(X < \$17,000\right) = \text{pnorm}(17000, \bar{x}, \frac{\sigma^2}{n}) = \text{pnorm}(17000, 18600, \frac{(1000)^2}{2})$$

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
```

(a) [easy] Write a model below for the amount of milk in each of the n bottles. Hint: use the normal model!

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = 0.081$$

$$\mathcal{N}\left(12, \frac{0.081}{21}\right)$$

(b) [easy] Find the MLE for σ^2 (compute the estimate as a number, not derive it analytically).

$$\begin{aligned}\hat{\sigma}^2_{,\text{MLE}} &= \frac{1}{n}\sum_{i=1}^{n}(x_i - 12)^2 \\ &= 1.94\end{aligned}$$

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

$$SSE = (n-1)\sigma^2 = (20)(1.94) = 38.8$$

$$\text{InvGamma}\left(\frac{21}{2}, \frac{38.8}{2}\right)$$

(d) [harder] Find a left-sided credible region for σ^2. This will give the upper bound for the machine's variance.

$$\begin{aligned}CR_{\sigma^2,95\%} &= \texttt{qinvgamma}(0.005, 10.5, 0.81) \\ &= 0.170\end{aligned}$$

(e) [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. Do this test.

$$\begin{aligned}H_0 &: \sigma^2 > 0.352 \\ H_a &: \sigma^2 \leq 0.352 \\ \mathbb{P}\left(\sigma^2 \leq 0.352\right) &= 1 - \texttt{pinvgamma}(0.352, 10.5, 1.94) \\ &= 0.038\end{aligned}$$

Accept H_0

(f) [harder] Find the probability the next bottle has more than 13oz of liquid.

$$\begin{aligned}
\mathbb{P}\left(X^* > 13 \mid X, \theta\right) &= 1 - \texttt{ptscale}(13, n, \mu, \hat{\sigma}^2_{,\text{MLE}}) \\
\mathbb{P}\left(X^* > 13 \mid X, \theta\right) &= 1 - \texttt{ptscale}(13, 21, 12, 1.94) \\
&= T_{50}(12, 1.94)
\end{aligned}$$