SOLUTIONS TO TAKE HOME EXAM FOR BAYESIAN INFERENCE IN THEORY AND PRACTISE 2007-03-27

MATTIAS VILLANI

Question 1a. *"Assume that you want to investigate the proportion (θ) of defective items manufactured at a production line. Your collegue takes a random sample of* 30 *items. Three were defective in the sample. Assume a uniform prior for θ. Compute the posterior of θ."*

Solution 1a. Data: $x_1, ..., x_{30}|\theta \overset{iid}{\sim} Bern(\theta)$. Likelihood $p(x_1, ..., x_n|\theta) = \theta^s(1-\theta)^f$, where $s = 3$ and $f = 27$.

Uniform prior: $\theta \sim Beta(\alpha, \beta)$, with $\alpha = \beta = 1$. Prior density:

$$p(\theta) = \frac{1}{B(\alpha, \beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}.$$

Posterior density:

$$\begin{aligned} p(\theta|x_1, ..., x_n) &\propto p(x_1, ..., x_n|\theta)p(\theta) \\ &\propto \theta^s(1-\theta)^f\theta^{\alpha-1}(1-\theta)^{\beta-1} \\ &= \theta^{\alpha+s-1}(1-\theta)^{\beta+f-1}, \end{aligned}$$

which we recongize as being proportional to the Beta density with parameters $\alpha+s$ and $\beta+f$. That is: the posterior is given by the $Beta(\alpha+s, \beta+f)$ density. In this specific example (where $\alpha = \beta = 1$, $s = 3$, $f = 27$), the posterior is therefore the $Beta(4, 28)$ density. Figure 1 and table below were produced with my Excel spreadsheet *bernoullimodellen.xls*.

Prior (Beta)
a 1
b 1

Posterior (Beta)
a+y 4
b+n-y 28

Data
Number of trials (n) 30
Number of successes (y) 3

Prior and posterior summaries
Prior mean 0,5
Prior standard dev. 0,288675135
ML estimate 0,1
Posterior mean 0,125

Research department, Sveriges Riksbank, SE-103 37 Stockholm, Sweden and *Department of Statistics, Stockholm University. E-mail: mattias.villani@riksbank.se.*

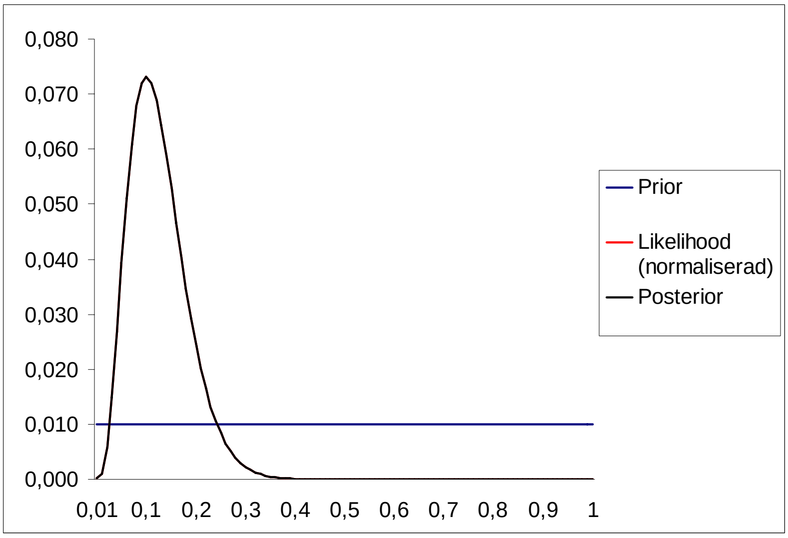

FIGURE 1. Figure 1. Prior and posterior for exercise 1a. Prior, likelihood and posterior for the Beta(1,1) prior.

Posterior standard dev. 0,057570773

Hint for 1b. Use the same approach as in 1a, but the binomial distribution instead of Bernoulli.

Question 1c. *"Your collegue now tells you that he did not decide on the sample size before the sampling was performed. His sampling plan was to keep on sampling items until he had found three defective ones. It just happend that the* $30th$ *item was the third one to be defective. Redo the posterior calculculation, this time under the new sampling scheme. Discuss the results."*

Solution 1c. The likelihood function is now given by the negative binomial distribution:

$$p(n|\theta) = \binom{n-1}{s-1}\theta^s(1-\theta)^f,$$

where I note that it is now n that is stochastic (in 1a it was s that was stochastic). But when we do a Bayesian analysis we are looking at the posterior $p(\theta|Data)$, where $Data$ is fixed. It does not matter if s or n was stochastic before we collected the data, what matters is the form of the likelihood as a function of θ. Evidently, the likelihood function, as a function of θ, is the same in 1a and 1b. Since the prior is assumed to be the same, so are the posteriors in 1a and 1b. The posterior of θ is therefore again the $Beta(4, 28)$ density. Bayesian inference therefore respects the *likelihood principle*: it is only the shape of the likelihood that should matter for inference.

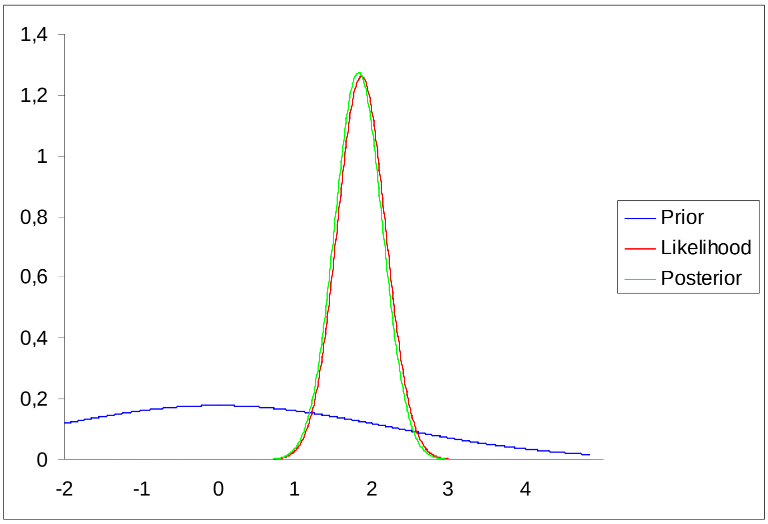

FIGURE 4. Prior-to-Posterior. Normal model. Exercise 2a.

Question 2a. *"Let $x_1, ..., x_{10} \overset{iid}{\sim} N(\theta, 1)$. Let the sample mean be $\bar{x} = 1.873$. Assume that $\theta \sim N(0, 5)$ apriori. Compute the posterior distribution of θ."*

Solution 2a. Remember the prior-to-posterior mapping for this model (Lecture 1):

$$\begin{aligned}
\theta &\sim N(\mu_0, \tau_0^2) \overset{x_1,...,x_n}{\Longrightarrow} \theta|x \sim N(\mu_n, \tau_n^2) \\
\frac{1}{\tau_n^2} &= \frac{n}{\sigma^2} + \frac{1}{\tau_0^2} \\
\mu_n &= w\bar{x} + (1-w)\mu_0 \\
w &= \frac{\frac{n}{\sigma^2}}{\frac{n}{\sigma^2} + \frac{1}{\tau_0^2}}.
\end{aligned}$$

Here we have $n = 10$, $\sigma^2 = 1$, $\mu_0 = 0$, $\tau_0^2 = 5$. This gives us

$$\begin{aligned}
w &= \frac{\frac{10}{1}}{\frac{10}{1} + \frac{1}{5}} = \frac{50}{51} = 0.98039 \\
\mu_n &= \frac{50}{51} \cdot 1.873 + \frac{1}{51} \cdot 0 = 1.8363 \\
\tau_n^2 &= \left(\frac{10}{1} + \frac{1}{5}\right)^{-1} = \frac{5}{51}.
\end{aligned}$$

From normalmodellen.xls you get the graph in Figure 4.

Question 2b. *"Assume now that you have a second sample $y_1, ..., y_{10} \overset{iid}{\sim} N(\theta, 2)$, where θ is the same quantity as in 2a. The sample mean in this second sample is $\bar{y} = 0.582$. Compute the posterior distribution of θ using both samples (the x's and the y's) under the assumption that the two samples are indepedent."*

Answer 2b. The easiest way to do this is use the posterior from the first sample as a prior for the second sample. That is, for this second sample we use the prior $\theta \sim N(1.836\,3, \frac{5}{51})$.

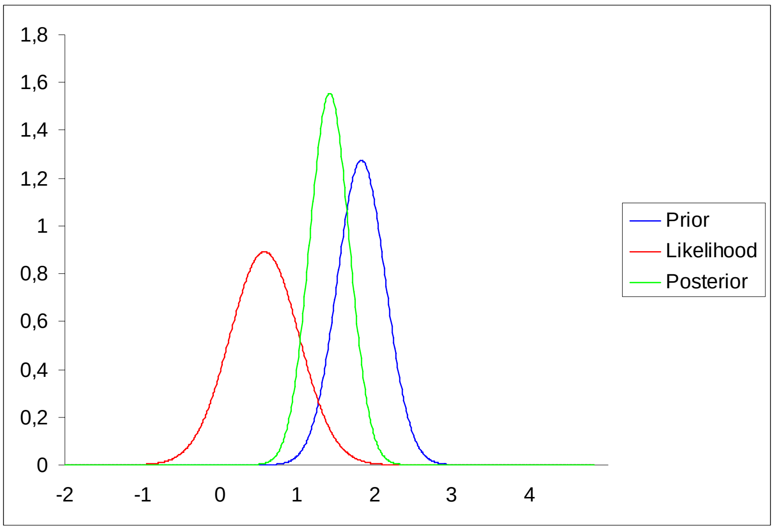

FIGURE 5. Prior-to-Posterior. Normal model. Exercise 2b.

We then get the posterior

$$\begin{aligned} w &= \frac{\frac{10}{2}}{\frac{10}{2}+\frac{1}{\frac{5}{51}}} = \frac{25}{76} \\ \mu_n &= \frac{25}{76}\cdot 0.582 + (1-\frac{25}{76})\cdot 1.8363 = 1.4237 \\ \tau_n^2 &= \left(\frac{10}{2}+\frac{1}{\frac{5}{51}}\right)^{-1} = \frac{5}{76}. \end{aligned}$$

Figure 5. Shows the results.

Exercise 2c. *"You have now managed to obtain a third sample* $z_1, ..., z_{10} \overset{iid}{\sim} N(\theta, 3)$*, with mean* $\bar{z} = 1.221$*. Unfortunatly, the measuring device for this latter sample was defective: any measurement above* 3 *was recorded as* 3*. There were* 2 *such measurements. Compute the posterior distribution based on all three samples (*x, y *and* z*)."*

Solution 2c. This is more complicated. Let us do as before, using the posterior from the first two samples (obtained in 2b) as the prior. The prior is therefore $\theta \sim N(1.4237, \frac{5}{76})$. Now, the mean of the eight measurements which were correctly recorded is $(1.221 \cdot 10 - 3 \cdot 2)/8 = 0.77625$. Let's first use these 8 observations to update the posterior. We then add the information from the two measurements that where truncated at 3. The 8 correctly recorded observations gives the following updating of the $N(1.4237, \frac{5}{76})$ prior:

$$\begin{aligned} w &= \frac{\frac{8}{3}}{\frac{8}{3}+\frac{1}{\frac{5}{76}}} = \frac{10}{67} \\ \mu_n &= \frac{10}{67}\cdot 0.77625 + (1-\frac{10}{67})\cdot 1.4237 = 1.3271 \\ \tau_n^2 &= \left(\frac{8}{3}+\frac{1}{\frac{5}{76}}\right)^{-1} = \frac{15}{268} = 0.05597. \end{aligned}$$

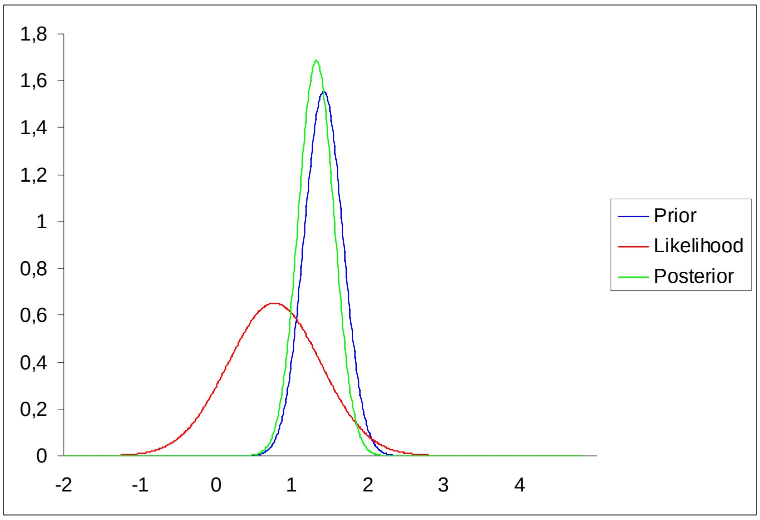

FIGURE 6. Prior-to-Posterior. Normal model. Exercise 2b, but using only the eight correctly recorded observations (and the 20 observations from 2a and 2b).

Note that most of the weight is now given to the prior (which is the posterior from 2a and 2b combined). This is reasonable since the 8 new observations have relatively large variance ($\sigma^2 = 3$). Figure 6 displays the prior-to-posterior mapping.

OK, now we are ready for the final piece of information: the two truncated observations. We do not know their exact values, but we do know that they were at least equal to or larger than 3. This is important information which we cannot ignore. The likelihood of these two observations (let's call them z_1 and z_2) is

$$\begin{aligned} p(z_1, z_2|\theta) &= \Pr(z_1 \geq 3)\Pr(z_2 \geq 3) = \left[1 - \Phi(\frac{3-\theta}{\sqrt{3}})\right]\left[1 - \Phi(\frac{3-\theta}{\sqrt{3}})\right] \\ &= \left[1 - \Phi(\frac{3-\theta}{\sqrt{3}})\right]^2. \end{aligned}$$

This likelihood function is graphed in Figure 7.

The posterior density from all the data combines this likelihood with the posterior from the 28 correctly recorded observations (10+10+8) that we obtained earlier. This 28-observation posterior is now playing the role of the prior. The full posterior is therefore proportional to

$$\exp\left[-\frac{1}{2 \cdot 0.05597}(\theta - 1.3271)^2\right]\left[1 - \Phi(\frac{3-\theta}{\sqrt{3}})\right]^2.$$

I have evaluated this expression in Excel. The result is given in Figure 8. Note how the two truncated observations has some, but not substantial, effect on the prior. The prior is pulled up towards larger θ-values. Compare Figure 8 with Figure 4, to see how much we have learned about θ from these 30 observations. Quite a bit.

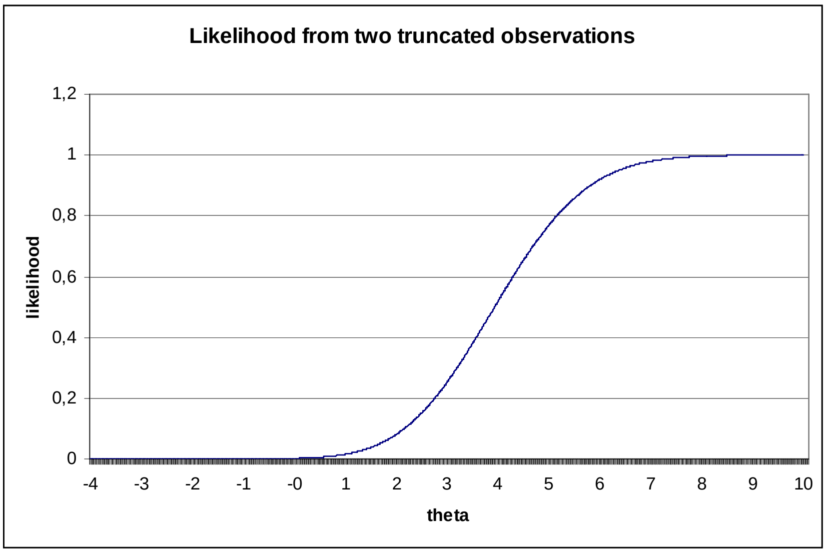

FIGURE 7. Likelihood from the two truncated observations.

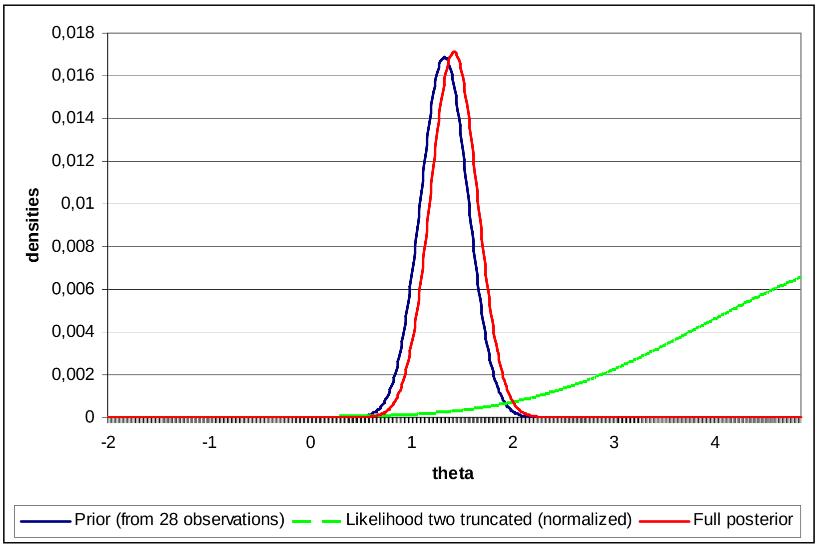

FIGURE 8. Prior-to-posterior mapping from adding the two truncated observations.

Question 3. *"You want to estimate the annual cost of your mobile phone bill. In the month of january 2007, you made* 23 *calls, with average duration of* 114 *seconds per call. The call rate is 5 SEK per minute (hence, every second costs 5/60 SEK, there are no initial costs. Same cost day and night). Assume an exponential distribution for the duration of the calls and that every month has 30 days. Use Bayesian methods to compute the probability that the annual call cost for the year* 2007 *exceeds 3000 SEK. Make whatever assumptions you find necessary/plausible."*

Solution 3. Let c_i denote the call cost in month i of 2007. Let n_i be the number of calls in this month and $\bar{d}_i$ the average duration of calls in the same month. The cost for month i