LINKÖPING UNIVERSITY   2017-04-01
Division of Statistics and Machine Learning   Bayesian Learning
Department of Computer and Information Science   Advanced level course
Mattias Villani

Mathematical Exercises 1

Teacher: Mattias Villani

Try to solve the problems before class. Don't worry if you fail, the important thing is trying.

You should not hand in any solutions.

This part of the course is not obligatory and is not graded.

---

1. WHAT MATTERS IS THE LIKELIHOOD

   (a) Assume that you want to investigate the proportion (θ) of defective items manufactured at a production line. Your colleague takes a random sample of 30 items. There were three defective items in the sample. Assume a uniform prior for θ. Compute the posterior of θ.

   (b) Your colleague 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 happened that the 30th item was the third one to be defective. Redo the posterior calculation, under the new sampling scheme. Compare the results with Problem 1(a). [Hint: negative binomial distribution]

2. NORMAL WITH A NORMAL IS NORMAL

   (a) Let $x_1, ..., x_n \overset{iid}{\sim} N(\theta, \sigma^2)$ and assume that σ^2 is known. Assume a uniform prior:

   $$p(\theta) \propto c.$$

   Derive the posterior distribution of θ.

   (b) Assume now a normal prior

   $$\theta \sim N(\mu_0, \tau_0^2).$$

   Derive the posterior distribution of θ.

3. It ends with an abnormal twist

   (a) 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.

   (b) 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 3a. 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.

   (c) You have now managed to obtain a third sample $z_1, ..., z_{10} \overset{iid}{\sim} N(\theta, 3)$, with mean $\bar{z} = 1.221$. Unfortunately, the measuring device for this latter sample was defective: any measurement above 3 was recorded as 3. There were two such measurements. Compute the posterior distribution based on all three samples (x, y and z). [Hint: in this case the posterior distribution is not a known distribution (it is not normal for example). It is enough to give an expression for the (unnormalized) posterior. You can also plot this over a grid on your computer, if you like.]

4. Solving this problem may take exponentially long time

   (a) Let $x_1, ..., x_n \overset{iid}{\sim} Expon(\theta)$. We use the parametrization of the exponential distribution where if $E(X) \sim Expon(\theta)$ then $E(X) = 1/\theta$.

   (b) Show that the conjugate prior for the exponential model is $\theta \sim Gamma(\alpha, \beta)$. Derive the posterior distribution for θ

Have fun!

- Mattias