Bertil Wegmann
Dept of Computer and Information Science
Linköping University

Bayesian Learning

Mathematical Exercises 1

Try to solve the problems before class and don't worry if you fail. The important thing is to try.
You should not hand in any solutions.
This part of the course is not obligatory and is not graded.

---

1. LIKELIHOOD OF A PROPORTION

   (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 and tells you for each item whether or not it was defective. So she records the data as $x_1 = 0, x_2 = 1, ..., x_n = 0$, where $x_i = 1$ if the item is defective and $x_i = 0$ otherwise. There were three defective items in the sample. Assume a uniform prior for θ. Compute the posterior of θ.

   (b) Assume your colleague only told you that there were three defective items in the sample of 30 items, but she did not tell you which specific items that were defective. Assume a uniform prior for θ. Compute the posterior of θ.

   (c) Your colleague now tells you that she did not decide on the sample size before the sampling was performed. Her sampling plan was to keep on sampling items until she 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. IT'S NORMAL!

   (a) Let $x_1, ..., x_n \overset{iid}{\sim} \mathrm{N}(\theta, \sigma^2)$ 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. Gradually add normal observations

(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 independent.

(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. A special case of the gamma distribution

(a) Let $x_1, ..., x_n \overset{iid}{\sim} \text{Expon}(\theta)$. We use the parametrization of the exponential distribution where if $X \sim \text{Expon}(\theta)$ then $E(X) = 1/\theta$. Show that the conjugate prior for the exponential model is $\theta \sim \text{Gamma}(\alpha, \beta)$. Derive the posterior distribution for θ.

Good luck!

Best, Bertil