CSE 515T (Spring 2019) Assignment 1

Due Wednesday, 6 February 2019

1. (Barber, originally adopted from David Spiegelhalter) A secret government agency has developed a scanner which determines whether a person is a terrorist. The scanner is fairly reliable; 95% of all scanned terrorists are identified as terrorists, and 95% of all upstanding citizens are identified as such. An informant tells the agency that exactly one passenger of 100 aboard an aeroplane in which you are seated is a terrorist. The police haul off the plane the first person for which the scanner tests positive. What is the probability that this person is a terrorist?

   Additionally, if the police were to scan all passengers, how many positive detections should we expect?

2. Suppose k has a geometric distribution with unknown success probability θ

   $$\Pr(k \mid \theta) = (1-\theta)^{k-1}\theta, \qquad k = 1, 2, \dots$$

   The geometric distribution is appropriate for modeling the number of independent Bernoilli trials required, each with success probability θ, before observing the first "success."

   Let the prior for θ be a beta distribution:

   $$p(\theta \mid \alpha, \beta) = \frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{B(\alpha, \beta)} \qquad 0 < \theta < 1$$

   where B is the beta function. Show that, given an observation k, the posterior $p(\theta \mid k, \alpha, \beta)$ is a beta distribution with updated parameters (α', β').

   Hint: it may help to look up properties of the Beta function, e.g., on Wikipedia.

3. Suppose that in the last question, we received a sample of n observations $\{k_1, k_2, \dots, k_n\}$. What is the posterior $p(\theta \mid k_1, k_2, \dots, k_n, \alpha, \beta)$? What is the posterior mean? The posterior mode?

   In light of this and the previous question, can you give an interpretation of the prior parameters α and β? What happens in the limit as $n \to \infty$?

4. (Scenario quoted from Morey, et al.) A 10-meter-long research submersible with several people on board has lost contact with its surface support vessel. The submersible has a rescue hatch exactly halfway along its length, to which the support vessel will drop a rescue line. Because the rescuers only get one rescue attempt, it is crucial that when the line is dropped to the craft in the deep water that the line be as close as possible to this hatch. The researchers on the support vessel do not know where the submersible is, but they do know that it forms distinctive bubbles. These bubbles could form anywhere along the craft's length, independently, with equal probability, and float to the surface where they can be seen by the support vessel.

   We wish to perform inference about the location of the rescue hatch given observed bubbles; call this location θ.

   A common "trick" when wishing to express absolute prior ignorance of a parameter is to use a so-called *uninformative* prior. In this case, we will consider the uninformative "prior"

$p(\theta) = 1$. This prior does not normalize, but we will see that it does not lead to major problems.

(a) Suppose the researchers observe the locations of exactly two bubbles, x_1 and x_2. Write down an appropriate likelihood for these data given θ and derive the posterior distribution for the location of the hatch, $p(\theta \mid x_1, x_2)$, using the uninformative prior described above.

(b) Now find a 50% Bayesian credible interval for θ given (x_1, x_2). Plot the width of this interval as a function of $|x_1 - x_2|$. Is this the relationship you would expect?

5. (Maximum-likelihood estimation.) Suppose you flip a coin with unknown bias θ; $\Pr(x = \mathrm{H} \mid \theta) = \theta$, three times and observe the outcome HHH. What is the maximum likelihood estimator for θ? Do you think this is a good estimator? Would you want to use it to make predictions?

   Consider a Bayesian analysis of θ with a beta prior $p(\theta \mid \alpha, \beta) = \mathcal{B}(\theta; \alpha, \beta)$. What is the posterior mean of θ? What is the posterior mode? Consider $(\alpha, \beta) = (50, 50)$. Plot the posterior density in this case. Is the maximum likelihood estimator a good summary of the distribution?

6. (Effect of weird priors.) Let us consider the following set of observations. We flip a coin independently $n = 1\,000$ times and observe $x = 900$ successes. Call the unknown bias of the coin $\theta \in (0, 1)$.

   For each of the prior distributions $p(\theta)$ below, please:

   - plot the prior distribution $p(\theta)$ over the range $0 < \theta < 1$
   - plot the posterior distribution given the above data, $p(\theta \mid \mathcal{D})$, over the range $0 < \theta < 1$
   - report the posterior mean, $\mathbb{E}[\theta \mid \mathcal{D}] = \int \theta\, p(\theta \mid \mathcal{D})\, \mathrm{d}\theta$.

   (a) A uniform prior on θ, which can be realized by selecting the beta distribution with $\alpha = \beta = 1$:
   $$p(\theta) = \mathcal{B}(\theta; \alpha = 1, \beta = 1).$$

   (b) A prior with extreme bias toward small values of θ:
   $$p(\theta) = \mathcal{B}(\theta; \alpha = 1, \beta = 100).$$

   (c) A prior that has no support on values greater than $\theta = {}^1\!/_2$:
   $$p(\theta) = \begin{cases} 2 & \theta < {}^1\!/_2; \\ 0 & \theta \geq {}^1\!/_2. \end{cases}$$

7. (Optimal *Price is Right* bidding.) Suppose you have a standard normal belief about an unknown parameter θ, $p(\theta) = \mathcal{N}(\theta; 0, 1^2)$. You are asked to give a point estimate $\hat{\theta}$ of θ, but are told that there is a heavy penalty for guessing too high. The loss function is
   $$\ell(\hat{\theta}, \theta; c) = \begin{cases} (\theta - \hat{\theta})^2 & \hat{\theta} < \theta; \\ c & \hat{\theta} \geq \theta \end{cases},$$
   where $c > 0$ is a constant cost for overestimating. What is the Bayesian estimator in this case? How does it change as a function of c? Plot the optimal action as a function of c for $0 < c < 10$. Hint: you may need to minimize certain expressions you encounter numerically as an analytic solution may not be available.