HOMEWORK 7, STAT 251

Remember to submit your R code along with the completed homework. You must include any requested plots and make sure to include your final answers to all questions in the writeup.

(1) The Orem City Department of Public Works is responsible for repairing potholes in city streets. Although not exactly right, it is reasonable to assume that the number of potholes that need to be repaired in a city block follows a Poisson(θ) distribution. If the prior belief about θ is that it follows a gamma distribution with shape=.5 and rate=0.5,

    (a) According to the prior distribution, what is the expected number of potholes per block?

    (b) What is the prior probability that θ exceeds 2.7?

(2) After a particularly heavy winter, they repaired the following number of potholes in each block along a 20 block stretch of State St: 3 3 7 1 4 6 6 7 3 1 5 5 5 3 3 0 3 1 2 2. We will assume that these data have a Poisson(θ) distribution. If Orem City's prior belief about θ is as in the previous problem,

    (a) What is the posterior distribution of θ?

    (b) What is their belief now, after seeing the data, about the expected number of potholes they will repair in each block of State St?

    (c) What is a 95% (posterior) credible interval for the parameter of the Poisson distribution?

    (d) What is the updated (i.e., posterior) probability that θ is greater than 2.7?

(3) The Utah Highway Patrol believes the number of accidents per week along a 50 mile stretch of I-15 follows a Poisson(θ) distribution. Suppose their current belief about θ is that it follows a gamma distribution with shape=5.1 and **scale**=.8.

    (a) In our class we discussed two parametrizations of the gamma distribution. Here we consider the parametrization in terms of shape and rate, rather than in terms of shape and scale. For their prior beliefs, what would the shape and rate inputs be for the gamma distribution? $shape =$, $rate =$.

    (b) Derive the prior predictive distribution and use it to find the probability that in the next week there will be at least 3 accidents? What does this probability say about the appropriateness (or lack there of) of the prior distribution used by the Utah Highway Patrol?

    (c) If in the next week there were 5 accidents, what is their updated distribution for θ? Plot the prior distribution and posterior distribution on the same graph. (Prior in gray, posterior in black.)

    (d) In the month following the week with 5 accidents, there were 2, 0, 6, and 3 accidents per week. What is the new posterior for θ now? Plot it in red on the same graph as in part b.

    (e) Derive the posterior predictive distribution and use it to find the predictive probability that in the next week, there will be at least 3 accidents?

(4) The City of Orem is interested in knowing if a new paving method will reduce the number of needed pothole repairs along State St. Since they know that the severity of the winter weather also affects the number of potholes, they collected data over a ten-year period prior to when the new paving method was used, and for 6 years after. For the ten years prior to the new method (the "before" period), the number of potholes in a "winter month" was as follows: 27 30 25 29 27 23 29 24 27 36 33 34 30 25 29 27 33 33 27 34 28 43 31 24 36 28 26 29 30 20 33. For the 6 years after the new method was used (the "after" period) the number of potholes in a "winter month" was as follows: 26 32 21 32 21 24 36 26 24 25 30 25 20 14 26 28 21 21. Use a gamma(shape=5, rate=0.2) prior for both cases (i.e., for θ during the before period and for θ during the after period).

    (a) What is the posterior probability that the (population) rate of potholes in a "winter month" was higher before the new paving method was used than it was after the new paving method was used?

    (b) What are the mean and median of the posterior distribution of $d = \theta_{before} - \theta_{after}$?

    (c) Report and interpret in context a 95% (posterior) credible interval for $d = \theta_{before} - \theta_{after}$?