University of California, Santa Cruz
Department of Applied Mathematics and Statistics
Baskin School of Engineering
Classical and Bayesian Inference - AMS 132

Right/wrong assumptions and Bayesian intervals

1. Suppose that $X_1, ..., X_n$ are random variables describing the monthly income (in thousands of dollars) of people in the city of Santa Cruz. From a sample of 25 person, the following monthly incomes were observed

   `1.10, 0.87, 0.84, 2.62, 17.37, 7.38, 3.24, 5.74, 0.88, 8.34, 4.57, 7.43, 26.54, 6.33, 6.21, 11.26, 3.93, 2.02, 3.53, 14.19, 3.85, 1.76, 3.40, 0.64, 0.36`

   The major has decided that a bonus will be given if the mean monthly income is less that 4 thousand dollars.

   Right/wrong assumptions:

   (a) What do you think about assuming a normal distribution for the monthly income? Would you propose other sampling distribution?

   (b) Suppose that regardless of a), someone assumes that the monthly income follows a normal distribution. Compute a symmetric 90% confidence interval for the mean monthly income. What do you think the mayor will do based on the available information?

   (c) Now, assume that $X_1, ..., X_n$ follow an exponential distribution with parameter λ. Find and compute a symmetric 90% confidence interval for the mean monthly income. What do you think the mayor will do based on the available information?
   Note: for this, use the fact that $2\sum_{i=1}^{n} X_i/\lambda \sim \chi^2_{(2n)}$.

   Bayesian intervals:

   (d) Consider a gamma prior distribution for λ with parameters $\alpha = 1$ and $\beta = 0.25$. Find two values c and d such that the prior probability of the mean monthly income being between these two values is 0.9.
   Note: for this, use the fact that is $X \sim Gamma(a, b)$, then $2Xb \sim \chi^2_{(2a)}$.

   (e) Assuming that $X_1, ..., X_n$ follow an exponential distribution with parameter λ and λ follows a gamma prior distribution with parameters $\alpha = 1$ and $\beta = 0.25$. Find two values c and d such that the posterior probability of the mean monthly income being between these two values is 0.9.
   Here (c, d) is called 90% credible region.

   (f) What happens with the 90% credible region found in e) if $\alpha = \beta = 0$?

(g) Consider the following pair of hyper parameters for λ: $(\alpha = 1, \beta = 0.25)$, $(\alpha = 1, \beta = 1)$, and $(\alpha = 1, \beta = 0.01)$. Compute the prior and posterior credible regions for each set of hyper parameters. See how the posterior credible regions change and compare them the 90% confidence interval