

Bayesian Statistics

Note: here the $\text{Gamma}(m, \lambda)$ distribution refers to the distribution with density $p(x) \propto x^{m-1} \exp(-\lambda x)$

1. Read the article at <http://www.nytimes.com/2011/01/11/science/11esp.html>. The following is a quotation from this article:

Consider the following experiment. Suppose there was reason to believe that a coin was slightly weighted toward heads. In a test, the coin comes up heads 527 times out of 1,000. Is this significant evidence that the coin is weighted? Classical analysis says yes. With a fair coin, the chances of getting 527 or more heads in 1,000 flips is less than 1 in 20, or 5 percent, the conventional cutoff. To put it another way: the experiment finds evidence of a weighted coin with 95 percent confidence. Yet many statisticians do not buy it. One in 20 is the probability of getting any number of heads above 526 in 1,000 throws. That is, it is the sum of the probability of flipping 527, the probability of flipping 528, 529 and so on. But the experiment did not find all of the numbers in that range; it found just one 527. It is thus more accurate, these experts say, to calculate the probability of getting that one number 527 if the coin is weighted, and compare it with the probability of getting the same number if the coin is fair. Statisticians can show that this ratio cannot be higher than about 4 to 1, according to Paul Speckman, a statistician, who, with Jeff Rouder, a psychologist, provided the example.

- i) Formalize and confirm the statement “Statisticians can show that this ratio cannot be higher than about 4 to 1”.
 - ii) In the above setting, if your prior probability that the coin is weighted is 0.5, what is the largest your posterior probability (that the coin is weighted) could be?
2. Consider x_1, x_2, \dots, x_n i.i.d. $\sim \text{Poisson}(\mu)$. Suppose the prior distribution on μ is $\text{Gamma}(m, \lambda)$.

- (a) Derive, analytically, the posterior distribution for μ . Explain the role of conjugacy here. Show that this posterior distribution depends on the data only through the sample mean \bar{x} (and n).
 - (b) Plot the posterior distribution for μ given data with $\bar{x} = 2.8$, $n = 10$ with the prior $m = 1$, $\lambda = 1$. Compute the posterior mean, the posterior median, and give a 90% Credible interval for μ .
 - (c) The posterior distribution can be interpreted in terms of “long-run” frequency along the follow lines: if one repeated a certain experiment, many times, then of the times we see “data like this” the posterior distribution tells us how often the parameter would fall in any given region. Make this statement more explicit (ie say what we mean by “a certain experiment” and “data like this”) and illustrate it by a simulation. (Use the data and prior above).
3. Suppose that, given σ , x is Normal with mean 0 and variance σ^2 , and that the prior for σ is such that $1/\sigma^2$ has a $\text{Gamma}(m, l)$ distribution. Find the posterior distribution of σ^2 . [Hint: the algebra is easier if you compute the posterior for the “precision”, $\tau := 1/\sigma^2$, rather than for the variance σ^2 . For this reason Bayesian analyses often consider the precision rather than the variance.]
4. Suppose $X_i|\theta_i \sim N(\theta_i, 1)$ (for $i = 1, \dots, n$ with n large). Then $X_i \pm 1.96$ is the conventional 95% CI for θ_i .
- i) Show that, for any fixed value of θ_i , the conventional 95% CI has the correct frequentist coverage properties. i.e. the interval will contain θ_i in 95% of trials.
 - ii) Suppose that the true values of θ_i are uniformly distributed on $[-a, a]$ for some a . Derive the posterior distribution for $\theta_i|X_i, a$. Is it true that, for any fixed x , $\Pr(\theta_i \in x \pm 1.96|X_i = x, a) \rightarrow 0.95$ as $a \rightarrow \infty$?
 - ii) Now suppose that the true values of θ_i are normally distributed: $\theta_i \sim N(0, 1)$. Suppose that our interest focuses on which of the θ_i are positive. We will say we are “confident” that θ_i is positive if its 95% CI contains only positive values. Among all the occasions i in which we are confident that θ_i is positive, what proportion will we be wrong? Provide a mathematical argument, and provide R code that confirms your result by simulation.

5. In Homework 1 you were asked to complete the exercise in the commented text at the end of `exercises/seeb/train_test.R`.
- (a) Swap your code and results with someone else from the class (henceforth referred to as your “partner”). [Your partner should be someone with whom you have not previously compared code/results for this question - if you need help identifying a partner, let me know.]
 - (b) Run your partner’s code yourself, and check you get the same answer that they report. If not, communicate with your partner to resolve the difference, and report the outcome.
 - (c) Compare the results of your code with the results of your partner’s code. Are they the same? If not, communicate with your partner to resolve the difference, and report the outcome.
 - (d) Compare the way that your partner tackled the problem with the way you tackled the problem. Comment on any similarities and differences. Is their code easier to read or harder to read than yours? Did they use any R tricks or functions that you were unfamiliar with? If possible improve your code (and correct it if necessary) as a result of what you have learned. Hand in your improved code and its output.