Name: Solutions

Bayesian Statistics, 22S:138
Midterm 1, Fall 2011

Show any computations that you carry out. Use the back of your exam paper if you run out of space.

1. Suppose that the following statements are true for a Midwestern town:

   - The probability that the leaves on the first maple trees in the town will begin turning color before Oct. 1 in any given year is 0.60. $P(L) = 0.60$
   - The probability that the first snowfall will occur before Nov. 15, given the first maple leaves started turning color before Oct. 1, is 0.50. $P(S|L) = 0.50$
   - The probability that the first snowfall will occur before Nov. 15, given the first maple leaves started turning color on or after Oct. 1, is 0.35. $P(S|\bar{L}) = 0.35$

   (a) What is the probability that the first snowfall will occur before Nov. 15? (Numeric answer; show your work.)

$$P(S) = P(S \cap L) + P(S \cap \bar{L})$$
$$= P(S|L)P(L) + P(S|\bar{L})P(\bar{L})$$
$$= 0.50(0.60) + (0.35)(1-0.60) = \boxed{0.44}$$

   (b) If I tell you that the first snowfall occurred on Dec. 1 in this town last year, what is the probability that the first maple leaves started turning color before Oct. 1 last year? (Numeric answer; show your work.)

So $P(\bar{S}$ occurred$)$ is

$$P(L|\bar{S}) = \frac{P(L \cap \bar{S})}{P(\bar{S})} = \frac{1-P(S|L)\,P(L)}{P(\bar{S})}$$
$$= \frac{.50(.60)}{1-0.44}$$
$$= \boxed{0.54}$$

2. You have read that body weights of a particular species of beetles follow a normal distribution with population mean $\mu = 8$ grams. You choose to accept that information as true. You wish to use Bayesian methods to infer about the population variance σ^2 of body weights in this species.

   (a) You wish to use a conjugate prior. Which parametric family is conjugate for the variance of a normal density with mean assumed known? (Just give the name of the family.)

   inverse gamma

   (b) Your prior knowledge about the unknown σ^2 is equivalent to the information in a previous dataset with 2 observations: $y_{old,1} = 6$ grams, and $y_{old,2} = 9$ grams. What prior parameters should you use in the conjugate prior? (Numeric answers; show how you got them.) I need $IG\left(\frac{n_0}{2}, \frac{n_0 \sigma_0^2}{2}\right)$

   $n_0 = 2$

   $\sigma_0^2 = [(y_{1,old} - \mu)^2 + (y_{2,old} - \mu)^2]/2$

   $= [(6-8)^2 + (9-8)^2]/2 = \frac{5}{2}$

   $IG\left(\frac{2}{2}, \frac{5}{2}\right) = IG\left(1, \frac{5}{2}\right)$

   Note: The results of part (b) are needed for the rest of this problem. Use your own answer to part (b) if you have one. If you were unable to do part (b), continue from here using an Inverse Gamma(2,2) prior.

   (c) You collect your new data by catching 10 beetles of the species of interest, weighing each one, and then releasing them. (No animals were harmed in this problem.) Here are three summary statistics calculated from your new data:

      i. $\bar{y} = \frac{\sum_{i=1}^{10} y_i}{10} = 7.39$ grams

      ii. $s^2 = \frac{\sum_{i=1}^{10}(y_i - \bar{y})^2}{9} = 0.629$

      iii. $\sum_{i=1}^{10}(y_i - \mu)^2 = \sum_{i=1}^{10}(y_i - 8)^2 = 9.40$

   Circle the statistic above that is a sufficient statistic for σ^2 when the normal mean μ is assumed known.

(d) What is the posterior density $p(\sigma^2|\mathbf{y})$? Just name the family and give its parameters. Show how you got the numeric values of the parameters, but you do not have to go through the entire derivation.

$$p(\sigma^2/y) \text{ is } IG\left(\frac{n_0}{2}+\frac{n}{2}, \frac{n_0\sigma_0^2}{2}+\frac{\Sigma(y_i-\mu)^2}{2}\right)$$
$$= IG\left(1+\frac{10}{2}, \frac{5}{2}+\frac{9.40}{2}\right) = IG(6, 7.2)$$

(e) What is the posterior mean of σ^2? (Numeric answer; show your work.)

$$\frac{\beta}{\alpha-1} = \frac{7.2}{5} = 1.44$$

(f) What line of R code would you use to calculate the endpoints of a 90% equal tail posterior credible set for σ^2?

1/qgamma(c(0.95, 0.05), 6, 7.2)

(g) Suppose the 90% credible set turned out to be (1.1, 3.5). The correct interpretation of the Bayesian 90% credible set is:

i. In 90% of random samples, this procedure produces an interval that traps the true value of σ^2.

ii. σ^2 moves around randomly. 90% of the time it is between 1.1 and 3.5, and the other 10% of the time it is not.

iii. For a person who agreed with the prior, there is 90% probability that the true σ^2 lies in (1.1, 3.5).

iv. 90% of weights of individual beetles are in (1.1, 3.5).

v. None of the above.

(h) You wish to do a second Bayesian analysis with a noninformative prior and to compare the results of the two analyses. Performing more than one Bayesian analysis with different priors is called (circle one):

i. robustness

ii. sensitivity analysis

iii. conservatism

iv. a waste of time

v. none of the above

(i) The 90% equal tail credible set obtained with the noninformative prior is most likely to be (circle one):

i. wider than the 90% credible obtained with the first (informative) prior

ii. of the same width as the 90% credible obtained with the first (informative) prior

iii. narrower than the 90% credible obtained with the first (informative) prior

iv. no way to know in advance which credible set will be wider

3. Y is a random variable with density proportional to the following expression:

$$p(y) \;\propto\; \frac{\lambda^2}{\Gamma(2)} y e^{-\lambda y}, \qquad y > 0$$

(a) What density is this? (Name the parametric family and give the values of the parameter or parameters.)

Gamma(2, λ)

(b) What parametric family would be the conjugate prior for the unknown parameter λ in the likelihood given in part (a)? Just name the parametric family.

We need a family in which the random variable appears in the same mathematical form as λ does in part (a), i.e. raised to a power and in a negative exponent.

This would be $\boxed{\text{Gamma.}}$