STA 250: Statistics

HW 5

Due Wed Oct 09 2013

1. Prove conjugacy of the following pdf families to the corresponding statistical model, i.e., in each case show that if we pick any arbitrary member of the family as our prior pdf $\xi(\theta)$, we get as a posterior pdf $\xi(\theta|x)$ that is also a member of the family, for any arbitrary observation x [see Appendix for details on the Gamma and Pareto pdfs]. Clearly define prior to posterior update formulas.

Model	Pdf family
$X_1, \cdots, X_n \overset{\text{IID}}{\sim} \mathsf{Poisson}(\mu)$, $\mu \in (0, \infty)$	$\{\mathsf{Gamma}(a, b) : a > 0, b > 0\}$
$X_1, \cdots, X_n \overset{\text{IID}}{\sim} \mathsf{Uniform}(0, \theta)$, $\theta \in (0, \infty)$	$\{\mathsf{Pareto}(a, b) : a > 0, b > 0\}$

2. Consider the model: $X^*|(\mu, \sigma^2) \sim \mathsf{Normal}(\mu, \sigma^2)$, $(\mu, \sigma^2) \sim \mathsf{N}\chi^{-2}(m, k, r, s^2)$.
   (a) What is the conditional distribution of X^* given σ^2? [Hint. Use Result 1 from Notes 8. Also use the "normal convolution result" that if $Z \sim \mathsf{Normal}(a, b^2)$ and $U|(Z = z) \sim \mathsf{Normal}(z, d^2)$ then $U \sim \mathsf{Normal}(a, b^2 + d^2)$]
   (b) What is the distribution of (X^*, σ^2)? [Hint. Use Result 1 again]
   (c) Argue that $(X^* - m)/\{s\sqrt{1 + 1/k}\} \sim t(r)$. [Hint. Use Result 2]

3. As a continuation of problem 2, consider the model $X^*, Y^* \overset{\text{IID}}{\sim} \mathsf{Normal}(\mu, \sigma^2)$, $(\mu, \sigma^2) \sim \mathsf{N}\chi^{-2}(m, k, r, s^2)$ and argue that $(X^* - Y^*)/\{s\sqrt{2}\} \sim t(r)$.

4. In the NSW study we have data $X = (X_1, \cdots, X_n)$ on the earning difference of $n = 185$ individuals enrolled in a job training program. Also consider a future observable X^*. Suppose we model $X_1, \cdots, X_n, X^* \overset{\text{IID}}{\sim} \mathsf{Normal}(\mu, \sigma^2)$ and assign (μ, σ^2) the $\mathsf{N}\chi^{-2}(0, 4, 2, 9000^2)$ prior. Use 2(c) above to answer the following:
   (a) Calculate a 95% prior predictive range for X^* (i.e., talk about X^* before we have observed X).
   (b) Calculate a 95% posterior predictive range for X^* when observations on X show $n = 185$, $\bar{X} = 4253.57$ and $s_X = 8926.985$.

5. Again consider the setting of problem 4 above but now suppose we assign (μ, σ^2) the reference prior $\xi(\mu, \sigma^2) = 1/\sigma^2$. Calculate a 95% posterior predictive range for X^* with same observations as in 4(b).

6. The female birth rate in the general population is taken to be 48.5%. Placenta previa is a rare pregnancy condition affecting 0.5% of all labors and it may influence female birth rate. We have data X giving the count of female births among $n = 980$ placenta previa births, modeled as $X \sim \mathsf{Binomial}(n, p)$, $p \in (0, 1)$. We will look at choosing a prior on p and carrying out a Bayesian analysis when with observation $X = 437$.

   (a) Before seeing the data, an expert believes that p is equally likely to be larger or smaller than 48.5%. When asked to think about p if she was told $p > 48.5\%$, she says she believes p equally likely to be larger or smaller than 52%. Find a prior pdf $\mathsf{Beta}(a, b)$ on p matching these beliefs by carrying out the following steps:

      i. For large $a, b > 0$ it is known that $\mathsf{Beta}(a, b) \approx \mathsf{Normal}(c, d^2)$ where $c = a/(a+b)$ and $d^2 = c(1-c)/(a+b+1)$. First identify a $\mathsf{Normal}(c, d^2)$ pdf that matches the expert's beliefs. Identify the corresponding $\mathsf{Beta}(a, b)$.

      ii. Work with the $\mathsf{Beta}(a, b)$ you identified above and indicate (with appropriate figures) how well it matches the expert's beliefs.

   (b) Use the $\mathsf{Beta}(a, b)$ pdf you identified in part (a) to produce a 95% posterior range for p.

   (c) Now use Jefreys' prior on p and produce a 95% posterior range for p. How similar/dissimilar is this interval to the expert's interval?