STA 114: Statistics

HW 5

Due Wed Oct 12 2011

1. A machine goes through 4 hazard levels θ, coded 0 through 3 (from low hazard to high hazard) with use over time. The hazard level can be measured by frequency of hazardous incidents X, again coded 0 through 3 (low frequency to high frequency). Suppose X is modeled with pmfs $f(x|\theta)$, $\theta \in \Theta = \{0, 1, 2, 3\}$ as given by the rows of the following table.

θ	$f(0\|\theta)$	$f(1\|\theta)$	$f(2\|\theta)$	$f(3\|\theta)$
0	$\frac{4}{10}$	$\frac{3}{10}$	$\frac{2}{10}$	$\frac{1}{10}$
1	0	$\frac{3}{6}$	$\frac{2}{6}$	$\frac{1}{6}$
2	0	0	$\frac{2}{3}$	$\frac{1}{3}$
3	0	0	0	1

   (a) For the discrete uniform prior pmf on θ, i.e., $\xi(0) = \xi(1) = \xi(2) = \xi(3) = 1/4$, and $\xi(\theta) = 0$ otherwise, calculate the posterior probability $P(\theta = 0|x)$ for each of $x = 0, 1, 2, 3$.

   (b) Consider the ML set $A_{1/2}(x) = \{\theta \in \Theta : L_x(\theta) \geq \frac{1}{2} L_x(\hat{\theta}_{\text{MLE}}(x))\}$. We saw in HW 3 (problem 1) that

$$A_{1/2}(0) = \{0\}, \quad A_{1/2}(1) = \{0, 1\}, \quad A_{1/2}(2) = \{1, 2\}, \quad A_{1/2}(3) = \{3\}.$$

   If we decide to report $A_{1/2}(x)$ as a plausible set of θ values given observation $X = x$, what quantitative calibrations can you attach to the reported set based on your Bayesian model with the discrete uniform prior (answer separately for each $x = 0, 1, 2, 3$)?

   (c) The machine under consideration is old, and is more likely to be in a high hazard level than a low one. Give a prior pmf that is consistent with this belief and repeat (a) and (b) for this prior pmf.

2. Consider two observations $X_1, X_2 \overset{\text{IID}}{\sim} g(x_i|\theta)$, $\theta \in (-\infty, \infty)$, where $g(y|\theta)$ is the following pmf: $g(y|\theta) = 0.5$ if $y = \theta + 1$ or $y = \theta - 1$, and $g(y|\theta) = 0$ for all other y. Suppose θ is assigned the $\mathsf{Normal}(0, 1)$ prior pdf. Given observed data $x = (x_1, x_2)$, do we get a posterior pdf or a posterior pmf $\xi(\theta|x)$ for θ? Explain. Give precise expressions for $\xi(\theta|x)$ for an arbitrary x.

3. Prove conjugacy of the following pdf families to the corresponding statistical model, .i.e., in each case show that any arbitrary member of the family $\xi(\theta)$, when used as a prior pdf, produces a posterior pdf $\xi(\theta|x)$ that is also a member of the family, for every observation x.

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

4. A set of n lactic acid measurements are modeled as $X_1, \cdots, X_n \overset{\text{IID}}{\sim} \mathsf{Normal}(\mu, \sigma^2)$, $\mu \in (-\infty, \infty)$, $\sigma = 1/3$.

   (a) Identify the posterior pdf $\xi(\mu|x)$ for each choice of the prior pdf $\xi(\mu) = \mathsf{Normal}(0, 10^2)$, $\mathsf{Normal}(5, 10^2)$, $\mathsf{Normal}(5, 1)$ when $x =$ (0.86, 1.53, 1.57, 1.81, 0.99, 1.09, 1.29, 1.78, 1.29, 1.58).
   (b) For each choice identify the intervals $[\mu_{\alpha/2}, \mu_{1-\alpha/2}]$ and $[\mu_{\alpha/2}(x), \mu_{1-\alpha/2}(x)]$, for $\alpha = 0.05$ and x as above.
   (c) For each choice calculate the posterior probability that μ is at least 1.

5. In a survey of $n = 500$ randomly chosen college students, 200 students said they support a certain federal policy. Let $p \in [0, 1]$ denote the fraction of all students in the college who support the policy.

   (a) Identify the posterior pdf $\xi(p|200)$ for each choice of the prior pdf $\xi(p) = \mathsf{Uniform}(0, 1)$, $\mathsf{Beta}(10, 10)$, $\mathsf{Beta}(0.5, 0.5)$.
   (b) For each choice identify the intervals $[p_{\alpha/2}, p_{1-\alpha/2}]$ and $[p_{\alpha/2}(x), p_{1-\alpha/2}(x)]$, for $\alpha = 0.05$ and x as above.
   (c) For each choice calculate the posterior probability that p exceeds 50%.

6. Annual hurricane counts from a set of n consecutive years are modeled as $X_1, \cdots, X_n \overset{\text{IID}}{\sim} \mathsf{Poisson}(\mu)$, $\mu \in (0, \infty)$.

   (a) Identify the posterior pdf $\xi(\mu|x)$ for each choice of the prior pdf $\xi(p) = \mathsf{Gamma}(1, 1/10)$, $\mathsf{Gamma}(10, 1)$, $\mathsf{Gamma}(1/100, 1/100)$ based on data (12, 14, 15, 12, 16, 14, 27, 10, 14, 16).
   (b) For each choice identify the intervals $[\mu_{\alpha/2}, \mu_{1-\alpha/2}]$ and $[\mu_{\alpha/2}(x), \mu_{1-\alpha/2}(x)]$, for $\alpha = 0.05$ and x as above.
   (c) For each choice calculate the posterior probability that μ is smaller than 15.

7. A set of n time intervals (in minutes) between two successive eruptions of a geyser is modeled as $X_1, \cdots, X_n \overset{\text{IID}}{\sim} \mathsf{Exponential}(\lambda)$, $\lambda \in (0, \infty)$.

   (a) Identify the posterior pdf $\xi(\lambda|x)$ for each choice of the prior pdf $\xi(\lambda) = \mathsf{Gamma}(1, 100)$, $\mathsf{Gamma}(10, 1000)$, $\mathsf{Gamma}(1/100, 1/100)$ based on data (79, 54, 74, 62, 85, 55, 88, 85, 51, 85)
   (b) For each choice identify the intervals $[\lambda_{\alpha/2}, \lambda_{1-\alpha/2}]$ and $[\lambda_{\alpha/2}(x), \lambda_{1-\alpha/2}(x)]$, for $\alpha = 0.05$ and x as above.
   (c) For each choice calculate the posterior probability that λ is smaller than 1/50.

8. Suppose the duration (in seconds) of a smile of a certain eight week old baby follows a $\mathsf{Uniform}(0, \theta)$ distribution with $\theta > 0$ unknown. The data available consists of n observed durations $X_1, X_2, \cdots, X_n$ from the baby.

   (a) Identify the posterior pdf $\xi(\theta|x)$ for each choice of the prior pdf $\xi(\theta) = \mathsf{Pareto}(2, 15)$, $\mathsf{Pareto}(5, 12)$, $\mathsf{Pareto}(15, 14)$ based on the observations (10.4, 19.6, 12.8, 14.8, 1.3, 0.7, 5.8, 6.9, 8.9, 9.4).
   (b) For each choice identify the intervals $[\theta_{\alpha/2}, \theta_{1-\alpha/2}]$ and $[\theta_{\alpha/2}(x), \theta_{1-\alpha/2}(x)]$, for $\alpha = 0.05$ and x as above.
   (c) For each choice calculate the posterior probability that θ exceeds 30.