Statistical Theory MT 2009

Problems 3: Solution sketches

1. Suppose $x_1, \dots, x_n$ are values sampled independently from a Poisson distribution with mean λ. The prior density of λ is $Gamma(\alpha, \beta)$, i.e. $\pi(\lambda) = C(\alpha, \beta)e^{-\beta\lambda}\lambda^{\alpha-1}$, for $\lambda > 0$. What is the normalising constant for the prior density? Find the posterior density of λ.

   Solution: We recall that the Gamma function is defined by

   $$\Gamma(x) = \int_0^\infty u^{x-1}e^{-u}du.$$

   The prior density of λ is

   $$\pi(\lambda) = C(\alpha, \beta)e^{-\beta\lambda}\lambda^{\alpha-1}$$

   and the normalising constant is

   $$C(\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}, \alpha > 0, \beta > 0.$$

   The posterior density of λ having observed the data $x = (x_1, \dots, x_n)$ is found using the relationship

   $$\pi(\lambda|x) \propto f(x|\lambda)\pi(\lambda),$$

   where

   $$f(x|\lambda) = \prod_{i=1}^n \frac{\lambda^{x_i}e^{-\lambda}}{x_i!} = \frac{\lambda^{\sum x_i}e^{-n\lambda}}{\prod x_i!},$$

   and

   $$\pi(\lambda) = C(\alpha, \beta)e^{-\beta\lambda}\lambda^{\alpha-1}.$$

   It follows that

   $$\pi(\lambda|x) \propto \lambda^{\sum x_i + \alpha - 1}e^{-(\beta+n)\lambda},$$

   after removing terms that do not involve λ. We recognise that the posterior distribution has the same form as the prior, i.e. it has a Gamma(α_1, β_1)-distribution, where

   $$\alpha_1 = \sum x_i + \alpha, \quad \beta_1 = \beta + n.$$

   Consequently the normalising constant is $\frac{{\beta_1}^{\alpha_1}}{\Gamma(\alpha_1)}$.

2. Let $X_1, \dots, X_n$ be a random sample of $\mathcal{N}(\theta, \sigma^2)$, where σ^2 known, and assume the prior distribution $\pi(\theta) \sim \mathcal{N}(\mu, \tau^2)$, where μ and τ^2 are known. Show that the posterior distribution of θ given the data $\mathbf{x}$ is $\pi(\theta|x) \sim \mathcal{N}\left(\frac{b}{a}, \frac{1}{a}\right)$ with

   $$a = \frac{n}{\sigma^2} + \frac{1}{\tau^2}; \qquad b = \frac{1}{\sigma^2}\sum x_i + \frac{\mu}{\tau^2}.$$

   Solution:

   We use the following **fact**: If $\alpha > 0$ and if

   $$p(x) \propto \exp(-\alpha x^2 - \beta x), \quad -\infty < x < \infty,$$

then p is normal.

From lectures:

$$f(x_1, \ldots, x_n|\theta) = (2\pi\sigma^2)^{-\frac{n}{2}} exp \left\{ -\frac{1}{2} \sum_{i=1}^{n} \frac{(x_i - \theta)^2}{\sigma^2} \right\}$$

so

$$\pi(\theta|\mathbf{x}) \quad \propto \quad exp \left\{ -\frac{1}{2} \left(\sum_{i=1}^{n} \frac{(x_i - \theta)^2}{\sigma^2} + \frac{(\theta - \mu)^2}{\tau^2} \right) \right\} =: e^{-\frac{1}{2}M}$$

So the posterior is normal. Calculate

$$M \quad = \quad a \left(\theta - \frac{b}{a} \right)^2 - \frac{b^2}{a} + c$$

where

$$a \quad = \quad \frac{n}{\sigma^2} + \frac{1}{\tau^2}; \quad b = \frac{1}{\sigma^2} \sum x_i + \frac{\mu}{\tau^2}; \quad c = \frac{1}{\sigma^2} \sum x_i^2 + \frac{\mu^2}{\tau^2}$$

Direct verification gives the result.

3. * Suppose that, as in Question 2, $X_1, \ldots, X_n$ is a random sample of $\mathcal{N}(\theta, \sigma^2)$, where σ^2 known, and assume the prior distribution $\pi(\theta) \sim \mathcal{N}(\mu, \tau^2)$, where μ and τ^2 are known. Show that the predictive distribution for x is $\mathcal{N}(\mu, \sigma^2 + \tau^2)$.

Solution: For the (prior) predictive distribution, is it easy to see that this is again normal;

$$\begin{aligned} p(x) &= \int f(x|\theta)\pi(\theta)d\theta \\ &\propto \int \exp\left(-\frac{1}{2\sigma^2}(x-\theta)^2 - \frac{1}{2\tau^2}(\theta-\mu)^2 \right) d\theta \\ &= e^{-\frac{1}{2\sigma^2}x^2} \int \exp\left(-\alpha\theta^2 - \beta(x-\gamma\mu)\theta \right) d\theta, \end{aligned}$$

where $\alpha = \alpha(\sigma^2, \tau^2, \mu)$ and $\beta = \beta(\sigma^2, \tau^2, \mu)$ are some functions of σ^2, τ^2, μ, and $\gamma = \gamma(\sigma^2, \tau^2)$ is a function of σ^2 and τ^2. The integral above will yield a normal integral when the square is completed, yielding that

$$p(x) \quad \propto \quad \exp\left\{ -\frac{1}{2\sigma^2}x^2 - \delta(x-\gamma\mu)^2 \right\},$$

where $\delta = \delta(\sigma^2, \tau^2, \mu)$ is some function of σ^2, τ^2, μ. Thus we recognize the normal distribution. To calculate mean and variance:

$$E(X) \quad = \quad \int E(X|\theta)\pi(\theta)d\theta = \int \theta\pi(\theta)d\theta = \mu$$

and

$$E(X^2) \quad = \quad \int E(X^2|\theta)\pi(\theta)d\theta = \int (\sigma^2 + \theta^2)\pi(\theta)d\theta = \sigma^2 + (\tau^2 + \mu^2)$$

and so

$$Var(X) = \sigma^2 + \tau^2 + \mu^2 - \mu^2 = \sigma^2 + \tau^2$$

as required.

4. Suppose that, given the success probability p, x is negative binomial with parameters 10 and p, i.e. x is the number of Bernoulli trials until there are 10 successes. Suppose that the prior for p is Beta(1/2, 1/2). Give the predictive distribution for x and the posterior density of p.

   Solution: We find the predictive distribution in the usual way. We will deal with a slightly more general case where x is negative binomial with parameters n and p. The predictive distribution is then

$$p(x) = \int_0^1 f(x|p)\pi(p)dp,$$

   where

$$f(x|p) = \binom{x-1}{n-1} p^n (1-p)^{x-n}, \quad x = n, n+1, \ldots$$

   and

$$\pi(p) = \frac{\Gamma(1)}{\Gamma(1/2)\Gamma(1/2)} p^{-1/2}(1-p)^{-1/2}, \quad 0 < p < 1,$$

   so that

$$\begin{aligned} p(x) &= \frac{\Gamma(1)}{\Gamma(1/2)\Gamma(1/2)} B\left(n+1/2, x-n+1/2\right) \binom{x-1}{n-1} \\ &= \frac{\Gamma(x-n+1/2)\Gamma(n+1/2)}{\pi x \Gamma(x-n+1)\Gamma(n)}, \end{aligned}$$

   for $x = n, n+1, \ldots$, is the predictive distribution.

   The posterior is

$$\begin{aligned} \pi(p|x) &\propto \binom{x-1}{n-1} p^n (1-p)^{x-n} p^{-1/2}(1-p)^{-1/2} \\ &\propto p^{n-1/2}(1-p)^{x-n-1/2}, \end{aligned}$$

   which we recognize as the $Beta(n+1/2, x-n+1/2)$-distribution.

5. The first apparently reliable datings of particular rock strata were obtained from the K/Ar method (comparing the proportions of Potassium 40 and Argon 40 in the rocks) in the 1960's, and these resulted in an estimate of 370±20 million years.

   In the late 1970's a newer method gave an age of 421±8 million years.

   Suppose that the K/Ar method results in a belief for the age which is normally distributed with mean 370 and standard deviation 20 million years. Suppose that the model for the newer method is that the observed age will be normally distributed with mean the true age and standard deviation 8 million years. How are the initial beliefs revised in the light of the results of the new method?

   Solution:

   From Problem 2, the posterior distribution of the normal mean is $\mathcal{N}(\theta_1, \phi_1) = \mathcal{N}\left(\frac{b}{a}, \frac{1}{a}\right)$ with

$$\begin{aligned} a &= \frac{1}{\sigma^2} + \frac{1}{\tau^2} \\ b &= \frac{x}{\sigma^2} + \frac{\mu}{\tau^2}. \end{aligned}$$

   In our case, $\mu = 370, x = 420, \tau = 20, \sigma = 8$, so that $\theta_1 = 413.1$ and $\phi_1 = 55.172$.