LTTC: Applied Bayesian Methods

You may use the following notation and results:

The **Beta distribution**, Beta(α, β), has probability density function

$$p(y \mid \alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1}(1-y)^{\beta-1} \ , \quad 0 < y < 1 \ ,$$

where $\Gamma(.)$ is the Gamma function, and mean $\frac{\alpha}{\alpha+\beta}$.

The **Binomial distribution**, Binomial(n, θ), has probability mass function

$$p(y \mid \theta) = \binom{n}{y} \theta^y (1-\theta)^{n-y} \ , \quad y = 0, 1, \ldots, n \ .$$

The **Gamma distribution**, Gamma(α, β), has probability density function

$$p(y \mid \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\beta y} \ , \quad y > 0 \ ,$$

and mean $\frac{\alpha}{\beta}$.

The **Negative Binomial distribution**, NegBin(r, θ), has probability mass function

$$p(y \mid \theta) = \binom{y+r-1}{y} \theta^r (1-\theta)^y \ , \quad y = 0, 1, 2, \ldots \ ,$$

and mean $\frac{r(1-\theta)}{\theta}$.

The **Normal distribution**, Normal(θ, τ^{-1}), has probability density function

$$p(y \mid \theta, \tau) = \sqrt{\frac{\tau}{2\pi}} \exp\left[-\frac{\tau}{2}(y-\theta)^2\right] \ , \quad -\infty < y < \infty \ .$$

The **Poisson distribution**, Poisson(λ), has probability mass function

$$p(y \mid \lambda) = \frac{\lambda^y}{y!} e^{-\lambda} \ , \quad y = 0, 1, 2, \ldots$$

Turn Over

1 (a) "Before reacting to a percentage you have to think what it is really telling you and to do that you need to put it in context. Take, for example, the statistic that 99% of deaths in the first four weeks of life occur in developing countries. Although that sounds horrifying, around 90 per cent of all births take place in developing countries. And so the chances of a baby dying in its first four weeks are 'only' 11 times greater there – bad enough, in all conscience." – Christina Pagel (from *Sense about Science and Straight Statistics*).

Show that "the chances of a baby dying in its first four weeks are 'only' 11 times greater there".

(b) In a population of women, we think the probability that a woman is pregnant is λ. For a new test for pregnancy, θ is the probability that a pregnant woman tests positive and η is the probability that a non-pregnant woman tests positive.

A woman decides to use the test to see whether she is pregnant. She tests positive. If she wants to repeat the test, what is the predictive probability that her second test will be negative? (Repeat tests on the same woman are conditionally independent given her true pregnancy status.)

(c) Suggest an optimal point estimator $T(y)$ for θ that minimises the posterior expected loss, $E_{\theta|y}[L(\theta, T(y))] = \int L(\theta, T(y))p(\theta|y)d\theta$, under zero-one loss. Justify your suggestion carefully.

(d) The Negative Binomial distribution, $\text{NegBin}(r, \theta)$, describes the distribution of the number of failures (denoted by Y) before the r-th success in an experiment that consists of a sequence of independent and identically distributed (i.i.d.) Bernoulli trials, where each trial has a probability θ of success. Suppose we collect a sample of n independent observations Y_i, $i = 1, \ldots, n$, with $Y_i \mid \theta \sim \text{NegBin}(r, \theta)$, where θ is unknown and the value of r is known.

Derive Jeffreys' prior for θ.

(e) Suppose we have collected n independent observations Y_i from a $\text{Normal}(\mu, \tau^{-1})$ distribution with unknown mean μ and unknown precision τ; that is,

$$Y_i \sim \text{Normal}(\mu, \tau^{-1}) \ , \quad \text{for} \quad i = 1, \ldots, n \ .$$

Two independent 'non-informative' priors are assumed for μ and τ:

$$\begin{aligned} \mu &\sim \text{Normal}(0, 10^6) \ , \\ \tau &\sim \text{Gamma}(0.001, 0.001) \ . \end{aligned}$$

Let $\mathbf{y}$ denote the values $(y_1, \ldots, y_n)$ of the observations. Derive two full-conditional distributions $p(\mu|\tau, \mathbf{y})$ and $p(\tau|\mu, \mathbf{y})$, and write down one of them as a Normal distribution and the other as a Gamma distribution with corresponding distribution parameters.

2 The data given in the table below are the numbers of pump failures, $Y_i = y_i$ for $i = 1, \ldots, 10$, observed in t_i thousands of hours for 10 different systems in a nuclear power plant.

	i-th system									
	1	2	3	4	5	6	7	8	9	10
y_i	5	1	5	14	3	19	1	1	4	22
t_i	94.3	15.7	62.9	125.8	5.2	31.4	1.0	1.0	2.1	10.5

An engineer is interested in modelling the numbers of failures Y_i and in statistical inference about the system-specific pump failure rates (i.e. the number of failures per thousand hours) θ_i . So you, as a Bayesian, have designed a simple, appropriate hierarchical Bayesian model for her.

(a) Construct a plausible 3-stage hierarchical model by writing down all the probabilistic distributions for the model that you have designed.

(b) Suppose, using a Gibbs sampler, you have obtained a sample, $(\theta_6^{(M+1)}, \ldots, \theta_6^{(N)})$, from the posterior distribution of θ_6. Explain how you would choose M and N. Explain how to use this sample to estimate the posterior median, the central 90% posterior credible interval, and the 50% highest posterior density (HPD) region of θ_6.

(c) Decribe how you could use a 'batching' method to estimate the Monte Carlo standard error (MCSE) of the posterior mean of θ_6, using the same sample $(\theta_6^{(M+1)}, \ldots, \theta_6^{(N)})$ as that in part (b).

(d) Explain how you could use your model and a Gibbs sampler to obtain samples from the predictive distribution of $\tilde{Y}$ in $\tilde{t} = 50$ thousands of hours for a new system in the power plant.

END OF PAPER