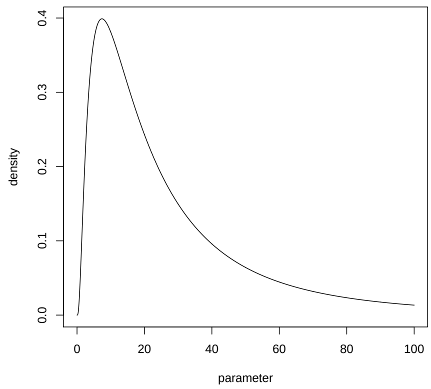
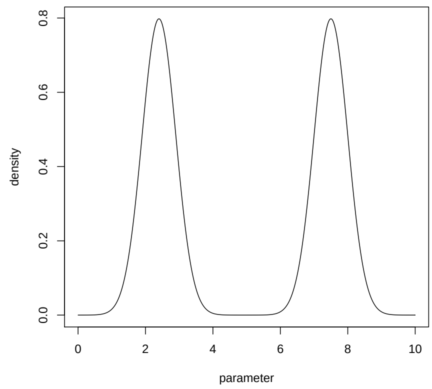
Bayesian Inference: Tutorial 3

1. (From 2009 exam - number in brackets correspond to number of marks - the total exam is out of 50 marks.) Let $X_1, \ldots, X_n$ denote independent random variables that depend on two parameters θ_1 and θ_2. Suppose that the posterior marginal distributions for the two parameters θ_1 and θ_2 are given in Figures 1(a) and (b) respectively.

Figure 1: Marginal posterior densities of parameters θ_1 and θ_2



(a) θ_1



(b) θ_2

Comment on the suitability of summarising these posterior distributions using the following summary statistics:

(a) posterior mean;

(b) posterior median;

(c) 95% symmetric credible interval;

(d) 95% highest posterior density interval.

[6]

Suggest another summary statistic that could be used in the summarising of the joint posterior distribution $\pi(\theta_1, \theta_2|\boldsymbol{x})$, where $\boldsymbol{x}$ denotes the observed data, to provide information on the posterior relationship between θ_1 and θ_2. [1]

2. (From December 2014 exam - number in brackets correspond to number of marks - the total exam is out of 50 marks.) A chemist is interested in the maximum possible yield produced by a certain chemical process. Due to the large variability in the data, he assumes that, given a scalar θ, each yield x_i, $i = 1, ..., n$, is independent of the other yields and follows a uniform distribution $U(0, \theta)$, so that,

$$f(x_i|\theta) = \frac{1}{\theta}, \qquad 0 < x_i < \theta.$$

Before the chemist sees any data, he assumes a Pareto prior distribution for θ, so that,

$$p(\theta) = \begin{cases} \frac{a x_0^a}{\theta^{a+1}} & \theta \geq x_0 \\ 0 & \text{otherwise,} \end{cases}$$

where $a > 0$ and $x_0 > 0$ are known parameters for the prior Pareto distribution. The mean of a Pareto distribution is given by, $\frac{ax_0}{a-1}$, for $a > 1$, whilst the median of a Pareto distribution is given by, $x_0 \times 2^{1/a}$.

a) Calculate the posterior distribution for θ. If the posterior is of standard form, state the distributional form (including the associated parameters of the distribution). [4]

b) Consider a Pareto prior distribution with parameters, $a = 2, x_0 = 0.1$. Consider also observed data $\boldsymbol{x} = \{x_1, x_2, x_3\} = \{3, 10, 17\}$. Obtain the posterior distribution and indicate how the expert's beliefs have changed after observing the data, using point summaries. Briefly discuss your findings. [2]

c) Consider an alternative Uniform prior $p(\theta) = U(0, 10)$. Without performing any calculations, discuss the scenario where $p(\theta) = U(0, 10)$, and the observed data are $\boldsymbol{x} = \{3, 9, 12\}$. [3]

3. We observe data $\boldsymbol{x} = \{x_1, \ldots, x_m\}$, from a Multinomial distribution, such that,

$$\boldsymbol{X} \sim MN(N, \boldsymbol{p}),$$

we wish to make inference on the parameters $\boldsymbol{p} = \{p_1, \ldots, p_m\}$. The prior on the unknown parameters $\boldsymbol{p}$, is specified to be of the form,

$$\boldsymbol{p} \sim Dir(\alpha_1, \ldots, \alpha_m).$$

What is the corresponding posterior distribution for the parameters $\boldsymbol{p}$? What is the posterior mean of p_i, $i = 1, \ldots, m$?
Note that a list of common probability distributions is provided in appendix A of the lecture notes

4. (From May 2009 exam - number in brackets correspond to number of marks - the total was 50 marks.) Let $X_1, \ldots, X_n$ be independent and identically distributed $N(\mu, \sigma^2)$ random variables, where σ^2 is known. We specify the prior,

$$\mu \sim TN(0, \tau^2),$$

with probability density function,

$$p(\mu) = \frac{\sqrt{2}}{\sqrt{\pi\tau^2}} \exp\left(-\frac{\mu^2}{2\tau^2}\right) \qquad 0 \le \mu < \infty.$$

(a) Calculate the posterior distribution of μ, given observed data $x_1, \ldots, x_n$. [5]

(b) For a given dataset, a practitioner obtains a posterior mean for μ of 3.5, posterior median of 2.96 and 95% highest posterior density interval of $(-1.5, 8.5)$. Without looking at their observed data, describe how the statistician automatically realises that there has been at least one mistake made within the analysis. [2]

(c) Discuss any disadvantages of specifying a $TN(0, \tau^2)$ prior on μ. [2]

Note that the following distributional information was also provided. A random variable X has a (positive) **Truncated-Normal distribution** $TN(\mu, \sigma^2)$, if its probability density function (p.d.f.) is

$$\frac{1}{\left(1 - \Phi\left(-\frac{\mu}{\sigma}\right)\right)\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \qquad 0 \le x < \infty,$$

where Φ denotes the cumulative distribution function of the standard Normal distribution. This result follows from,

$$\frac{N(\mu,\sigma^2)}{\int_0^\infty N(\mu,\sigma^2)dx} = \frac{N(\mu,\sigma^2)}{P(X>0|\mu,\sigma)} = \frac{N(\mu,\sigma^2)}{P(\frac{X-\mu}{\sigma} > \frac{-\mu}{\sigma}|\mu,\sigma)} = \frac{N(\mu,\sigma^2)}{1-P(Z<\frac{-\mu}{\sigma})},$$

where $Z \sim N(0,1)$.

5. (To discuss in groups if possible) You are presented with the following prior beliefs regarding an unknown parameter of interest $\theta \in \mathcal{R}$:

   (a) Prior information: average of 50 and bounds of [20,80]

   (b) Prior information: average of 50 and bounds of [25,100]

   Use this prior information to propose a sensible prior distribution on θ in each scenario.