MATH 11177: Bayesian Theory
Worksheet 3

(15 marks total)

Note: $I(\cdot)$ is an indicator function that equals 1 when the condition $\cdot$ is true and equals 0 when the condition if false.

1. (3 marks) Let $y_1, \ldots, y_n$ be *iid* Normal(μ, σ^2). Assume μ is known. Show that Jeffreys' prior for σ^2 is of the form $\pi(\sigma^2) \propto \frac{1}{\sigma^2}$, for $\sigma^2 > 0$.

2. (3 marks) Let $y_1, \ldots, y_n$ be *iid* Poisson(θ). Show that Jeffreys' prior for θ is proportional to $\frac{1}{\sqrt{\theta}}$.

3. (5 marks) A chemist is interested in the maximum possible yield of paracetamol that is produced by a new chemical manufacturing process. Due to the large variability in the data, they assume that, given a scalar θ, each yield y_i, $i = 1, \ldots, n$, is independent of the other yields and follows a uniform distribution $U(0, \theta)$. Thus θ is an upper bound on the yield of paracetamol. The sampling distribution for the yield is

$$f(y_i|\theta) = \frac{1}{\theta}\ I(0 < y_i \leq \theta),\ \ i = 1, \ldots, n$$

Before the chemist sees any data, they assume a Pareto prior distribution for θ, which is

$$\pi(\theta) = \frac{\alpha\beta^\alpha}{\theta^{\alpha+1}}\ I(\theta \geq \beta) \tag{1}$$

where $\alpha > 0$ and $\beta > 0$ are known parameters for the prior Pareto distribution, specified by the chemist[1]. Note that the mean of a Pareto distribution is $\frac{\alpha\beta}{\alpha-1}$, for $\alpha > 1$, and the median is $\beta \times 2^{1/\alpha}$. Examples of Pareto densities are shown in Figure 1.

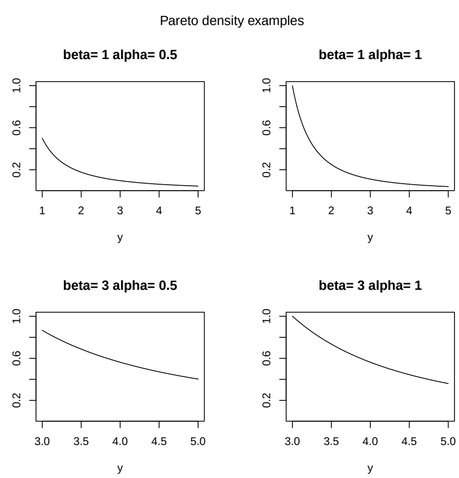

Figure 1: Pareto density examples

(a) Verify that $\pi(\theta)$ in Eq'n 1 is a probability density function (pdf).

[1]Note that the parameter β is setting a lower bound on the value of θ.

(b) Verify that the posterior distribution of θ, $p(\theta|y_1, \ldots, y_n)$, is a Pareto distribution (i.e., the Pareto is a conjugate prior).

(c) The chemist specifies a Pareto prior distribution with parameters, $\alpha = 2, \beta = 0.1$ and carried out a production run with n=3 batches of paracetamol. The observed yields were $= \{y_1, y_2, y_3\} = \{3, 10, 17\}$. Obtain the posterior distribution and indicate how the expert's beliefs have changed after observing the data using the mean and median values (prior vs posterior). Briefly discuss your findings.

(d) Suppose instead that the chemist specified the alternative prior $\theta \sim U(0, 15)$. Discuss the implications with regard to the posterior distribution for the given observed data.

4. (4 marks) We observe data $= \{y_1, \ldots, y_m\}$, from a Multinomial distribution, i.e.,

$$\sim \text{Multinomial}(N, p_1, p_2, \ldots, p_m),$$

The Multinomial is a *multivariate extension of the Binomial distribution* and the pmf is the following:

$$\Pr(y_1, \ldots, y_m) = \frac{N!}{y_1! y_2! \ldots y_m!} p_1^{y_1} p_2^{y_2} \ldots p_m^{y_m}$$

where y_i can equal 0, 1, 2, ..., or N subject to the constraint that $\sum_{i=1}^{m} y_i = N$. Further the p_i's are the probability of "event" i happening, where $0 \leq p_i \leq 1$ for all p_i, subject to the constraint that $\sum_{i=1}^{m} p_i = 1$.

We wish to make inference on the parameters $= \{p_1, \ldots, p_m\}$.

A Dirichlet prior on is specified:

$$\sim \text{Dirichlet}(\alpha_1, \ldots, \alpha_m).$$

The Dirichlet is a *multivariate extension of the Beta distribution*[2]. The pdf for a Dirichlet is the following:

$$\pi(p_1, \ldots, p_m) = \frac{\Gamma(\alpha_O)}{\prod_{i=1}^{m} \Gamma(\alpha_i)} \prod_{i=1}^{m} p_i^{\alpha_i - 1},$$

where $\alpha_i > 0$ and $\alpha_O = \sum_{i=1}^{m} \alpha_i$.

Further note:

$$E[p_i] = \frac{\alpha_i}{\alpha_O} \qquad Var[p_i] = \frac{\alpha_i(\alpha_O - \alpha_i)}{\alpha_O^2(\alpha_O + 1)}$$

(a) Determine the corresponding posterior distribution for the parameters .

(b) Calculate the posterior means for p_i.

(c) Application: We throw a 6-sided die 60 times and record the number of times that we observe the number $i = 1, \ldots, 6$, which we denote by y_i. Let p_i denote the associated probability of throwing the number $i = 1, \ldots, 6$ and set $= \{p_1, \ldots, p_6\}$. We observe the data $= \{10, 12, 12, 8, 7, 11\}$. We specify the following Dirichlet prior on the vector of p_i's:

$$Dir(\alpha_1, \ldots, \alpha_6) = Dir(1, 1, 1, 1, 1, 1)$$

Calculate the posterior means for p_1, p_2, ..., p_6.

[2]Given what you know about the relationship between the Binomial(n, p) sampling distribution and a Beta prior for p, you might be able to guess what the relationship between a Multinomial$(n, p_1, \ldots, p_m)$ sampling distribution and a Dirichlet prior for $p_1, \ldots, p_m$.