MAS3301 Bayesian Statistics

M. Farrow
School of Mathematics and Statistics
Newcastle University

Semester 2, 2008-9

9 Conjugate Priors II: More uses of the beta distribution

9.1 Geometric observations

9.1.1 Model

Suppose that we will observe $X_1, \ldots, X_n$ where these are mutually independent given the value of θ and

$$X_i \sim \text{geometric}(\theta)$$

where

$$\Pr(X_i = j \mid \theta) = (1-\theta)^{j-1}\theta$$

for $j = 1, 2, \ldots$.

Then the likelihood is

$$\begin{aligned} L(\theta;\ x) &= \prod_{i=1}^{n} \theta(1-\theta)^{x_i - 1} \\ &\propto \theta^n (1-\theta)^{S-n} \end{aligned}$$

where $S = \sum_{i=1}^{n} x_i$.

A conjugate prior is a *beta* distribution which has a pdf proportional to

$$\theta^{a-1}(1-\theta)^{b-1}$$

for $0 < \theta < 1$.

The posterior pdf is proportional to

$$\theta^{a-1}(1-\theta)^{b-1} \times \theta^n (1-\theta)^{S-n} = \theta^{a+n-1}(1-\theta)^{b+S-n-1}.$$

This is proportional to the pdf of a beta$(a+n,\ b + \sum x_i - n)$ distribution.

9.1.2 Example

Consider the problem posed in Question 3 of Problems 1 (2.7) except that we will give the probability of a defective item a beta prior.
A machine is built to make mass-produced items. Each item made by the machine has a probability p of being defective. Given the value of p, the items are independent of each other. The prior distribution for p is a beta(a, b) distribution. The machine is tested by counting the number of items made before a defective is produced. Find the posterior distribution of p given that the first defective item is the thirteenth to be made.

Here $n = 1$ and $x = 13$ so the posterior distribution is beta$(a+1, b+12)$. For example, suppose that $a = b = 1$ so that the prior distribution is uniform. Then the posterior distribution is beta$(2, 13)$. The prior mean is 0.5. The posterior mean is $2/(2+13) = 0.1333$. The prior variance is $1/12 = 0.0833333$ giving a standard deviation of 0.2887. The posterior variance is 0.007222 giving a standard deviation of 0.08498.
Note that, in the original question, we did not consider values of p greater than 0.05 to be possible. We could introduce such a constraint and still give p a continuous prior. We will return to this question later.

9.2 Negative binomial observations

9.2.1 Model

Suppose that we will observe $X_1, \ldots, X_n$ where these are mutually independent given the value of θ and

$$X_i \sim \text{negative} - \text{binomial}(r_i, \theta)$$

where $r_1, \ldots, r_n$ are known or will be observed and

$$\Pr(X_i = j \mid \theta) = \begin{pmatrix} j-1 \\ r_i - 1 \end{pmatrix} (1-\theta)^{j-r_i} \theta^{r_i}$$

for $j = r_i, r_i + 1, \ldots$.

Then the likelihood is

$$\begin{aligned} L(\theta;\ x) &= \prod_{i=1}^{n} \begin{pmatrix} x_i - 1 \\ r_i - 1 \end{pmatrix} \theta^{r_i} (1-\theta)^{x_i - r_i} \\ &\propto \theta^R (1-\theta)^{S-R} \end{aligned}$$

where $S = \sum_{i=1}^n x_i$ and $R = \sum_{i=1}^n r_i$.

A conjugate prior is a *beta* distribution which has a pdf proportional to

$$\theta^{a-1}(1-\theta)^{b-1}$$

for $0 < \theta < 1$.

The posterior pdf is proportional to

$$\theta^{a-1}(1-\theta)^{b-1} \times \theta^R (1-\theta)^{S-R} = \theta^{a+R-1}(1-\theta)^{b+S-R-1}.$$

This is proportional to the pdf of a beta$(a + \sum r_i,\ b + \sum x_i - \sum r_i)$ distribution.

9.2.2 Example

Consider the problem posed in Question 3 of Problems 1 (2.7) except that we will give the probability of a defective item a beta prior and we will continue observing items until we have found 5 defectives.
A machine is built to make mass-produced items. Each item made by the machine has a probability p of being defective. Given the value of p, the items are independent of each other. The prior distribution for p is a beta(a, b) distribution. The machine is tested by counting the number of items made before five defectives are produced. Find the posterior distribution of p given that the fifth defective item is the 73rd to be made.

Here $n = 1$, $r = 5$ and $x = 73$ so the posterior distribution is beta$(a + 5, b + 68)$. For example, suppose that $a = b = 1$ so that the prior distribution is uniform. Then the posterior distribution is beta$(6, 69)$. The prior mean is 0.5. The posterior mean is $6/(6+69) = 0.08$. The prior variance is $1/12 = 0.0833333$ giving a standard deviation of 0.2887. The posterior variance is 0.0009684 giving a standard deviation of 0.03112.

10 Conjugate Priors III: Use of the gamma distribution

10.1 Gamma distribution

The gamma distribution is a conjugate prior for a number of models, including Poisson and exponential data. In a later lecture we will also see that it has a role in the case of normal data.

If $\theta \sim \text{gamma}(a, b)$ then it has pdf given by

$$f(\theta) = \begin{cases} 0 & (\theta \leq 0) \\ \frac{b^a \theta^{a-1} e^{-b\theta}}{\Gamma(a)} & (0 < \theta < \infty) \end{cases}$$

where $a > 0$, $b > 0$ and $\Gamma(u)$ is the *gamma function*

$$\Gamma(u) = \int_0^\infty w^{u-1} e^{-w} \, dw = (u-1)\Gamma(u-1).$$

If $\theta \sim \text{gamma}(a, b)$ then the mean and variance of θ are

$$\mathrm{E}(\theta) = \frac{a}{b}, \qquad \mathrm{var}(\theta) = \frac{a}{b^2}.$$

Proof:

$$\begin{aligned} \mathrm{E}(\theta) &= \int_0^\infty \theta \frac{b^a \theta^{a-1} e^{-b\theta}}{\Gamma(a)} \, d\theta \\ &= \frac{\Gamma(a+1)}{b\Gamma(a)} \int_0^\infty \frac{b^{a+1} \theta^{a+1-1} e^{-b\theta}}{\Gamma(a+1)} \, d\theta \\ &= \frac{\Gamma(a+1)}{b\Gamma(a)} = \frac{a}{b}. \end{aligned}$$

Similarly

$$\mathrm{E}(\theta^2) = \frac{\Gamma(a+2)}{b^2\Gamma(a)} = \frac{(a+1)a}{b^2}.$$

So

$$\begin{aligned} \mathrm{var}(\theta) &= \mathrm{E}(\theta^2) - [\mathrm{E}(\theta)]^2 \\ &= \frac{(a+1)a}{b^2} - \frac{a^2}{b^2} = \frac{a}{b^2}. \end{aligned}$$

Notes:

1. We call a the *shape parameter* or *index* and b the *scale parameter*. Sometimes people use $c = b^{-1}$ instead of b so the pdf becomes

$$\frac{c^{-1}(\theta/c)^{a-1} e^{-\theta/c}}{\Gamma(a)}.$$

2. The coefficient of variation, that is the standard deviation divided by the mean, is $(\sqrt{a}/b^2)/(a/b) = 1/\sqrt{a}$.

3. For $a = 1$ the distribution is exponential. For large a it is more symmetric and closer to a normal distribution.

4. A χ^2_ν distribution is a $\text{gamma}(\nu/2,\ 1/2)$ distribution. Thus, if $X \sim \text{gamma}(a, b)$ and $Y = 2bX$, then $Y \sim \chi^2_{2a}$.

10.2 Poisson observations

10.2.1 Model

Suppose that we will observe $X_1, \ldots, X_n$ where these are mutually independent given the value of θ and

$$X_i \sim \text{Poisson}(\theta).$$

Then the likelihood is

$$\begin{aligned} L(\theta;\ x) &= \prod_{i=1}^{n} \frac{e^{-\theta}\theta^{x_i}}{x_i!} \\ &\propto \theta^{S} e^{-n\theta} \end{aligned}$$

where $S = \sum_{i=1}^{n} x_i$.

The conjugate prior is a *gamma* distribution which has a pdf proportional to

$$\theta^{a-1}e^{-b\theta}$$

for $0 < \theta < \infty$.

The posterior pdf is proportional to

$$\theta^{a-1}e^{-b\theta} \times \theta^{S}e^{-n\theta} = \theta^{a+S-1}e^{-(b+n)\theta}.$$

This is proportional to the pdf of a gamma$(a + \sum x_i,\ b + n)$ distribution.

10.2.2 Example

An ecologist counts the numbers of centipedes in each of twenty one-metre-square quadrats. The numbers are as follows.

```
14 13  7 10 15 15  2 13 13 11 10 13  5 13  9 12  9 12  8  7
```

It is assumed that the numbers are independent and drawn from a Poisson distribution with mean θ. The prior distribution for θ is a gamma distribution with mean 20 and standard deviation 10. This gives $a/b = 20$ and $a/b^2 = 100$. Therefore $a = 4$ and $b = 0.2$. From the data $S = \sum x_i = 211$ and $n = 20$.
The posterior distribution for θ is gamma$(4 + 211,\ 0.2 + 20)$. That is gamma(215,20.2). The posterior mean is $215/20.2 = 10.64$, the posterior variance is $215/20.2^2 = 0.5269$ and the posterior standard deviation is $\sqrt{0.5269} = 0.726$.

10.3 Exponential observations

10.3.1 Model

Suppose that we will observe $X_1, \ldots, X_n$ where these are mutually independent given the value of θ and

$$X_i \sim \text{exponential}(\theta).$$

Then the likelihood is

$$\begin{aligned} L(\theta;\ x) &= \prod_{i=1}^{n} \theta e^{-\theta x_i} \\ &= \theta^n e^{-S\theta} \end{aligned}$$

where $S = \sum_{i=1}^{n} x_i$.

The conjugate prior is a *gamma* distribution which has a pdf proportional to

$$\theta^{a-1} e^{-b\theta}$$

for $0 < \theta < \infty$.

The posterior pdf is proportional to

$$\theta^{a-1} e^{-b\theta} \times \theta^n e^{-S\theta} = \theta^{a+n-1} e^{-(b+S)\theta}.$$

This is proportional to the pdf of a gamma$(a+n,\ b+\sum x_i)$ distribution.

10.3.2 Example

A machine continuously produces nylon filament. From time to time the filament snaps. Suppose that the time intervals, in minutes, between snaps are random, independent and have an exponential(θ) distribution. Our prior distribution for θ is a gamma$(6, 1800)$ distribution. This gives a prior mean of $6/1800 = 0.0033$, a prior variance of $6/1800^2 = 1.85 \times 10^{-6}$ and a prior standard deviation of $\sqrt{1.85 \times 10^{-6}} = 0.00136$.

Note that the mean time between snaps is $1/\theta$. We say that this mean has an *inverse gamma* prior since its inverse has a gamma prior.

The prior mean for $1/\theta$, when $\theta \sim \text{gamma}(a, b)$ with $a > 1$ is

$$\begin{aligned}
\mathrm{E}\left(\frac{1}{\theta}\right) &= \int_0^\infty \frac{\theta^{-1} b^a \theta^{a-1} e^{-b\theta}}{\Gamma(a)} \, d\theta \\
&= \frac{b^a \Gamma(a-1)}{b^{a-1}\Gamma(a)} \int_0^\infty \frac{b^{a-1}\theta^{a-1-1} e^{-b\theta}}{\Gamma(a-1)} \, d\theta \\
&= \frac{b}{a-1}
\end{aligned}$$

Similarly, provided $a > 2$,

$$\mathrm{E}\left(\frac{1}{\theta^2}\right) = \frac{b^2}{(a-1)(a-2)}$$

so

$$\mathrm{var}\left(\frac{1}{\theta}\right) = \frac{b^2}{(a-1)(a-2)} - \left(\frac{b}{a-1}\right)^2 = \frac{b^2}{(a-1)^2(a-2)}.$$

In our case the prior mean of the time between snaps is $1800/5 = 360$ (i.e. 6 hours), the prior variance is $1800^2/(5^2 \times 4) = 32400$ and the prior standard deviation is $\sqrt{32400} = 180$ (i.e. 3 hours).

Forty intervals are observed with lengths as follows.

55	30	231	592	141	139	695	56	803	642	1890	208	246	183	38	486
264	1091	368	222	662	150	2	133	417	418	743	216	138	306	201	145
804	193	66	577	773	268	388	861								

So $S = \sum x_i = 15841$ and $n = 40$. Thus our posterior distribution for θ is $\text{gamma}(6+40, 1800+15841)$. That is it is $\text{gamma}(46, 17641)$. This gives a posterior mean of $46/17641 = 0.00261$, a posterior variance of $46/17641^2 = 1.478 \times 10^{-7}$ and a posterior standard deviation of $\sqrt{1.478 \times 10^{-7}} = 3.845 \times 10^{-4}$. The posterior mean for the time between snaps is $17641/45 = 392.0$, the posterior variance of the mean time between snaps is $17641^2/(45^2 \times 44) = 3492.76$ and the posterior standard deviation is $\sqrt{3492.76} = 59.1$.

10.4 Finding a hpd interval

Suppose that we want to find a hpd interval for a scalar parameter θ which has a unimodal posterior pdf $f^{(1)}(\theta)$ and we want this interval to have probability P. Then we require the following two conditions.

$$\begin{aligned} \int_{k_1}^{k_2} f^{(1)}(\theta)\, d\theta &= P \\ f^{(1)}(k_1) &= f^{(1)}(k_2) \end{aligned}$$

where k_1 is the lower limit and k_2 is the upper limit of the interval.

For example, consider the centipedes example in section 10.2.2. We can use a R function like the following.

```
function(p,a,b,delta=0.0001,maxiter=30)
{tail<-1-p
 delta<-delta^2
 lower1<-0.0
 lower2<-qgamma(tail,a,b)
 i<-0
 progress<-matrix(rep(0,(3*maxiter)),ncol=3)
 write.table("Lower   Upper   Difference",file="")
 while((diff^2>delta) & (i<maxiter))
    {low<-(lower1+lower2)/2
     lprob<-pgamma(low,a,b)
     ldens<-dgamma(low,a,b)
     uprob<-lprob+p
     upp<-qgamma(uprob,a,b)
     udens<-dgamma(upp,a,b)
     diff<-(udens-ldens)/(udens+ldens)
     i<-i+1
     if (diff<0) lower2<-low else lower1<-low
     progress[i,1]<-low
     progress[i,2]<-upp
     progress[i,3]<-diff
     }
 progress<-signif(progress[1:i,],6)
 write.table(progress,file="",col.names=FALSE)
 result<-c(low,upp)
 result
 }
```

This function uses the fact that the lower limit of a $100\alpha\%$ interval has to be between 0.0 and the lower $100(1-\alpha)\%$ point of the gamma distribution. It then uses "interval halving" to find the lower limit such that the difference between the densities at the upper and lower limits is zero.

Here we use it to find out hpd interval.

```
> hpdgamma(0.95,215,20.2)
"   Lower   Upper   Difference"
1 4.73919 11.865 1
2 7.10879 11.865 0.999997
3 8.29359 11.8669 0.977787
4 8.88599 11.9082 0.664075
5 9.18219 12.0306 0.140152
6 9.33029 12.2034 -0.286223
7 9.25624 12.0986 -0.0541013
8 9.21921 12.0613 0.0470541
```

```
9 9.23773 12.079 -0.00244887
10 9.22847 12.07 0.0225619
11 9.2331 12.0745 0.0101224
12 9.23541 12.0767 0.00385341
13 9.23657 12.0779 0.000706445
14 9.23715 12.0785 -0.000870166
15 9.23686 12.0782 -8.15991e-05
[1]  9.23686 12.07817
```

So, we have found our 95% hpd interval to three decimal places as $9.237 < \theta < 12.078$.

10.5 Problems 3

1. In a small survey, a random sample of 50 people from a large population is selected. Each person is asked a question to which the answer is either "Yes" or "No." Let the proportion in the population who would answer "Yes" be θ. Our prior distribution for θ is a beta$(1.5, 1.5)$ distribution. In the survey, 37 people answer "Yes."

   (a) Find the prior mean and prior standard deviation of θ.

   (b) Find the prior probability that $\theta < 0.6$.

   (c) Find the likelihood.

   (d) Find the posterior distribution of θ.

   (e) Find the posterior mean and posterior standard deviation of θ.

   (f) Plot a graph showing the prior and posterior probability density functions of θ on the same axes.

   (g) Find the posterior probability that $\theta < 0.6$.

   Notes:

   The probability density function of a beta(a, b) distribution is $f(x) = kx^{a-1}(1-x)^{b-1}$ where k is a constant.

   If $X \sim$ beta(a, b) then the mean of X is

   $$\mathrm{E}(X) = \frac{a}{a+b}$$

   and the variance of X is

   $$\mathrm{var}(X) = \frac{ab}{(a+b+1)(a+b)^2}.$$

   If $X \sim$ beta(a, b) then you can use a command such as the following in R to find $\Pr(X < c)$.

   ```
   pbeta(c,a,b)
   ```

   To plot the prior and posterior probability densities you may use R commands such as the following.

   ```
   theta<-seq(0.01,0.99,0.01)
   prior<-dbeta(theta,a,b)
   posterior<-dbeta(theta,c,d)
   plot(theta,posterior,xlab=expression(theta),ylab="Density",type="l")
   lines(theta,prior,lty=2)
   ```

2. The populations, n_i, and the number of cases, x_i, of a disease in a year in each of six districts are given in the table below.

Population n	Cases x
120342	2
235967	5
243745	3
197452	5
276935	3
157222	1

We suppose that the number X_i in a district with population n_i is a Poisson random variable with mean $n_i\lambda/100000$. The number in each district is independent of the numbers in other districts, given the value of λ. Our prior distribution for λ is a gamma distribution with mean 3.0 and standard deviation 2.0.

(a) Find the parameters of the prior distribution.

(b) Find the prior probability that $\lambda < 2.0$.

(c) Find the likelihood.

(d) Find the posterior distribution of λ.

(e) Find the posterior mean and posterior standard deviation of λ.

(f) Plot a graph showing the prior and posterior probability density functions of λ on the same axes.

(g) Find the posterior probability that $\lambda < 2.0$.

Notes:

The probability density function of a gamma(a, b) distribution is $f(x) = kx^{a-1}\exp(-bx)$ where k is a constant.

If $X \sim$ gamma(a, b) then the mean of X is $\mathrm{E}(X) = a/b$ and the variance of X is $\mathrm{var}(X) = a/(b^2)$.

If $X \sim$ gamma(a, b) then you can use a command such as the following in R to find $\Pr(X < c)$.

```
pgamma(c,a,b)
```

To plot the prior and posterior probability densities you may use R commands such as the following.

```
lambda<-seq(0.00,5.00,0.01)
prior<-dgamma(lambda,a,b)
posterior<-dgamma(lambda,c,d)
plot(lambda,posterior,xlab=expression(lambda),ylab="Density",type="l")
lines(lambda,prior,lty=2)
```

3. Geologists note the type of rock at fixed vertical intervals of six inches up a quarry face. At this quarry there are four types of rock. The following model is adopted.

The conditional probability that the next rock type is j given that the present type is i and given whatever has gone before is p_{ij}. Clearly $\sum_{j=1}^{4} p_{ij} = 1$ for all i.

The following table gives the observed (upwards) transition frequencies.

		To rock			
		1	2	3	4
From rock	1	56	13	24	4
	2	15	93	22	35
	3	20	25	153	11
	4	6	35	11	44

Our prior distribution for the transition probabilities is as follows. For each i we have a uniform distribution over the space of possible values of $p_{i1}, \ldots, p_{i4}$. The prior distribution of $p_{i1}, \ldots, p_{i4}$ is independent of that for $p_{k1}, \ldots, p_{k4}$ for $i \neq k$.

Find the matrix of posterior expectations of the transition probabilities.

Note that the integral of $x_1^{n_1} x_2^{n_2} x_3^{n_3} x_4^{n_4}$ over the region such that $x_j > 0$ for $j = 1, \ldots, 4$ and $\sum_{j=1}^{4} x_j = 1$, where $n_1, \ldots, n_4$ are positive is

$$\int_0^1 x_1^{n_1} \int_0^{1-x_1} x_2^{n_2} \int_0^{1-x_1-x_2} x_3^{n_3} (1 - x_1 - x_2 - x_3)^{n_4} .dx_3.dx_2.dx_1$$

$$= \frac{\Gamma(n_1+1)\Gamma(n_2+1)\Gamma(n_3+1)\Gamma(n_4+1)}{\Gamma(n_1+n_2+n_3+n_4+4)}$$

4. A biologist is interested in the proportion, θ, of badgers in a particular area which carry the infection responsible for bovine tuberculosis. The biologist's prior distribution for θ is a beta$(1, 19)$ distribution.

   (a) i. Find the biologist's prior mean and prior standard deviation for θ.

   ii. Find the cumulative distribution function of the biologist's prior distribution and hence find values θ_1, θ_2 such that, in the biologist's prior distribution, $\Pr(\theta < \theta_1) = \Pr(\theta > \theta_2) = 0.05$.

   (b) The biologist captures twenty badgers and tests them for the infection. Assume that, given θ, the number, X, of these carrying the infection has a binomial$(20, \theta)$ distribution. The observed number carrying the infection is $x = 2$.

   i. Find the likelihood function.

   ii. Find the biologist's posterior distribution for θ.

   iii. Find the biologist's posterior mean and posterior standard deviation for θ.

   iv. Use R to plot a graph showing the biologist's prior and posterior probability density functions for θ.

5. A factory produces large numbers of packets of nuts. As part of the quality control process, samples of the packets are taken and weighed to check whether they are underweight. Let the true proportion of packets which are underweight be θ and assume that, given θ, the packets are independent and each has probability θ of being underweight. A beta$(1, 9)$ prior distribution for θ is used.

   (a) The procedure consists of selecting packets until either an underweight packet is found, in which case we stop and note the number X of packets examined, or $m = 10$ packets are examined and none is underweight, in which we case we stop and note this fact.

   i. Find the posterior distribution for θ when $X = 7$ is observed.

   ii. Find the posterior distribution for θ when no underweight packets are found out of $m = 10$.

   (b) Now consider varying the value of m. Use R to find the posterior probability that $\theta < 0.02$ when no underweight packets are found out of

   i. $m = 10$,

   ii. $m = 20$,

   iii. $m = 30$.

6. The numbers of patients arriving at a minor injuries clinic in 10 half-hour intervals are recorded. It is supposed that, given the value of a parameter λ, the number X_j arriving in interval j has a Poisson distribution $X_j \sim$ Poisson(λ) and X_j is independent of X_k for $j \neq k$. The prior distribution for λ is a gamma(a, b) distribution. The prior mean is 10 and the prior standard deviation is 5.

   (a) i. Find the values of a and b.

ii. Let $W \sim \chi^2_{2a}$. Find values w_1, w_2 such that $\Pr(W < w_1) = \Pr(W > w_2) = 0.025$. Hence find values l_1, l_2 such that, in the prior distribution, $\Pr(\lambda < l_1) = \Pr(\lambda > l_2) = 0.025$.

iii. Using R (or otherwise) find a 95% prior highest probability density interval for λ.

iv. Compare these two intervals.

(b) The data are as follows.

```
9  12  16  12  16  11  18  13  12  19
```

i. Find the posterior distribution of λ.

ii. Using R (or otherwise) find a 95% posterior highest probability density interval for λ.

7. The numbers of sales of a particular item from an Internet retail site in each of 20 weeks are recorded. Assume that, given the value of a parameter λ, these numbers are independent observations from the Poisson(λ) distribution.

Our prior distribution for λ is a gamma(a, b) distribution.

(a) Our prior mean and standard deviation for λ are 16 and 8 respectively. Find the values of a and b.

(b) The observed numbers of sales are as follows.

```
14 19 14 21 22 33 15 13 16 19 27 22 27 21 16 25 14 23 22 17
```

Find the posterior distribution of λ.

(c) Using R or otherwise, plot a graph showing both the prior and posterior probability density functions of λ.

(d) Using R or otherwise, find a 95% posterior hpd interval for λ. *(Note: The R function* `hpdgamma` *is available from the Module Web Page).*

8. In a medical experiment, patients with a chronic condition are asked to say which of two treatments, A, B, they prefer. (You may assume for the purpose of this question that every patient will express a preference one way or the other). Let the population proportion who prefer A be θ. We observe a sample of n patients. Given θ, the n responses are independent and the probability that a particular patient prefers A is θ.

Our prior distribution for θ is a beta(a, a) distribution with a standard deviation of 0.25.

(a) Find the value of a.

(b) We observe $n = 30$ patients of whom 21 prefer treatment A. Find the posterior distribution of θ.

(c) Find the posterior mean and standard deviation of θ.

(d) Using R or otherwise, plot a graph showing both the prior and posterior probability density functions of θ.

(e) Using R or otherwise, find a symmetric 95% posterior probability interval for θ. *(Hint: The R command* `qbeta(0.025,a,b)` *will give the 2.5% point of a beta(a, b) distribution).*

9. The survival times, in months, of patients diagnosed with a severe form of a terminal illness are thought to be well modelled by an exponential(λ) distribution. We observe the survival times of n such patients. Our prior distribution for λ is a gamma(a, b) distribution.

(a) Prior beliefs are expressed in terms of the median lifetime, m. Find an expression for m in terms of λ.

(b) In the prior distribution, the lower 5% point for m is 6.0 and the upper 5% point is 46.2. Find the corresponding lower and upper 5% points for λ. Let these be k_1, k_2 respectively.

(c) Let $k_2/k_1 = r$. Find, to the nearest integer, the value of ν such that, in a χ^2_ν distribution, the 95% point divided by the 5% point is r and hence deduce the value of a.

(d) Using your value of a and one of the percentage points for λ, find the value of b.

(e) We observe $n = 25$ patients and the sum of the lifetimes is 502. Find the posterior distribution of λ.

(f) Using the relationship of the gamma distribution to the χ^2 distribution, or otherwise, find a symmetric 95% posterior interval for λ.

Note: The R command `qchisq(0.025,nu)` *will give the lower 2.5% point of a* χ^2 *distribution on* `nu` *degrees of freedom.*

Homework 3

Solutions to Questions **7, 8, 9** *of Problems 3 are to be submitted in the Homework Letterbox no later than 4.00pm on Monday March 9th.*