MAS3301 Bayesian Statistics
Problems 5 and Solutions

Semester 2

2008-9

Problems 5

1. *(Some of this question is also in Problems 4).* I recorded the attendance of students at tutorials for a module. Suppose that we can, in some sense, regard the students as a sample from some population of students so that, for example, we can learn about the likely behaviour of next year's students by observing this year's. At the time I recorded the data we had had tutorials in Week 2 and Week 4. Let the probability that a student attends in both weeks be θ_{11}, the probability that a student attends in week 2 but not Week 4 be θ_{10} and so on. The data are as follows.

Attendance	Probability	Observed frequency
Week 2 and Week 4	θ_{11}	$n_{11} = 25$
Week 2 but not Week 4	θ_{10}	$n_{10} = 7$
Week 4 but not Week 2	θ_{01}	$n_{01} = 6$
Neither week	θ_{00}	$n_{00} = 13$

Suppose that the prior distribution for $(\theta_{11}, \theta_{10}, \theta_{01}, \theta_{00})$ is a Dirichlet distribution with density proportional to

$$\theta_{11}^3 \theta_{10} \theta_{01} \theta_{00}^2.$$

(a) Find the prior means and prior variances of $\theta_{11},\ \theta_{10},\ \theta_{01},\ \theta_{00}$.

(b) Find the posterior distribution.

(c) Find the posterior means and posterior variances of $\theta_{11},\ \theta_{10},\ \theta_{01},\ \theta_{00}$.

(d) Using the R function `hpdbeta` which may be obtained from the Web page (or otherwise), find a 95% posterior hpd interval, based on the exact posterior distribution, for θ_{00}.

(e) Find an approximate 95% hpd interval for θ_{00} using a normal approximation based on the posterior mode and the partial second derivatives of the log posterior density.

Compare this with the exact hpd interval.

Hint: To find the posterior mode you will need to introduce a Lagrange multiplier.

(f) The population mean number of attendances out of two is $\mu = 2\theta_{11} + \theta_{10} + \theta_{01}$. Find the posterior mean of μ and an approximation to the posterior standard deviation of μ.

2. Samples are taken from twenty wagonloads of an industrial mineral and analysed. The amounts in ppm (parts per million) of an impurity are found to be as follows.

```
44.3 50.2 51.7 49.4 50.6 55.0 53.5 48.6 48.8 53.3
59.4 51.4 52.0 51.9 51.6 48.3 49.3 54.1 52.4 53.1
```

We regard these as independent samples from a normal distribution with mean μ and variance $\sigma^2 = \tau^{-1}$.

Find a 95% posterior hpd interval for μ under each of the following two conditions.

(a) The value of τ is known to be 0.1 and our prior distribution for μ is normal with mean 60.0 and standard deviation 20.0.

(b) The value of τ is unknown. Our prior distribution for τ is a gamma distribution with mean 0.1 and standard deviation 0.05. Our conditional prior distribution for μ given τ is normal with mean 60.0 and precision 0.025τ (that is, standard deviation $\sqrt{40}\tau^{-1/2}$).

3. We observe a sample of 30 observations from a normal distribution with mean μ and precision τ. The data, $y_1, \ldots, y_{30}$, are such that

$$\sum_{i=1}^{30} y_i = 672 \qquad \text{and} \qquad \sum_{i=1}^{30} y_i^2 = 16193.$$

(a) Suppose that the value of τ is known to be 0.04 and that our prior distribution for μ is normal with mean 20 and variance 100. Find the posterior distribution of μ and evaluate a posterior 95% hpd interval for μ.

(b) Suppose that we have a gamma$(1, 10)$ prior distribution for τ and our conditional prior distribution for μ given τ is normal with mean 20 and variance $(0.1\tau)^{-1}$. Find the marginal posterior distribution for τ, the marginal posterior distribution for μ and the marginal posterior 95% hpd interval for μ.

4. The following data come from the experiment reported by MacGregor *et al.* (1979). They give the supine systolic blood pressures (mm Hg) for fifteen patients with moderate essential hypertension. The measurements were taken immediately before and two hours after taking a drug.

Patient	1	2	3	4	5	6	7	8
Before	210	169	187	160	167	176	185	206
After	201	165	166	157	147	145	168	180
Patient	9	10	11	12	13	14	15	
Before	173	146	174	201	198	148	154	
After	147	136	151	168	179	129	131	

We are interested in the effect of the drug on blood pressure. We assume that, given parameters μ, τ, the changes in blood pressure, from before to after, in the n patients are independent and normally distributed with unknown mean μ and unknown precision τ. The fifteen differences are as follows.

```
-9  -4 -21  -3 -20 -31 -17 -26 -26 -10 -23 -33 -19 -19 -23
```

Our prior distribution for τ is a gamma$(0.35, 1.01)$ distribution. Our conditional prior distribution for μ given τ is a normal $N(0,\ [0.003\tau]^{-1})$ distribution.

(a) Find the marginal posterior distribution of τ.

(b) Find the marginal posterior distribution of μ.

(c) Find the marginal posterior 95% hpd interval for μ.

(d) Comment on what you can conclude about the effect of the drug.

5. The lifetimes of certain components are supposed to follow a Weibull distribution with known shape parameter $\alpha = 2$. The probability density function of the lifetime distribution is

$$f(t) = \alpha\rho^2 t \exp[-(\rho t)^2]$$

for $0 < t < \infty$.

We will observe a sample of n such lifetimes where n is large.

(a) Assuming that the prior density is nonzero and reasonably flat so that it may be disregarded, find an approximation to the posterior distribution of ρ. Find an approximate 95% hpd interval for ρ when $n = 300$, $\sum \log(t) = 1305.165$ and $\sum t^2 = 3161776$.

(b) Assuming that the prior distribution is a gamma(a, b) distribution, find an approximate 95% hpd interval for ρ, taking into account this prior, when $a = 2$, $b = 100$, $n = 300$, $\sum \log(t) = 1305.165$ and $\sum t^2 = 3161776$.

6. Given the value of λ, the number X_i of transactions made by customer i at an online store in a year has a Poisson(λ) distribution, with X_i independent of X_j for $i \neq j$. The value of λ is unknown. Our prior distribution for λ is a gamma(5,1) distribution.

We observe the numbers of transactions in a year for 45 customers and

$$\sum_{i=1}^{45} x_i = 182.$$

(a) Using a χ^2 table (i.e. without a computer) find the lower 2.5% point and the upper 2.5% point of the prior distribution of λ.

(These bound a 95% symmetric prior credible interval).

(b) Find the posterior distribution of λ.

(c) Using a normal approximation to the posterior distribution, based on the posterior mean and variance, find a 95% symmetric posterior credible interval for λ.

(d) Find an expression for the posterior predictive probability that a customer makes m transactions in a year.

(e) As well as these "ordinary customers," we believe that there is a second group of individuals. The number of transactions in a year for a member of this second group has, given θ, a Poisson(θ) distribution and our beliefs about the value of θ are represented by a gamma(1,0.05) distribution.

A new individual is observed who makes 10 transactions in a year. Given that our prior probability that this is an ordinary customer is 0.9, find our posterior probability that this is an ordinary customer.

Hint: You may find it best to calculate the logarithms of the predictive probabilities before exponentiating these. For this you might find the R function `lgamma` useful. It calculates the log of the gamma function. Alternatively it is possible to do the calculation using the R function `dnbinom`.

(N.B. In reality a slightly more complicated model is used in this type of application).

7. The following data give the heights in cm of 25 ten-year-old children. We assume that, given the values of μ and τ, these are independent observations from a normal distribution with mean μ and variance τ^{-1}.

66	66	69	61	58	53	78	71	49	57	54	61	49
64	63	60	53	51	65	70	55	55	74	70	42	

(a) Assuming that the value of τ^{-1} is known to be 64 and our prior distribution for μ is normal with mean 55 and standard deviation 5, find a 95% hpd interval for the height in cm of another ten-year-old child drawn from the same population.

(b) Assume now that the value of τ is unknown but we have a prior distribution for it which is a gamma(2,128) distribution and our conditional prior distribution for μ given τ is normal with mean 55 and variance $(2.56\tau)^{-1}$. Find a 95% hpd interval for the height in cm of another ten-year-old child drawn from the same population.

8. A random sample of $n = 1000$ people was chosen from a large population. Each person was asked whether they approved of a proposed new law. The number answering "Yes" was $x = 372$. (For the purpose of this exercise all other responses and non-responses are teated as simply "Not Yes"). Assume that x is an observation from the binomial(n, p) distribution where p is the unknown proportion of people in the population who would answer "Yes."

   Our prior distribution for p is a uniform distribution on $(0, 1)$.

   Let $p = \Phi(\theta)$ so $\theta = \Phi^{-1}(p)$ where $\Phi(y)$ is the standard normal distribution function and $\Phi^{-1}(z)$ is its inverse.

   (a) Find the maximum likelihood estimate of p and hence find the maximum likelihood estimate of θ.

   (b) Disregarding the prior distribution, find a large-sample approximation to the posterior distribution of θ.

   (c) Using your approximate posterior distribution for θ, find an approximate 95% hpd interval for θ.

   (d) Use the exact posterior distribution for p to find the actual posterior probability that θ is inside your approximate hpd interval.

   Notes:
   - The standard normal distribution function $\Phi(x) = \int_{-\infty}^{x} \phi(u)\ du$ where $\phi(u) = (2\pi)^{-1/2} \exp\{-u^2/2\}$.
   - Let l be the log-likelihood. Then

   $$\frac{dl}{d\theta} = \frac{dl}{dp}\frac{dp}{d\theta}$$

   and

   $$\begin{aligned} \frac{d^2 l}{d\theta^2} = \frac{d}{d\theta}\left\{\frac{dl}{dp}\frac{dp}{d\theta}\right\} &= \frac{d}{d\theta}\left\{\frac{dl}{dp}\right\}\frac{dp}{d\theta} + \frac{dl}{dp}\frac{d^2 p}{d\theta^2} \\ &= \frac{d^2 l}{dp^2}\left(\frac{dp}{d\theta}\right)^2 + \frac{dl}{dp}\frac{d^2 p}{d\theta^2} \end{aligned}$$

   - Derivatives of p :

   $$\begin{aligned} \frac{dp}{d\theta} &= \phi(\theta) \\ \frac{d^2 p}{d\theta^2} &= -\theta\phi(\theta) \end{aligned}$$

   - You can evaluate $\Phi^{-1}(u)$ using R with
   ```
   qnorm(u,0,1)
   ```
   and $\phi(u)$ is given by
   ```
   dnorm(u,0,1)
   ```

9. The amounts of rice, by weight, in 20 nominally 500g packets are determined. The weights, in g, are as follows.

496	506	495	491	488	492	482	495	493	496
487	490	493	495	492	498	491	493	495	489

   Assume that, given the values of parameters μ, τ, the weights are independent and each has a normal $N(\mu,\ \tau)$ distribution.

   The values of μ and τ are unknown. Our prior distribution is as follows. We have a gamma$(2,\ 9)$ prior distribution for τ and a $N(500,\ (0.005\tau)^{-1})$ conditional prior distribution for μ given τ.

(a) Find the posterior probability that $\mu < 495$.

(b) Find the posterior predictive probability that a new packet of rice will contain less than 500g of rice

10. A machine which is used in a manufacturing process jams from time to time. It is thought that the frequency of jams might change over time as the machine becomes older. Once every three months the number of jams in a day is counted. The results are as follows.

Observation i	1	2	3	4	5	6	7	8
Age of machine t_i (months)	3	6	9	12	15	18	21	24
Number of jams y_i	10	13	24	17	20	22	20	23

Our model is as follows. Given the values of two parameters α, β, the number of jams y_i on a dat when the machine has age t_i months has a Poisson distribution

$$y_i \sim \text{Poisson}(\lambda_i)$$

where

$$\log_e(\lambda_i) = \alpha + \beta t_i.$$

Assume that the effect of our prior distribution on the posterior distribution is negligible and that large-sample approximations may be used.

(a) Let the values of α and β which maximise the likelihood be $\hat{\alpha}$ and $\hat{\beta}$. Assuming that the likelihood is differentiable at its maximum, show that these satisfy the following two equations

$$\begin{aligned}\sum_{i=1}^{8}(\hat{\lambda}_i - y_i) &= 0 \\ \sum_{i=1}^{8} t_i(\hat{\lambda}_i - y_i) &= 0\end{aligned}$$

where

$$\log_e(\hat{\lambda}_i) = \hat{\alpha} + \hat{\beta} t_i$$

and show that these equations are satisfied (to a good approximation) by

$$\hat{\alpha} = 2.552 \quad \text{and} \quad \hat{\beta} = 0.02638.$$

(You may use R to help with the calculations, but show your commands).

You may assume from now on that these values maximise the likelihood.

(b) Find an approximate symmetric 95% posterior interval for $\alpha + 24\beta$.

(c) Find an approximate symmetric 95% posterior interval for $\exp(\alpha + 24\beta)$, the mean jam-rate per day at age 24 months.

(You may use R to help with the calculations, but show your commands).

Homework 5

Solutions to Questions 9 and 10 of Problems 5 are to be submitted in the Homework Letterbox no later than 4.00pm on **Tuesday** *May 5th.*

Reference

MacGregor, G.A., Markandu, N.D., Roulston, J.E. and Jones, J.C., 1979. Essential hypertension: effect of an oral inhibitor of angiotensin-converting enzyme. *British Medical Journal*, **2**, 1106-1109.

Solutions

1. (a) Prior distribution is Dirichlet(4,2,2,3).
So $A_0 = 4 + 2 + 2 + 3 = 11$.
The prior means are
$$\frac{a_{0,i}}{A_0}.$$
The prior variances are
$$\frac{a_{0,i}}{(A_0+1)A_0} - \frac{a_{0,i}^2}{A_0^2(A_0+1)}.$$
Prior means:
$$\begin{array}{lcll} \theta_{11}: & \frac{4}{11} & = & \underline{0.3636} \\ \theta_{10}: & \frac{2}{11} & = & \underline{0.1818} \\ \theta_{01}: & \frac{2}{11} & = & \underline{0.1818} \\ \theta_{00}: & \frac{3}{11} & = & \underline{0.2727} \end{array}$$
Prior variances:
$$\begin{array}{lcll} \theta_{11}: & \frac{4}{12\times 11} - \frac{4^2}{11^2\times 12} & = & \underline{0.019284} \\ \theta_{10}: & \frac{2}{12\times 11} - \frac{2^2}{11^2\times 12} & = & \underline{0.012397} \\ \theta_{01}: & \frac{2}{12\times 11} - \frac{2^2}{11^2\times 12} & = & \underline{0.012397} \\ \theta_{00}: & \frac{3}{12\times 11} - \frac{3^2}{11^2\times 12} & = & \underline{0.016529} \end{array}$$
<u>NOTE:</u> Suppose that we are given prior means $m_1, \dots, m_4$ and one prior standard deviation s_1. Then
$$a_{0i} = m_i A_0$$
and
$$s_1^2 = \frac{m_1 A_0}{(A_0+1)A_0} - \frac{m_1^2 A_0^2}{A_0^2(A_0+1)} = \frac{m_1(1-m_1)}{A_0+1}.$$
Hence
$$A_0 + 1 = \frac{m_1(1-m_1)}{s_1^2} = \frac{0.3636(1-0.3636)}{0.019284} = 12.$$
Hence $A_0 = 11$ and $a_{0i} = 11m_i$. For example $a_{01} = 11 \times 0.3636 = 4$.

(b) Posterior distribution is Dirichlet(4+25, 2+7, 2+6, 3+13). That is Dirichlet(29,9,8,16).

(c) Now $A_1 = 29 + 9 + 8 + 16 = 62$.
The posterior means are
$$\frac{a_{1,i}}{A_1}.$$
The posterior variances are
$$\frac{a_{1,i}}{(A_1+1)A_1} - \frac{a_{1,i}^2}{A_1^2(A_1+1)}.$$

Posterior means:

$$\begin{array}{lccc} \theta_{11}: & \dfrac{29}{62} & = & \underline{0.4677} \\ \theta_{10}: & \dfrac{9}{62} & = & \underline{0.1452} \\ \theta_{01}: & \dfrac{8}{62} & = & \underline{0.1290} \\ \theta_{00}: & \dfrac{16}{62} & = & \underline{0.2581} \end{array}$$

Posterior variances:

$$\begin{array}{lccc} \theta_{11}: & \dfrac{29}{63 \times 62} - \dfrac{29^2}{62^2 \times 63} & = & \underline{0.003952} \\ \theta_{10}: & \dfrac{9}{63 \times 62} - \dfrac{9^2}{62^2 \times 63} & = & \underline{0.001970} \\ \theta_{01}: & \dfrac{8}{63 \times 62} - \dfrac{8^2}{62^2 \times 63} & = & \underline{0.001784} \\ \theta_{00}: & \dfrac{16}{63 \times 62} - \dfrac{16^2}{62^2 \times 63} & = & \underline{0.003039} \end{array}$$

(d) Posterior distribution for θ_{00} is beta(16, $62-16$). That is beta(16,46).
Using the R command `hpdbeta(0.95,16,46)` gives $0.15325 < \theta_{00} < 0.36724$.

(e) The log posterior density is (apart from a constant)

$$\sum_{j=1}^{4}(a_{1,j}-1)\log\theta_j.$$

Add $\lambda(\sum_{j=1}^{j}\theta_j - 1)$ to this and differentiate wrt θ_j then set the derivative equal to zero. This gives

$$\frac{a_{1,j}-1}{\hat{\theta}_j} + \lambda = 0$$

which leads to

$$\hat{\theta}_j = -\frac{(a_{1,j}-1)}{\lambda}.$$

However $\sum_{j=1}^{4}\theta_j = 1$ so

$$-\sum_{j=1}^{4}\frac{(a_{1,j}-1)}{\lambda} = 1$$

so

$$\lambda = -\sum_{j=1}^{4}(a_{1,j}-1)$$

and

$$\hat{\theta}_j = \frac{a_{1,j}-1}{\sum a_{1,k} - 4}.$$

Hence the posterior mode for θ_{00} is

$$\hat{\theta}_{00} = \frac{15}{58} = \underline{0.2586}.$$

The second derivatives of the log likelihood are

$$\frac{\partial^2 l}{\partial \theta_j^2} = -\frac{a_{1,j} - 1}{\theta_j^2} \qquad \text{and} \qquad \frac{\partial^2 l}{\partial \theta_j \partial \theta_k} = 0.$$

Since the mixed partial second derivatives are zero, the information matrix is diagonal and the posterior variance of θ_j is approximately

$$\frac{\hat{\theta}_j^2}{a_{1,j} - 1} = \frac{(a_{1,j} - 1)^2}{(a_{1,j} - 1)(\sum a_{1,k} - 4)^2} = \frac{(a_{i,j} - 1}{(\sum a_{1,k} - 4)^2}.$$

The posterior variance of θ_{00} is approximately

$$\frac{15}{58^2} = \underline{0.00445898}.$$

The approximate 95% hpd interval is $0.2586 \pm 1.96\sqrt{0.00445898}$. that is

$$\underline{0.12772 < \theta_{00} < 0.38948}.$$

This is a little wider than the exact interval.

(f) Approximation based on posterior mode and curvature:
Posterior modes:

$$\theta_{11} : \frac{28}{58} \qquad \theta_{10} : \frac{8}{58} \qquad \theta_{01} : \frac{7}{58}$$

So, approx. posterior mean of μ is

$$2 \times \frac{28}{58} + \frac{8}{58} + \frac{7}{58} = \frac{71}{58} = \underline{1.22414}.$$

Approx. posterior variances:

$$\theta_{11} : \frac{28}{58^2} \qquad \theta_{10} : \frac{8}{58^2} \qquad \theta_{01} : \frac{7}{58^2}$$

Since the (approx.) covariances are all zero, the approx. posterior variance of μ is

$$4 \times \frac{28}{58^2} + \frac{8}{58^2} + \frac{7}{58^2} = \frac{127}{58^2} = 0.0377527$$

so approx. standard deviation is

$$\sqrt{0.0377527} = \underline{0.1943}.$$

N.B. There is an alternative exact calculation, as follows, which is also acceptable.
Posterior mean:

$$2 \times \frac{29}{62} + \frac{9}{62} + \frac{8}{62} = \underline{1.20968}.$$

Posterior covariances:

$$-\frac{a_{1,j}a_{1,k}}{A_1^2(A_1 + 1)}$$

$$\begin{aligned}
\text{var}(\mu) &= 4\text{var}(\theta_{11}) + \text{var}(\theta_{10}) + \text{var}(\theta_{01}) \\
&\quad +4\text{covar}(\theta_{11}, \theta_{10}) + 4\text{covar}(\theta_{11}, \theta_{01}) + 2\text{covar}(\theta_{10}, \theta_{01}) \\
&= 4\left(\frac{29}{63 \times 62} - \frac{29^2}{63 \times 62^2}\right) + \left(\frac{9}{63 \times 62} - \frac{9^2}{63 \times 62^2}\right) + \left(\frac{8}{63 \times 62} - \frac{8^2}{63 \times 62^2}\right) \\
&\quad -4\left(\frac{29 \times 9}{63 \times 62^2}\right) - 4\left(\frac{29 \times 8}{63 \times 62^2}\right) - 2\left(\frac{9 \times 8}{63 \times 62^2}\right) \\
&= \frac{133}{63 \times 62} - \frac{5625}{63 \times 62^2} = 0.0108229.
\end{aligned}$$

So the standard deviation is $\underline{0.1040}$. The difference is quite big!

2. From the data

$$\sum_{i=1}^{n} y_i = 1028.9 \qquad\qquad \sum_{i=1}^{n} y_i^2 = 53113.73$$

$$\bar{y} = 51.445 \qquad s_n^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2 = \frac{1}{n}\left\{53113.73 - \frac{1}{20}1028.9^2\right\} = 9.09848$$

(a) Prior mean: $M_0 = 60.0$
Prior precision: $P_0 = 1/20^2 = 0.0025$
Data precision: $P_d = n\tau = 20 \times 0.1 = 2$
Posterior precision: $P_1 = P_0 + P_d = 2.0025$
Posterior mean:

$$M_1 = \frac{0.0025 \times 60.0 + 2 \times 51.445}{2.0025} = 51.4557$$

Posterior std. dev.:

$$\sqrt{\frac{1}{2.0025}} = 0.706665$$

95% hpd interval: $51.4557 \pm 1.96 \times 0.706665$. That is

$$\underline{50.0706 < \mu < 52.8408}$$

(b) Prior $\tau \sim \text{gamma}(d/2,\ dv/2)$ where

$$\frac{d/2}{dv/2} = \frac{1}{v} = 0.1$$

so $v = 10$ and

$$\sqrt{\frac{d/2}{(dv/2)^2}} = \frac{1}{v}\sqrt{\frac{2}{d}} = 0.05$$

so $\sqrt{2/d} = 0.5$ so $2/d = 0.25$ so $d = 8$.
Hence

$$\begin{aligned}
d_0 &= 8 \\
v_0 &= 10 \\
c_0 &= 0.025 \\
m_0 &= 60.0 \\
m_1 &= \frac{c_0 m_0 + n\bar{y}}{c_0 + n} = \frac{0.025 \times 60.0 + 1028.9}{20.025} = 51.4557 \\
c_1 &= c_0 + n = 20.025 \\
d_1 &= d_0 + n = 28 \\
r^2 &= \frac{1}{n}\sum(y_i - m_0)^2 = (\bar{y} - m_0)^2 + s_n^2 \\
&= (51.445 - 60)^2 + 9.09848 = 82.2865 \\
v_d &= \frac{c_0 r^2 + n s_n^2}{c_0 + n} = 9.189846 \\
v_1 &= \frac{d_0 v_0 + n v_d}{d_0 + n} = 9.42132
\end{aligned}$$

95% hpd interval:

$$M_1 \pm t_{28}\sqrt{\frac{v_1}{c_1}}$$

That is

$$M_1 \pm 2.048 \times 0.68591$$

That is

$$\underline{50.051 < \mu < 52.860}$$

3. (a) We have

$$\begin{aligned}
P_0 &= 0.01 \\
P_d &= n\tau = 30 \times 0.04 = 1.2 \\
P_1 &= 0.01 + 1.2 = 1.21 \\
M_0 &= 20 \\
\bar{y} &= 22.4 \\
M_1 &= \frac{P_0 M_0 + P_d \bar{y}}{P_1} = \frac{0.01 \times 20 + 1.2 \times 22.4}{1.21} = 22.380
\end{aligned}$$

Posterior:

$$\mu \sim N(22.380,\ 1.21^{-1}). \qquad \text{That is} \qquad \mu \sim N(22.380,\ 0.8264).$$

95% hpd interval: $22.380 \pm 1.96\sqrt{0.8264}$. That is

$$\underline{20.60 < \mu < 24.16}$$

(b) We have

$$\begin{aligned}
d_0 &= 2 \\
d_1 &= d_0 + n = 32 \\
v_0 &= 10 \\
c_0 &= 0.1 \\
c_1 &= c_0 + n = 30.1 \\
m_0 &= 20 \\
\bar{y} &= 22.4 \\
m_1 &= \frac{c_0 m_0 + n\bar{y}}{c_0 + n} = \frac{0.1 \times 20 + 30 \times 22.4}{30.1} = 22.392 \\
s_n^2 &= \frac{1}{n}\left\{\sum_{i=1}^{n} y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)^2\right\} = 38.0067 \\
(\bar{y} - m_0)^2 &= (22.4 - 20)^2 = 5.76 \\
r^2 &= (\bar{y} - m_0)^2 + s_n^2 = 43.7667 \\
v_d &= \frac{c_0 r^2 + n s_n^2}{c_0 + n} = 38.0258 \\
v_1 &= \frac{d_0 v_0 + n v_d}{d_0 + n} = 36.27419
\end{aligned}$$

Marginal posterior distribution for τ: $d_1 v_1 \tau = 11670.77\tau \sim \chi^2_{32}$.
Marginal posterior distribution for μ:

$$\frac{\mu - m_1}{\sqrt{v_1/c_1}} = \frac{\mu - 22.392}{\sqrt{36.27419/30.1}} \sim t_{32}$$

95% hpd interval:

$$m_1 \pm 2.037\sqrt{\frac{v_1}{c_1}}.$$

That is

$$22.392 \pm 2.037\sqrt{\frac{36.27419}{30.1}}.$$

That is

$$\underline{20.16 < \mu < 24.63}.$$

4. Data:

$$n = 15 \qquad \sum y = -284 \qquad \sum y^2 = 6518$$

$$\bar{y} = -18.9333 \qquad s_n^2 = \frac{1}{15}\left\{6518 - \frac{284^2}{15}\right\} = \frac{1140.9333}{15} = 76.06222$$

Calculate posterior:

$$\begin{aligned}
d_0 &= 0.7 \\
v_0 &= 2.02/0.7 = 2.8857 \\
c_0 &= 0.003 \\
m_0 &= 0 \\
d_1 &= d_0 + 15 = 15.7 \\
(\bar{y} - m_0)^2 &= \bar{y}^2 = 358.4711 \\
r^2 &= (\bar{y} - m_0)^2 + s_n^2 = 434.5333 \\
v_d &= \frac{c_0 r^2 + n s_n^2}{c_0 + n} = 76.1339 \\
v_1 &= \frac{d_0 v_0 + n v_d}{d_0 + n} = 72.8681 \\
c_1 &= c_0 + 15 \\
m_1 &= \frac{c_0 m_0 + n\bar{y}}{c_0 + n} = \frac{-284}{15.003} = 18.9295
\end{aligned}$$

(a) Marginal posterior distribution for τ is gamma$(d_1/2,\ d_1 v_1/2)$. That is

$$\underline{\text{gamma}(7.85,\ 572.014)}.$$

(Alternatively $d_1 v_1 \tau \sim \chi^2_{d_1}$. That is $1144.025\tau \sim \chi^2_{15.7}$).

(b) Marginal posterior for μ:

$$\frac{\mu - m_1}{\sqrt{v_1/c_1}} = \frac{\mu - 18.9295}{\sqrt{4.8569}} \sim t_{15.7}$$

(c) We can use R for the critical points of $t_{15.7}$: ± 2.1232

```
qt(0.975,15.7)
```

95% interval: $-18.9295 \pm 2.1232\sqrt{4.8569}$. That is

$$\underline{-23.61 < \mu < -14.25}.$$

(d) Comment, e.g., since zero is well outside the 95% interval it seems clear that the drug reduces the blood pressure.

5. (a) The likelihood is

$$\begin{aligned} L &= \prod_{i=1}^{n} 2\rho^2 t_i \exp[-(\rho t_i)^2] \\ &= 2^n \rho^{2n} \left(\prod_{i=1}^{n} t_i \right) \exp[-\rho^2 \sum_{i=1}^{n} t_i^2] \end{aligned}$$

The log likelihood is

$$l = n \log 2 + 2n \log \rho + \sum_{i=1}^{n} \log(t_i) - \rho^2 \sum_{i=1}^{n} t_i^2.$$

So

$$\frac{\partial l}{\partial \rho} = \frac{2n}{\rho} - 2\rho \sum_{i=1}^{n} t_i^2$$

and, setting this equal to zero at the mode $\hat{\rho}$, we find

$$\begin{aligned} n &= \hat{\rho}^2 \sum_{i=1}^{n} t_i^2 \\ \hat{\rho}^2 &= \frac{n}{\sum t_i^2} \\ \hat{\rho} &= \sqrt{\frac{n}{\sum t_i^2}} = \sqrt{\frac{300}{3161776}} = 0.0097408. \end{aligned}$$

The second derivative is

$$\frac{\partial^2 l}{\partial \rho^2} = -\frac{2n}{\rho^2} - 2\sum_{i=1}^{n} t_i^2$$

so the posterior variance is approximately

$$\frac{1}{2(n/\hat{\rho}^2 + \sum t_i^2)} = \frac{1}{2(\sum t_i^2 + \sum t_i^2)} = 7.90695 \times 10^{-8}.$$

Our 95% hpd interval is therefore $0.0097408 \pm 1.96\sqrt{7.90695 \times 10^{-8}}$. That is

$$\underline{0.009190 < \rho < 0.010292}.$$

(b) The prior density is proportional to $\rho^1 e^{-100\rho}$ so the log prior density is $\log \rho - 100\rho$ plus a constant. The log posterior is therefore $g(\rho)$ plus a constant where

$$g(\rho) = (2n+1)\log \rho - 100\rho - \rho^2 \sum_{i=1}^{n} t_i^2.$$

So

$$\frac{\partial g}{\partial \rho} = \frac{2n+1}{\rho} - 100 - 2\rho \sum_{i=1}^{n} t_i^2.$$

Setting this equal to zero at the mode $\hat{\rho}$ we find

$$2\sum_{i=1}^{n} t_i^2 \hat{\rho}^2 + 100\hat{\rho} - (2n+1) = 0.$$

This quadratic equation has two solutions but one is negative and ρ must be positive so

$$\hat{\rho} = \frac{-100 + \sqrt{100^2 + 8(\sum t_i^2)(2n+1)}}{4\sum t_i^2} = 0.0097410.$$

The second derivative is

$$\frac{\partial^2 g}{\partial \rho^2} = -\frac{2n+1}{\rho^2} - 2\sum_{i=1}^{n} t_i^2$$

so the posterior variance is approximately

$$\frac{1}{2[(n+1/2)/\hat{\rho}^2 + \sum t_i^2]} = 7.900519 \times 10^{-8}.$$

Our 95% hpd interval is therefore $0.0097410 \pm 1.96\sqrt{7.900519 \times 10^{-8}}$. That is

$$\underline{0.009190 < \rho < 0.010292}.$$

So the prior makes no noticeable difference in this case.

6. (a) $\lambda \sim \text{gamma}(5, 1)$ so $2\lambda \sim \text{gamma}(5, 1/2)$, i.e. gamma$(10/2,\ 1/2)$, i.e. χ^2_{10}.
   From tables, 95% interval, $3.247 < 2\lambda < 20.48$. That is

$$\underline{1.6235 < \lambda < 10.24}$$

   (b) Prior density prop. to $\lambda^{5-1}e^{-\lambda}$.
   Likelihood

$$L = \prod_{i=1}^{45} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!} = \frac{e^{-45\lambda}\lambda^{\sum x_i}}{\prod x_i!} \propto e^{-45\lambda}\lambda^{182}.$$

   Posterior density prop. to $\lambda^{187-1}e^{-46\lambda}$. This is a

$$\underline{\text{gamma}(187,\ 46)}$$

   distribution.

   (c) Posterior mean:

$$\frac{147}{46} = 4.0652$$

   Posterior variance:

$$\frac{187}{46^2} = 0.088774$$

   Posterior sd:

$$\sqrt{\frac{187}{46^2}} = 0.29728$$

   95% interval $4.0652 \pm 1.96 \times 0.29728$. That is

$$\underline{3.4826 < \lambda < 4.6479}$$

   (d) Joint prob. of $\lambda,\ X = m$:

$$\frac{46^{187}}{\Gamma(187)}\lambda^{187-1}e^{-46\lambda}\frac{\lambda^m e^{-\lambda}}{m!} = \frac{46^{187}}{\Gamma(187)}\frac{\Gamma(187+m)}{47^{187+m}}\frac{1}{m!}\frac{47^{187+m}}{\Gamma(187+m)}\lambda^{187+m-1}e^{-47\lambda}$$

Integrate out λ:

$$\begin{aligned}
\Pr(X=m) &= \frac{46^{187}}{47^{187+m}} \frac{\Gamma(187+m)}{\Gamma(187)m!} \\
&= \frac{(186+m)!}{186!m!} \left(\frac{46}{47}\right)^{187} \left(\frac{1}{47}\right)^{m} \\
&= \left(\begin{array}{c} 186+m \\ m \end{array} \right) \left(\frac{46}{47}\right)^{187} \left(\frac{1}{47}\right)^{m}
\end{aligned}$$

(e) Joint probability (density) of $\theta,\ X=m$:

$$0.05e^{-0.05\theta}\frac{\theta^m e^{-\theta}}{m!} = \frac{0.05}{m!}\frac{\Gamma(1+m)}{1.05^{m+1}}\frac{1.05^{m+1}}{\Gamma(1+m)}\theta^{m+1-1}e^{-1.05\theta}$$

Integrate out θ :

$$\Pr(X=m) = \frac{0.05}{1.05^{m+1}}\frac{\Gamma(1+m)}{m!} = \left(\frac{0.05}{1.05}\right)\left(\frac{1}{1.05}\right)^m$$

Log posterior probs:
"Ordinary":

$$\begin{aligned}
\log(P_1) &= \log[\Gamma(187+10)] - \log[\Gamma(187)] - \log[\Gamma(11)] + 187\log(46/47) + 10\log(1/47) \\
&= \log[\Gamma(197)] - \log[\Gamma(187)] - \log(\Gamma(11)] + 187\log(46) - 197\log(47) \\
&= -5.079796
\end{aligned}$$

```
> lgamma(197) - lgamma(187) - lgamma(11) + 187*log(46) - 197*log(47)
[1] -5.079796
```

"Type 2":

$$\begin{aligned}
\log(P_2) &= \log(0.05/1.05) + 10\log(1/1.05) \\
&= \log(0.05) - 11\log(1.05) \\
&= -3.532424
\end{aligned}$$

Hence the predictive probabilities are as follows.

"Ordinary": $P_1 = \exp(-5.079796) = 0.006221178$
"Type 2": $P_2 = \exp(-3.532424) = 0.02923396$

Hence the posterior probability that this is an ordinary customer is

$$\frac{9\times 0.006221178}{9\times 0.006221178 + 1\times 0.02923396} = \underline{0.65698}$$

7.

8.

9. Prior:

$$\tau \sim \text{gamma}\left(\frac{4}{2},\ \frac{18}{2}\right) \quad \text{so} \quad d_0 = 4,\ d_0 v_0 = 18,\ v_0 = 4.5.$$

$$\mu \mid \tau \sim N(500,\ (0.005\tau)^{-1}) \quad \text{so} \quad m_0 = 500,\ c_0 = 0.005.$$

Data:

$$\sum y = 9857, \qquad n = 20, \quad \bar{y} = \frac{9857}{20} = 492.85$$

$$\sum y^2 = 4858467, \quad s_n^2 = \frac{1}{n}\left\{\sum y^2 - n\bar{y}^2\right\} = \frac{444.55}{20} = 22.2275$$

Posterior:

$$\begin{aligned}
c_1 &= c_0 + n = 20.005 \\
m_1 &= \frac{c_0 m_0 + n\bar{y}}{c_0 + n} = 492.8518 \\
d_1 &= d_0 + n = 24 \\
r^2 &= (\bar{y} - m_0)^2 + s_n^2 = 73.35 \\
v_d &= \frac{c_0 r^2 + n s_n^2}{c_0 + n} = 22.2403 \\
v_1 &= \frac{d_0 v_0 + n v_d}{d_0 + n} = 19.2836
\end{aligned}$$

(2 marks)

(a)

$$\frac{\mu - 492.8518}{\sqrt{19.2836/20.005}} \sim t_{24}$$

$$\begin{aligned}
\Pr(\mu < 495) &= \Pr\left(\frac{\mu - 492.8518}{\sqrt{19.2836/20.005}} < \frac{495 - 492.8518}{\sqrt{19.2836/20.005}}\right) \\
&= \Pr(t_{24} < 2.1990) = \underline{0.9807}
\end{aligned}$$

(Eg. use R: `pt(2.1880,24)`).

(3 marks)

(b)

$$c_p = \frac{c_1}{c_1 + 1} = \frac{20.005}{21.005} = 0.9524$$

$$\frac{Y - 492.8518}{\sqrt{19.2836/0.9524}} \sim t_{24}$$

$$\begin{aligned}
\Pr(Y < 500) &= \Pr\left(\frac{Y - 492.8518}{\sqrt{19.2836/0.9524}} < \frac{500 - 492.8518}{\sqrt{19.2836/0.9524}}\right) \\
&= \Pr(t_{24} < 1.5886) = \underline{0.9374}
\end{aligned}$$

(Eg. use R: `pt(1.5886,24)`).

(3 marks)

10. (a) Likelihood:

$$L = \prod_{i=1}^{8} \frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!}$$

Log likelihood:

$$\begin{aligned} l &= -\sum \lambda_i + \sum y_i \log \lambda_i - \sum \log(y_i!) \\ &= -\sum \lambda_i + \sum y_i(\alpha + \beta t_i) - \sum \log(y_i!) \end{aligned}$$

Derivatives:

$$\begin{aligned} \frac{\partial \lambda_i}{\partial \alpha} &= \frac{\partial}{\partial \alpha} e^{\alpha+\beta t_i} = \lambda_i \\ \frac{\partial \lambda_i}{\partial \beta} &= \frac{\partial}{\partial \beta} e^{\alpha+\beta t_i} = \lambda_i t_i \\ \frac{\partial l}{\partial \alpha} &= -\sum \frac{\partial \lambda_i}{\partial \alpha} + \sum y_i = -\sum \lambda_i + \sum y_i = -\sum(\lambda_i - y_i) \\ \frac{\partial l}{\partial \beta} &= -\sum \frac{\partial \lambda_i}{\partial \beta} + \sum y_i t_i = -\sum \lambda_i t_i + \sum y_i t_i = -\sum t_i(\lambda_i - y_i) \end{aligned}$$

At the maximum

$$\frac{\partial l}{\partial \alpha} = \frac{\partial l}{\partial \beta} = 0.$$

Hence $\hat{\alpha}$ and $\hat{\beta}$ satisfy the given equations.
Calculations in R:

```
> y<-c(10,13,24,17,20,22,20,23)
> t<-seq(3,24,3)
> lambda<-exp(2.552+0.02638*t)
> sum(lambda-y)
[1] 0.001572513
> sum(t*(lambda-y))
[1] -0.003254096
```

These values seem close to zero but let us try a small change to the parameter values:

```
> lambda<-exp(2.55+0.0264*t)
> sum(lambda-y)
[1] -0.2522928
> sum(t*(lambda-y))
[1] -3.606993
```

The results are now much further than zero suggesting that the given values are very close to the solutions.

(5 marks)

(b) Second derivatives:

$$\begin{aligned}
\frac{\partial^2 l}{\partial \alpha^2} &= -\sum \frac{\partial \lambda_i}{\partial \alpha} = -\sum \lambda_i \\
\frac{\partial^2 l}{\partial \beta^2} &= -\sum t_i \frac{\partial \lambda_i}{\partial \alpha} = -\sum t_i^2 \lambda_i \\
\frac{\partial^2 l}{\partial \alpha \partial \beta} &= -\sum \frac{\partial \lambda_i}{\partial \beta} = -\sum t_i \lambda_i
\end{aligned}$$

Variance matrix:

$$V = -\begin{pmatrix} \frac{\partial^2 l}{\partial \alpha^2} & \frac{\partial^2 l}{\partial \alpha \partial \beta} \\ \frac{\partial^2 l}{\partial \alpha \partial \beta} & \frac{\partial^2 l}{\partial \beta^2} \end{pmatrix}$$

Numerically using R:

```
>  lambda<-exp(2.552+0.02638*t)
> d2<-matrix(nrow=2,ncol=2)
> d2[1,1]<- -sum(lambda)
> d2[1,2]<- - sum(t*lambda)
> d2[2,1]<- - sum(t*lambda)
> d2[2,2]<- - sum((t^2)*lambda)
> V<- - solve(d2)
> V
             [,1]          [,2]
[1,]  0.038194535 -0.0021361807
[2,] -0.002136181  0.0001449430
```

The mean of $\alpha + 24\beta$ is $2.552 + 24 \times 0.02638 = 3.18512$.
The variance is $0.038194535 + 24^2 * 0.0001449430 + 2 \times 1 \times 24 \times (-0.0021361807) = 0.01914501$.
Alternative matrix-based calculation in R:

```
> dim(m)<-c(1,2)
> v<-m%*%V%*%t(m)
> v
           [,1]
[1,] 0.01914501
```

The approximate 95% interval is

$$3.18512 \pm 1.96\sqrt{0.01914501}$$

That is

$$\underline{2.9139 < \alpha + 24\beta < 3.4563}$$

(5 marks)

(c) The interval for λ_{24} is

$$e^{2.9139} < e^{\alpha+24\beta} < e^{3.4563}.$$

That is

$$\underline{18.429 < \lambda_{24} < 31.700}.$$

(2 marks)