Question 1

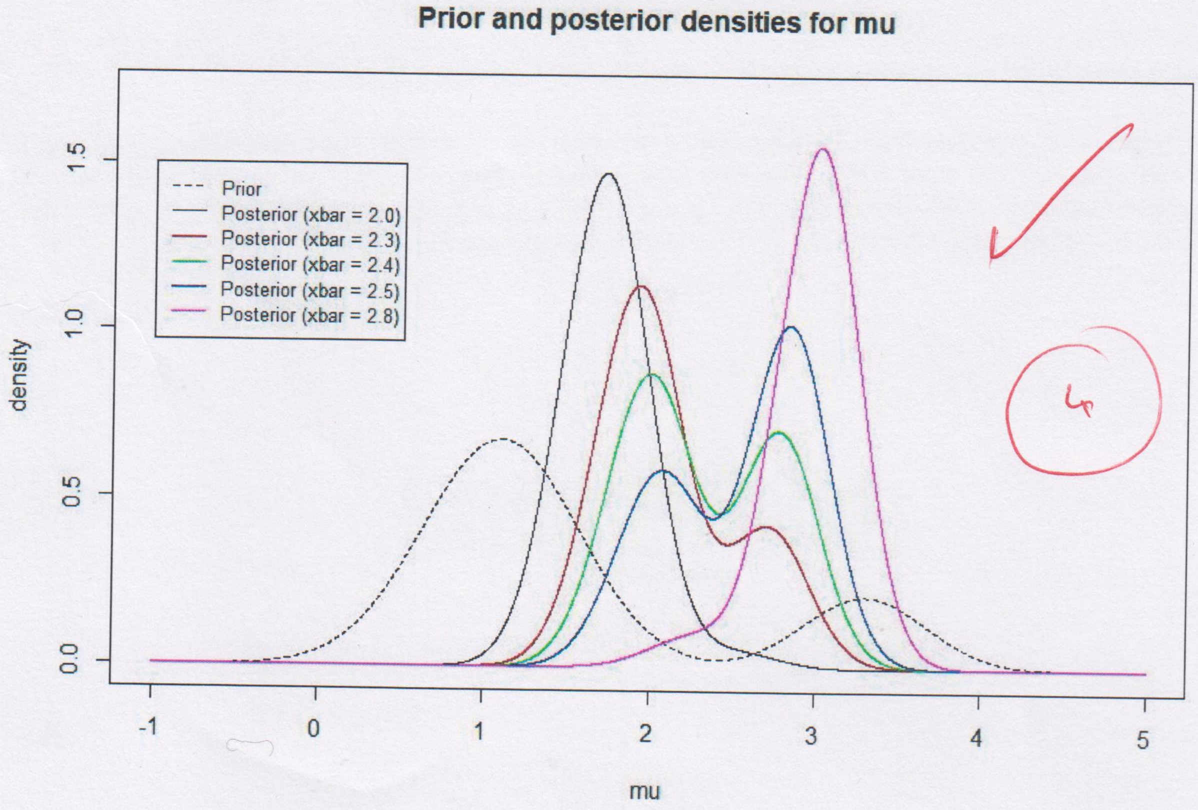
a) Results obtained are summarised in Table 1 below (3 d.p.).

$\bar{x}$	Normal mixture	Mean	Sd	$Pr(\mu > 2.5 \mid \boldsymbol{x})$
No data	$0.2N(3.3, 0.37^2) + 0.8N(1.1, 0.47^2)$	1.540	0.989	0.198
2.0	$0.028N(2.549, 0.240^2) + 0.972N(1.720, 0.262^2)$	1.743	0.296	0.018
2.3	$0.249N(2.722, 0.240^2) + 0.751N(1.926, 0.262^2)$	2.125	0.430	0.216
2.4	$0.423N(2.780, 0.240^2) + 0.577N(1.995, 0.262^2)$	2.327	0.463	0.387
2.5	$0.615N(2.838, 0.240^2) + 0.385N(2.064, 0.262^2)$	2.540	0.452	0.584
2.8	$0.939(3.011, 0.240^2) + 0.061N(2.270, 0.262^2)$	2.966	0.300	0.935

Table 1: The summary table for prior and posterior mixture distributions.

15

b) Graph for prior and posterior densities for μ is provided below.



4

Figure 1: Prior density (dashed) and posterior densities for different values of $\bar{x}$ (colours).

c) Observations are summarised below:

- For the prior distribution with the smallest value of $\bar{x}, \bar{x} = 2$, the posterior distribution is almost unimodal in shape. As $\bar{x}$ becomes larger, the right hand mode becomes more pronounced, as we see for $\bar{x} = 2.3, 2.4$. For these values, $p_1^* < p_2^*$. Beyond this point, the weights change such that $p_1^* > p_2^*$ and we see the left hand mode becomes larger and the right hand becomes smaller. This pattern continues as $\bar{x}$ increases and for $\bar{x} = 2.8$ the distribution is almost unimodal once more. In this sense, we see a form of symmetry as we change the value of $\bar{x}$ – we go from almost unimodal, to decidedly bimodal for $\bar{x} = 2.4$, where $p_1^* \approx p_2^*$ and then back to a more unimodal distribution beyond this point.
- We see that all of the posterior curves lie further to the right than the prior curve (dashed). Therefore we could say that the mean value has been shifted to the right.
- After comparing prior and posterior distributions for μ we observe that the variance has decreased, meaning that we are more precise in our beliefs after observing the data. This is evident in the way that the posterior curves are less 'stretched out' than the prior curve.
- We see that the $Pr(\mu > 2.5)$ under the prior distribution is 19.8%. Then, for $\bar{x} = 2.0$, this reduces further to only 1.8%. However if our observed value for $\bar{x}$ is 2.3 or greater, our belief that $\mu > 2.5$ increases compared with the prior. For example, we see that $Pr(\mu > 2.5)$, is 93.5%. For these large observed values of $\bar{x}$ the data has been very informative in that we have gone from having very little belief that $\mu > 2.5$ to being quite certain that this is in fact the case.

6

d)

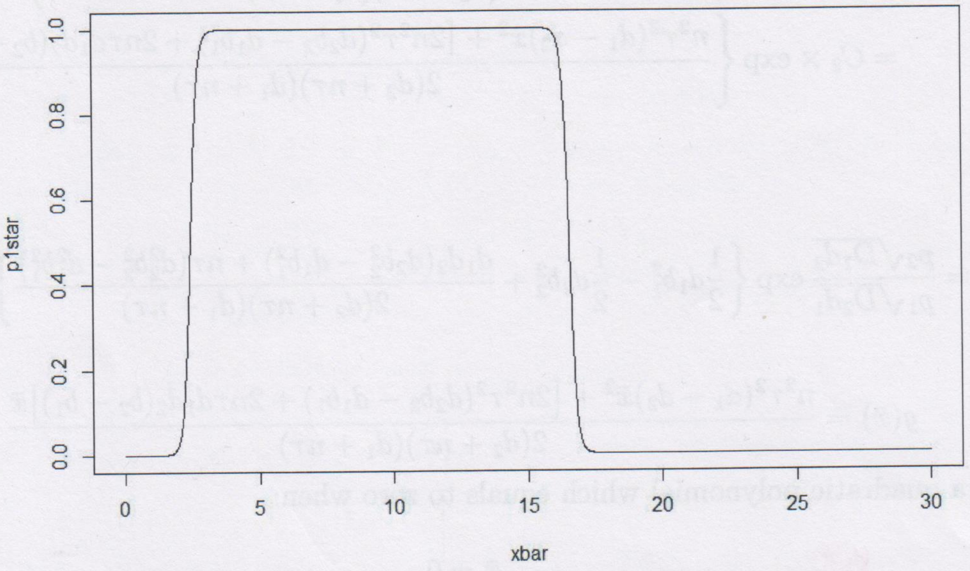


Figure 2: Plot of the posterior weight p_1^* against $\bar{x}$.

Comments:

- Low values $\bar{x}$ corresponds to small values of p_1^*. In this case we are in favour of the second component with smallest mean.
- As $\bar{x}$ increases, so does the probability of the first component with the larger mean.
- When $\bar{x}$ increases to about 16.5, the second component with smaller prior mean becomes more likely again.
- Once p_1^* reaches it's maximum at about $\bar{x} = 4$ we see that it then remains uniform till about $\bar{x} = 16.5$ where p_1^* then decreases.

Underlying Mathematics:
From question 11 we know that

$$(p_1^*)^{-1} - 1 = \frac{p_2\sqrt{D_1 d_2}}{p_1\sqrt{D_2 d_1}} \exp\left\{\frac{1}{2}\left[D_2 B_2^2 - d_2 b_2^2 - D_1 B_1^2 + d_1 b_1^2\right]\right\}$$

To find the relationship between p_1^* and $\bar{x}$, we need to find a function of p_1^* with respect to $\bar{x}$, so by extracting constant C which does not depend on $\bar{x}$, we have:

$$\begin{aligned}
(p_1^*)^{-1} - 1 &= C_1 \times \exp\left\{\frac{1}{2}D_2 B_2^2 - \frac{1}{2}D_1 B_1^2\right\} \\
&= C_1 \times \exp\left\{\frac{1}{2}\left[\frac{(d_2 b_2 + n\tau\bar{x})^2}{d_2 + n\tau} - \frac{(d_1 b_1 + n\tau\bar{x})^2}{d_1 + n\tau}\right]\right\} \\
&= C_1 \times \exp\left\{\frac{(d_2 b_2 + n\tau\bar{x})^2(d_1 + n\tau) - (d_1 b_1 + n\tau\bar{x})^2(d_2 + n\tau)}{2(d_2 + n\tau)(d_1 + n\tau)}\right\} \\
&= C_2 \times \exp\left\{\frac{n^2\tau^2(d_1 - d_2)\bar{x}^2 + \left[2n^2\tau^2(d_2 b_2 - d_1 b_1) + 2n\tau d_1 d_2(b_2 - b_1)\right]\bar{x}}{2(d_2 + n\tau)(d_1 + n\tau)}\right\}
\end{aligned}$$

where

$$C_2 = \frac{p_2\sqrt{D_1 d_2}}{p_1\sqrt{D_2 d_1}} \exp\left\{\frac{1}{2}d_1 b_1^2 - \frac{1}{2}d_2 b_2^2 + \frac{d_1 d_2(d_2 b_2^2 - d_1 b_1^2) + n\tau(d_2^2 b_2^2 - d_1^2 b_1^2)}{2(d_2 + n\tau)(d_1 + n\tau)}\right\} >> 0$$

Let

$$g(\bar{x}) = \frac{n^2\tau^2(d_1 - d_2)\bar{x}^2 + \left[2n^2\tau^2(d_2 b_2 - d_1 b_1) + 2n\tau d_1 d_2(b_2 - b_1)\right]\bar{x}}{2(d_2 + n\tau)(d_1 + n\tau)}$$

which is a quadratic polynomial which equals to zero when:

$$\bar{x} = 0$$

$$\text{or} \qquad \bar{x} = -\frac{2n\tau(d_2 b_2 - d_1 b_1) + 2d_1 d_2(b_2 - b_1)}{n\tau(d_1 - d_2)} \approx 19$$

Let

$$A = C_2 \times \exp\{g(\bar{x})\}$$

Therefore,

$$p_1^* = (A+1)^{-1}$$

By changing the value of $\bar{x}$ in the range $(0, 30)$, we have:

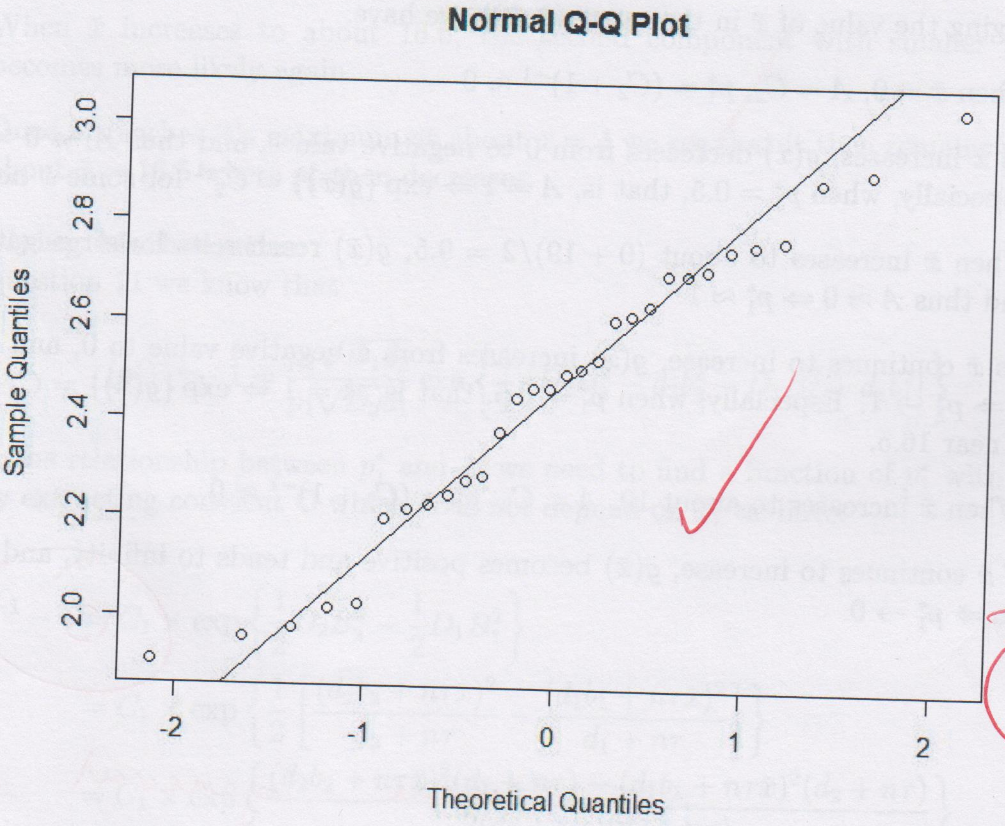
- When $\bar{x} = 0$, $A = C_2$, $p_1^* = (C_2+1)^{-1} \approx 0$
- As $\bar{x}$ increases, $g(\bar{x})$ decreases from 0 to negative values, and thus $A \to 0 \Rightarrow p_1^* \to 1$. Especially, when $p_1^* = 0.5$, that is, $A = 1 \Rightarrow \exp\{g(\bar{x})\} = C_2^{-1}$ for some $\bar{x}$ near 2.5.
- When $\bar{x}$ increases to about $(0+19)/2 = 9.5$, $g(\bar{x})$ reaches its lowest negative value, and thus $A \approx 0 \Rightarrow p_1^* \approx 1$
- As $\bar{x}$ continues to increase, $g(\bar{x})$ increases from a negative value to 0, and thus $A \to 0 \Rightarrow p_1^* \to 1$. Especially, when $p_1^* = 0.5$, that is, $A = 1 \Rightarrow \exp\{g(\bar{x})\} = C_2^{-1}$ for some $\bar{x}$ near 16.5.
- When $\bar{x}$ increases to about 19, $A = C_2$, $p_1^* = (C_2+1)^{-1} \approx 0$
- If $\bar{x}$ continues to increase, $g(\bar{x})$ becomes positive and tends to infinity, and thus $A \to \infty \Rightarrow p_1^* \to 0$.

15

more general proof possible.

Question 2

a) We see that there is a good fit to the line which supports the suitability of the Normal distribution as a model for the variation in enzyme measurements.



4

Figure 3: Normal Q-Q Plot.

b) Recalling that for the data we have $b = 2.6$, $c = 1$, $g = 5$, $h = 0.4$, hence

$$\mu \sim t_{2g}\left(b, \frac{h}{gc}\right) \implies \mu \sim t_{10}\,(2.6, 0.08)$$
$$\tau \sim Ga(g, h) \implies \tau \sim Ga(5, 0.4)$$
$$\sigma \sim Inv\text{-}Chi(g, h) \implies \sigma \sim Inv\text{-}Chi(5, 0.4)$$

Prior mean and Sd for μ:

$$E(\mu) = 2.6$$

$$Var(\mu) = \frac{10 \cdot 0.08}{10 - 2} = 0.1 \implies Sd(\mu) = \sqrt{0.1} = \frac{\sqrt{10}}{10} \approx 0.3162(4dp)$$

Prior mean and Sd for τ:

$$E(\tau) = \frac{5}{0.4} = 12.5.$$

$$Var(\tau) = \frac{5}{0.4^2} = 31.25 \implies Sd(\tau) = \sqrt{31.25} = \frac{5\sqrt{5}}{2} \approx 5.5902(4dp)$$

Prior mean and Sd for σ:

$$E(\sigma) = \sqrt{0.4}\,\Gamma(5 - 1/2)\,/\Gamma(5)$$

We know that:

$$\Gamma\left(\frac{9}{2}\right) = \frac{7}{2} \cdot \frac{5}{2} \cdot \frac{3}{2} \cdot \frac{1}{2}\Gamma\left(\frac{1}{2}\right) = \frac{105}{16}\sqrt{\pi}$$

Therefore,

$$\begin{aligned} E(\sigma) &= \frac{2 \cdot 105 \cdot \sqrt{\pi}}{\sqrt{10} \cdot 2 \cdot 3 \cdot 4 \cdot 16} \\ &= \frac{35\sqrt{\pi}}{64\sqrt{10}} \\ &= 0.306523 \\ &\approx 0.3065(4dp) \end{aligned}$$

$$Var(\sigma) = \frac{0.4}{4} - 0.306523^2 = 0.00604365 \implies Sd(\sigma) = \sqrt{0.00604365} = 0.07774092 \approx 0.0777(4dp$$

6

c) We know that:

$$b = 2.6, \quad c = 1, \quad g = 5, \quad h = 0.4,$$

and also:

$$n = 30, \quad \bar{x} = 2.459333, \quad s = 0.3058502.^{1}$$

This gives:

$$B = \frac{bc + n\bar{x}}{c + n} = \frac{2.6 \cdot 1 + 30 \cdot 2.459333}{1 + 30} = 2.463871$$

$$C = c + n = 31$$

$$G = g + \frac{n}{2} = 5 + \frac{30}{2} = 20$$

$$H = h + \frac{cn(\bar{x} - b)^2}{2(c + n)} + \frac{ns^2}{2} = 0.4 + \frac{1 \cdot 30 \cdot (2.459333 - 2.6)^2}{2(1 + 30)} + \frac{30 \cdot 0.3058502^2}{2} = 1.812739926$$

[1](s has been calculated using the built in R command for standard deviation)

So, we have:

$$\begin{pmatrix}\mu\\ \tau\end{pmatrix} \mid \underline{x} \sim NGa(2.463871, 31, 20, 1.812739926)$$

Marginal posteriors are:

$$\mu \mid \boldsymbol{x} \sim t_{2G}\left(B, \frac{H}{GC}\right) \implies \mu \mid \boldsymbol{x} \sim t_{40}\,(2.463871, 0.002923774)$$

$$\tau \mid \boldsymbol{x} \sim Ga(G, H) \implies \tau \mid \boldsymbol{x} \sim Ga(20, 1.812739926)$$

$$\sigma \mid \boldsymbol{x} \sim Inv\text{-}Chi(G, H) \implies \sigma \mid \boldsymbol{x} \sim Inv\text{-}Chi(20, 1.812739926)$$

Posterior mean and Sd for μ:

$$E(\mu \mid \boldsymbol{x}) = 2.463871 \approx 2.4639(4dp)$$

$$Var(\mu \mid \boldsymbol{x}) = \frac{40 \cdot 0.002923774}{40-2} = 0.003077657$$

$$\implies Sd(\mu \mid \boldsymbol{x}) = \sqrt{0.003077657} = 0.05547663 \approx 0.0555(4dp)$$

Posterior mean and Sd for τ:

$$E(\tau \mid \boldsymbol{x}) = \frac{20}{1.812739926} = 11.033022286 \approx 11.0330(4dp)$$

$$Var(\tau \mid \boldsymbol{x}) = \frac{20}{1.812739926^2} = 6.086379 \implies Sd(\tau \mid \boldsymbol{x}) = \sqrt{6.086379} = 2.467059 \approx 2.4671(4dp)$$

Posterior mean and Sd for σ:

$$E(\sigma \mid \boldsymbol{x}) = \sqrt{1.812739926}\,\Gamma(20 - 1/2)\,/\Gamma(20)$$

We know that:

$$\begin{aligned}
\Gamma\left(\frac{39}{2}\right) &= \frac{37}{2}\cdot\frac{35}{2}\cdot\frac{33}{2}\cdot\frac{31}{2}\cdots\frac{1}{2}\Gamma\left(\frac{1}{2}\right)\\
&= \frac{37}{2}\cdot\frac{35}{2}\cdot\frac{33}{2}\cdot\frac{31}{2}\cdots\frac{1}{2}\sqrt{\pi}\\
&= 1.564177\times 10^{16}\sqrt{\pi}\\
&= 2.772432\times 10^{16}
\end{aligned}$$

We can also calculate:

$$\Gamma(20) = 19! = 1.216451 \times 10^{17}$$

Therefore,

$$\begin{aligned}E(\sigma \mid \boldsymbol{x}) &= \sqrt{1.812739926} \cdot \frac{2.772432 \times 10^{16}}{1.216451 \times 10^{17}} \\ &= 0.3068556 \\ &\approx 0.3069(4dp)\end{aligned}$$

$$\begin{aligned}Var(\sigma \mid \boldsymbol{x}) &= \frac{1.812739926}{19} - 0.306523^2 \\ &= 0.001451014997\end{aligned}$$

$$\implies Sd(\sigma \mid \boldsymbol{x}) = \sqrt{0.001451014997} = 0.03809219077 \approx 0.0381(4dp)$$

3

error carried.

We can summarise our findings in Table 2 below.

	μ	$\mu \mid \boldsymbol{x}$	τ	$\tau \mid \boldsymbol{x}$	σ	$\sigma \mid \boldsymbol{x}$
Expectation	2.6000	2.4639	12.5000	11.0330	0.3065	0.3069
Sd	0.3162	0.0555	5.5902	2.4671	0.0777	0.0381

Table 2: Summary of changes in prior and posterior mean and SD for μ, τ and σ

d) Firstly, we plot the marginal prior and posterior densities for μ, τ and σ

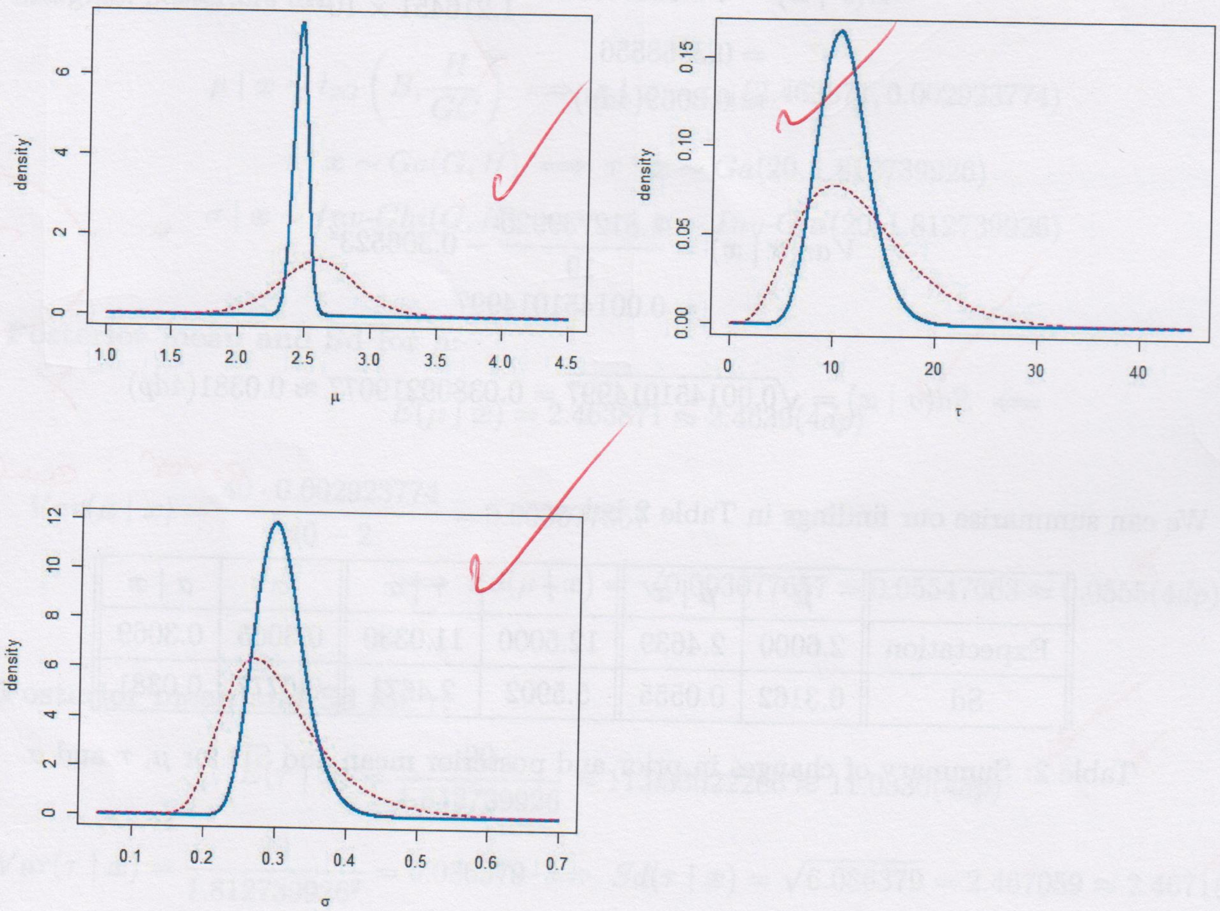


Figure 4: Prior (red dashed) and posterior (blue solid) densities for μ, τ and σ.

Next, we produce contour plots of the joint prior and posterior densities for $(\mu, \tau)^T$

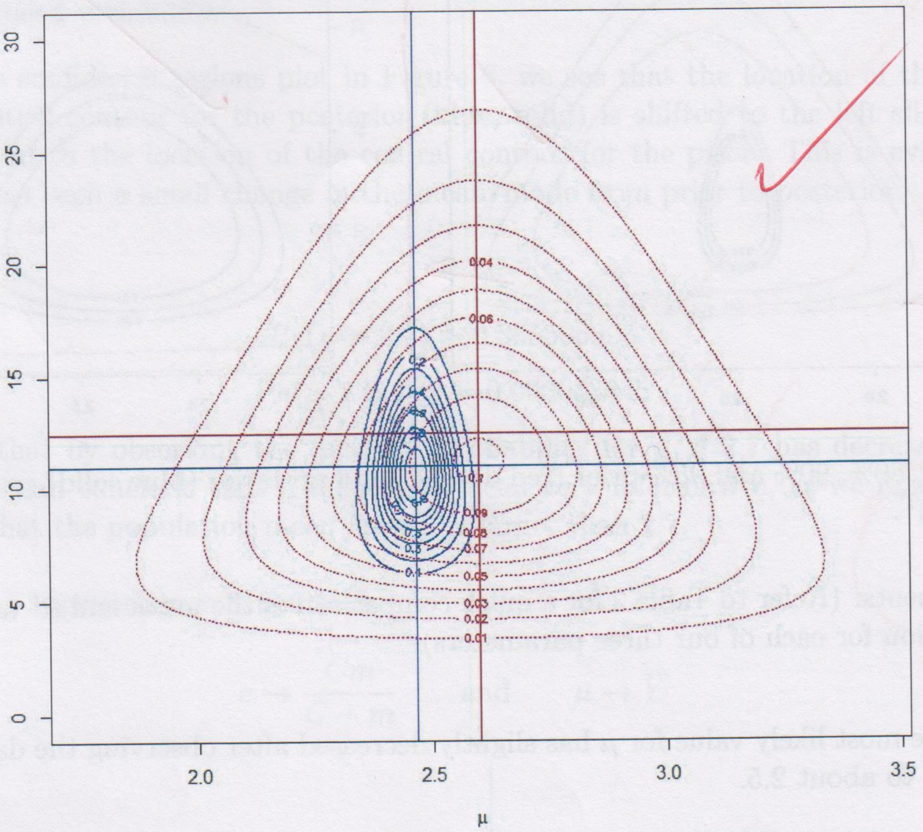


Figure 5: Contour plots of the prior (red dashed) and posterior (blue solid) densities for $(\mu, \tau)^T$.

e) Plots for prior and posterior 80%, 90% and 95% confidence regions intervals are illustrated below.

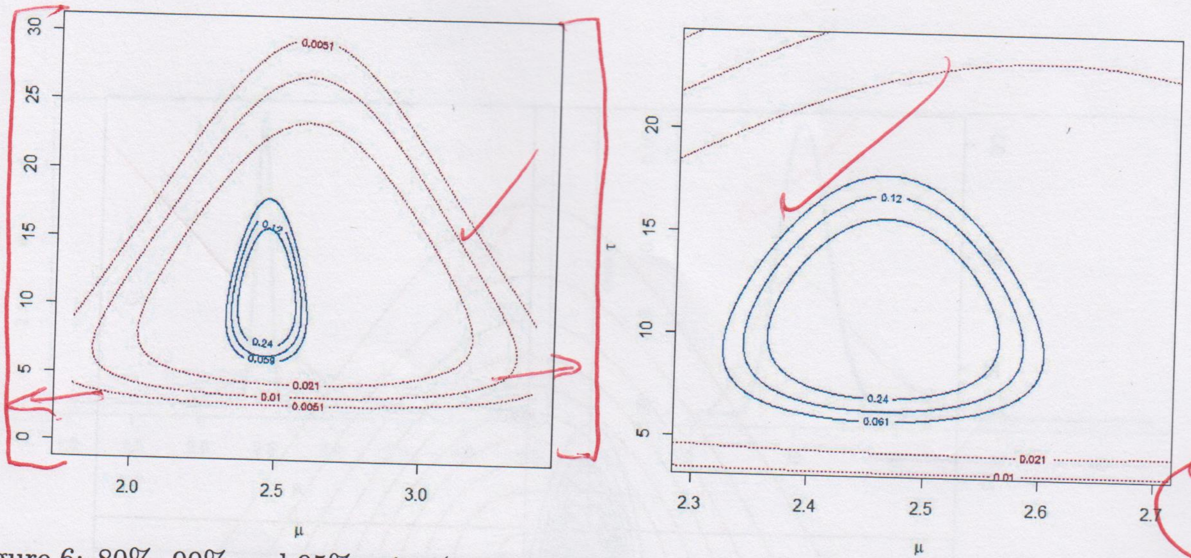


Figure 6: 80%, 90% and 95% prior (red dashed) and posterior (blue solid) confidence regions for $(\mu, \tau)^T$.

f) Comments: (Refer to Table 2 for a quick comparison of the mean and sd values from prior to precision for each of our three parameters)
For μ:

- The most likely value for μ has slightly decreased after observing the data, from about 2.6 to about 2.5.
- There is a considerable increase in the precision for μ, from $\frac{1}{0.3162^2} = 10.0018$ to $\frac{1}{0.0555^2} = 324.6490$, which makes the distribution more condensed. We could say that our beliefs about μ have been 'focused' greatly, having observed the data.

For τ:

- The most likely value for τ has slightly increased after observing the data, from about 10.4 to about 11.
- Distribution looks more condensed, which indicates that the variability has decreased. Indeed the variance has decreased from 5.5902^2 to 2.4671^2.

For σ:

- The most likely value for σ has slightly increased after observing the data, from about 0.27 to about 0.3.
- Distribution looks slightly more condensed, which indicates that the variability has decreased. Again, referring to table two we do indeed have a decrease in variance from 0.0777^2 to 0.0381^2.

Dependence structure and confidence regions:

- We see that the contours for the posterior distribution are more elliptical than those for the prior, which indicates a change in the dependence structure.

- The posterior is much more tightly concentrated than the prior, that is, the posterior has reduced variability.

- For the confidence regions plot in Figure 6, we see that the location of the centre of the central contour for the posterior (blue, solid) is shifted to the left slightly when compared to the location of the central contour for the prior. This is evidence that there has been a small change in the mean/mode from prior to posterior. 6

g)

$$Pr(\mu > 2.7) = 0.365506905$$
$$Pr(\mu > 2.7|\underline{x}) = 0.000043345$$

We can see that by observing the data the probability for $\mu > 2.7$ has decreased greatly. Therefore we can conclude that the data has been very informative, as we now have very little belief that the population mean level μ is larger than 2.7. 1

h) By making appropriate substitutions:

$$c \to \frac{Cm}{C+m} \quad \text{and} \quad \mu \to \overline{Y}$$

$$\overline{Y} \mid \mu, \tau \sim N\left(\mu, \frac{1}{m\tau}\right)$$

$$\mu \mid \tau \sim N\left(b, \frac{1}{c\tau}\right), \qquad \tau \sim Ga\left(g, h\right).$$

$$\begin{pmatrix}\mu\\ \tau\end{pmatrix} \sim NGa(b, c, g, h) \implies \mu \sim t_{2g}\left(b, \frac{h}{gc}\right)$$

$$\overline{Y} \mid \underline{x}, \tau \sim N\left(B, \frac{C+m}{C\tau m}\right) \equiv N\left(B, \frac{1}{\tau\frac{Cm}{C+m}}\right)$$

$$\tau \mid \underline{x} \sim Ga(G, H)$$

$$\begin{pmatrix}\overline{Y}\\ \tau\end{pmatrix} \mid \underline{x} \sim NGa\left(B, \frac{Cm}{C+m}, G, H\right)$$

$$\implies \overline{Y} \mid \underline{x} \sim t_{2G}\left(B, \frac{H(C+m)}{GCm}\right)$$

2

Therefore,

$$E[\overline{Y} \mid \underline{x}] = B = 2.4639 \ (4 \ d.p.)$$

i) The plot of the predictive distribution of $\overline{Y}$ for the case $m = 20$ is illustrated below:

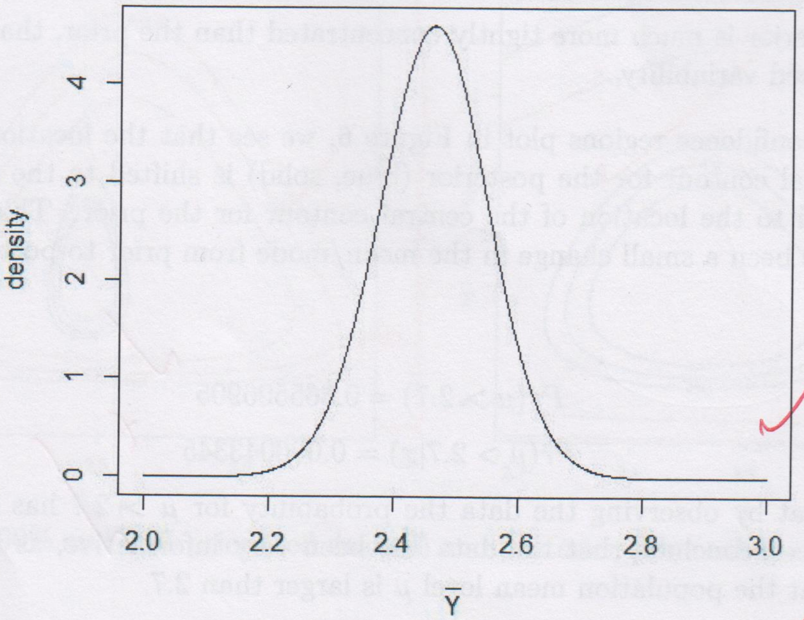


Figure 7: Predictive density for $\overline{Y}$

The 95% prediction interval for $\bar{Y}$ is $(2.289359, 2.638383)$.

j) From Hint 1 and Hint 2, we can know that

$$mV\tau|\tau \sim \chi^2_{m-1} \equiv Ga\left(\frac{m-1}{2}, \frac{1}{2}\right)$$

and thus,

$$V|\tau \sim Ga\left(\frac{m-1}{2}, \frac{m\tau}{2}\right)$$

Using the definition of the predictive density on page 50, we know that

$$f(V|\underline{x}) = \int_0^\infty f(v|\tau)\,\pi(\tau|\underline{x})\,d\tau$$

$$= \int_0^\infty \frac{\left(\frac{m\tau}{2}\right)^{\frac{m-1}{2}} v^{\frac{m-1}{2}-1} e^{-\frac{m\tau}{2}v}}{\Gamma\left(\frac{m-1}{2}\right)} \times \frac{H^G\,\tau^{G-1}\,e^{-H\tau}}{\Gamma(G)}\,d\tau$$

$$= \frac{\left(\frac{m}{2}\right)^{\frac{m-1}{2}} v^{\frac{m-1}{2}-1}\,H^G\,\Gamma\left(\frac{m-1}{2}+G\right)}{\Gamma\left(\frac{m-1}{2}\right)\Gamma(G)\,\left(H+\frac{mv}{2}\right)^{\frac{m-1}{2}+G}} \int_0^\infty \frac{\left(H+\frac{mv}{2}\right)^{\frac{m-1}{2}+G} \tau^{\frac{m-1}{2}+G-1} e^{-\left(H+\frac{mv}{2}\right)\tau}}{\Gamma\left(\frac{m-1}{2}+G\right)}\,d\tau$$

1

We can see that the above integral is just the density of $Ga\left(\frac{m-1}{2}+G, H+\frac{mv}{2}\right)$, and thus equals to 1. Since V being the variance of the future sample is greater than 0, we can assume that V has a scaled F-distribution, then all we need to do is to compare the rest part of the above equation with Hint 3. Rearranging, we have:

$$f(V|\underline{x}) = \frac{\Gamma\left(\frac{m-1}{2}+G\right)}{\Gamma\left(\frac{m-1}{2}\right)\Gamma(G)} \times \frac{\left(\frac{m}{2}\right)^{\frac{m-1}{2}} v^{\frac{m-1}{2}-1} H^G}{\left(H+\frac{mv}{2}\right)^{\frac{m-1}{2}+G}}$$

$$= \frac{1}{B\left(\frac{m-1}{2},\frac{2G}{2}\right)} \times v^{\frac{m-1}{2}-1} \times \frac{\left(\frac{m}{2}\right)^{\frac{m-1}{2}} H^G}{\left(H+\frac{mv}{2}\right)^{\frac{m-1}{2}+G}} \qquad (*)$$

Now we have $v_1 = m-1$, $v_2 = 2G$, the next step is to find the value of a. Substituting v_1 and v_2 in Hint 3 and comparing with equation $(*)$, we have

$$\frac{\left(\frac{m}{2}\right)^{\frac{m-1}{2}} H^G}{\left(H+\frac{mv}{2}\right)^{\frac{m-1}{2}+G}} = \left(\frac{m-1}{2Ga}\right)^{\frac{m-1}{2}} \left(1+\frac{(m-1)v}{2Ga}\right)^{-\left(\frac{m-1}{2}+G\right)}$$

Working on both sides of the equation we get

$$LHS = \left(\frac{\frac{m}{2}}{H+\frac{mv}{2}}\right)^{\frac{m-1}{2}} \left(\frac{H}{H+\frac{mv}{2}}\right)^{G}$$

$$= \left(\frac{2H}{m}+v\right)^{-\frac{m-1}{2}} \left(1+\frac{mv}{2H}\right)^{-G}$$

$$RHS = \left(\frac{2Ga}{m-1}\left[1+\frac{(m-1)v}{2Ga}\right]\right)^{-\frac{m-1}{2}} \left(1+\frac{(m-1)v}{2Ga}\right)^{-G}$$

$$= \left(\frac{2Ga}{m-1}+v\right)^{-\frac{m-1}{2}} \left(1+\frac{(m-1)v}{2Ga}\right)^{-G}$$

From LHS and RHS we can easily see that

$$\frac{2Ga}{m-1} = \frac{2H}{m}$$

thus

$$a = \frac{(m-1)H}{mG}$$

Therefore, we know that V does have a scaled F-distribution, that is

$$V|\underline{x} \sim aF_{v_1,v_2}$$

where

$$v_1 = m-1, \qquad v_2 = 2G, \qquad a = \frac{(m-1)H}{mG}$$

9

k) When m=20 we can determine a 95% equi-tailed PI for V to be $(0.03690867, 0.1796467)$ given by the scaled F distribution on 19 and 40 degrees of freedom. Therefore we can calculate the confidence interval for $\sqrt{V}$ (the SD of the future sample) to be: (0.1921163, 0.4238475). As the confidence intervals for the variance and SD are small we would make the assumption that our distribution for the future sample is appropriate. ✗

wrong, but quite close

1