MAST30027 Modern Applied Statistics Assignment 4

James La Fontaine 1079860

Tutorial: Wed 1-2PM, Yidi Deng

Question 1

(a)

1.

$x_i \sim N(75, \frac{1}{\tau})$ for $i = 1, \dots, 100$

$\Rightarrow f(x_i \mid \tau) \propto \tau^{1/2} e^{-\frac{\tau}{2}(x_i - 75)^2}$

$\tau \sim \text{Gamma}(2,1)$

$\Rightarrow f(\tau) = \frac{1}{\Gamma(2)} \tau e^{-\tau} = \tau e^{-\tau}$

a) $P(\tau \mid X_1, \dots, X_n) \propto P(X_1, \dots, X_n \mid \tau)\, P(\tau)$

$\propto \tau e^{-\tau} \prod_{i=1}^{100} \tau^{1/2} e^{-\frac{\tau}{2}(x_i - 75)^2}$

$\propto \tau^{51} e^{-\tau} e^{\frac{-\tau}{2} \sum_{i=1}^{100} (x_i - 75)^2}$

$\propto \tau^{51} e^{-\tau \left[1 + \frac{\sum_{i=1}^{100} (x_i - 75)^2}{2}\right]}$

```
> (posterior_rate = 1 + (sum((X-75)^2)) / 2)
>
> 1805.65
```

$\Longrightarrow (\tau \mid X_1, ..., X_n) \sim \text{Gamma}(52, 1805.65)$

(b)

Let $\beta = 1 + \frac{\sum_{i=1}^{100}(x_i - 75)^2}{2}$

b) $P(\tilde{X} \mid X) = \int P(\tilde{X} \mid \tau) P(\tau \mid X) \, d\tau$

$$= \int \frac{\tau^{1/2}}{\sqrt{2\pi}} e^{-\tau/2 \, (\tilde{x} - 75)^2} \frac{\beta^{52}}{\Gamma(52)} \tau^{51} e^{-\tau\beta} \, d\tau$$

$$= \frac{\beta^{52}}{\sqrt{2\pi}\,\Gamma(52)} \int \tau^{\frac{103}{2}} e^{-\frac{\tau}{2}(\tilde{x} - 75)^2 - \tau\beta} \, d\tau$$

$$= \frac{\Gamma\left(\frac{105}{2}\right)\beta^{52}}{\sqrt{2\pi}\,\Gamma(52)} \int \tau^{\frac{105}{2} - 1} e^{-\tau\left(\frac{(\tilde{x} - 75)^2}{2} + \beta\right)} \frac{1}{\Gamma\left(\frac{105}{2}\right)} \, d\tau$$

$$= \frac{\Gamma\left(\frac{105}{2}\right)\beta^{52}}{\sqrt{2\pi}\,\Gamma(52)\left(\frac{(\tilde{x} - 75)^2}{2} + \beta\right)^{\frac{105}{2}}} \int \frac{\left(\frac{(\tilde{x} - 75)^2}{2} + \beta\right)^{\frac{105}{2}}}{\Gamma\left(\frac{105}{2}\right)} \tau^{\frac{105}{2} - 1} e^{-\tau\left(\frac{(\tilde{x} - 75)^2}{2} + \beta\right)} \, d\tau$$

$$= \frac{\Gamma\left(\frac{105}{2}\right)\beta^{52}}{\sqrt{2\pi}\,\Gamma(52)\left(\frac{(\tilde{x} - 75)^2}{2} + \beta\right)^{\frac{105}{2}}}$$ as the integral of the Gamma pdf $= 1$

$$= \frac{\Gamma\left(\frac{105}{2}\right)\beta^{52}}{\Gamma\left(\frac{104}{2}\right)\sqrt{2\pi}} \left(\frac{(\tilde{x} - 75)^2}{2} + \beta\right)^{-\frac{105}{2}}$$

$$= \frac{\Gamma\left(\frac{105}{2}\right)\beta^{52}}{\Gamma\left(\frac{104}{2}\right)\sqrt{2\pi}} \left(\beta\left(\frac{(\tilde{x} - 75)^2}{2\beta} + 1\right)\right)^{-\frac{105}{2}}$$

$$= \frac{\Gamma\left(\frac{105}{2}\right)\beta^{52}}{\Gamma\left(\frac{104}{2}\right)\sqrt{2\pi}} \beta^{-\frac{105}{2}} \left(1 + \frac{1}{104} \frac{104(\tilde{x} - 75)^2}{\beta}\right)^{-\frac{105}{2}}$$

$$= \frac{\Gamma\left(\frac{105}{2}\right)}{\Gamma\left(\frac{104}{2}\right)\sqrt{2\pi\beta}} \left(1 + \frac{1}{104} \frac{52(\tilde{x} - 75)^2}{\beta}\right)^{-\frac{105}{2}}$$

$$= \frac{\Gamma\left(\frac{105}{2}\right)}{\Gamma\left(\frac{104}{2}\right)\sqrt{104\pi\frac{\beta}{52}}} \left(1 + \frac{1}{104} \frac{52(\tilde{x} - 75)^2}{\beta}\right)^{-\frac{105}{2}}$$

$\Rightarrow (\tilde{x} \mid x) \sim t(\nu, a, b)$ where $\nu = 104$,
$a = 75$,
$b = \frac{\beta}{52}$

from (a)
$b = \frac{1805.65}{52}$
$= 34.724$

Question 2

(a)

2. $x_1, \dots, x_{100}$ and $y_1, \dots, y_{150}$ i.i.d

$$x_i \sim N(\mu_1, 1^2)$$
$$y_i \sim N\left(\mu_2, \left(\tfrac{1}{\sqrt{2}}\right)^2\right)$$
$$\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix} \sim N(\vec{\mu}, \vec{\Sigma}) \quad \text{with } \vec{\mu} = \begin{pmatrix}0\\0\end{pmatrix} \text{ and } \vec{\Sigma} = \begin{pmatrix}\frac{3}{5} & -\frac{2}{5}\\ -\frac{2}{5} & \frac{3}{5}\end{pmatrix}$$
$$f_{\vec{\mu},\Sigma}(\vec{x}) = \frac{1}{2\pi|\vec{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2}(\vec{x}-\vec{\mu})^T \Sigma^{-1}(\vec{x}-\vec{\mu})\right)$$

a) Derive $p(\mu_1 | \mu_2, x_1, \dots, x_{100}, y_1, \dots, y_{150})$
and $p(\mu_2 | \mu_1, x_1, \dots, x_{100}, y_1, \dots, y_{150})$

$$\Sigma^{-1} = \begin{pmatrix}3 & 2\\ 2 & 3\end{pmatrix}$$

$$p(\mu_1, \mu_2) \propto e^{\left(-\frac{1}{2}(\mu_1\ \mu_2)\begin{pmatrix}3 & 2\\ 2 & 3\end{pmatrix}\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}\right)}$$
$$= e^{\left(-\frac{1}{2}(3\mu_1 + 2\mu_2,\ 2\mu_1 + 3\mu_2)\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}\right)}$$
$$= e^{\left(-\frac{1}{2}(3\mu_1^2 + 4\mu_1\mu_2 + 3\mu_2^2)\right)}$$

$$p(\mu_1 | \mu_2, X, Y) \propto P(X, Y, \mu_1, \mu_2) \ [\text{drop terms without } \mu_1]$$
$$\propto e^{-\frac{1}{2}\left(\sum_{i=1}^{100}(x_i - \mu_1)^2 + 3\mu_1^2 + 4\mu_1\mu_2\right)}$$
$$\propto e^{-\frac{1}{2}\left(\sum_{i=1}^{100}(x_i^2 - 2x_i\mu_1 + \mu_1^2) + 3\mu_1^2 + 4\mu_1\mu_2\right)}$$
$$\propto e^{-\frac{1}{2}\left(100\mu_1^2 - 2\mu_1\sum_{i=1}^{100}x_i + 3\mu_1^2 + 4\mu_1\mu_2\right)}$$
$$\propto e^{-\frac{1}{2}\left(103\mu_1^2 - 2\mu_1\left(\sum_{i=1}^{100}x_i - 2\mu_2\right)\right)}$$
$$\propto e^{-\frac{103}{2}\left(\mu_1^2 - 2\mu_1\left(\frac{\sum_{i=1}^{100}x_i - 2\mu_2}{103}\right)\right)}$$
$$\propto e^{-\frac{1}{2}\frac{\left(\mu_1 - \frac{\sum_{i=1}^{100}x_i - 2\mu_2}{103}\right)^2}{\frac{1}{103}}}$$
$$\Rightarrow (\mu_1 | \mu_2, X, Y) \sim N\left(\frac{\sum_{i=1}^{100}x_i - 2\mu_2}{103}, \frac{1}{103}\right)$$

$$p(\mu_2 | \mu_1, X, Y) \propto P(X, Y, \mu_1, \mu_2) \ [\text{drop terms without } \mu_2]$$
$$\propto e^{-\left(\sum_{i=1}^{150}(y_i - \mu_2)^2 + 2\mu_1\mu_2 + \frac{3}{2}\mu_2^2\right)}$$
$$\propto e^{-\left(\sum_{i=1}^{150}(y_i^2 - 2y_i\mu_2 + \mu_2^2) + \frac{3}{2}\mu_2^2 + 2\mu_1\mu_2\right)}$$
$$\propto e^{-\left(150\mu_2^2 - 2\mu_2\sum_{i=1}^{150}y_i + \frac{3}{2}\mu_2^2 + 2\mu_1\mu_2\right)}$$
$$\propto e^{-\left(\frac{303}{2}\mu_2^2 - 2\mu_2\left(\sum_{i=1}^{150}y_i - \mu_1\right)\right)}$$
$$\propto e^{-\frac{303}{2}\left(\mu_2^2 - 4\mu_2\left(\frac{\sum_{i=1}^{150}y_i - \mu_1}{303}\right)\right)}$$
$$\propto e^{\frac{-\frac{1}{2}\left(\mu_2 - 2\frac{\sum_{i=1}^{150}y_i - \mu_1}{303}\right)^2}{\frac{1}{303}}}$$
$$\Rightarrow (\mu_2 | \mu_1, X, Y) \sim N\left(2\frac{\sum_{i=1}^{150}y_i - \mu_1}{303}, \frac{1}{303}\right)$$

(f)

2.

f) $\log q_1(\mu_1) \propto E_{\mu_2}\left[-\frac{1}{2}\left(\sum_{i=1}^{100}(x_i-\mu_1)^2+3\mu_1^2+4\mu_1\mu_2\right)\right]$

$\log q_2(\mu_2) \propto E_{\mu_1}\left[-\left(\sum_{i=1}^{150}(y_i-\mu_2)^2+\frac{3}{2}\mu_2^2+2\mu_1\mu_2\right)\right]$

from (a)

$\Rightarrow q_1(\mu_1)$: pdf of $N(\mu_1^*, \sigma_1^{2*})$, $\mu_1^* = \dfrac{\sum_{i=1}^{100} x_i - 2E_{\mu_2}(\mu_2)}{103}$, $\sigma_1^{2*} = \frac{1}{103}$

$\Rightarrow q_2(\mu_2)$: pdf of $N(\mu_2^*, \sigma_2^{2*})$, $\mu_2^* = 2\dfrac{\sum_{i=1}^{150} y_i - E_{\mu_1}(\mu_1)}{303}$, $\sigma_2^{2*} = \frac{1}{303}$

$E_{\mu_2}(\mu_2) = \mu_2^*$

$E_{\mu_1}(\mu_1) = \mu_1^*$

(g)

2.

g) $ELBO\left(q^*_{\mu_1}(\mu_1), q^*_{\mu_2}(\mu_2)\right) = E_{\mu_1\mu_2}\left[\log p(X, Y, \mu_1, \mu_2) - \log\left(q^*_{\mu_1}(\mu_1) q^*_{\mu_2}(\mu_2)\right)\right]$

$$= E_{\mu_1,\mu_2}\left[\log(p(X|\mu_1)p(Y|\mu_2)p(\mu_1,\mu_2)) - \log q^*_{\mu_1}(\mu_1) - \log q^*_{\mu_2}(\mu_2)\right]$$

$$= E_{\mu_1,\mu_2}\left[\log\left(p(X|\mu_1)p(Y|\mu_2)p(\mu_1,\mu_2)\right)\right] - E_{\mu_1\mu_2}\left[\log q^*_{\mu_1}(\mu_1)\right] - E_{\mu_1,\mu_2}\left[\log q^*_{\mu_2}(\mu_2)\right]$$

$$E_{\mu_1,\mu_2}\left[\log q^*_{\mu_1}(\mu_1)\right] \propto -\frac{1}{2}\log\sigma_1^{2*} - \frac{E_{\mu_1}\left[(\mu_1-\mu_1^*)^2\right]}{2\sigma_1^{2*}}$$

$$\propto -\frac{1}{2}\log\sigma_1^{2*} - \frac{\sigma_1^{2*}}{2\sigma_1^{2*}}$$

$$\propto -\frac{1}{2}\log\sigma_1^{2*}$$

$$\Rightarrow E_{\mu_1\mu_2}\left[\log q^*_{\mu_2}(\mu_2)\right] \propto -\frac{1}{2}\log\sigma_2^{2*}$$

$$\propto E_{\mu_1\mu_2}\left[-\frac{1}{2}\sum_{i=1}^{100}(x_i-\mu_1)^2 - \sum_{i=1}^{150}(y_i-\mu_2)^2 - \frac{3}{2}\mu_1^2 - 2\mu_1\mu_2 - \frac{3}{2}\mu_2^2\right] + \frac{1}{2}\log\sigma_1^{2*} + \log\sigma_2^{2*}$$

from (a)

$$\propto -\frac{103}{2}\left(\sigma_1^{2*}+\mu_1^{2*}\right) + \mu_1^*\sum_{i=1}^{100}x_i - \frac{303}{2}\left(\sigma_2^{2*}+\mu_2^{2*}\right)$$

$$+ 2\mu_2^*\sum_{i=1}^{150}y_i - 2\mu_1\mu_2 + \log\sigma_1^{2*} + \log\sigma_2^{2*} + \text{const.}$$

$$\propto -\frac{103}{2}\left(\mu_1^{2*}\right) + \mu_1^*\sum_{i=1}^{100}x_i - \frac{303}{2}\left(\mu_2^{2*}\right) + 2\mu_2^*\sum_{i=1}^{150}y_i - 2\mu_1^*\mu_2^* + \text{const.}$$