AMS 206B Homework #2 **Mary Silva**
UCSC 5 February 2019

1. Let $X_1, ..., X_n$ be iid samples such that $X_i|\theta \sim N(\theta, \sigma^2)$, where σ^2 known and θ is unknown. Also let your prior for θ be a mixture of conjugate priors, i.e.,

$$\pi(\theta) = \sum_{l=1}^{K} w_l \phi(\theta|\mu_l, \tau^2)$$

where $\phi(\theta|\mu_l, \tau^2)$ denotes the Gaussian density with mean μ_l and variance τ and mixture weights $0 < w_l < 1$ for all $l = 1, ..., k$ with $\sum_{l=1}^{K} w_l = 1$.

(a) For the posterior distribution for θ based on this prior.
Solution: First we consider the l- components of the Gaussian mixture prior, then $m_l(\boldsymbol{x})$ is the l-th marginal.

$$\begin{aligned} f(\theta, \boldsymbol{x}) \propto f(\boldsymbol{x}|\theta)\pi(\theta) &= f(\boldsymbol{x}|\theta) \sum_{l=1}^{K} w_l \pi_l(\theta) \\ &= \sum_{l=1}^{K} w_l f(\boldsymbol{x}|\theta)\pi_l(\theta) \\ &= \sum_{l=1}^{K} w_l f_l(\theta|\boldsymbol{x}) m_l(\boldsymbol{x}) \end{aligned}$$

For any component, the posterior is

$$f_l(\theta|\boldsymbol{x}) = \phi\left(\theta \middle| \left(\frac{1}{\tau^2} + \frac{n}{\sigma^2}\right)^{-1} \left(\frac{n\bar{x}}{\sigma^2} + \frac{\mu_l}{\tau^2}\right), \left(\frac{1}{\tau^2} + \frac{n}{\sigma^2}\right)^{-1}\right)$$

The l-th component marginal

$$m_l(\boldsymbol{x}) = \phi_n\left(\boldsymbol{x} \middle| \mathbf{1_n}\mu_l, \mathbf{1_n}\mathbf{1_n}'\tau^2 + \boldsymbol{I_n}\sigma^2\right)$$

Thus,

$$f(\theta|\boldsymbol{x}) = \frac{\sum_{l=1}^{K} w_l \phi_n\left(\boldsymbol{x} \middle| \mathbf{1_n}\mu_l, \mathbf{1_n}\mathbf{1_n}'\tau^2 + \boldsymbol{I_n}\sigma^2\right) \phi\left(\theta \middle| \left(\frac{1}{\tau^2} + \frac{n}{\sigma^2}\right)^{-1} \left(\frac{n\bar{x}}{\sigma^2} + \frac{\mu_l}{\tau^2}\right), \left(\frac{1}{\tau^2} + \frac{n}{\sigma^2}\right)^{-1}\right)}{\sum_{l=1}^{K} w_l \phi_n\left(\boldsymbol{x} \middle| \mathbf{1_n}\mu_l, \mathbf{1_n}\mathbf{1_n}'\tau^2 + \boldsymbol{I_n}\sigma^2\right)}$$

(b) Find the posterior mean.
Solution:

$$E(\theta|\boldsymbol{x}) = \frac{\sum_{l=1}^{K} w_l \phi_n\left(\boldsymbol{x} \middle| \mathbf{1_n}\mu_l, \mathbf{1_n}\mathbf{1_n}'\tau^2 + \boldsymbol{I_n}\sigma^2\right) \left(\frac{1}{\tau^2} + \frac{n}{\sigma^2}\right)^{-1} \left(\frac{n\bar{x}}{\sigma^2} + \frac{\mu_l}{\tau^2}\right)}{\sum_{l=1}^{K} w_l \phi_n\left(\boldsymbol{x} \middle| \mathbf{1_n}\mu_l, \mathbf{1_n}\mathbf{1_n}'\tau^2 + \boldsymbol{I_n}\sigma^2\right)}$$

(c) Find the prior predictive distribution.
Solution:

$$\begin{aligned}
m(\boldsymbol{x}) &= \int_{-\infty}^{\infty} f(\boldsymbol{x}|\theta) \sum_{l=1}^{K} w_l \pi_l(\theta) d\theta \\
&= \sum_{l=1}^{K} w_l \int_{-\infty}^{\infty} f(\boldsymbol{x}|\theta) \pi_l(\theta) \\
&= \sum_{l=1}^{K} w_l m_l(\boldsymbol{x}) \\
&= \sum_{l=1}^{K} w_l \phi_n \left(\boldsymbol{x} \middle| \boldsymbol{1_n} \mu_l, \boldsymbol{1_n}\boldsymbol{1_n}'\tau^2 + \boldsymbol{I_n}\sigma^2 \right)
\end{aligned}$$

(d) Find the posterior predictive distribution.
Solution:

$$\begin{aligned}
f(y|\boldsymbol{x}) &\propto \int_{-\infty}^{\infty} f(y|\theta)\pi(\theta|x) d\theta \\
&\propto \sum_{l=1}^{K} w_l \int_{-\infty}^{\infty} f(y|\theta) \phi \left(\theta \middle| \left(\frac{1}{\tau^2} + \frac{n}{\sigma^2} \right)^{-1} \left(\frac{n\bar{x}}{\sigma^2} + \frac{\mu_l}{\tau^2} \right), \left(\frac{1}{\tau^2} + \frac{n}{\sigma^2} \right)^{-1} \right) d\theta \\
&= \sum_{l=1}^{K} w_l m_l(\boldsymbol{x}) \\
&= \sum_{l=1}^{K} w_l \phi_n \left(\boldsymbol{x} \middle| \boldsymbol{1_n}\mu_l, \boldsymbol{1_n}\boldsymbol{1_n}'\tau^2 + \boldsymbol{I_n}\sigma^2 \right) \phi \left(y \middle| \left(\frac{1}{\tau^2} + \frac{n}{\sigma^2} \right)^{-1} \left(\frac{n\bar{x}}{\sigma^2} + \frac{\mu_l}{\tau^2} \right), \sigma^2 + \left(\frac{1}{\tau^2} + \frac{n}{\sigma^2} \right)^{-1} \right) \\
&= w_l^*
\end{aligned}$$

Therefore

$$f(y|\boldsymbol{x}) = \frac{w_l^*}{\sum_{l=1}^{K} w_l \phi_n \left(\boldsymbol{x} \middle| \boldsymbol{1_n}\mu_l, \boldsymbol{1_n}\boldsymbol{1_n}'\tau^2 + \boldsymbol{I_n}\sigma^2 \right)}$$

2. Let $X_1, ..., Xn$ be an iid sample s.t. $X_i|\theta \sim N(\theta, 1)$. Suppose that you know $\theta > 0$ and you want your prior to reflect hat. Hence, you let your prior be

$$\pi(\theta) = \frac{1}{\sqrt{2\pi\tau^2}\Phi(\mu/\tau)} \exp\left\{ -\frac{(\theta-\mu)^2}{2\tau^2} \right\} I_{[0,\infty)(\theta)}$$

(a) Find the posterior distribution for θ
Solution: Since $\bar{x}$ is a sufficient statistic, we use sufficiency principle:

$$\begin{aligned}
\pi(\theta|\bar{x}) &\propto f(\bar{x}|\theta) \\
&= \exp\left\{ -\frac{(\bar{x}-\theta)^2}{2/n} \right\} \exp\left\{ -\frac{(\theta-\mu)^2}{2\tau^2} \right\} I_{[0,\infty)}(\theta)
\end{aligned}$$

which is the kernel for $N\left(\left(n + \frac{1}{\tau^2}\right)^{-1} \left(n\bar{x} + \frac{\mu}{\tau^2}\right), \left(n + \frac{1}{\tau^2}\right)^{-1} \right)$ truncated at 0. Let $\tau^{2*} = \left(n + \frac{1}{\tau^2}\right)^{-1}$ and $\mu^2 = \left(n + \frac{1}{\tau^2}\right)^{-1} \left(n\bar{x} + \frac{\mu}{\tau^2}\right)$

(b) Find the prior predictive distribution.
Solution:

$$
\begin{aligned}
m(\boldsymbol{x}) &= \int_{-\infty}^{\infty} f(\boldsymbol{x}|\theta)\pi(\theta)d\theta \\
&= \frac{1}{(2\pi)^{n/2}}\frac{1}{(2\pi\tau^2)^{n/2}}\frac{1}{\Phi(\mu/\tau)}\exp\left\{-\frac{\sum_{i=1}^n x_i^2}{2}-\frac{\mu^2}{2\tau^2}\right\} \\
&\times \int_0^\infty \exp\left\{-\frac{1}{2}\left(n+\frac{1}{\tau^2}\right)^{-1}\left(\theta-\left(n+\frac{1}{\tau^2}\right)^{-1}\left(n\bar{x}+\frac{\mu}{\tau^2}\right)\right)^2\right\}d\theta \\
&= \frac{1}{(2\pi)^{n/2}}\frac{\tau^*}{\tau}\frac{\Phi(\mu^*/\tau^{2*})}{\phi(\mu/\tau)}\exp\left\{-\frac{1}{2}\left[\sum_{i=1}^n x_i^2+\frac{\mu^2}{\tau^2}-\left(n+\frac{1}{\tau^2}\right)^{-1}\left(n\bar{x}+\frac{\mu}{\tau^2}\right)\right]\right\}
\end{aligned}
$$

Typing this was a BAD idea... I kept losing track of my τ's and τ^2's and τ^*'s :(

3. Let $X_1, ..., X_n$ be an iid sample from a truncated normal with unknown mean θ and variance 1. If $\theta \sim N(\mu, \tau^2)$, find the posterior for θ.
Solution:

$$
\begin{aligned}
\pi(\theta|\bar{x}) &\propto f(\bar{x}|\theta)\pi(\theta) \\
&= \frac{1}{\Phi(\theta)^n}\exp\left\{-\frac{1}{2}\left(n+\frac{1}{\tau^2}\right)^{-1}\left(\theta-\left(n+\frac{1}{\tau^2}\right)^{-1}\left(n\bar{x}+\frac{\mu}{\tau^2}\right)\right)^2\right\}
\end{aligned}
$$

This is not conjugate prior since there is the term $\frac{1}{\Phi(\theta)^n}$.

4. $x|\theta \sim Bin(n,\theta), \theta \sim Beta\left(\frac{\sqrt{n}}{2}, \frac{\sqrt{n}}{2}\right)$

(a) Give associated posterior:

$$\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{m(x)}$$

Similar to what we did in class, we obtain

$$\pi(\theta|x) = \frac{\theta^{x+\frac{\sqrt{n}}{2}-1}(1-\theta)^{n+\frac{\sqrt{n}}{2}-x-1}}{\mathcal{B}\left(\frac{\sqrt{n}}{2}+x, \frac{\sqrt{n}}{2}+n-x\right)}$$

which is a Beta distribution with shape $\frac{\sqrt{n}}{2}+x$, and scale$\frac{\sqrt{n}}{2}+n-x$.

(b) What is the estimator that minimizes the posterior expected loss if $\mathcal{L}(\delta,\theta) = (\theta-\delta)^2$?Call such estimator $\delta^\pi(x)$ and show that its associated risk $R(\theta,\delta^\pi)$ is constant.

The Bayes estimate under the squared error loss is the posterior mean:

$$\delta^\pi(x) = E(\theta|x) = \frac{\frac{\sqrt{n}}{2}+x}{n+\sqrt{n}}$$

Then the risk function can be defined as

$$\begin{aligned}
R(\theta,\delta) &= Var^{\pi}(\delta^{\pi}) + Bias^{\pi 2} \\
&= \frac{n\theta(1-\theta)}{(\sqrt{n}+n)^2} + \frac{\frac{\sqrt{n}}{2} - \sqrt{n}\theta}{(\sqrt{n}+n)^2} \\
&= \frac{n\theta - n\theta^2 + (\frac{n}{4} - n\theta + n\theta^2)}{(\sqrt{n}+n)^2} \\
&= \frac{n}{4(\sqrt{n}+n)^2} = \text{a constant}
\end{aligned}$$

(c) Let $\delta_0^{\pi} = x/n$. Find the risk for this estimator.

$$\begin{aligned}
R(\theta,\delta_0) &= E_\theta[\mathcal{L}(\theta,\delta_0)] \\
&= E_\theta[(\theta - x/n)^2] \\
&= E[\theta^2 - \frac{2x\theta}{n} + \frac{x^2}{n^2}] \\
&= \theta^2 - \frac{2\theta}{n}EX + \frac{1}{n^2}EX^2 \\
&= \theta^2 - 2\theta^2 + \frac{1}{n^2}((n\theta)^2 + n\theta(1-\theta)) \\
&= \theta^2 - 2\theta^2 + \frac{(\theta^2 + \theta(1-\theta))}{n} \\
&= -\theta^2 + \frac{\theta}{n} \\
&= \frac{\theta(1-\theta)}{n}
\end{aligned}$$

Therefore

$$R(\theta,\delta^{\pi}(x)) = \frac{1}{4(\sqrt{n}+n)^2} \leq \frac{\theta(1-\theta)}{n} = R(\theta,\delta_0^{\pi}(x))$$

5. $x \sim N(\theta,1), \theta \sim N(0,n)$. Let $\delta^{\pi}(x)$ be the estimator that minimizes the posterior expected loss under squared error. Show that the Bayes risk $r(\pi,^{\pi}) = \frac{n}{n+1}$.
From class we know that

$$\theta|x \sim N\left(\left(1+\frac{1}{n}\right)^{-1}x, \left(1+\frac{1}{n}\right)^{-1}\right)$$