Worksheet 3

Foundations of Bayesian Methodology

Wenjie Tu

Spring Semester 2022

Exercise 3 (Conjugate Bayes: analytical derivation)

Assumptions:

$$\begin{aligned} y_1, \cdots, y_n \mid m &\overset{i.i.d}{\sim} \mathcal{N}(m, \kappa^{-1}) \\ m &\sim \mathcal{N}(\mu, \lambda^{-1}) \end{aligned}$$

3 (a)

Prior distribution:

$$f(m) = \sqrt{\frac{\lambda}{2\pi}} \exp\left(-\frac{\lambda}{2}(m-\mu)^2\right)$$

The prior predictive distribution of one future observation y assuming that no observations have been collected yet:

$$\begin{aligned}
f(y) &= \int_{-\infty}^{\infty} f(y, m) \mathrm{d}m \quad \textcolor{red}{\text{marginalize over its prior distribution}} \\
&= \int_{-\infty}^{\infty} f(y \mid m) f(m) \mathrm{d}m \\
&= \int_{-\infty}^{\infty} \sqrt{\frac{\kappa}{2\pi}} \exp\left(-\frac{\kappa}{2}(y-m)^2\right) \sqrt{\frac{\lambda}{2\pi}} \exp\left(-\frac{\lambda}{2}(m-\mu)^2\right) \mathrm{d}m \\
&= \frac{\sqrt{\kappa\lambda}}{2\pi} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\left(\kappa(y^2 - 2ym + m^2) + \lambda(m^2 - 2m\mu + \mu^2)\right)\right) \mathrm{d}m \\
&= \frac{\sqrt{\kappa\lambda}}{2\pi} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\left((\kappa+\lambda)\left(m - \frac{\kappa y + \lambda\mu}{\kappa+\lambda}\right)^2 - \frac{(\kappa y + \lambda\mu)^2}{\kappa+\lambda} + \kappa y^2 + \lambda\mu^2\right)\right) \mathrm{d}m \\
&= \frac{\sqrt{\kappa\lambda}}{2\pi} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\left((\kappa+\lambda)\left(m - \frac{\kappa y + \lambda\mu}{\kappa+\lambda}\right)^2 + \frac{\kappa\lambda(y-\mu)^2}{\kappa+\lambda}\right)\right) \mathrm{d}m \\
&= \frac{\sqrt{\kappa\lambda}}{2\pi} \exp\left(-\frac{\kappa\lambda(y-\mu)^2}{2(\kappa+\lambda)}\right) \int_{-\infty}^{\infty} \exp\left(-\frac{\kappa+\lambda}{2}\left(m - \frac{\kappa y + \lambda\mu}{\kappa+\lambda}\right)^2\right) \mathrm{d}m \\
&= \frac{\sqrt{\kappa\lambda}}{2\pi} \sqrt{\frac{2\pi}{\kappa+\lambda}} \exp\left(-\frac{\kappa\lambda(y-\mu)^2}{2(\kappa+\lambda)}\right) \underbrace{\int_{-\infty}^{\infty} \sqrt{\frac{\kappa+\lambda}{2\pi}} \exp\left(-\frac{\kappa+\lambda}{2}\left(m - \frac{\kappa y + \lambda\mu}{\kappa+\lambda}\right)^2\right) \mathrm{d}m}_{=1} \\
&= \frac{\sqrt{\kappa\lambda}}{2\pi} \sqrt{\frac{2\pi}{\kappa+\lambda}} \exp\left(-\frac{\kappa\lambda(y-\mu)^2}{2(\kappa+\lambda)}\right) \\
&= \sqrt{\frac{1}{2\pi\left(\frac{1}{\lambda} + \frac{1}{\kappa}\right)}} \exp\left(-\frac{(y-\mu)^2}{2\left(\frac{1}{\lambda} + \frac{1}{\kappa}\right)}\right)
\end{aligned}$$

The prior predictive distribution of one future observation y is

$$\mathcal{N}(\mu, \lambda^{-1} + \kappa^{-1})$$

3 (b)

In worksheet 2, we have already derived the posterior distribution:

$$m \mid y_1, \cdots, y_n \sim \mathcal{N}\left(\frac{\kappa n \bar{y} + \lambda \mu}{n\kappa + \lambda}, \frac{1}{n\kappa + \lambda}\right)$$

Denote:

$$\mu_{\text{post}} = \frac{\kappa n \bar{y} + \lambda \mu}{n\kappa + \lambda}$$
$$\lambda_{\text{post}} = n\kappa + \lambda$$

$$\begin{aligned} f(m \mid y_1, \cdots, y_n) &= \sqrt{\frac{n\kappa + \lambda}{2\pi}} \exp\left(-\frac{n\kappa + \lambda}{2}\left(m - \frac{\kappa n \bar{y} + \lambda \mu}{n\kappa + \lambda}\right)^2\right) \\ &= \sqrt{\frac{\lambda_{\text{post}}}{2\pi}} \exp\left(-\frac{\lambda_{\text{post}}}{2}\left(m - \mu_{\text{post}}\right)^2\right) \end{aligned}$$

The posterior predictive distribution of one future observation y_{n+1} given that $y_1, \cdots, y_n$ have been observed:

$$\begin{aligned} f(y_{n+1} \mid y_1, \cdots, y_n) &= \int_{-\infty}^{\infty} f(y_{n+1}, m \mid y_1, \cdots, y_n) \mathrm{d}m \quad \textcolor{red}{\text{marginalize over its posterior distribution}} \\ &= \int_{-\infty}^{\infty} f(y_{n+1} \mid m, y_1, \cdots, y_n) f(m \mid y_1, \cdots, y_n) \mathrm{d}m \\ &= \int_{-\infty}^{\infty} f(y_{n+1} \mid m) f(m \mid y_1, \cdots, y_n) \mathrm{d}m \\ &= \int_{-\infty}^{\infty} \sqrt{\frac{\kappa}{2\pi}} \exp\left(-\frac{\kappa}{2}(y_{n+1} - m)^2\right) \underbrace{\sqrt{\frac{\lambda_{\text{post}}}{2\pi}} \exp\left(-\frac{\lambda_{\text{post}}}{2}\left(m - \mu_{\text{post}}\right)^2\right)}_{\text{posterior distribution}} \mathrm{dm} \\ &\vdots \quad \text{repeat the same steps as for the prior predictive distribution} \\ &= \sqrt{\frac{1}{2\pi\left(\frac{1}{\lambda_{\text{post}}} + \frac{1}{\kappa}\right)}} \exp\left(-\frac{(y - \mu_{\text{post}})^2}{2\left(\frac{1}{\lambda_{\text{post}}} + \frac{1}{\kappa}\right)}\right) \end{aligned}$$

$$y_{n+1} \mid y_1, \cdots, y_n \sim \mathcal{N}(\mu_{\text{post}}, \lambda_{\text{post}}^{-1} + \kappa^{-1})$$