Assignment2
Conjuagate Priors

Weiqi Xie, Xinyang Li

January 30, 2024

1 Q1.1

To show the posterior distribution of the mean μ from the given information, we can use Bayes' theorem.

The prior distribution for the mean is,

$$N(\mu \mid \mu_0, \sigma_0^2)$$

and the likelihood of the N i.i.d. data is,

$$N(x_i \mid \mu, \sigma^2)$$

for $i \in N$, the observed data points are assumed to be independent and identically distributed.

The following is the derivation

Prior distribution of μ

$$P(\mu) = N(\mu|\mu_0, \sigma_0^2)$$

Likelihood function of the data

$$P(\{x_i\}|\mu) = \prod_{i=1}^{N} N(x_i|\mu, \sigma^2)$$

Posterior distribution of μ

$$P(\mu|\{x_i\}) \propto P(\{x_i\}|\mu) \cdot P(\mu)$$

$$P(\mu|\{x_i\}) \propto \prod_{i=1}^{N} N(x_i|\mu, \sigma^2) \cdot N(\mu|\mu_0, \sigma_0^2)$$

$$\propto \exp\left(-\frac{1}{2}\left[\frac{1}{\sigma^2}\sum_{i=1}^{N}(x_i - \mu)^2 + \frac{1}{\sigma_0^2}(\mu - \mu_0)^2\right]\right)$$

$$\propto \exp\left(-\frac{1}{2}\left[\frac{1}{\sigma^2}\sum_{i=1}^{N}(x_i^2 - 2x_i\mu + \mu^2) + \frac{1}{\sigma_0^2}(\mu^2 - 2\mu\mu_0 + \mu_0^2)\right]\right)$$

Simplify the expression

$$\propto \exp\left(-\frac{1}{2}\left[\frac{N}{\sigma^2}\mu^2 - 2\mu\left(\frac{\sum_{i=1}^{N} x_i}{\sigma^2}\right) + \frac{1}{\sigma^2}\sum_{i=1}^{N} x_i^2 + \frac{1}{\sigma_0^2}\mu^2 - 2\mu\left(\frac{\mu_0}{\sigma_0^2}\right) + \frac{\mu_0^2}{\sigma_0^2}\right]\right)$$

$$\propto \exp\left(-\frac{1}{2}\left[\left(\frac{N}{\sigma^2} + \frac{1}{\sigma_0^2}\right)\mu^2 - 2\mu\left(\frac{\sum_{i=1}^{N} x_i}{\sigma^2} + \frac{\mu_0}{\sigma_0^2}\right) + \frac{1}{\sigma^2}\sum_{i=1}^{N} x_i^2 + \frac{\mu_0^2}{\sigma_0^2}\right]\right)$$

Now, we can recognize this expression is a function of normal distribution. Completing the square results in a Gaussian distribution. The terms in the exponential are proportional to a Gaussian distribution with mean $\frac{\frac{\sum_{i=1}^{N} x_i}{\sigma^2} + \frac{\mu_0}{\sigma_0^2}}{\frac{N}{\sigma^2} + \frac{1}{\sigma_0^2}}$ and variance $\frac{1}{\frac{N}{\sigma^2} + \frac{1}{\sigma_0^2}}$

Therefore, the posterior distribution for the mean μ is

$$N\left(\mu \,\middle|\, \frac{\frac{\sum_{i=1}^{N} x_i}{\sigma^2} + \frac{\mu_0}{\sigma_0^2}}{\frac{N}{\sigma^2} + \frac{1}{\sigma_0^2}}, \frac{1}{\frac{N}{\sigma^2} + \frac{1}{\sigma_0^2}}\right)$$

Comparing this with the provided expression, we can see that the derived posterior distribution matches the given form.

$$N(\mu|\frac{\sigma_0^2}{N\sigma^2 + \sigma_0^2}\bar{x} + \frac{\sigma^2}{N\sigma^2 + \sigma_0^2}\mu_0, \left(\frac{1}{\sigma_0^2} + \frac{N}{\sigma^2}\right)^{-1})$$

2 Q1.2

To show the posterior distribution of the mean μ from the given information, we can use Bayes' theorem.

Like Q1.1, we can know the likelihood function for a normal distribution with known mean μ and unknown variance σ^2 is given by:

$$p(\{x_i\}|\mu, \sigma^2) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right)$$

The prior distribution for the variance is given by the inverse-gamma distribution:

$$p(\sigma^2|\alpha, \beta) \propto (\sigma^2)^{-\alpha-1} \exp\left(-\frac{\beta}{\sigma^2}\right)$$

The posterior distribution is proportional to the product of the likelihood and the prior distribution:

$$\begin{aligned} p(\sigma^2|\{x_i\}, \mu, \alpha, \beta) &\propto p(\{x_i\}|\mu, \sigma^2) \cdot p(\sigma^2|\alpha, \beta) \\ &\propto \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right) \cdot (\sigma^2)^{-\alpha-1} \exp\left(-\frac{\beta}{\sigma^2}\right) \end{aligned}$$

Taking simplifying,

$$\begin{aligned} \log p(\sigma^2|\{x_i\}, \mu, \alpha, \beta) \propto &-\frac{1}{2}\left(N\log(2\pi) + N\log(\sigma^2) + \sum_{i=1}^{N} \frac{(x_i-\mu)^2}{\sigma^2}\right) \\ &-(\alpha+1)\log(\sigma^2) - \frac{\beta}{\sigma^2} \\ \log p(\sigma^2|\{x_i\}, \mu, \alpha, \beta) \propto &-\frac{1}{2}(N+2\alpha+2)\log(\sigma^2) \\ &+\frac{1}{2\sigma^2}\sum_{i=1}^{N}(x_i-\mu)^2 - \frac{1}{2}\left(\frac{2\beta}{\sigma^2} + N\log(2\pi) + 2\beta\right) \end{aligned}$$

This can be expressed in the form $\frac{1}{2}a(\sigma^2 - b)^2$ by completing the square, where $a = -(N+2\alpha+2)$ and b is the term inside the square. Solving for b, we find:

$$b = \frac{\sum_{i=1}^{N}(x_i-\mu)^2}{(N+2\alpha+2)}$$

Substituting b back into the expression for a, we get:

$$a = -\frac{1}{\sigma^2}\sum_{i=1}^{N}(x_i-\mu)^2$$

$$\begin{aligned} \log p(\sigma^2|\{x_i\}, \mu, \alpha, \beta) \propto &-\frac{1}{2}(N+2\alpha+2)\log(\sigma^2) \\ &+\frac{1}{2\sigma^2}\sum_{i=1}^{N}(x_i-\mu)^2 - \frac{1}{2}\left(\frac{2\beta}{\sigma^2} + N\log(2\pi) + 2\beta\right) \end{aligned}$$

Finally, we have the posterior distribution:

$$p(\sigma^2|\{x_i\}, \mu, \alpha, \beta) \propto (\sigma^2)^{-\frac{1}{2}(N+2\alpha+2)-1} \exp\left(-\frac{\sum_{i=1}^{N}(x_i-\mu)^2}{2\sigma^2} - \frac{\beta}{\sigma^2}\right)$$

Comparing this to the standard form of the inverse-gamma distribution $IG(\sigma^2|\alpha', \beta')$,

$$\alpha' = \frac{1}{2}(N+2\alpha+2), \quad \beta' = \frac{\sum_{i=1}^{N}(x_i-\mu)^2}{2} + \beta$$

Therefore, the posterior distribution is $IG(\sigma^2|\alpha', \beta')$ with parameters $\alpha' = \frac{1}{2}(N+2\alpha+2)$ and $\beta' = \frac{\sum_{i=1}^{N}(x_i-\mu)^2}{2} + \beta$.

3 Q1.3

To show the posterior distributions of the mean μ and the precision $\tau = \frac{1}{\sigma^2}$, we can use Bayes' theorem
The likelihood function for the normal distribution is given by

$$p(x \mid \mu, \tau) = \prod_{i=1}^{N} \frac{\tau}{\sqrt{2\pi}} \exp\left(-\frac{\tau}{2}(x_i - \mu)^2\right)$$

The prior distributions are given as

$$\begin{aligned} p(\mu \mid \mu_0, \tau_0) &= \frac{\tau_0}{\sqrt{2\pi}} \exp\left(-\frac{\tau_0}{2}(\mu - \mu_0)^2\right) \\ p(\tau \mid \alpha, \beta) &= \frac{\beta^\alpha}{\Gamma(\alpha)} \tau^{\alpha-1} \exp\left(-\beta\tau\right) \end{aligned}$$

Compute the posterior distributions.
Posterior distribution for μ

$$p(\mu \mid x, \tau) \propto p(x \mid \mu, \tau) \cdot p(\mu \mid \mu_0, \tau_0) \propto \prod_{i=1}^{N} \frac{\tau}{\sqrt{2\pi}} \exp\left(-\frac{\tau}{2}(x_i - \mu)^2\right) \cdot \frac{\tau_0}{\sqrt{2\pi}} \exp\left(-\frac{\tau_0}{2}(\mu - \mu_0)^2\right)$$

$$\propto -\frac{1}{2}\left(\tau \sum_{i=1}^{N} x_i^2 + \tau_0\mu_0^2 - 2\tau\mu_0 \sum_{i=1}^{N} x_i + (\tau + \tau_0)\mu^2\right)$$

This expression is quadratic in μ and can be recognized as the kernel of a Gaussian distribution. The posterior distribution for μ is therefore a Gaussian distribution with mean and precision given by

$$\begin{aligned} \mu_{\text{post}} &= \frac{\tau \sum_{i=1}^{N} x_i + \tau_0\mu_0}{\tau + \tau_0} \\ (\tau_{\text{post}})^{-1} &= \tau + \tau_0 \end{aligned}$$

This gives the posterior distribution for μ as

$$N(\mu \mid \mu_{\text{post}}, (\tau_{\text{post}})^{-1}) = N\left(\mu \mid \frac{\tau \sum_{i=1}^{N} x_i + \tau_0\mu_0}{\tau + \tau_0}, (\tau + \tau_0)^{-1}\right)$$

Posterior Distribution for τ

$$p(\tau \mid x, \mu) \propto p(x \mid \mu, \tau) \cdot p(\tau \mid \alpha, \beta) \propto \prod_{i=1}^{N} \frac{\tau}{\sqrt{2\pi}} \exp\left(-\frac{\tau}{2}(x_i - \mu)^2\right) \cdot \frac{\beta^\alpha}{IG(\alpha)} \tau^{\alpha-1} \exp\left(-\beta\tau\right)$$

Taking the logarithm and simplifying

$$\propto (\alpha - 1)\log(\tau) - \beta\tau - \tau \sum_{i=1}^{N} (x_i - \mu)^2$$

This expression is proportional to the kernel of a Gamma distribution. The posterior distribution for τ is therefore a Gamma distribution with shape and rate parameters given by

$$\begin{aligned} \alpha_{\text{post}} &= \alpha + \frac{N}{2} \\ \beta_{\text{post}} &= \beta + \frac{1}{2} \sum_{i=1}^{N} (x_i - \mu)^2 + \frac{N\tau_0}{2}(\tau + \tau_0)(\bar{x} - \mu_0)^2 \end{aligned}$$

This gives the posterior distribution for τ as

$$IG(\tau \mid \alpha_{\text{post}}, \beta_{\text{post}}) = IG\left(\tau \mid \alpha + \frac{N}{2}, \beta + \frac{1}{2} \sum_{i=1}^{N} (x_i - \mu)^2 + \frac{N\tau_0}{2}(\tau + \tau_0)(\bar{x} - \mu_0)^2\right)$$