week3

November 20, 2021

1 3주차

1.1 Posterior Predictive Distribution

- After the data y have been observed, we can predict an unknown observable $\tilde{y}$
- The posterior predictive distribution of a future observation, $\tilde{y}$ is:

$$\begin{aligned} p(\tilde{y}|y) &= \int p(\tilde{y}, \theta|y) d\theta \\ &= \int p(\tilde{y}|\theta, y) p(\theta|y) d\theta \\ &= \int p(\tilde{|}\theta) p(\theta|y) d\theta \end{aligned}$$

- Assumed y and $\tilde{y}$ are conditional independent given θ
- prior predictive distribution before y observed

1.1.1 Example 1 : Binomial model

$$y_i \overset{iid}{\sim} Bern(\theta)$$

$$Y \sim Bin(n, \theta), 0 \leq \theta \leq 1$$

$$\begin{aligned} &\theta \sim Unif(0,1) \\ &\Rightarrow \theta|y \sim Beta(y+1, n-y+1) \end{aligned}$$

* Posterior predictive distribution for $\tilde{y} = 1$

$$P(\tilde{y} = 1|y) = \frac{y+1}{n+2}$$

which is known as `Laplace's law of succession`.

$$p(\tilde{y} = 1|y) = \int_0^1 p(\tilde{y} = 1|\theta) p(\theta|y) d\theta = \int_0^1 \theta p(\theta|y) d\theta = E[\theta|y] = \frac{y+1}{n+2}$$

$$y = 0 \text{ (all failures) } \Rightarrow p(\tilde{y} = 1|y) = \frac{1}{n+2}$$

$$y = 1 \text{ (all successes) } \Rightarrow p(\tilde{y} = 1|y) = \frac{n+1}{n+2}$$

- prior predictive distribution

$$p(\tilde{y}=1) = \int_0^1 p(\tilde{y}=1|\theta)p(\theta)d\theta = \int_0^1 \theta d\theta = \frac{1}{2}$$

1.1.2 Example 2 : Poisson Model

- Data model: $y_i \overset{iid}{\sim} Poisson(\theta), i = 1, ..., n$
- Prior distribution : $\theta \sim Gamma(\alpha, \beta)$
- Posterior distribution of θ given y

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{y_i!}\theta^{y_i}e^{-\theta} = (\prod_{i=1}^{n} \frac{1}{y_i!}\theta^{\sum y_i}e^{-n\theta}$$

- MLE for θ:

$$logL(\theta) = log(\prod \frac{1}{y_i!} + \sum y_i log\theta - n\theta$$

$$\frac{\partial logL}{\partial \theta} = \frac{\sum y_i}{\theta} - n = 0$$

$$\Rightarrow \hat{\theta}_{ML} = \frac{1}{n}\sum_{i=1}^{n} y_i = \bar{y}$$

- Posterior Distribution

$$\begin{aligned} p(\theta|y) &\propto p(y|\theta)p(\theta) \\ &= [\prod_{i=1}^{n} \frac{1}{y_i!}\theta^{y_i}e^{-\theta}]\frac{1}{\Gamma(\alpha)\beta^{\alpha}}\theta^{\alpha-1}e^{-\frac{\theta}{\beta}} \\ &\propto \theta^{\sum y_i e^{-n\theta}} e^{\alpha-1}e^{-\frac{\theta}{\beta}} \\ &= e^{\sum y_i+\alpha-1}e^{-(n+\frac{1}{\beta})\theta} \end{aligned}$$

$$\int_0^{\infty} c\theta^{\sum y_i+\alpha-1}e^{-(n+\frac{1}{\beta})\theta}d\theta = 1$$

$$\Rightarrow \theta|y \sim Gamma(\sum y_i + \alpha, [n+\frac{1}{\beta}]^{-1})$$

$$\Rightarrow p(\theta|y) = \frac{(n+\frac{1}{\beta})^{\sum y_i+\alpha}}{\Gamma(\sum y_i+\alpha)}e^{\sum y_i+\alpha-1}e^{-(n+\frac{1}{\beta})\theta}$$

$$\begin{aligned} E[\theta|y] &= \frac{\sum y_i + \alpha}{n+\frac{1}{\beta}} = \hat{\theta}_{Bayes} \\ &= \frac{n}{n+\frac{1}{\beta}}(\frac{\sum y_i}{n}) + \frac{\frac{1}{\beta}}{n+\frac{1}{\beta}}(\alpha\beta) \\ &\Rightarrow \text{ Weighted average of sample mean and prior mean} \end{aligned}$$

  - $n \uparrow \Rightarrow E[\theta|y] \rightarrow \hat{\theta}_{ML}$
  - $n \downarrow \Rightarrow E[\theta|y] \rightarrow \alpha\beta$ (prior mean)

1.2 Posterior Predictive Distribution of Poisson Model

- Posterior predictive distribution, $p(\tilde{y}|y)$:

$$\begin{aligned}
p(\tilde{y}|y_1,...,y_n) &= \int_0^\infty p(\tilde{y}|\theta)p(\theta|y_1,...,y_n)d\theta \\
&= \int_0^\infty \frac{1}{\tilde{y}!}e^{-\theta}\theta^{\tilde{y}}\frac{(n+\frac{1}{\beta})^{\sum y_i+\alpha}}{\Gamma(\sum y_i+\alpha)}\theta^{\sum y_i+\alpha-1}e^{-(n+\frac{1}{\beta})\theta}d\theta \\
&= [\frac{1}{\tilde{y}!}\frac{(n+\frac{1}{\beta})^{\sum y_i+\alpha}}{\Gamma(\sum y_i+\alpha)}]\int_0^\infty \theta^{\tilde{y}+\sum y_i+\alpha-1}e^{-(n+\frac{1}{\beta}+1)\theta}d\theta \\
&= [\frac{1}{\tilde{y}!}\frac{(n+\frac{1}{\beta})^{\sum y_i+\alpha}}{\Gamma(\sum y_i+\alpha)}][\frac{\Gamma(\tilde{y}+\sum y_i+\alpha)}{(n+\frac{1}{\beta}+1)^{\tilde{y}+\sum y_i+\alpha}}] \\
&= \frac{\Gamma(\tilde{y}+\sum y_i+\alpha)}{\Gamma(\sum y_i+\alpha)\tilde{y}!}(\frac{n+\frac{1}{\beta}}{n+\frac{1}{\beta}+1})^{\sum y_i+\alpha}(\frac{1}{n+\frac{1}{\beta}+1})^{\tilde{y}}
\end{aligned}$$

$$\Rightarrow \theta|y \sim NegBin(\sum y_i+\alpha, \frac{n+\frac{1}{\beta}}{n+\frac{1}{\beta}+1})$$

$\tilde{y} =$ the number failures until r th successes

1.3 Normal Model with a Single Observation

- Normal model with unknown mena θ and known variance σ^2

$$y \sim N(\theta, \sigma^2)$$

- Prior distribution : $\theta \sim N(\mu : \tau^2)$
- Posterior distribution of θ given y

$$\begin{aligned}
p(\theta|y) &\propto p(y|\theta)p(\theta) \\
&= [\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2\sigma^2}(y-\theta)^2}][\frac{1}{\sqrt{2\pi}\tau}e^{-\frac{1}{2\tau^2}(\theta-\mu)^2}] \\
&\propto exp[-\frac{1}{2\sigma^2}(y^2-2y\theta+\theta^2)-\frac{1}{2\tau^2}(\theta^2-2\theta\mu+\mu^2)] \\
&\propto exp[-\frac{1}{2\sigma^2}(\theta^2-2y\theta)-\frac{1}{2\tau^2}(\theta-2\mu\theta)] \\
&= exp[-\frac{1}{2}(\frac{1}{\sigma^2}+\frac{1}{\tau^2})\theta^2+2\frac{1}{2}(\frac{1}{\sigma^2}+\frac{\mu}{\tau^2})\theta] \\
&= exp[-\frac{1}{2}(\frac{1}{\sigma^2}+\frac{1}{\tau^2})(\theta^2-2\frac{\frac{y}{\sigma^2}+\frac{\mu}{\tau^2}}{\frac{1}{\sigma^2}+\frac{1}{\tau^2}}\theta)] \\
&\propto exp[-\frac{1}{2}(\frac{1}{\sigma^2}+\frac{1}{\tau^2})(\theta-\frac{\frac{y}{\sigma^2}+\frac{\mu}{\tau^2}}{\frac{1}{\sigma^2}+\frac{1}{\tau^2}})^2]
\end{aligned}$$

$$\theta|y \sim N(\frac{\frac{y}{\sigma^2}+\frac{\mu}{\tau^2}}{\frac{1}{\sigma^2}+\frac{1}{\tau^2}}, [\frac{1}{\sigma^2}+\frac{1}{\tau^2}]^{-1})$$

$$E[\theta|y] = \frac{\frac{1}{\sigma^2}}{\frac{1}{\sigma^2}+\frac{1}{\tau^2}}y + \frac{\frac{1}{\tau^2}}{\frac{1}{\sigma^2}+\frac{1}{\tau^2}}\mu$$

- τ^2 means prior variance
- If $\tau^2 \uparrow \Rightarrow$ Little information $\Rightarrow E[\theta|y] \to y$ (sample mean)
- If $\tau^2 \downarrow \Rightarrow$ Much information $\Rightarrow E[\theta|y] \to \mu$

- posterior variance $= [\frac{1}{\sigma^2} + \frac{1}{\tau^2}]^{-1}$
- precision $= \frac{1}{\text{variance}}$ (reciprocal of variance)
- $\frac{1}{\sigma^2}$: precision of data model
- $\frac{1}{\tau^2}$: precision of prior
- posterior precision = prior precision + data precision

1.4 Normal Model with Multiple Observations

- Normal model with unknown mean θ and known variance σ^2

$$y_i \overset{iid}{\sim} N(\theta, \sigma^2), i = 1, ..., n$$

- Prior distribution : $\theta \sim N(\mu, \tau^2)$
- Posterior distribution of θ given $y_1, ..., y_n$

$$\begin{aligned} p(\theta|y) \propto p(y|\theta)p(\theta) & \\ & \propto [\prod_{i=1}^{n} e^{-\frac{1}{2\sigma^2}(y_i-\theta)^2}]e^{-\frac{1}{2\tau^2}(\theta-\mu)^2} \\ & = e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\theta)^2}e^{-\frac{1}{2\tau^2}(\theta-\mu)^2} \\ & \propto exp[-\frac{1}{2}(\frac{n}{\sigma^2}+\frac{1}{\tau^2})(\theta - \frac{\frac{\sum y_i}{\sigma}+\frac{\mu}{\tau^2}}{\frac{n}{\sigma^2}+\frac{1}{\tau^2}})^2] \end{aligned}$$

$$\Rightarrow \theta|y \sim N(\frac{\frac{\sum y_i}{\sigma^2}+\frac{\mu}{\tau^2}}{\frac{n}{\sigma^2}+\frac{1}{\tau^2}}, [\frac{n}{\sigma^2}+\frac{1}{\tau^2}]^{-1})$$

- Posterior mean

$$E[\theta|y] = \frac{\frac{n}{\sigma^2}}{\frac{n}{\sigma^2}+\frac{1}{\tau^2}}(\frac{\sum y_i}{n}) + \frac{\frac{1}{\tau^2}}{\frac{n}{\sigma^2}+\frac{1}{\tau^2}}(\mu)$$

$$n \uparrow \Rightarrow E[\theta|]y \to \hat{\theta}_{ML} = \bar{y}$$

$$\tau^2 \uparrow \text{(no information)} \Rightarrow E[\theta|y] \to \bar{y}$$

$$\tau^2 \downarrow \text{(much information)} \Rightarrow E[\theta|y] \to \mu$$

  - Posterior precision = sample precision + prior precision