Bayesian_Statistics_HW1

陈子睿 15220212202842

3/26/24

1 Different priors for Gamma distribution

Consider $y_1, \ldots, y_n \sim Gamma(a, a\theta)$ where a is given and $\theta > 0$.

(a) The Jeffrey's prior is given by:

$$\begin{aligned}
\pi(\theta) &\propto \sqrt{I(\theta)} \\
&= \sqrt{-E_Y\left[\frac{\partial^2}{\partial^2\theta}\log f(Y|\theta)\right]} \\
&= \sqrt{-E_Y\left[\frac{\partial^2}{\partial^2\theta}\left(a\log(a\theta) - \log(\Gamma(a)) + (a-1)\log X - a\theta X\right)\right]} \\
&= \sqrt{-E_Y\left[-\frac{a}{\theta^2}\right]} \\
&= \frac{\sqrt{a}}{\theta}.
\end{aligned}$$

(b) A possible conjugate prior of θ is

$$a\theta \sim Gamma(\alpha, \beta).$$

Hence

$$\theta \sim Gamma(\alpha, \frac{\beta}{a}).$$

Then we have the prior density

$$\pi(\theta) = \frac{(\frac{\beta}{a})^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\frac{\beta}{a}\theta},$$

and the density of y

$$p(y) = \frac{(a\theta)^a}{\Gamma(a)}y^{a-1}e^{-a\theta y}.$$

The posterior density is:

$$\begin{aligned}
\pi(\theta|y_1, \ldots, y_n) &\propto p(y_1, \ldots, y_n|\theta)\pi(\theta) \\
&\propto \left[\frac{(a\theta)^a}{\Gamma(a)}\right]^n \left(\prod_{i=1}^n y_i^{a-1}e^{-a\theta y_i}\right)\frac{(\frac{\beta}{a})^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\frac{\beta}{a}\theta} \\
&\propto \theta^{\alpha+na-1}e^{-(a\sum_i y_i + \frac{\beta}{a})\theta} \\
&\sim Gamma(\alpha + na, a\textstyle\sum_i y_i + \frac{\beta}{a}).
\end{aligned}$$

(c) Now we assume that a is unknown but fixed for using an empirical Bayes approach. By the definition of Gamma distribution, for $z \sim Gamma(\alpha, \beta)$, we have that

$$E(z) = \frac{\alpha}{\beta},$$

$$Var(z) = \frac{\alpha}{\beta^2}.$$

Therefore, $E(y_i) = a/a\theta = 1/\theta$, $Var(y_i) = a/(a\theta)^2 = 1/a\theta^2$ for all $i = 1, \dots, n$. Hence we can define $\hat{a}(\mathbf{y})$ as:

$$\hat{a}(\mathbf{y}) = \frac{\bar{y}^2}{s^2}.$$

By some calculation, we can conclude that

$$E[\hat{a}(\mathbf{y})] = E\left[\frac{\bar{y}^2}{s^2}\right] = a.$$

Then the empirical Bayes prior is

$$\pi(\theta|\hat{a}(\mathbf{y})) = \frac{(\frac{\beta}{\hat{a}(\mathbf{y})})^{\alpha}}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\frac{\beta}{\hat{a}(\mathbf{y})}\theta}$$

Since the empirical Bayes approach is given by

$$\begin{aligned}
\pi(\theta|y_1, \dots, y_n, \hat{a}(\mathbf{y})) &\propto p(y_1, \dots, y_n|\theta)\pi(\theta|\hat{a}(\mathbf{y})) \\
&\propto \theta^{\alpha+n\hat{a}(\mathbf{y})-1} e^{-(\hat{a}(\mathbf{y})\sum_i y_i + \frac{\beta}{\hat{a}(\mathbf{y})})\theta} \\
&= \theta^{\alpha+n\bar{y}^2/s^2-1} e^{-(\bar{y}^2/s^2\sum_i y_i + \frac{\beta}{\bar{y}^2/s^2})\theta} \\
&\sim Gamma(a + n\frac{\bar{y}^2}{s^2}, \frac{\bar{y}^2}{s^2}\textstyle\sum_i y_i + \frac{\beta}{\bar{y}^2/s^2}).
\end{aligned}$$

2 A coin toss question in the uniform prior case

Since the prior knowledge about p is that $0 \le p \le 1$, the uniform prior of p can be considered as

$$f(p) = \mathbf{1}_{[0,1]}.$$

The likelihood of the experiment is

$$f(x|p) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Therefore the posterior is given by

$$\begin{aligned}
f(p|x) &\propto f(x|p)f(p) \\
&= \binom{n}{k} p^k (1-p)^{n-k} \mathbf{1}_{[0,1]} \\
&\propto p^{10}(1-p)^{12-10} \\
&\sim Beta(11, 3).
\end{aligned}$$

Now, suppose $y \sim Binomial(5, p)$, then the posterior prediction distribution is

$$
\begin{aligned}
f(y|x) &= \int f(y|x,p) f(p|x) dp \\
&= \int \binom{5}{m} p^m (1-p)^{5-m} \frac{1}{B(11,3)} p^{10} (1-p)^2 dp \\
&= \int \binom{5}{m} \frac{1}{B(11,3)} p^{10+m} (1-p)^{7-m} dp \\
&= \binom{5}{m} \frac{B(10+m+1, 8-m)}{B(11,3)} \int \frac{1}{B(10+m+1, 8-m)} p^{10+m} (1-p)^{7-m} dp \\
&= \binom{5}{m} \frac{B(10+m+1, 8-m)}{B(11,3)}
\end{aligned}
$$

Therefore the prediction of the number of tails if we toss the coin 5 times more is given by

$$
\begin{aligned}
E(y|x) &= \int m \binom{5}{m} \frac{B(10+m+1, 8-m)}{B(11,3)} dm \\
&= \sum_{m=0}^{5} m \binom{5}{m} \frac{B(10+m+1, 8-m)}{B(11,3)}
\end{aligned}
$$

By the program below

```
y <- 0
n <- 5
m <- 0
for (m in 0:5) {
  y<- y + m * choose(n,m) * beta(10+m+1,8-m) / beta(11,3)
}
y
```

we have

```
[1] 3.928571.
```

3 Posterior predictive distribution of univariate normal with unknown mean

We're given that $X_1, \ldots, X_n|\mu \sim$ i.i.d $N(\mu, \sigma_1^2)$ and $\mu \sim N(\mu_0, \sigma_0^2)$. Then

$$\begin{aligned}
f(\mu|X_1, \ldots, X_n) &\propto f(\mu) f(X_1, \ldots, X_n|\mu) \\
&= f(\mu) \prod_{i=1}^{n} f(X_i|\mu) \\
&\propto exp\left\{-\frac{1}{2\sigma_0^2}(\mu - \mu_0)^2\right\} \prod_{i=1}^{n} exp\left\{-\frac{1}{2\sigma_1^2}(X_i - \mu)^2\right\} \\
&\propto exp\left\{-\frac{1}{2}\left[\frac{1}{\sigma_0^2}(\mu - \mu_0)^2 + \frac{1}{\sigma_1^2}\sum_{i=1}^{n}(X_i - \mu)^2\right]\right\} \\
&\sim N(\mu_n, \sigma_n^2).
\end{aligned}$$

where

$$\mu_n = \frac{\frac{1}{\sigma_0^2}\mu_0 + \frac{n}{\sigma_1^2}\bar{x}}{\frac{1}{\sigma_0^2} + \frac{n}{\sigma_1^2}} \text{ and } \frac{1}{\sigma_n^2} = \frac{1}{\sigma_0^2} + \frac{n}{\sigma_1^2}.$$

The the posterior predictive distribution $f(X_{n+1}|X_1, \ldots, X_n)$ is given by

$$\begin{aligned}
f(X_{n+1}|X_1, \ldots, X_n) &= \int f(X_{n+1}|X_1, \ldots, X_n, \mu) f(\mu|X_1, \ldots, X_n) d\mu \\
&\propto \int exp\left\{-\frac{1}{2\sigma_1^2}(X_{n+1} - \mu)^2\right\} exp\left\{-\frac{1}{2\sigma_n^2}(\mu - \mu_n)^2\right\} d\mu \\
&\propto exp\left\{-\frac{1}{2(\sigma_1^2 + \sigma_n^2)}(X_{n+1} - \mu_n)^2\right\} \\
&\sim N(\mu_n, \sigma_1^2 + \sigma_n^2).
\end{aligned}$$

4 Dirichlet distribution

(a) The Dirichlet prior has form

$$f(\theta|\alpha) = \frac{1}{B(\alpha)} \theta_{11}^{4-1} \theta_{01}^{2-1} \theta_{10}^{2-1} \theta_{00}^{3-1}.$$

where $\theta = (\theta_{11}, \theta_{01}, \theta_{10}, \theta_{00})'$, $\alpha = (4, 2, 2, 3)'$. By the property of Dirichlet distribution Dirichlet distribution, we have that

$$E(\theta_{11}) = \frac{4}{4+2+2+3} = \frac{4}{11}, E(\theta_{01}) = \frac{2}{4+2+2+3} = \frac{2}{11},$$

$$E(\theta_{10}) = \frac{2}{4+2+2+3} = \frac{2}{11}, E(\theta_{00}) = \frac{3}{4+2+2+3} = \frac{3}{11}.$$

and

$$Var(\theta_{11}) = \frac{4 \times (11-4)}{11^2 \times 12} = \frac{7}{363}, Var(\theta_{01}) = \frac{2 \times (11-2)}{11^2 \times 12} = \frac{3}{242},$$

$$Var(\theta_{10}) = \frac{2 \times (11-2)}{11^2 \times 12} = \frac{3}{242}, Var(\theta_{00}) = \frac{3 \times (11-3)}{11^2 \times 12} = \frac{2}{121}.$$

(b) The posterior is

$$\begin{aligned}
f(\theta|y) &\propto f(\theta)f(y|\theta) \\
&\propto \theta_{11}^3\theta_{01}\theta_{10}\theta_{00}^2 \binom{55}{y_{11}, y_{01}, y_{10}, y_{00}} \theta_{11}^{y_{11}}\theta_{01}^{y_{01}}\theta_{10}^{y_{10}}\theta_{00}^{y_{00}} \\
&\propto \theta_{11}^{3+y_{11}}\theta_{01}^{1+y_{01}}\theta_{10}^{1+y_{10}}\theta_{00}^{2+y_{00}} \\
&\sim Dirichlet(4+y_{11}, 2+y_{01}, 2+y_{10}, 3+y_{00}) \\
&= Dirichlet(29, 8, 9, 16)
\end{aligned}$$

(c) In the same case as (a), we have

$$E(\theta_{11}|y) = \frac{29}{62}, E(\theta_{01}|y) = \frac{4}{31}, E(\theta_{10}|y) = \frac{9}{62}, E(\theta_{00}|y) = \frac{8}{31},$$

$$Var(\theta_{11}|y) = \frac{319}{80724}, Var(\theta_{01}|y) = \frac{12}{6727}, Var(\theta_{10}|y) = \frac{53}{26908}, Var(\theta_{00}|y) = \frac{184}{60543}.$$

(d) Using R, we have

```
library(MCMCpack)
alpha <- c(29, 8, 9, 16)
# generating Dirichlet Distribution sample
set.seed(123)
samples <- rdirichlet(100000, alpha)
ci_95 <- apply(samples, 2, function(x) quantile(x, probs = c(0.025, 0.975)))
print(ci_95)
           [,1]       [,2]       [,3]      [,4]
2.5%  0.3454844 0.05858723 0.06968283 0.1584641
97.5% 0.5913356 0.22236531 0.24273000 0.3735698
```

The 95% posterior interval, based on the exact posterior distribution is given by the output.

5 Log-normal distribution

(a) The prior distribution of $log\theta$ is $N(2.5, 0.5^2)$. For convenience, we denote $\psi = log\theta \sim N(2.5, 0.5^2)$Then the posterior distribution of ψ is

$$\begin{aligned}
f(\psi|x) &\propto f(x|\psi)f(\psi) \\
&\sim N(\mu_n, \sigma_n^2)
\end{aligned}$$

with

$$\mu_n = \frac{\frac{n}{\sigma_1^2}\bar{x} + \frac{\mu_0}{\sigma_0^2}}{\frac{n}{\sigma_1^2} + \frac{1}{\sigma_0^2}} = \frac{\frac{10}{0.05^2} \times 3.035 + \frac{2.5}{0.5^2}}{\frac{10}{0.05^2} + \frac{1}{0.5^2}} = 3.034466,$$

and

$$\sigma_n^2 = \frac{1}{\frac{10}{0.05^2} + \frac{1}{0.5^2}} = 0.01580349^2.$$

(b) A symmetric 95% posterior interval for $log\theta$ is

$$\left[\mu_n - z_{\alpha/2}\sigma_n, \mu_n + z_{\alpha/2}\sigma_n\right] = [3.003492, 3.06544].$$

(c) A symmetric 95% posterior interval for θ is

$$\left[e^{\mu_n - z_{\alpha/2}\sigma_n}, e^{\mu_n + z_{\alpha/2}\sigma_n}\right] = [20.15579, 21.4439].$$

(d) The posterior probability that $\theta < 20$ is equivalent to $log\theta < log20$, by using

```
pnorm(log(20), mean = mean, sd = sd, lower.tail = TRUE, log.p = FALSE)
[1] 0.007123774
```

Thus

$$P(\theta < 20) = P(log\theta < log20) = 0.007123774.$$