Homework 9

PkGu

12/1/2021

P[263] 4.

(1). For model $\mathcal{M}_{\alpha,\beta}$:

$$\begin{cases} \mathbf{x}_n \sim B(n,p) \\ p \sim Beta(\alpha,\beta) \end{cases}$$

We have marginal likelihood (or model evidence) $\mathbf{x}|\alpha,\beta$

$$P_{\alpha,\beta}(t) = \binom{n}{t}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{\Gamma(\alpha+t)\Gamma(n+\beta-t)}{\Gamma(n+\alpha+\beta)}$$

where $t = \sum_1^n x_i$. By definition,

$$\begin{aligned} -\frac{1}{2}BIC_{\alpha,\beta} &= \log P_{\alpha,\beta}(t|\hat{p}) - \frac{1}{2}\log n \\ &= t\log\hat{p} + (n-t)\log(1-\hat{p}) + \log\binom{n}{t} - \frac{1}{2}\log n \end{aligned}$$

where $\hat{p} = \hat{p}_{MLE} = \frac{t}{n}$.

(2). $-\frac{1}{2}BIC$ are used to approximately compute $\log P_{\alpha,\beta}(t)$. Taking $\alpha = 2, \beta = 4$, we try to approximately compute the difference between them, which is defined by $\mathbb{E}[|\log P(t) - (-\frac{1}{2}BIC)|]$. For simplicity in computation, we use Stirling formula

$$\log(\Gamma(m)) = m\log m - m - \frac{1}{2}\log m + \frac{1}{2}\log(2\pi) + o(1)$$

So

$$\begin{aligned} \log P(t) &= \log\Gamma(n+1) - \log\Gamma(t+1) - \log\Gamma(n-t+1) \\ &\quad + \log\Gamma(t+2) + \log\Gamma(n-t+4) - \log\Gamma(n+6) + \log(20) \\ &\approx n\log(n) - t\log(t) - (n-t)\log(n-t) \\ &\quad + \frac{1}{2}(\log(n) - \log(t) - \log(n-t)) \\ &\quad + (t+2)\log(t+2) + (n-t+4)\log(n-t+4) - (n+6)\log(n+6) \\ &\quad + \frac{1}{2}(\log(n+6) - \log(t+2) - \log(n-t+4)) \\ &\quad + \log(20) \\ &\approx \log(t) + 3\log(n-t) - 5\log(n) + \log(20) \end{aligned}$$

$$
\begin{aligned}
-\frac{1}{2}BIC &= t\log(t) + (n-t)\log(n-t) - n\log(n) \\
&\quad + \log\Gamma(n+1) - \log\Gamma(t+1) - \log\Gamma(n-t+1) \\
&\quad - \frac{1}{2}\log(n) \\
&\approx -\frac{1}{2}\log(t) - \frac{1}{2}\log(n-t)
\end{aligned}
$$

```
n <- 500
t <- numeric(n)
m <- 1000
err <- numeric(m)
rti <- numeric(m)
p <- rbeta(n,2,4)
for (i in 1:m){
  for (j in 1:n){
    t[j] <- rbinom(1,1,p[j])
  }
  tsum <- sum(t)
  err[i] <- abs((3/2)*log(tsum)+(7/2)*log(n-tsum)-5*log(n)+log(20))
}
print(mean(err))
```

```
## [1] 0.0709796
```

The result shows that in this case the BIC has accuracy $(1 \pm \epsilon)$ where ϵ can take 10% when $n > 100$, 7% when $n > 10000$.

P[263] 5.

(1). The model evidences are respectively

$$
\begin{aligned}
f(\mathbf{x}|M_0) &= \int_{\mathbb{R}} f(\mathbf{x}|\mu) f(\mu|M_0) d\mu \\
&= \int_{\mathbb{R}} (\frac{1}{2\pi})^{\frac{31}{2}} e^{-\frac{30(\bar{x}-\mu)^2 + \sum(x_i-\bar{x})^2}{2}} e^{-\frac{(\mu-1)^2}{2}} d\mu \\
&= (\frac{1}{2\pi})^{\frac{31}{2}} e^{-\frac{\sum(x_i-\bar{x})^2}{2}} \int_{\mathbb{R}} e^{-\frac{\mu^2 - 2\frac{30\bar{x}+1}{31}\mu + \frac{(30\bar{x}^2+1)}{31}}{\frac{2}{31}}} d\mu \\
&= (\frac{1}{2\pi})^{\frac{31}{2}} e^{-\frac{\sum(x_i-\bar{x})^2}{2}} e^{-\frac{31(\bar{x}-1)^2}{60}} \int_{\mathbb{R}} e^{-\frac{((\mu-\frac{30\bar{x}+1}{31})^2}{\frac{2}{31}}} d\mu \\
&= (2\pi)^{-15}(31)^{-\frac{1}{2}} e^{-\frac{\sum(x_i-\bar{x})^2}{2}} e^{-\frac{31(\bar{x}-1)^2}{60}}
\end{aligned}
$$

$$
\begin{aligned}
f(\mathbf{x}|M_1) &= \int_{\mathbb{R}} f(\mathbf{x}|\mu) f(\mu|M_1) d\mu \\
&= \int_{-1}^{1} (2\pi)^{-15} e^{-\frac{30(\bar{x}-\mu)^2 + \sum (x_i-\bar{x})^2}{2}} \frac{1}{2} d\mu \\
&= (2\pi)^{-15} e^{-\frac{\sum (x_i-\bar{x})^2}{2}} \frac{1}{2} \int_{-1}^{1} e^{-\frac{(\bar{x}-\mu)^2}{2\cdot\frac{1}{30}}} d\mu \\
&= (2\pi)^{-15} e^{-\frac{\sum (x_i-\bar{x})^2}{2}} \frac{1}{2\sqrt{30}} \int_{-\sqrt{30}-\sqrt{30}\bar{x}}^{\sqrt{30}-\sqrt{30}\bar{x}} e^{-\frac{\nu 2}{2}} d\nu \\
&= (2\pi)^{-15} e^{-\frac{\sum (x_i-\bar{x})^2}{2}} (\frac{\pi}{60})^{\frac{1}{2}} (\Phi(\sqrt{30}-\sqrt{30}\bar{x}) - \Phi(-\sqrt{30}-\sqrt{30}\bar{x}))
\end{aligned}
$$

So the Bayes factor is

$$
BF_{01} = (\frac{60}{31\pi})^{\frac{1}{2}} e^{-\frac{31(\bar{x}-1)^2}{60}} \Big/ \Big(\Phi(\sqrt{30}-\sqrt{30}\bar{x}) - \Phi(-\sqrt{30}-\sqrt{30}\bar{x}) \Big)
$$

(2).

Importance sampling method: Since $f(\mathbf{x}|M_0) = \mathbb{E}_{\mu|M_0}[f(\mathbf{x}|\mu)])$, and

$$
\pi(\mu|M_0) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(\mu-1)^2}{2}}
$$

$$
\pi(\mu|M_1) = \frac{1}{2}\delta_{(-1,1)}
$$

$$
f(\mathbf{x}|\mu) \propto e^{15(\bar{x}-\mu)^2}
$$

```
m <- 10
n <- 1000

wght1 <- function(mu){
  if(mu>-1 & mu<1){
    w <- sqrt(2*pi)*exp((mu-1)^2/2)/2
    return(w)
  }
  else{
    return(0)
  }
}

for(i in 1:m){
  w1 <- numeric(n)
  xbar <- rnorm(1,1,1/30)
  mu0 <- rnorm (n,1,1)
  f0 <- exp(15*(xbar-mu0))
  exp0 <- mean(f0)
  for(j in 1:n){
    w1[j] <- wght1(mu0[j])
  }
  wf1 <- exp(15*(xbar-mu0))*w1
```

```r
  exp1 <- mean(wf1)/mean(w1)
  print(exp0/exp1)
}
```

```
## [1] 5380473
## [1] 24472022
## [1] 828273.2
## [1] 10980867
## [1] 76612.27
## [1] 2128224
## [1] 1589853
## [1] 12568.71
## [1] 15374.21
## [1] 36376123814
```

MCMC method: We use random walk Monte Carlo algorithm here… An example has shown in Homework 7. But I can't understand why should I arrange such a complicated approximate algorithm for sampling from these simple distributions (a normal distribution and a uniform one)…