第二次作业

舒文炫 PB18000019

目录

4.10 题

(1)

解: 由题可知,θ 的先验分布为

$$\pi(\theta) = \frac{100^{10}}{\Gamma(10)}\theta^{-11}e^{-\frac{100}{\theta}}$$

给定 $X_1, ..., X_n$ 后的后验分布

$$\begin{aligned}\pi(\theta|X) &\propto \pi(\theta)f(X|\theta) \\ &\propto \frac{100^{10}}{\Gamma(10)}\theta^{-11}e^{-\frac{100}{\theta}}(\frac{1}{\theta})^5 e^{-\frac{\sum_{i=1}^5 x_i}{\theta}} I_{(0.\infty)}(\theta) \\ &\propto \theta^{-16}e^{-\frac{100+\sum_{i=1}^5 x_i}{\theta}} I_{(0.\infty)}(\theta)\end{aligned}$$

也即 θ 的后验分布为 $\Gamma^{-1}(15, 100 + \sum_{i=1}^5 x_i)$, 带入 $x_1, ..., x_5$ θ 的后验分布为 $\Gamma^{-1}(15, 153)$ 由逆伽马分布 $\Gamma(\alpha, \beta)$ 的性质 θ 的后验期望估计为 $\mu^\pi(x) = \frac{\beta}{\alpha-1} = \frac{153}{14} = 10.93$ 后验方差估计为 $V^\pi(x) = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)} = \frac{153^2}{14^2 * 13} = 9.19$

(2)

将 $\Gamma^{-1}(\alpha,\beta)=\frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{-\alpha-1}e^{-\frac{\beta}{\theta}}$ 对 θ 求导可得

$$\begin{aligned}\frac{d\Gamma^{-1}(\alpha,\beta)}{d\theta}&=\frac{\beta^\alpha}{\Gamma(\alpha)}[(-\alpha-1)\theta^{-\alpha-2}e^{-\frac{\beta}{\theta}}+\theta^{-\alpha-1}e^{-\frac{\beta}{\theta}}\beta\theta^{-2}]\\&=\frac{\beta^\alpha}{\Gamma(\alpha)}e^{-\frac{\beta}{\theta}}\theta^{-\alpha-2}(-\alpha-1+\beta\theta^{-1})\end{aligned}$$

令上式为 0，得到 $\delta(x)=\frac{\beta}{\alpha+1}$ 即为众数估计，带入值 $\alpha=15,\beta=153$, 得到众数估计 $\frac{153}{16}=9.56$ 其后验均方误差

$$PMSE(\delta(x))=V^\pi(x)+(\mu^\pi(x)-\delta(x))^2$$

带入数据可得 $PMSE(\delta(x))=9.19+(10.93-9.56)^2=11.07$ 即后验众数估计的后验均方误差为 11.07.

4.22 题

(1)

解若先验分布假设为 $N(\mu,\tau^2)$ 由 θ 对称单峰，中位数为 0, 可知其均值为 0，也即 $\mu=0$ 再由 $\frac{1}{4}$ 分位数为-1,即 $P(\theta\le-1)=0.25$ 也即 $P(\frac{\theta}{\tau}\le-\frac{1}{\tau})=0.25$

由 $\frac{\theta}{\tau}\sim N(0,1)$

```
print(qnorm(0.25))
```

```
## [1] -0.6744898
```

```
print(1/qnorm(0.75))
```

```
## [1] 1.482602
```

```
print((1/qnorm(0.75))^2)
```

```
## [1] 2.198109
```

$\frac{1}{4}$ 分位数为-0.674

从而 $\tau = \frac{1}{0.67} = 1.483$

θ 的先验分布为 $N(0, 2.198)$ θ 的后验分布

$$\begin{aligned}\pi(\theta|x) &\propto f(x|\theta)\pi(\theta)\\ &\propto e^{-\frac{(x-\theta)^2}{2}}e^{-\frac{(\theta-\mu)^2}{2\tau^2}}\\ &\propto e^{-\frac{(\tau^2+1)\theta^2-(2x\tau^2+2\mu)\theta+\tau^2x+\mu^2}{2\tau^2}}\\ &\propto e^{-\frac{(\theta-\frac{x\tau^2+\mu}{\tau^2+1})^2}{\frac{2\tau^2}{\tau^2+1}}}\end{aligned}$$

θ 的后验分布为 $N(\frac{x\tau^2+\mu}{\tau^2+1}, \frac{\tau^2}{\tau^2+1})$

```
tau2=(1/qnorm(0.75))^2
print(6*tau2/(tau2+1))
```

```
## [1] 4.123892
```

```
print(tau2/(tau2+1))
```

```
## [1] 0.6873153
```

代入数据为 $N(4.124, 0.687)$

下面求 90% 的 HPD 可信区间，由于后验分布正态，是单峰对称的，只需要求 c, 使得 $P(\frac{|\theta-\mu 1|}{\tau 1} > c|x) = 0.05$, 那么 c 为 $N(0,1)$ 的 0.95 分位数

```
c=qnorm(0.95)
print(c)
```

```
## [1] 1.644854
```

带入得所求区间为 $[-c\tau_1+\mu_1, c\tau_1+\mu_1]$

```
c=qnorm(0.95)
tau2=(1/qnorm(0.75))^2
tau1=sqrt(tau2/(tau2+1))
mu1=6*tau2/(tau2+1)
print(-c*tau1+mu1)
```

```
## [1] 2.760234
```

```
print(c*tau1+mu1)
```

```
## [1] 5.487549
```

即 90% 的 HPD 区间为 $[2.76, 5.48]$

(2)

解柯西分布 $C(\alpha,\lambda)$ 的密度函数为 $\frac{\alpha}{\pi(\lambda^2+(x-\alpha)^2)}$

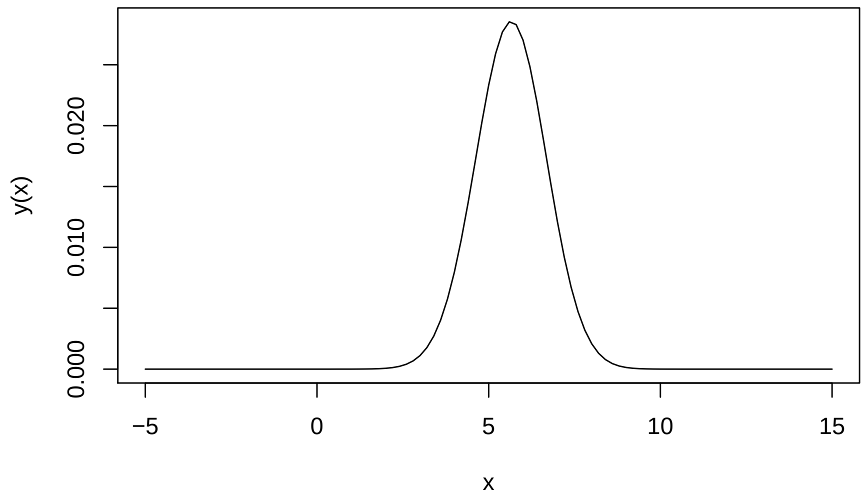
由中位数为 0，知 $\alpha=0$, 且我们有 $\frac{\theta}{\lambda}\sim C(0,1)$, 查表得 $C(0,1)$ 的 0.25 分位数即为-1，从而可得 $\lambda=1$ 即先验分布为 $C(0,1)$.

下面求后验

$$
\begin{aligned}
\pi(\theta|x) &\propto f(x|\theta)\pi(\theta)\\
&\propto e^{-\frac{(x-\theta)^2}{2}}\frac{1}{1+\theta^2}
\end{aligned}
$$

该分布函数不是常规的分布，考虑进行数值计算。

```
y<-function(x)
  exp(-(6-x)^2/2)/(1+x^2)
curve(y,-5,15)
```



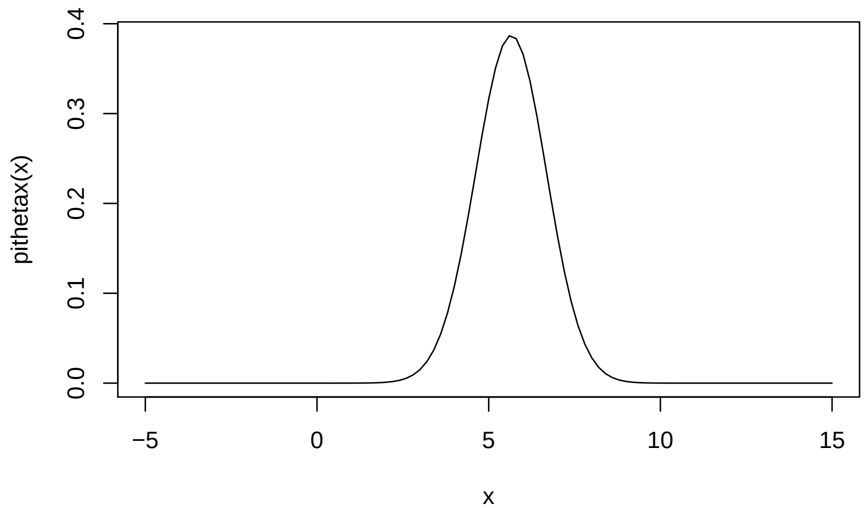

```
print(integrate(y,-Inf,Inf))
```

```
## 0.07384742 with absolute error < 1.8e-05
```

从而我们得到了后验分布约为 $\frac{e^{-\frac{(6-\theta)^2}{2}}}{0.0738(1+\theta^2)}$

画出图像为

```
pithetax<-function(x)# 后验密度
  exp(-(6-x)^2/2)/(1+x^2)/0.0738
curve(pithetax,-5,15)
```




```
print(integrate(pithetax,-Inf,Inf))
```

```
## 1.000643 with absolute error < 9.3e-05
```

```
opty<-function(x)# 求该函数的极小，对应原函数的极大
  -exp(-(6-x)^2/2)/(1+x^2)/0.0738
qaq<-optimize(opty,c(-5,15))
print(qaq)
```

```
## $minimum
## [1] 5.657194
##
## $objective
## [1] -0.3871336
```

观察这个图像发现是单峰的，峰值在 x=5.657194 处出现，且尾部概率很小。

```
hpd<-function(p,f){
  area=0
  k=1
```

```r
  k1=0 #f 不会小于 0
  k2=0.4 #f 不会大于 0.4
  table<-cbind(k,area)
  repeat{
    k=(k1+k2)/2
    fk<-function(x)
      exp(-(6-x)^2/2)/(1+x^2)/0.0738-k
    x1=uniroot(fk,c(5.66,15))
    x2=uniroot(fk,c(-5,5.66))
    area=integrate(f,x2$root,x1$root)$value
    if(area>p)k1=k else k2=k
    table1<-cbind(k,area)
    table<-rbind(table,table1)
    if(abs(area-p)<10^-5)
      break
  }

  print(table)

  return (c(x2$root,x1$root))
}
pithetax<-function(x)
  exp(-(6-x)^2/2)/(1+x^2)/0.0738
interval<-hpd(0.95,pithetax)
```

```
##                k      area
##  [1,] 1.00000000 0.0000000
##  [2,] 0.20000000 0.7493724
##  [3,] 0.10000000 0.9001759
##  [4,] 0.05000000 0.9572440
##  [5,] 0.07500000 0.9301538
##  [6,] 0.06250000 0.9440689
##  [7,] 0.05625000 0.9507496
```

```
##  [8,] 0.05937500 0.9474297
##  [9,] 0.05781250 0.9490963
## [10,] 0.05703125 0.9499244
## [11,] 0.05664062 0.9503374
## [12,] 0.05683594 0.9501310
## [13,] 0.05693359 0.9500277
## [14,] 0.05698242 0.9499761
## [15,] 0.05695801 0.9500019
```

```r
print(interval)
```

```
## [1] 3.623661 7.662668
```

这里函数的思想是闭区间套，让 k 逐渐收敛到我们需要的值。从而我们得到了可信水平 95% 的 HPD 可信集 $[3.62, 7.66]$

4.32 题

解由题 θ 的先验分布为

$$\pi(\theta) = \begin{cases} 0 & \theta < 5 \\ \frac{375}{\theta^4} & \theta \geq 5 \end{cases}$$

给定 θ 后，$f(x|\theta) = \frac{x}{\theta}$ 所以 θ 的后验分布为

$$\begin{aligned} \pi(\theta|x) &\propto f(x|\theta)\pi(\theta) \\ &\propto \frac{\prod_{i=1}^{5} x_i}{\theta^5} \frac{375}{\theta^4} I_{(\theta \geq 5)} I_{\theta > x_{(n)}} \end{aligned}$$

去掉常数部分得到 $\pi(\theta|x) \propto \frac{1}{\theta^9} I_{(\theta>14)}$ 也即 θ 的后验分布为 $Pa(14, 8)$ 下面我们求多重检验问题

$$P(H_1|x) = \int_0^{15} \frac{8\times 14^8}{\theta^9} I_{(\theta>14)} d\theta = 8 \times 14^8 (-\frac{1}{8\theta^8}\Bigg|_{\theta=14}^{\theta=15}) = 1 - (\frac{14}{15})^8 = 0.42$$

$$P(H_2|x) = \int_{15}^{20} \frac{8\times 14^8}{\theta^9} I_{(\theta>14)} d\theta = 8 \times 14^8 (-\frac{1}{8\theta^8}\Bigg|_{\theta=15}^{\theta=20}) = (\frac{14}{15})^8 - (\frac{14}{20})^8 = 0.52$$

$$P(H_3|x) = \int_{20}^{+\infty} \frac{8\times 14^8}{\theta^9} I_{(\theta>14)} d\theta = 8 \times 14^8 (-\frac{1}{8\theta^8}\Bigg|_{\theta=20}^{\theta=\infty}) = (\frac{14}{20})^8 - 0 = 0.06$$

从而 $P(H_2|x)$ 最大，即我们支持假设 H_2

4.34 题

解由题知，两架天平称重都服从正态分布 $N(\theta, 0.25)$，且钻石质量的先验分布为 $N(10,1)$, 从而我们得到钻石质量的后验分布为 $N(\mu_1, \eta_1^2)$, 其中

$$\mu_1 = \frac{\tau^2\bar{x} + \frac{\alpha_1^2}{n}\mu}{\frac{\alpha_1^2}{n} + \tau^2}$$
$$\eta_1^2 = \frac{\frac{\alpha_1^2}{n}\tau^2}{\frac{\alpha_1^2}{n} + \tau^2}$$

其中 $\tau^2 = 1, \alpha_1^2 = 0.25, n = 10, \mu = 10$, 代入得

```
tau2=1
alpha1=0.25
n=10
mu=10
x=c(9.45,10.62,9.40,10.12,9.85,10.92,10.93,9.85,9.81,10.28)
mx=mean(x)
mu1=(tau2*mx+alpha1/n*mu)/(tau2+alpha1/n)
eta1=(tau2*alpha1/n)/(tau2+alpha1/n)
print(mu1)
```

```
## [1] 10.12
```

```
print(eta1)
```

```
## [1] 0.02439024
```

可知 θ 的后验分布为 $N(10.12, 0.024)$ 在给定了 θ 的后验后，第二架天平测量值的后验分布

$$p_2(x|\theta) \propto e^{-\frac{(x-\theta)^2}{2\alpha_2^2}} e^{-\frac{(\theta-\mu_1)^2}{2\eta_1^2}}$$
$$\propto e^{-\frac{(x-\mu 1)^2}{2(\eta_1^2+\alpha_2^2)}}$$

即后验分布为 $N(\mu_1, \eta_1^2 + \alpha_2^2)$ 代入数据

```
eta2=eta1+alpha1
mu2=mu1
print(mu2)
```

```
## [1] 10.12
```

```
print(eta2)
```

```
## [1] 0.2743902
```

以后验均值作为第二架天平测量值的预测值，即为 10.12

95% 的预测区间为 $[\mu_2 - \eta_2 u_{0.025}, \mu_2 + \eta_2 u_{0.025}]$ 代入数据

```
print(mu2-sqrt(eta2)*qnorm(0.975))
```

```
## [1] 9.093326
```

```
print(mu2+sqrt(eta2)*qnorm(0.975))
```

```
## [1] 11.14667
```

从而后验预测区间为 $[9.09, 11.15]$