第八次作业

PB18000029 舒文炫

目录

7.2 题

考虑对 $N(\mu,1)$ 似然，我们有了从中抽出的样本 $x_1,...,x_n$，本题 $n=30$, 先用精确计算的方法对每个先验来求后验概率，那首先就要得到后验分布

对先验 $(1)N(0,2)$，这由于是正态分布的共轭分布，所以容易得到后验分布为正态分布，且后验均值和方差为

$$E(\theta|X_n)=\frac{\sigma^2/n}{\sigma^2/n+\tau^2}\mu+\frac{\tau^2}{\sigma^2/n+\tau^2}\bar{X}$$

$$Var(\theta|X_n)=\frac{\sigma^2\tau^2}{n\tau^2+\sigma^2}$$

这里代入 $\sigma^2=1,n=30,\mu=0,\tau^2=2$, 下面计算用 R 代码来展示

```
set.seed(1)## 先固定随机出一组随机数
Xn<-rnorm(30)## 从 N(0,1) 中抽出 30 个样本
sigma<-1
n<-30
mu<-0
tau<-2
## 一些参数
```

```r
postheta<-sigma/n/(sigma/n+tau)*mu+tau/(tau+sigma/n)*mean(Xn)
possigma<-sqrt(sigma*tau/(n*tau+sigma))
##-0.5~0.5
p1<-pnorm(0.5,postheta,possigma)-pnorm(-0.5,postheta,possigma)
##-0.2~0.6
p2<-pnorm(0.6,postheta,possigma)-pnorm(-0.2,postheta,possigma)
print(p1)
```

```
## [1] 0.9889847
```

```r
print(p2)
```

```
## [1] 0.9376435
```

这里可以看到两个后验概率第一个是 0.989，第二个是 0.938

下面取先验是 $N(1,2)$ 的情况，过程与上面相同，下面用 R 代码展示

```r
set.seed(1)## 先固定随机出一组随机数
Xn<-rnorm(30)## 从 N(0,1) 中抽出 30 个样本
sigma<-1
n<-30
mu<-1
tau<-2
## 一些参数
postheta<-sigma/n/(sigma/n+tau)*mu+tau/(tau+sigma/n)*mean(Xn)
possigma<-sqrt(sigma*tau/(n*tau+sigma))
##-0.5~0.5
p1<-pnorm(0.5,postheta,possigma)-pnorm(-0.5,postheta,possigma)
##-0.2~0.6
p2<-pnorm(0.6,postheta,possigma)-pnorm(-0.2,postheta,possigma)
print(p1)
```

```
## [1] 0.9864043
```

```
print(p2)
```

```
## [1] 0.9470493
```

可以看到两个后验概率第一个是 0.986，第二个是 0.947

下面取先验是 $U(-3,3)$ 的情况，这时后验分布不是正态分布了，但是我们得到了这个密度的核

$$\pi(\mu|X) \propto e^{\frac{(\mu-\bar{X})^2}{2\sigma^2/n}} I_{-3<\mu<3}$$

, 这里要求得密度函数，需要一点数值计算，下面计算过程用 R 代码展示

```
set.seed(1)## 先固定随机出一组随机数
Xn<-rnorm(30)## 从 N(0,1) 中抽出 30 个样本
sigma<-1
n<-30
## 一些参数
postheta<-mean(Xn)
possigma<-sqrt(sigma/n)
f<-function(x) exp(-(x-postheta)*(x-postheta)/(2*sigma/n))
## 密度函数前面的系数
a<-1/integrate(f,-3,3)$value
## 密度函数
f1<-function(x) a*exp(-(x-postheta)^2/(2*sigma/n))
##-0.5~0.5
p1<-integrate(f1,-0.5,0.5)$value
##-0.2~0.6
p2<-integrate(f1,-0.2,0.6)$value
print(p1)
```

```
## [1] 0.9881905
```

```
print(p2)
```

```
## [1] 0.9367854
```

可以看到两个后验概率第一个是 0.988，第二个是 0.937

下面使用正态逼近的方式来计算每个先验下的后验概率，根据定理，我们知道，当 n 趋于无穷是，后验分布趋于正态分布 $N(\tilde{\mu_n}, \tilde{I_n}^{-1})$, 其中 $\tilde{\mu_n}$ 为后验众数，$\tilde{I_n}$ 为费希尔信息阵。

当先验分布为 $N(0,2)$ 时，后验密度为正态分布，具体的前面算过了，这里后验众数即为后验均值，主要是算费希尔信息阵, 这里为了记号方便，记后验分布为 $N(\alpha, \beta^2)$

$$\begin{aligned}\frac{\partial ln\pi(\mu|X_n)}{\partial\mu} &= -\frac{\mu-\alpha}{\beta^2}\\ \tilde{I_n} &= -\frac{\partial^2 ln\pi(\mu|X_n)}{\partial\mu^2}\\ &= \frac{1}{\beta^2}\end{aligned}$$

从而可以得到，用正态逼近这里分布就是 $N(\alpha, \beta^2)$, 主要是本来后验就是正态，那么逼近正态等于计算出来的，很正常。

实际上结果和精确计算的没有区别，两个后验概率第一个是 0.989，第二个是 0.938

先验取为 N(1,2) 时也是一样，与精确计算的结果相同，两个后验概率第一个是 0.986，第二个是 0.947

先验取为 $U(-3,3)$ 这里就不一样了，由之前的计算，后验分布不是正态的, 密度函数是这样的形式

$$\pi(\mu|X_n) = ae^{\frac{(\mu-\alpha)^2}{2\beta^2}} I_{-3<\mu<3}$$

, 费希尔信息阵

$$\begin{aligned}\frac{\partial ln\pi(\mu|X_n)}{\partial\mu} &= -\frac{\mu-\alpha}{\beta^2}\\ \tilde{I_n} &= -\frac{\partial^2 ln\pi(\mu|X_n)}{\partial\mu^2}\\ &= \frac{1}{\beta^2}\end{aligned}$$

下面用 R 代码计算

```
set.seed(1)## 先固定随机出一组随机数
Xn<-rnorm(30)## 从 N(0,1) 中抽出 30 个样本
sigma<-1
n<-30
## 一些参数
postheta<-mean(Xn)
possigma<-sqrt(sigma/n)
## 之前算的后验密度函数
f1<-function(x) -a*exp(-(x-postheta)^2/(2*sigma/n))
opt<-optimize(f1,interval=c(-3,3))
## 正态逼近的后验均值
posmu<-opt$minimum
## 正态逼近的后验标准差
possigma<-sqrt(sigma/n)
##-0.5~0.5
p1<-pnorm(0.5,posmu,possigma)-pnorm(-0.5,posmu,possigma)
##-0.2~0.6
p2<-pnorm(0.6,posmu,possigma)-pnorm(-0.2,posmu,possigma)
print(p1)
```

```
## [1] 0.9881896
```

```
print(p2)
```

```
## [1] 0.936789
```

这里得到结果，两个后验概率第一个是 0.988，第二个是 0.937

如果考虑从 $N(1,1)$ 中抽样，因为分析和前面一样，这里我就给 R 代码以及运行结果。先看精确计算

先验分布 $N(0,2)$，R 代码如下

```
set.seed(1)## 先固定随机出一组随机数
Xn<-rnorm(30,1,1)## 从 N(1,1) 中抽出 30 个样本
sigma<-1
```

```
n<-30
mu<-0
tau<-2
## 一些参数
postheta<-sigma/n/(sigma/n+tau)*mu+tau/(tau+sigma/n)*mean(Xn)
possigma<-sqrt(sigma*tau/(n*tau+sigma))
##-0.5~0.5
p1<-pnorm(0.5,postheta,possigma)-pnorm(-0.5,postheta,possigma)
##-0.2~0.6
p2<-pnorm(0.6,postheta,possigma)-pnorm(-0.2,postheta,possigma)
print(p1)
```

```
## [1] 0.000908164
```

```
print(p2)
```

```
## [1] 0.005137106
```

两个后验概率第一个是 0.0009，第二个是 0.0051, 这里值都很小，因为先验是偏向 1 那边的。

先验分布 $N(1,2)$,R 代码如下

```
set.seed(1)## 先固定随机出一组随机数
Xn<-rnorm(30,1,1)## 从 N(1,1) 中抽出 30 个样本
sigma<-1
n<-30
mu<-1
tau<-2
## 一些参数
postheta<-sigma/n/(sigma/n+tau)*mu+tau/(tau+sigma/n)*mean(Xn)
possigma<-sqrt(sigma*tau/(n*tau+sigma))
##-0.5~0.5
p1<-pnorm(0.5,postheta,possigma)-pnorm(-0.5,postheta,possigma)
##-0.2~0.6
p2<-pnorm(0.6,postheta,possigma)-pnorm(-0.2,postheta,possigma)
```

```
print(p1)
```

```
## [1] 0.000665374
```

```
print(p2)
```

```
## [1] 0.003942008
```

两个后验概率第一个是 0.0007，第二个是 0.0059

先验分布 $U(-3,3)$,R 代码如下

```
set.seed(1)## 先固定随机出一组随机数
Xn<-rnorm(30,1,1)## 从 N(1,1) 中抽出 30 个样本
sigma<-1
n<-30
## 一些参数
postheta<-mean(Xn)
possigma<-sqrt(sigma/n)
f<-function(x) exp(-(x-postheta)*(x-postheta)/(2*sigma/n))
## 密度函数前面的系数
a<-1/integrate(f,-3,3)$value
## 密度函数
f1<-function(x) a*exp(-(x-postheta)^2/(2*sigma/n))
##-0.5~0.5
p1<-integrate(f1,-0.5,0.5)$value
##-0.2~0.6
p2<-integrate(f1,-0.2,0.6)$value
print(p1)
```

```
## [1] 0.0007107371
```

```
print(p2)
```

```
## [1] 0.004114432
```

两个后验概率第一个是 0.0007，第二个是 0.0041

使用正态逼近方法，前两个先验对应的结果是一样的

先验分布 $N(0,2)$ 时，两个后验概率第一个是 0.0009，第二个是 0.0051

先验分布 $N(0,2)$ 时，两个后验概率第一个是 0.0007，第二个是 0.0059

先验分布 $U(-3,3)$ 时，需要做一些计算, 过程和上面一样

```
set.seed(1)## 先固定随机出一组随机数
Xn<-rnorm(30,1,1)## 从 N(1,1) 中抽出 30 个样本
sigma<-1
n<-30
## 一些参数
postheta<-mean(Xn)
possigma<-sqrt(sigma/n)
## 之前算的后验密度函数
f1<-function(x) -a*exp(-(x-postheta)^2/(2*sigma/n))
opt<-optimize(f1,interval=c(-3,3))
## 正态逼近的后验均值
posmu<-opt$minimum
## 正态逼近的后验标准差
possigma<-sqrt(sigma/n)
##-0.5~0.5
p1<-pnorm(0.5,posmu,possigma)-pnorm(-0.5,posmu,possigma)
##-0.2~0.6
p2<-pnorm(0.6,posmu,possigma)-pnorm(-0.2,posmu,possigma)
print(p1)
```

```
## [1] 0.0007107388
```

```
print(p2)
```

```
## [1] 0.00411444
```

两个后验概率第一个是 0.0007，第二个是 0.0041

7.5 题

证明: 本题 X 的分布的参数空间是有限的，ppt 上对应有一个参数空间可数的定理，可以用在这里由题后验分布可计算如下：

$$\begin{aligned}\pi(\theta_r|x) &= \frac{f(x|\theta_i)\pi_i}{\sum_{i=1}^k f(x|\theta_i)\pi_i} \\ &= \frac{f(x|\theta_i)\pi_i/f(x|\theta_t)}{\sum_{i=1}^k f(x|\theta_i)\pi_i/f(x|\theta_t)} \\ &= \frac{exp(ln\frac{f(x|\theta_i)\pi_i}{f(x|\theta_t)})}{\sum_{i=1}^k exp(ln\frac{f(x|\theta_i)\pi_i}{f(x|\theta_t)})} \\ &= \frac{exp(ln\pi_i + S_{i,t})}{\sum_{i=1}^k exp(ln\pi_i + S_{i,t})}\end{aligned}$$

其中 $S_{i,t} = \sum_{j=1}^n \frac{lnf(x_j|\theta_i)}{lnf(x_j|\theta_t)}$ 对任意 t, 由大数定律得到

$$\lim_{n\to+\infty}\frac{1}{n}S_{i,t} = \int f(x|\theta_t)ln\frac{f(x|\theta_i)}{f(x|\theta_t)}dx$$

而由题，所有的 $f(x|\theta_i)$ 都不相同，且上面积分值在 $i = t$ 时为 0，故而 $S_{t,t} = 0$, 这里对任意 t 都对，当 $i \neq t$ 时, 由于分布 $f(x|\theta_i)$ 和分布 $f(x|\theta_t)$ 分布不同，KL 散度大于 0(这里用到了 KL 散度恒大于等于 0，且等于 0 当且仅当两个分布相等，这个定理之前课上也说过，这里就直接用了)，从而得到 $S_{i,t} = -\infty$, 从而可得

$$\lim_{n\to+\infty}\pi(\theta_t|x) = 1, \forall t \in 1, 2, ..., k$$

也即在每个 θ_i 处相合。