

STA623 - Bayesian Data Analysis - Practical 3 (Solutions)

Marc Henrion

2023-08-28

Practical 3

Notation

- X, Y, Z - random variables
- x, y, z - measured / observed values
- $\bar{X}, \bar{Y}, \bar{Z}$ - sample mean estimators for X, Y, Z
- $\bar{x}, \bar{y}, \bar{z}$ - sample mean estimates of X, Y, Z
- \hat{T}, \hat{t} - given a statistic T , estimator and estimate of T
- $P(A)$ - probability of an event A occurring
- $f_X(\cdot), f_Y(\cdot), f_Z(\cdot)$ - probability mass / density functions of X, Y, Z ; sometimes $p_X(\cdot)$ etc. rather than $f_X(\cdot)$
- $p(\cdot)$ - used as a shorthand notation for pmfs / pdfs if the use of this is unambiguous (i.e. it is clear which is the random variable)
- $X \sim F$ - X distributed according to distribution function F
- $E[X], E[Y], E[Z], E[T]$ - the expectation of X, Y, Z, T respectively

Exercise 1

Show that the Bayes estimator $\hat{\theta}_B$ for the quadratic loss function $\mathcal{C}(\theta - \hat{\theta}) = (\theta - \hat{\theta})^2$ is given by the posterior mean. In other words, show that:

$$E[\theta|y] = \arg \min_{\hat{\theta}} \int_y \int_{\Theta} \mathcal{C}(\theta - \hat{\theta}) p(\theta, y) d\theta dy$$

Exercise 1 (Solution)

We saw in lectures that we can use the multiplication rule and then optimise the inner integral only:

$$\hat{\theta}_B = \arg \min_{\hat{\theta}} \int \mathcal{C}(\theta - \hat{\theta}) p(\theta|y) d\theta \quad (1)$$

$$= \arg \min_{\hat{\theta}} \int (\theta - \hat{\theta})^2 p(\theta|y) d\theta \quad (2)$$

To find the minimum, we solve

$$\frac{d}{d\hat{\theta}} \int (\theta - \hat{\theta})^2 p(\theta|y) d\theta = 0 \quad (3)$$

$$\Leftrightarrow \int 2(\theta - \hat{\theta})(-1) p(\theta|y) d\theta = 0 \quad (4)$$

$$\Leftrightarrow \int (\theta - \hat{\theta}) p(\theta|y) d\theta = 0 \quad (5)$$

$$\Leftrightarrow \int \theta p(\theta|y) d\theta = \int \hat{\theta} p(\theta|y) d\theta \quad (6)$$

Note that $\int \hat{\theta} p(\theta|y) d\theta = \hat{\theta} \int p(\theta|y) d\theta = \hat{\theta}$ since the posterior distribution for θ is a probability distribution and needs to integrate to 1.

Therefore we see that

$$\hat{\theta}_B = \int \theta p(\theta|y) d\theta$$

This is the posterior mean $E[\theta|y] = \int \theta p(\theta|y) d\theta$.

Exercise 2

Suppose $\pi \sim \text{Beta}(2, 3)$ and $Y_1, \dots, Y_n \sim_{\text{iid}} \text{Bernoulli}(\pi)$. Further suppose we observe data y_1, \dots, y_n with $n = 25, k = \sum_i y_i = 16$.

Find the following:

- posterior distribution $p(\pi|k)$ and plot it, comparing it to the prior distribution
- posterior predictive distribution $p(\tilde{y}|y_1, \dots, y_n)$
- the 95% quantile-based Bayesian confidence interval for π
- the 95% HPD interval

Further, compute:

- $P(\pi > 0.5|k)$
- For the following 2 hypotheses: $H_1 : \pi \in [0.3, 0.5], H_2 : \pi \in [0.5, 0.7]$, compute the prior and posterior odds and calculate the Bayes factor.

Exercise 2 (Solution)

Posterior distribution

Recall:

$$\begin{cases} \text{prior } \Pi & \sim \text{Beta}(a, b) \\ \text{likelihood } Y|\Pi & \sim \text{Bin}(n, \pi) \end{cases}$$

$$\Rightarrow \text{posterior } \Pi|Y = k \sim \text{Beta}(a + k, b + n - k)$$

Here:

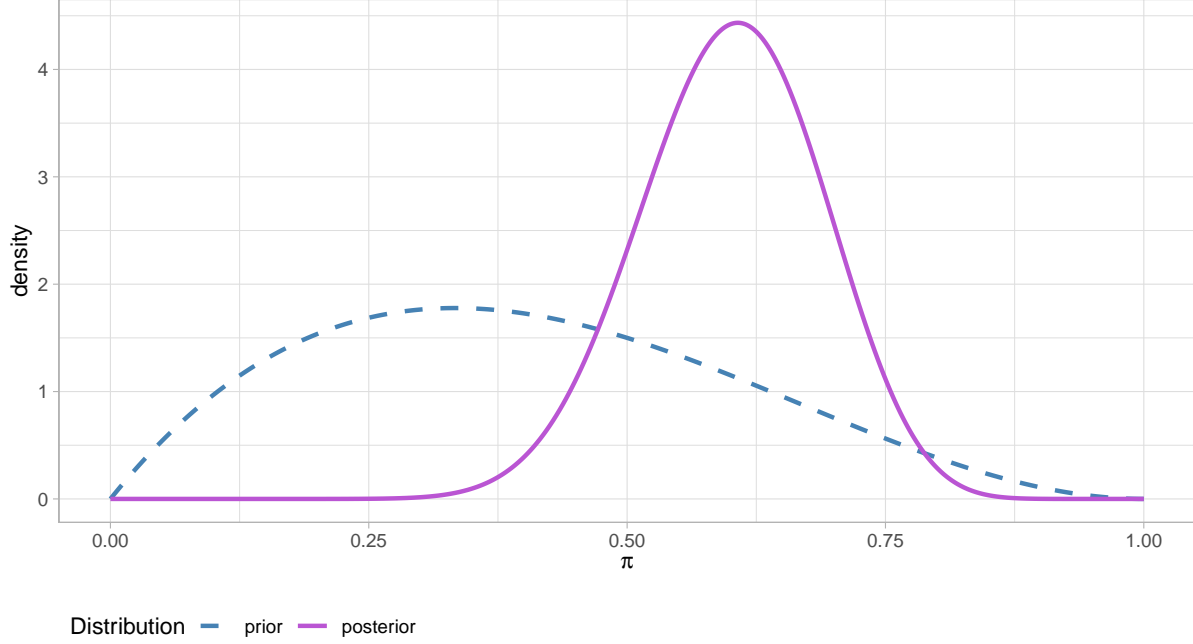
- $a = 2$
- $b = 3$
- $n = 25$
- $k = 16$

Hence, we have a $\text{Beta}(2+16, 3+9) = \text{Beta}(18, 12)$ distribution.

Alternatively, you can also see that

$$p(\pi|k=16) \propto \pi^{2-1}(1-\pi)^{3-1}\pi^{16}(1-\pi)^{25-16} = \pi^{17}(1-\pi)^{11} \\ \Rightarrow \Pi|Y=k \sim \text{Beta}(18, 12)$$

```
df<-data.frame(  
  type=factor(levels=c("prior","posterior"),c(rep("prior",1000),rep("posterior",1000))),  
  xx=rep(seq(0,1,length=1000),2)  
) %>%  
  mutate(  
    yy=case_when(  
      type=="prior"~dbeta(xx,2,3),  
      type=="posterior"~dbeta(xx,18,12))  
  )  
  
df %>%  
  ggplot(mapping=aes(x=xx,y=yy,col=type,lty=type)) +  
  geom_line(mapping=aes(y=yy),lwd=1) +  
  scale_color_manual(name="Distribution",values=c("steelblue","mediumorchid")) +  
  scale_linetype_manual(name="Distribution",values=c(2,1)) +  
  theme_light() + theme(legend.position = "bottom",legend.justification = "left") +  
  ylab("density") + xlab(expression(pi))
```



Posterior predictive distribution

$$p(\tilde{Y} = 1|y_1, \dots, y_n) = \int p(\tilde{y}, \pi|y_1, \dots, y_n) d\pi \quad (7)$$

$$= \int p(\tilde{y}|\pi) p(\pi|y_1, \dots, y_n) d\pi \quad (8)$$

$$= \int \pi \frac{\Gamma(30)}{\Gamma(18)\Gamma(12)} \pi^{17} (1-\pi)^{11} d\pi \quad (9)$$

$$= \frac{\Gamma(30)\Gamma(19)}{\Gamma(18)\Gamma(31)} \int \frac{\Gamma(31)}{\Gamma(19)\Gamma(12)} \pi^{18} (1-\pi)^{11} d\pi \quad (10)$$

$$= \frac{\Gamma(30)\Gamma(19)}{\Gamma(18)\Gamma(31)} = \frac{18}{30} = 3/5 = 0.6 \quad (11)$$

From this it follows that $p(\tilde{Y} = 0|y_1, \dots, y_n) = 1 - \frac{3}{5} = \frac{2}{5} = 0.4$.

Credible intervals

We have:

- $q_{0.025; \text{Beta}(18,12)} = 0.4226$
- $q_{0.975; \text{Beta}(18,12)} = 0.7648$

Therefore the 95% quantile based Bayesian confidence interval for π is given by $[0.42, 0.76]$

```
qbeta(c(0.025,0.975),18,12)
```

```
[1] 0.4226046 0.7647598
```

For the HPD we are going to use the function `hdi()` from the R package `HDInterval`:

```
library(HDInterval)
hdp<-hdi(qbeta,0.95,shape1=18,shape2=12)
print(hdp)
```

```
      lower      upper
0.4273464 0.7690367
attr(,"credMass")
[1] 0.95
```

From this we find that the 95% HDP interval for π is given by $[0.43, 0.77]$

Probability that π is larger than 0.5

We have

$$P(\pi > 0.5 | Y = k) = \int_{0.5}^1 p(\pi | k) d\pi \quad (12)$$

$$= \int_{0.5}^1 \frac{\Gamma(30)}{\Gamma(18)\Gamma(12)} \pi^{17} (1 - \pi)^{11} d\pi \quad (13)$$

$$= \approx 0.87 \quad (14)$$

```
integrate(dbeta,lower=0.5,upper=1,shape1=18,shape2=12)
```

```
0.8675346 with absolute error < 2.8e-09
```

Alternatively we could have worked with the posterior cdf:

```
1-pbeta(0.5,18,12)
```

```
[1] 0.8675346
```

Bayes factor

Prior odds

$$\frac{P(H_1)}{P(H_2)} = \frac{\int_{0.3}^{0.5} p_{\beta(3,2)}(\pi) d\pi}{\int_{0.5}^{0.7} p_{\beta(3,2)}(\pi) d\pi} = 1.48$$

Prosterior odds

$$\frac{P(H_1|k)}{P(H_2|k)} = \frac{\int_{0.3}^{0.5} p_{\beta(18,12)}(\pi) d\pi}{\int_{0.5}^{0.7} p_{\beta(18,12)}(\pi) d\pi} = 0.18$$

```
priorOdds<-  
  integrate(dbeta,0.3,0.5,shape1=2,shape2=3)$value /  
  integrate(dbeta,0.5,0.7,shape1=2,shape2=3)$value  
#priorOdds<-  
# (pbeta(0.5,2,3)-pbeta(0.3,2,3))/(pbeta(0.7,2,3)-pbeta(0.5,2,3))  
# alternative way of doing this  
priorOdds
```

```
[1] 1.482517
```

```
posteriorOdds<-  
  integrate(dbeta,0.3,0.5,shape1=18,shape2=12)$value /  
  integrate(dbeta,0.5,0.7,shape1=18,shape2=12)$value  
posteriorOdds
```

```
[1] 0.1789838
```

Bayes Factor

$$\text{BF} = \frac{\text{posterior odds}}{\text{prior odds}} = 0.18/1.48 = 0.12$$

```
bayesFactor<-posteriorOdds/priorOdds  
bayesFactor
```

```
[1] 0.1207296
```

Exercise 3

Suppose $\lambda \sim \text{Gamma}(5, 2)$ and $Y_1, \dots, Y_n \sim_{\text{iid}} \text{Poisson}(\lambda)$. Further suppose we observe data y_1, \dots, y_n with $n = 18, k = \sum_i y_i = 40$.

Find the following:

- posterior distribution $p(\lambda|y_1, \dots, y_n)$ and plot it, comparing it to the prior distribution
- posterior predictive distribution $p(\tilde{y}|y_1, \dots, y_n)$
- the 95% quantile-based Bayesian confidence interval for λ
- the 95% HPD interval

Further, compute:

- $P(\lambda \leq 1|y_1, \dots, y_n)$
- For the following 2 hypotheses: $H_1 : \lambda \in [0.75, 1.25], H_2 : \lambda \in [1.75, 2.25]$, compute the prior and posterior odds and calculate the Bayes factor.

Exercise 3 (Solution)

Posterior distribution

$$p(\lambda|y_1, \dots, y_n) \propto \lambda^{5-1} e^{-2\lambda} \lambda^{\sum_i y_i} e^{-n\lambda} \quad (15)$$

$$\propto \lambda^{5+40-1} e^{-(2+18)\lambda} \quad (16)$$

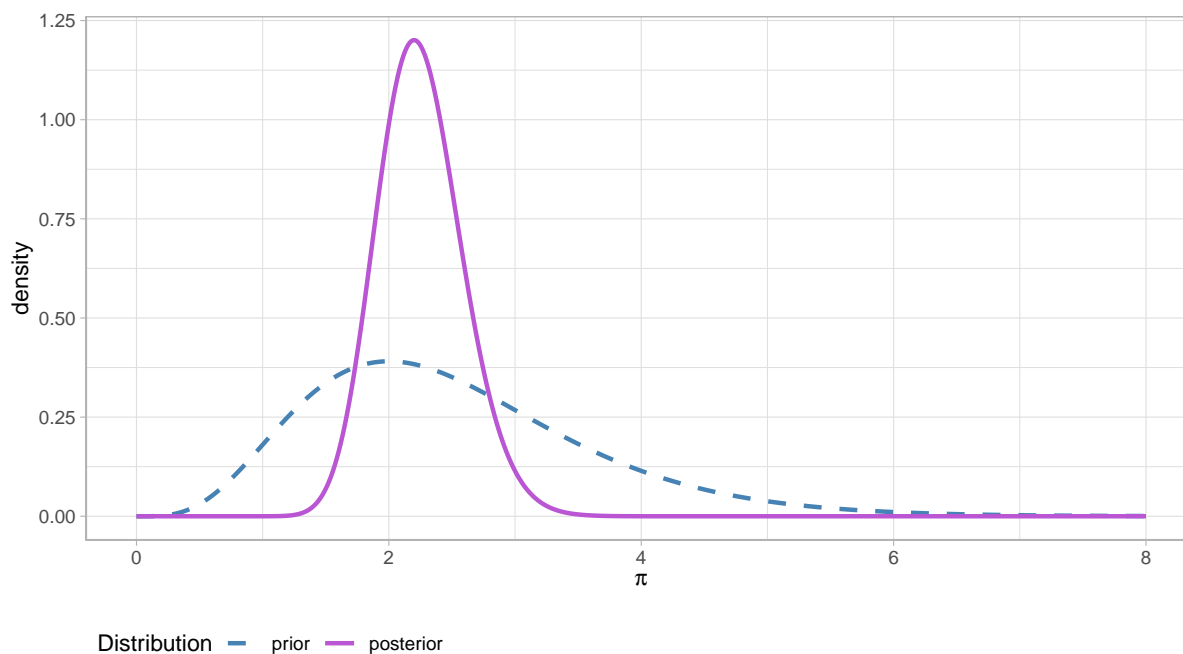
$$\Rightarrow \Lambda|y_1, \dots, y_n \sim \Gamma(45, 20) \quad (17)$$

```
df<-data.frame(
  type=factor(levels=c("prior","posterior"),c(rep("prior",1000),rep("posterior",1000))),
  xx=rep(seq(0,8,length=1000),2)
) %>%
  mutate(
    yy=case_when(
      type=="prior"~dgamma(xx,shape=5,rate=2),
      type=="posterior"~dgamma(xx,shape=45,rate=20))
  )

df %>%
  ggplot(mapping=aes(x=xx,y=yy,col=type,lty=type)) +
```



```
geom_line(mapping=aes(y=yy),lwd=1) +
scale_color_manual(name="Distribution",values=c("steelblue","mediumorchid")) +
scale_linetype_manual(name="Distribution",values=c(2,1)) +
theme_light() + theme(legend.position = "bottom",legend.justification = "left") +
ylab("density") + xlab(expression(pi))
```



Posterior predictive distribution

From Practical 1&2, Exercise 4, we know:

$$\tilde{Y}|y_1, \dots, y_n \sim \text{NegBin}(45, 20/21)$$

Credible intervals

We have:

- $q_{0.025;\Gamma(45,20)} = 1.6412$
- $q_{0.975;\Gamma(45,20)} = 2.9534$

Therefore the 95% quantile based Bayesian confidence interval for π is given by $[1.64, 2.95]$

```
qgamma(c(0.025,0.975),shape=45,rate=20)
```

```
[1] 1.641165 2.953397
```

For the HPD we are going to use the function `hdi()` from the R package `HDInterval`:

```
library(HDInterval)
hdp<-hdi(qgamma,0.95,shape=45,rate=20)
print(hdp)
```

```
      lower      upper
1.611342 2.917067
attr(,"credMass")
[1] 0.95
```

From this we find that the 95% HDP interval for λ is given by $[1.61, 2.92]$

We have

$$P(\lambda \leq 1 | y_1, \dots, y_n) = \int_0^1 p(\lambda | y_1, \dots, y_n) d\lambda \quad (18)$$

$$= \int_0^1 \frac{20^{45}}{\Gamma(45)} \lambda^{44} e^{-20\lambda} d\lambda \quad (19)$$

$$= 1.06 \cdot 10^{-6} \quad (20)$$

```
integrate(dgamma,lower=0,upper=1,shape=45,rate=20)
```

```
1.060263e-06 with absolute error < 4.4e-13
```

Bayes factor

Prior odds

$$\frac{P(H_1)}{P(H_2)} = \frac{\int_{0.75}^{1.25} p_{\gamma(5,2)}(\lambda) d\lambda}{\int_{1.75}^{2.25} p_{\Gamma(5,2)}(\lambda) d\lambda} = 0.2513$$

Prosterior odds

$$\frac{P(H_1|y_1, \dots, y_n)}{P(H_2|y_1, \dots, y_n)} = \frac{\int_{0.75}^{1.25} p_{\gamma(45,20)}(\lambda) d\lambda}{\int_{1.75}^{2.25} p_{\Gamma(45,20)}(\lambda) d\lambda} = 0.0004$$

```
priorOdds<-
  integrate(dgamma,0.75,1.25,shape=5,rate=2)$value /
  integrate(dgamma,1.25,2.25,shape=5,rate=2)$value
priorOdds
```

```
[1] 0.2513296
```

```
posteriorOdds<-
  integrate(dgamma,0.75,1.25,shape=45,rate=20)$value /
  integrate(dgamma,1.25,2.25,shape=45,rate=20)$value
posteriorOdds
```

```
[1] 0.0003851655
```

Bayes Factor

$$\text{BF} = \frac{\text{posterior odds}}{\text{prior odds}} = 0.004/0.2513 = 0.0015$$

```
bayesFactor<-posteriorOdds/priorOdds
bayesFactor
```

```
[1] 0.001532511
```