

# Lab 3

Yaqi Shi, 1003813180

2023-01-25

## Lab Exercises

### Question 1

Consider the happiness example from the lecture, with 118 out of 129 women indicating they are happy. We are interested in estimating  $\theta$ , which is the (true) proportion of women who are happy. Calculate the MLE estimate  $\hat{\theta}$  and 95% confidence interval.

#### Answer

We solve this question using a likelihood method. We have  $\theta$  being the proportion and we assume  $Y|\theta \sim \text{Bin}(n, \theta)$ . Here we have  $Y_1 = 118$  with  $n = 129$  and we want to calculate the MLE of  $\theta$ ,  $\hat{\theta}$

First we calculate the likelihood function  $L(\theta; Y)$

$$\begin{aligned} L(\theta; Y) &= \prod_{i=1}^1 f(y_i; \theta) \\ &= \binom{n}{y_1} \theta^{y_1} (1 - \theta)^{n - y_1} \end{aligned}$$

The log likelihood function is:

$$\begin{aligned} l(\theta; Y) &= \log(L(\theta; Y)) \\ &= \log \left( \binom{n}{y_1} \theta^{y_1} (1 - \theta)^{n - y_1} \right) \\ &= \log \binom{n}{y_1} + y_1 \log(\theta) + (n - y_1) \log(1 - \theta) \end{aligned}$$

Now to calculate the maximum likelihood estimators of  $\theta$ , we take the derivative with respect to  $\theta$

$$\begin{aligned} \frac{d}{d\theta} l(\theta; Y) &= \frac{d}{d\theta} \left( \log \binom{n}{y_1} + y_1 \log(\theta) + (n - y_1) \log(1 - \theta) \right) \\ &= 0 + \frac{y_1}{\theta} - \frac{n - y_1}{1 - \theta} \\ &= \frac{y_1}{\theta} - \frac{n - y_1}{1 - \theta} \end{aligned}$$

We obtain the MLE for  $\theta$  by setting  $\frac{d}{d\theta} l(\theta; Y)$  to zero, that is:

$$\begin{aligned} \frac{d}{d\theta} l(\hat{\theta}; Y) &= 0 \\ \frac{y_1}{\hat{\theta}} - \frac{n - y_1}{1 - \hat{\theta}} &= 0 \end{aligned}$$

$$\begin{aligned}\frac{y_1}{\hat{\theta}} &= \frac{n - y_1}{1 - \hat{\theta}} \\ y_1(1 - \hat{\theta}) &= (n - y_1)\hat{\theta} \\ y_1 - y_1\hat{\theta} &= n\hat{\theta} - y_1\hat{\theta} \\ \hat{\theta} &= \frac{y_1}{n}\end{aligned}$$

Thus we get, the maximum likelihood estimator of  $\theta$  is

$$\hat{\theta} = \frac{y_1}{n}$$

To calculate the 95% confidence interval, we need to calculate the following

$$\begin{aligned}I(\theta) &= -\frac{d^2}{d\theta^2}l(\theta; Y) \\ &= -\frac{d}{d\theta} \left( \frac{d}{d\theta}l(\theta; Y) \right) \\ &= -\frac{d}{d\theta} \left( \frac{y_1}{\theta} - \frac{n - y_1}{1 - \theta} \right) \\ &= - \left( -\frac{y_1}{\theta^2} - \frac{n - y_1}{(1 - \theta)^2} \right) \\ &= \frac{y_1}{\theta^2} + \frac{n - y_1}{(1 - \theta)^2}\end{aligned}$$

Then for the MLE  $\hat{\theta} = \frac{y_1}{n}$  we have:

$$\begin{aligned}I(\hat{\theta}) &= \frac{y_1}{\hat{\theta}^2} + \frac{n - y_1}{(1 - \hat{\theta})^2} \\ &= \frac{n^2}{y_1} + \frac{n^2}{n - y_1} \\ &= \frac{n * n * n}{y_1(n - y_1)} \\ &= \frac{n}{\hat{\theta}(1 - \hat{\theta})}\end{aligned}$$

Thus we have the se of this MLE estimate being

$$se(\hat{\theta}) = \sqrt{I(\hat{\theta})^{-1}} = \sqrt{\frac{\hat{\theta}(1 - \hat{\theta})}{n}}$$

Now we calculate the results in R

```
# Calculation
n <- 129
y_1 <- 118
(theta_hat <- y_1/n)

## [1] 0.9147287

se <- sqrt(theta_hat*(1-theta_hat)/n)
(ci <- c(theta_hat - 1.96*se, theta_hat + 1.96*se))

## [1] 0.8665329 0.9629244
```

Thus the MLE estimate  $\hat{\theta} = 118/129 = 0.915$  and 95% confidence interval is (0.867, 0.963)

## Question 2

Assume a Beta(1,1) prior on  $\theta$ . Calculate the posterior mean for  $\hat{\theta}$  and 95% credible interval.

### Answer

For Beta(1,1) prior, this is equivalent to uniform (0,1) distribution. Thus we have

$$\theta \sim U(0, 1)$$

$$p(\theta) = 1$$

Based on what we have in lecture, with uniform prior and binomial model we have

$$\theta|y \sim \text{Beta}(y + 1, n - y + 1)$$

Based on the beta distribution, we have

The posterior mean

```
y <- 118
n <- 129
post_mean <- (y+1)/(n+2)
post_mean
```

```
## [1] 0.9083969
```

$$E(\theta|y) = \frac{y + 1}{y + 1 + n - y + 1} = \frac{y + 1}{n + 2} = \frac{119}{131} = 0.908$$

The 95% credible interval

```
cred_int <- c(qbeta(0.025, (y+1), (n-y+1)), qbeta(0.975, (y+1), (n-y+1)))
cred_int
```

```
## [1] 0.8536434 0.9513891
```

$$\text{CredInt} = (\theta_{\frac{\alpha}{2}}, \theta_{1-\frac{\alpha}{2}}) = (\theta_{0.025}, \theta_{0.975}) = (0.854, 0.951)$$

## Question 3

Now assume a Beta(10,10) prior on  $\theta$ . What is the interpretation of this prior? Are we assuming we know more, less or the same amount of information as the prior used in Question 2?

### Answer

This means that we assume  $\theta$  follows a Beta(10,10) distribution. This means that before we obtain the data, we already know 10 successes(happy) and 10 failures(unhappy) results. It assumes 9 more success and 9 more failure compared to Q2 where we have Beta(1,1). Also, the prior assumes the proportion of happy female is more close to 0.5. This assumes we know more information as the prior used in the question 2, as the Q2 use a Beta(1,1) which is the uniform distribution.

With a Beta(10,10) being the prior distribution, we have the new posterior distribution being

$$\theta|y \sim \text{Beta}(y + 10, n - y + 10)$$

Based on the beta distribution, we have

The posterior mean

```
y <- 118
n <- 129
post_mean <- (y+10)/(n+20)
post_mean
```

```
## [1] 0.8590604
```

$$E(\theta|y) = \frac{y + 10}{y + 10 + n - y + 10} = \frac{y + 10}{n + 20} = \frac{128}{149} = 0.859$$

The 95% credible interval

```
cred_int <- c(qbeta(0.025, (y+10), (n-y+10)), qbeta(0.975, (y+10), (n-y+10)))
cred_int
```

```
## [1] 0.7990363 0.9099708
```

$$CredInt = (\theta_{\frac{\alpha}{2}}, \theta_{1-\frac{\alpha}{2}}) = (\theta_{0.025}, \theta_{0.975}) = (0.799, 0.910)$$

Compared the results in Q2, we can see it shifted towards 0.5 a bit.

## Question 4

Create a graph in ggplot which illustrates

- The likelihood (easiest option is probably to use `geom_histogram` to plot the histogram of appropriate random variables)
- The priors and posteriors in question 2 and 3 (use `stat_function` to plot these distributions)

Comment on what you observe.

### Answer

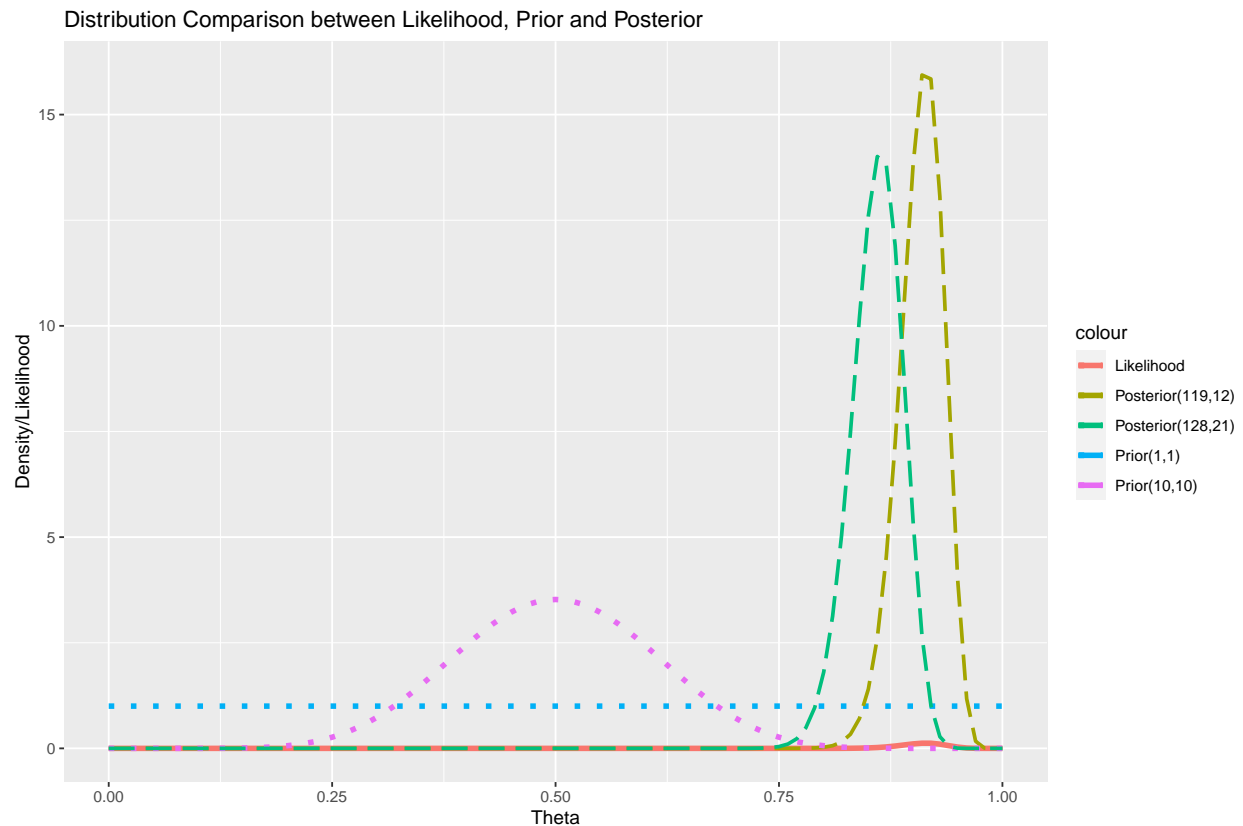
```
# Set up the basic of the plot
base <- ggplot() +
  xlim(0, 1) +
  labs(title="Distribution Comparison between Likelihood, Prior and Posterior",
        y="Density/Likelihood", x="Theta")

# Plot the likelihood
base <- base +
  geom_function(fun = dbinom, aes(colour = "Likelihood"),
               args = list(x = 118, size = 129),
               linetype = 1, lwd = 1.5)

# Plot the prior and posterior in Q2
base <- base +
  geom_function(fun = dbeta, aes(colour = "Prior(1,1)"),
               args = list(shape1 = 1, shape2 = 1),
               linetype = 3, lwd = 1.5) +
  geom_function(fun = dbeta, aes(colour = "Posterior(119,12)"),
               args = list(shape1 = 119, shape2 = 12),
               linetype = 5, lwd = 1)
```

```
# Plot the prior and posterior in Q3
base <- base +
  geom_function(fun = dbeta, aes(colour = "Prior(10,10)"),
    args = list(shape1 = 10, shape2 = 10),
    linetype = 3, lwd = 1.5) +
  geom_function(fun = dbeta, aes(colour = "Posterior(128,21)"),
    args = list(shape1 = 128, shape2 = 21),
    linetype = 5, lwd = 1)

# Display the plot
base
```



Comment:

Based on the plot, we can see that with a Beta(1,1) prior, both Bayesian method and likelihood method result in similar results. When we choose Beta(10,10) as a prior, we can see that the prior distribution are more concentrated between 0.25 and 0.75 and is symmetric around 0.5, which results in the posterior distribution shifted left towards 0.5 compared to the posterior distribution with a beta(1,1) prior. This plots indicates the influence of a informative prior compared to a uniform prior and compares the results against likelihood result.

## Question 5

(No R code required) A study is performed to estimate the effect of a simple training program on basketball free-throw shooting. A random sample of 100 college students is recruited into the study. Each student first

shoots 100 free-throws to establish a baseline success probability. Each student then takes 50 practice shots each day for a month. At the end of that time, each student takes 100 shots for a final measurement. Let  $\theta$  be the average improvement in success probability.  $\theta$  is measured as the final proportion of shots made minus the initial proportion of shots made.

Given two prior distributions for  $\theta$  (explaining each in a sentence):

- A noninformative prior, and
- A subjective/informative prior based on your best knowledge

**Answer**

If we denote the success probability before the study as  $0 \leq p_1 \leq 1$  and the success probability after the study as  $0 \leq p_2 \leq 1$ , then  $\theta = p_2 - p_1 \in [-1, 1]$  based on the value of the probability.

- A noninformative prior  $Unif(-\infty, \infty)$

This is noninformative because this does not give any more information  $\theta$ . It is a uniform distribution on an infinite interval.

- A informative prior based on your best knowledge  $Beta(1, 1)$

This is informative as we believe that after one month practice, students are expected to have improvement thus we limit the value of  $\theta$  to be non-negative. Other than this, it is hard to know how much students can improve, thus  $Beta(1,1)/Uniform(0,1)$  seems appropriate.