

# Lab 3

Qiaoyu (Terence) Liang

## 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.

Notice the likelihood is proportional to  $\theta^y(1-\theta)^{n-y}$ , so the log-likelihood is proportional to  $y \log \theta + (n-y) \log(1-\theta)$ . Thus, we have  $\frac{\partial l}{\partial \theta} = \frac{y}{\theta} - \frac{n-y}{1-\theta}$ . Solving  $\frac{\partial l}{\partial \theta} = 0$ , we get  $\hat{\theta} = \frac{y}{n} = \frac{118}{129} \approx 0.9147287$ . We can easily verify that  $\frac{\partial^2 l}{\partial \theta^2} < 0$ . Thus, the MLE estimate is indeed  $\hat{\theta} = \frac{y}{n} = \frac{118}{129} \approx 0.9147287$ .

In this case, we can calculate the Wald 95% confidence interval using  $\hat{\theta} \pm 1.96 \sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}}$ , then we have the 95% confidence interval (0.8665338, 0.9629236).

```
n <- 129
y <- 118
# MLE estimate
theta_hat <- y / n
theta_hat
```

```
## [1] 0.9147287
```

```
# 95% confidence interval
CI_L <- theta_hat - qnorm(0.975) * sqrt(theta_hat*(1-theta_hat)/n)
CI_R <- theta_hat + qnorm(0.975) * sqrt(theta_hat*(1-theta_hat)/n)
CI_L
```

```
## [1] 0.8665338
```

```
CI_R
```

```
## [1] 0.9629236
```

## Question 2

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

From the week 3 lecture slide, we know the posterior is  $\theta|y \sim \text{Beta}(y+1, n-y+1)$  when we have a a Beta(1,1) prior on  $\theta$ . Recall that the mean of Beta( $a, b$ ) is obtained as  $\frac{a}{a+b}$ . Thus, the posterior mean for  $\hat{\theta}$  is:

$$E(\theta|y) = \frac{y+1}{y+1+n-y+1} = \frac{y+1}{n+2} = \frac{118+1}{129+2} \approx 0.9083969.$$

We then can get 95% credible interval (0.8536434, 0.9513891).

```
# posterior mean
a <- y + 1
b <- n - y + 1
E <- a / (a + b)
E
```

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

```
# 95% credible interval
qbeta(c(0.025,0.975), y + 1, n - y + 1)
```

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

### 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?

A Beta(10,10) prior can be interpreted as: We observe 10 women who are happy and 10 women who are unhappy.

I think this Beta(10,10) prior is more informative than the Beta(1, 1) prior, where the latter is essentially the Uniform(0, 1) prior which gives no preference for the range of  $\theta$ . For Beta(10,10) prior, this prior indicates  $\theta$  is more likely to be near around 0.5 since the expected proportion of women aged 65+ who are happy is  $(10)/(10 + 10) = 0.5$ .

### Question 4

Create a graph in ggplot which illustrates

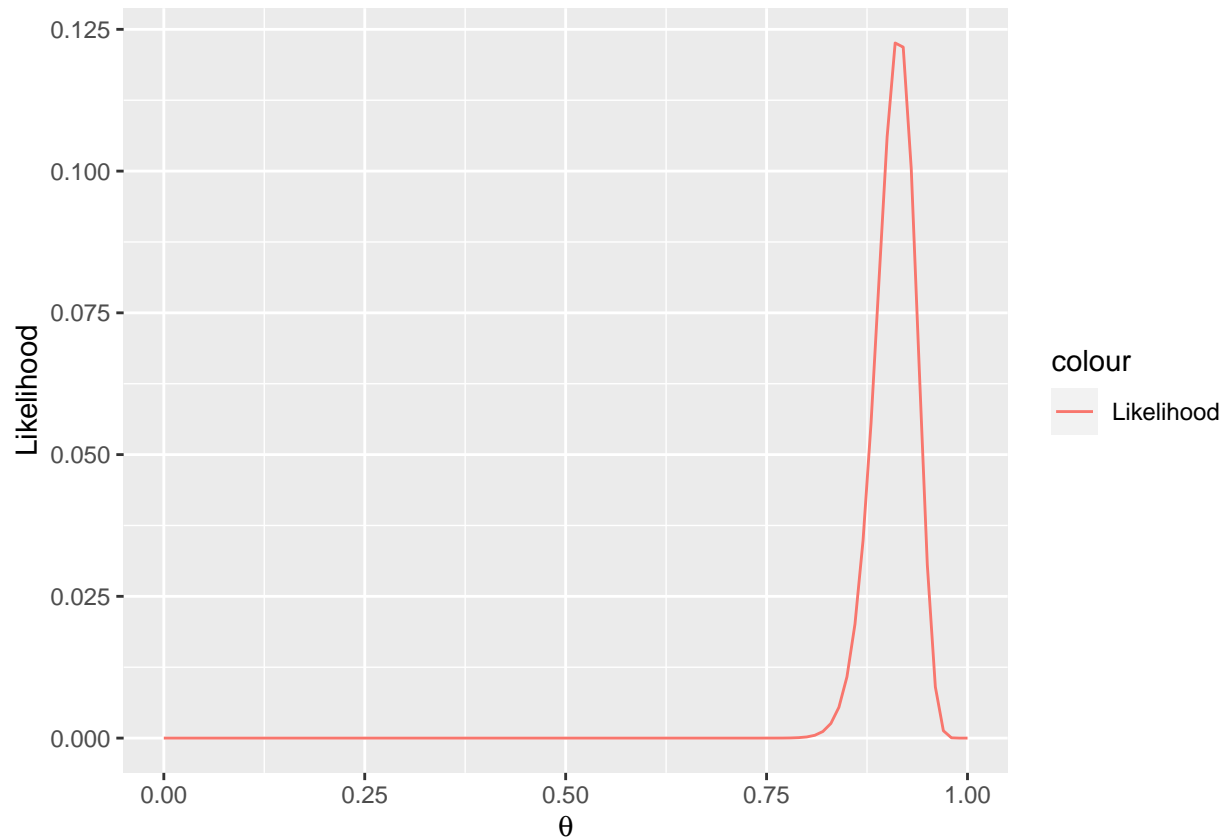
- The likelihood (easiest option is probably to use `geom_histogram` to plot the histogram of appropriate random variables)
- The priors and posteriors in question 2 and 3 (use `stat_function` to plot these distributions)

Comment on what you observe.

```
library(tidyverse)
```

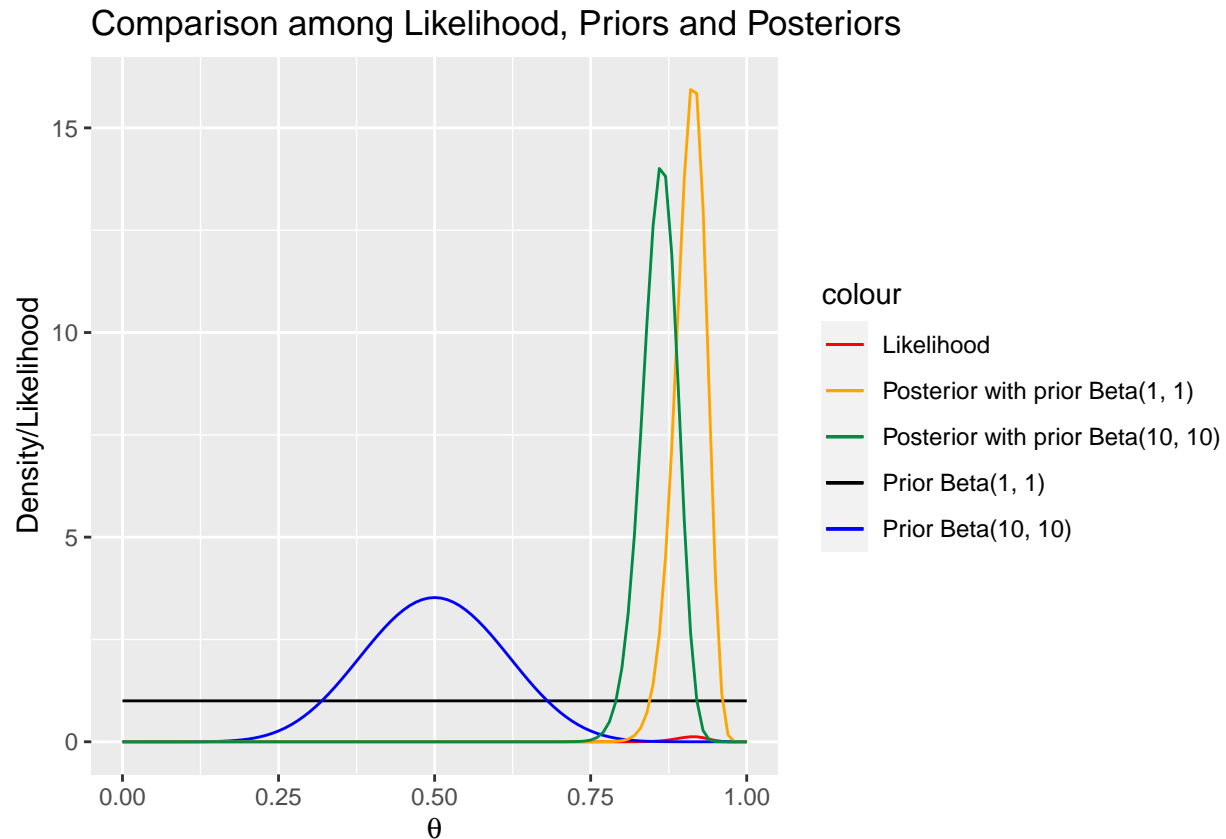
```
## -- Attaching packages ----- tidyverse 1.3.2 --
## v ggplot2 3.4.0      v purrr   1.0.1
## v tibble  3.1.8      v dplyr  1.0.10
## v tidyr   1.2.1      v stringr 1.5.0
## v readr   2.1.3      v forcats 0.5.2
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
theta <-
  ggplot() +
  xlim(0,1)+
  geom_function(aes(color="Likelihood"), fun = function(x) choose(n, y)*x^y*(1-x)^(n-y)) +
  labs(
    x = bquote(theta), y = "Likelihood")
theta
```



```
library(tidyverse)
base <-
  ggplot() +
  xlim(0,1)
cols <- c("Likelihood" = "red",
          "Prior Beta(1, 1)" = "black",
          "Prior Beta(10, 10)" = "blue",
          "Posterior with prior Beta(1, 1)" = "orange",
          "Posterior with prior Beta(10, 10)" = "springgreen4")
base +
  geom_function(aes(color="Likelihood"), fun = function(x) choose(n, y)*x^y*(1-x)^(n-y))+
  stat_function(aes(color="Prior Beta(1, 1)"), fun = dbeta,
    args = list(shape1 = 1, shape2 = 1))+
  stat_function(aes(color="Prior Beta(10, 10)"), fun = dbeta,
    args = list(shape1 = 10, shape2 = 10))+
  stat_function(aes(color="Posterior with prior Beta(1, 1)"), fun = dbeta,
    args = list(shape1 = y + 1, shape2 = n - y + 1))+
```

```
stat_function(aes(color="Posterior with prior Beta(10, 10)"), fun = dbeta,
              args = list(shape1 = y + 10, shape2 = n - y + 10))+
scale_color_manual("colour", values=cols)+
labs(
  x = bquote(theta), y = "Density/Likelihood",
  title = "Comparison among Likelihood, Priors and Posteriors")
```



Comment: From the above plot, it seems the estimates of  $\theta$  are quite similar based on the maximum likelihood estimation and Bayesian posterior mean estimation using Beta(1,1) prior. When we use Beta(10,10) prior, we can find the posterior corresponding to Beta(10,10) prior is dragged to the left compared with the likelihood and the posterior with prior Beta(1, 1) since Beta(10,10) prior concentrates more between 0.25 and 0.75. Thus, this plot somehow illustrates the different effects of an informative prior and a uniform prior. And, this plot also somehow demonstrates the relation between Bayesian estimation using a uniform prior and maximum likelihood estimation.

## 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

In this case, we notice that  $\theta$  ranges from  $-1$  to  $1$ , so a noninformative prior for  $\theta$  is a  $Unif(-1, 1)$  which covers the entire possible range of  $\theta$  and assigns the equal probability to every possible value in that range.

- A subjective/informative prior based on your best knowledge

Students are likely to improve after practicing for a month, so a subjective/informative prior can be a Normal prior  $N(0.2, 0.1)$  since an average improvement of 20% of the shots seems reasonable.