

Lab 3

Fred Haochen Song

25/01/23

Question 1

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

Let's calculate the MLE first: say $y_i = 0, 1$ where it is 1 if the i -th woman in the survey expressed that they are happy.

$$L(\theta; y) = \binom{n}{k} p^k (1-p)^{n-k}$$

Plug-in the number that we have, it will result in:

$$L(\theta; y) = \binom{129}{118} p^{118} (1-p)^{129-118}$$

and that result in a log-likelihood of:

$$l(\theta; y) = c + 118 \log(p) + 11 \log(1-p)$$

Taking derivative and setting the derivative to be 0, we solve for $\hat{\theta} = \hat{p}$:

$$\frac{118}{\hat{p}} - \frac{11}{1-\hat{p}} = 0$$

and we have:

$$\hat{\theta} = \hat{p} = \frac{118}{129} \approx 0.91$$

calculating the confidence interval we have is: $[0.87, 0.96]$

```

n <- 129
p_hat <- 118/129
sd <- sqrt(p_hat*(1-p_hat)/n)
CI_U <- p_hat + qnorm(0.975)*sd
CI_L <- p_hat - qnorm(0.975)*sd

CI_U

```

```
[1] 0.9629236
```

```
CI_L
```

```
[1] 0.8665338
```

This means: If repeated samples were taken and the 95% confidence interval was computed for each sample, 95% of the intervals would contain the population mean.

Question 2

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

Given that we have a $\text{Uniform}(0,1)$ prior, following the notes from the class, the posterior distribution follows a $\text{Beta}(\# \text{Success}+1, \# \text{Failure}+1)$ distribution

therefore given that

$$\text{Success} = 118$$

$$\text{Failure} = 11$$

we have that the posterior distribution $y|\theta \sim \text{Beta}(119, 12)$

Therefore the mean is $\frac{119}{119+12} \approx 0.91$

And a 95% credible interval can be calculated using `qbeta`:

```
qbeta(0.025, 119, 12)
```

```
[1] 0.8536434
```

```
qbeta(0.975, 119, 12)
```

```
[1] 0.9513891
```

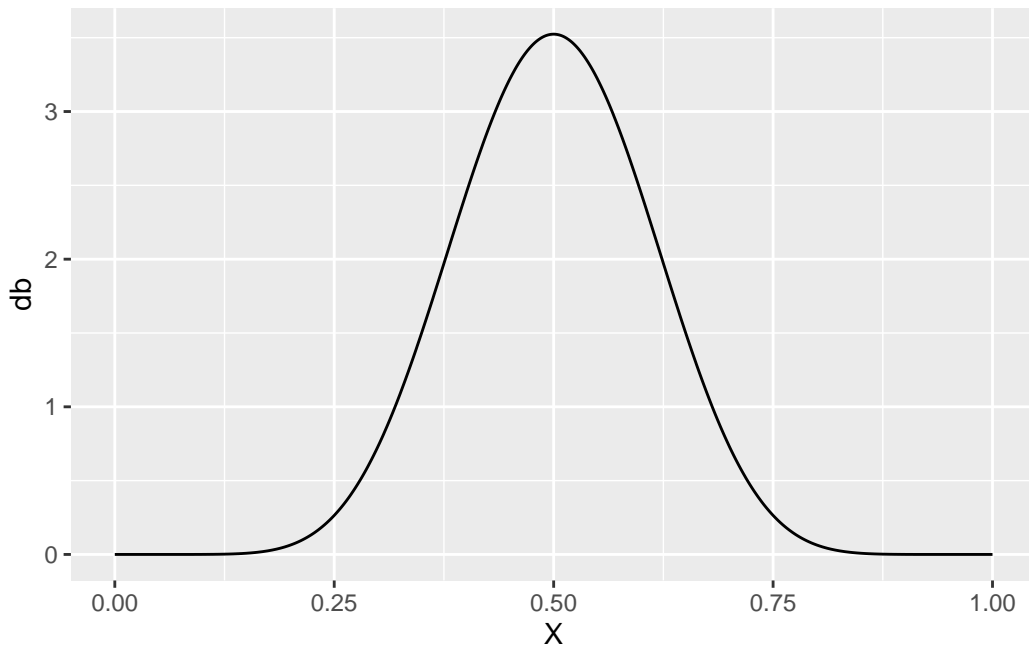
and hence conducted a 95% Credible interval is: [0.85, 0.95]. This means if I repeat the experiment over and over, the interval will contain the parameter 95% of the time.

Question 3

Now assume a Beta(10,10) prior on θ . 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?

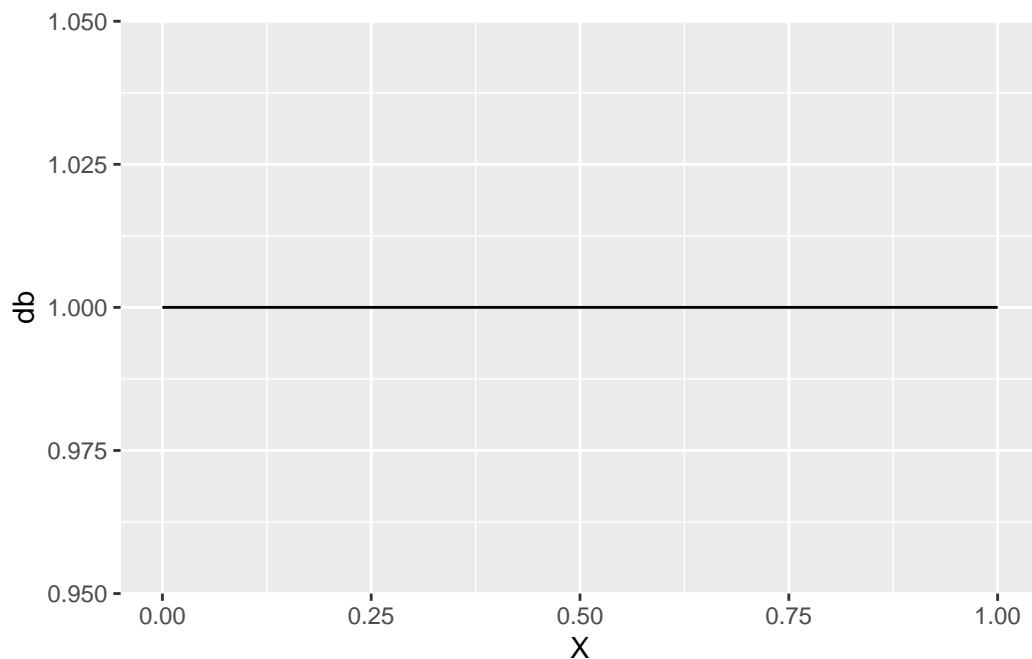
A Beta(10,10) distribution looks like the following on the interval from 0 to 1:

```
library(ggplot2)
X <- seq(0,1,length=1000)
db <- dbeta(X, 10, 10)
qplot(X, db, geom="line")
```



Compare to a Beta(1,1) which is a Uniform(0,1) distribution:

```
X <- seq(0,1,length=1000)
db <- dbeta(X, 1, 1)
qplot(X, db, geom="line")
```



This prior means that we have more information of the prior compare to assume it follows a $\text{uniform}(0,1)$ distribution. In fact, it indicates that we are given 9 successes and 9 failures as the prior information.

Now our new posterior distribution follows $\text{Beta}(129,22)$ instead of $\text{Beta}(119, 12)$

with a mean of $\frac{129}{129+22} \approx 0.85$

and a 95% credible interval of: $[0.79, 0.91]$

```
qbeta(0.025, 129,22)
```

```
[1] 0.7939893
```

```
qbeta(0.975, 129,22)
```

```
[1] 0.9057449
```

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.

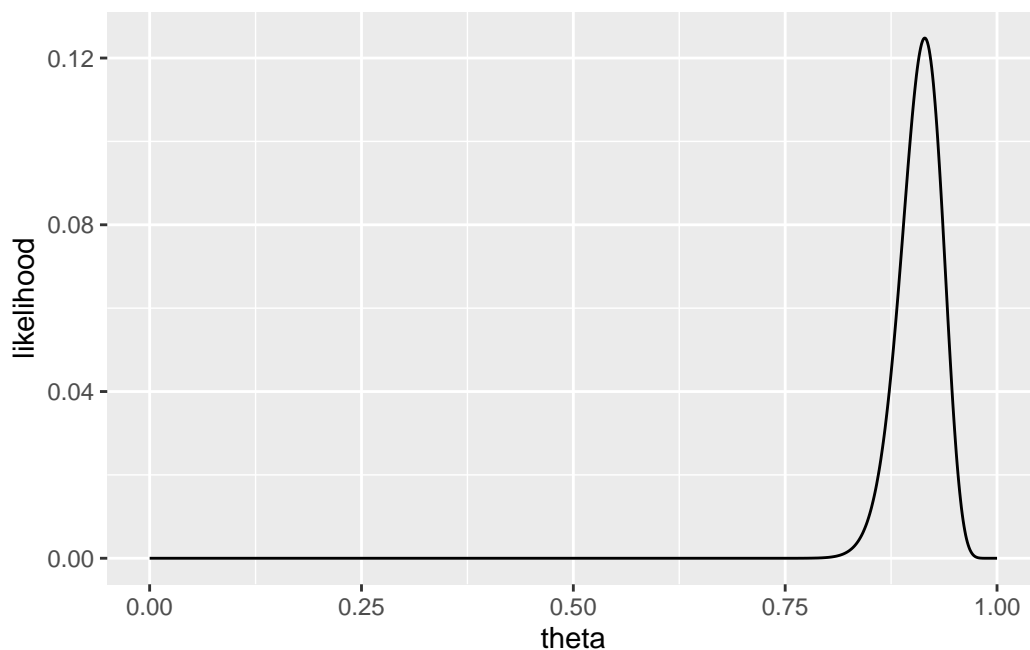
The likelihood graph:

As we see in Question 1, the likelihood function is of the format:

$$L(\theta; y) = \binom{129}{118} p^{118} (1-p)^{129-118}$$

plotting it we have:

```
theta <- seq(0,1,length=1000)
likelihood <- choose(129,118) * theta^118*(1-theta)^11
qplot(theta, likelihood, geom="line")
```

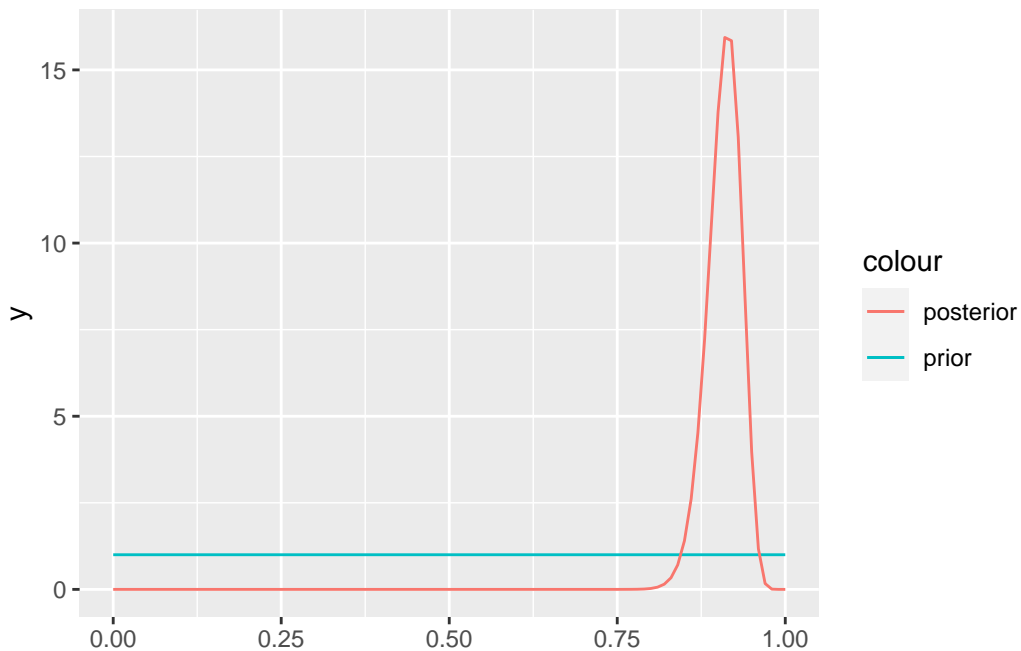


This shows that Theta will only tend to make a difference when the theta is around 0.8 to 0.97.

Priors and Posterior in Questions 2

As we have calculated, the Prior follows a Beta(1,1) and the Posterior follows a Beta(119,12) distribution. Plotting them together we have:

```
base <-  
  ggplot() +  
  xlim(0, 1)  
base +  
  stat_function(aes(colour = "prior"), fun = dbeta, args = list(shape1 = 1, shape2 = 1)) +  
  stat_function(aes(colour = "posterior"), fun = dbeta, args = list(shape1 = 119, shape2 = 12))
```

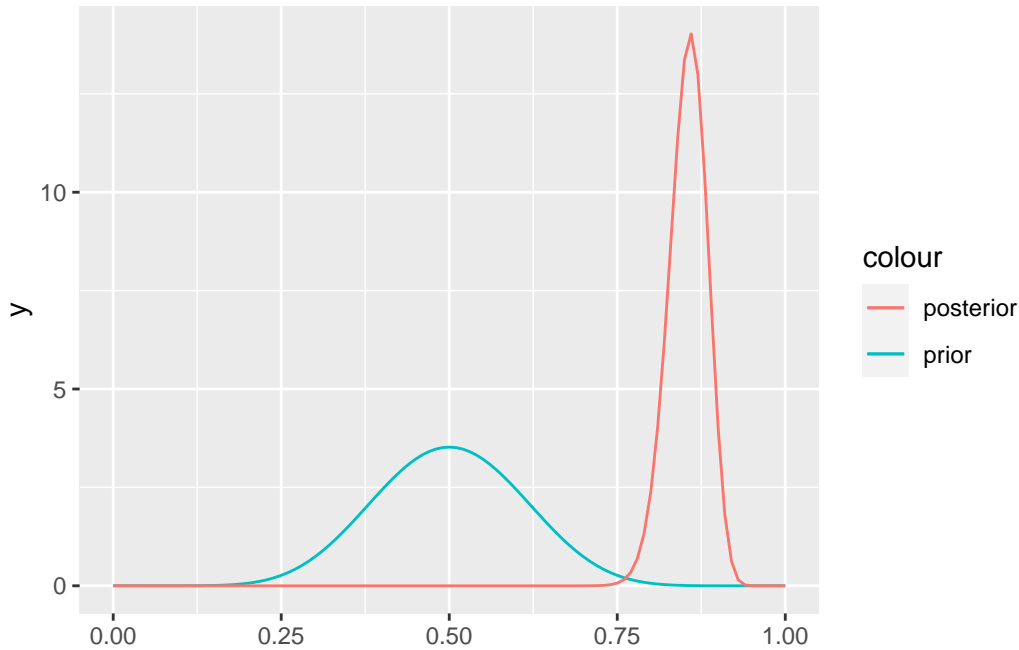


It can be seen that with an uniform assumption on the prior, we reach a similar result as from the frequentist MLE approach.

Priors and Posterior in Questions 3

As we have calculated, the Prior follows a Beta(10,10) and the Posterior follows a Beta(129,22) distribution. Plotting them together we have:

```
base <-
  ggplot() +
  xlim(0, 1)
base +
  stat_function(aes(colour = "prior"), fun = dbeta, args = list(shape1 = 10, shape2 = 10))
  stat_function(aes(colour = "posterior"), fun = dbeta, args = list(shape1 = 129, shape2 =
```



As we see from all the above three graphs, it seems like the results of posterior follows a similar distribution for different priors/frequentist, this is likely due to the fact that we have enough data to estimate the posterior distribution well enough.

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 θ be the average improvement in success probability. θ is measured as the final proportion of shots made minus the initial proportion of shots made.

Given two prior distributions for θ (explaining each in a sentence):

- A noninformative prior, and

for a non-informative prior, we can assume upon that $\int p(\theta) = 1$, which means that the integration of the underlying distribution of differences between the success probability of the initial and final states is 1.

- A subjective/informative prior based on your best knowledge

Imagine that it is going to be quite hard for θ to be very large or very small, (one means that the person improves drastically, the other one means no improvement at all), given that in mind, a normal assumption on the prior might be possible, we could assume a semi-informative prior of $N(0,1)$.