Lab 3

Ichiro Hashimoto

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.

Answer

Recall that the log-likelihood $l(y\,;\theta) = \log P(Y = y\,;\theta) = \log[\binom{n}{y}\theta^y(1-\theta)^{n-y}] = y\log\theta + (n-y)\log(1-\theta) + (\text{term independent of }\theta)$. We can solve $\frac{dl}{d\theta} = 0$ by $\hat{\theta} = \frac{y}{n}$. Here, we have $n = 129, y = 118$. So the MLE estimate $\hat{\theta}$ is estimated as

```
y <- 118
n <- 129
theta_hat <- y/n
theta_hat
```

[1] 0.91

Confidence interval is obained as $\theta \in [\frac{y-1.96\hat{\sigma}}{n}, \frac{y+1.96\hat{\sigma}}{n}]$, where $\hat{\sigma}$ is the standard deviation estimated at $\hat{\theta}$, and hence estimated as

```
sigma_hat <- sqrt(n*theta_hat*(1-theta_hat))
CI <- c((y-1.96*sigma_hat)/n, (y + 1.96*sigma_hat)/n)
CI
```

[1] 0.87 0.96

Answer

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

Solution

Observe that $p(\theta|y) \propto P(y|\theta)p(\theta) \propto \theta^y(1-\theta)^{n-y} = \theta^{y+1-1}(1-\theta)^{n-y+1-1}$. Therefore, $p(\theta|y) = \text{Beta}(y+1, n-y+1)$. Recall that the mean of $\text{Beta}(a,b)$ is obtained as $\frac{a}{a+b}$. Therefore, the posterior mean can be estimated as

```
a <- y+1
b <- n-y+1
m <- a/(a+b)
m
```

[1] 0.91

Since the posterior is $\text{Beta}(y+1, n-y+1)$ distributed, we can use 'qbeta' function to obtain its 95% credible interval as follows

```
c <- c(0.025, 0.975)
CI2 <- qbeta(c, shape1 = y+1, shape2 = n-y+1)
CI2
```

[1] 0.85 0.95

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?

Answer

θ has the peak density at $1/2$ and the density converges to zero as $\theta \to 0$ and $\to 1$. Therefore, this prior indicates θ is more likely to be located near $1/2$. Compared to the flat prior distribution in Question 2, we know more information about θ with this prior.
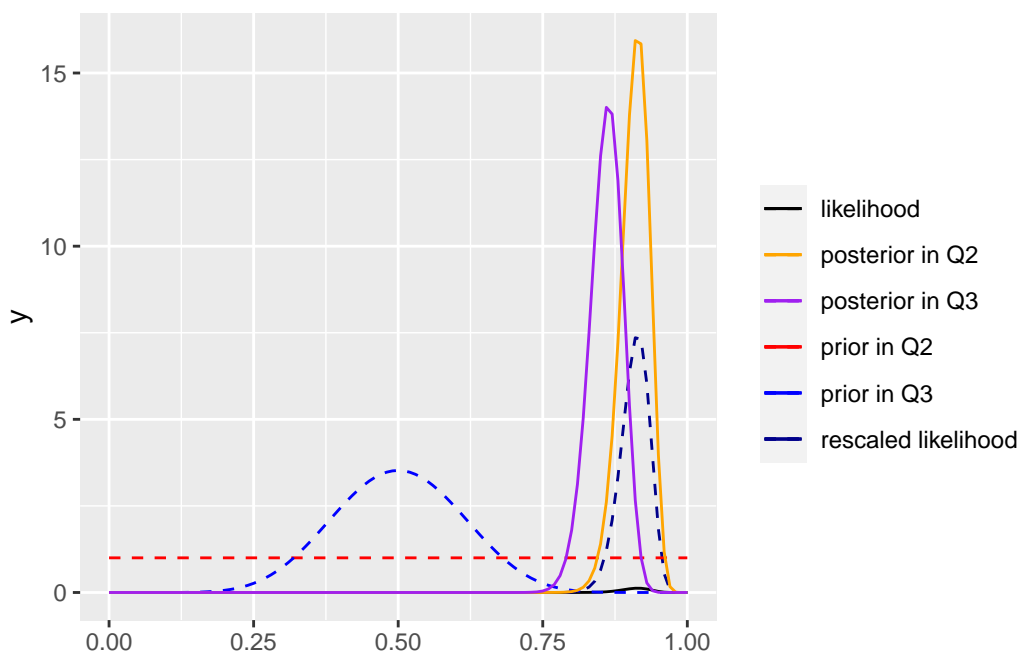
Question 4

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

The graph is provided as follows

From these plots, you can see that with the prior in Q2, which is completely flat, the corresponding posterior is almost identically distributed as the likelihood. On the other hand, you may observe that the posterior with the prior in Q3 has its density mass pulled toward the density mass of the prior in Q3.

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

Answer

Suppose the baseline success probability is establieshed as p. Then, a possible choice of a non informative prior is the uniform distribution on $[0, 1 - p]$.

- A subjective/informative prior based on your best knowledge

Answer

It is reasonable to consider that it is very less likely to observe zero improvement. It is also reasonable to consider that success rate will never get close to 1 with just a month of practice. So, a prior distribution with its peak at somewhere between 0 and p and converges to zero when $\theta \to 0$ or $1 - p$ would be more informative. One way of providing such distribution is by Beta(a, b) distribution rescaled to fit on $[0, 1 - p]$.