Week 3: Intro to Bayes

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: The likelihood $L(\theta) \sim \theta^{118}(1-\theta)^{11}$, to obtain the MLE, set the derivative of the likelihood w.r.t. θ to 0,

$\frac{dL(\theta)}{d\theta} = 118\theta^{117}(1-\theta)^{11} - 11\theta^{118}(1-\theta)^{10} = 0,$

The MLE of θ is $\hat{\theta} = \frac{118}{129} \approx 0.915$. Its 95% confidence interval is,

```r
mle <- 118/129
n <- 129
z <- 1.96
se = sqrt(mle*(1-mle)/n)
c(mle-z*se, mle+z*se)
```

```
[1] 0.8665329 0.9629244
```

Question 2

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

Answer: The beta distribution is a conjugate prior of the binomial likelihood. The posterior distribution given a Beta(1,1) prior is, $p(\theta|y) \sim \text{Beta}(y+1, n-y+1)$. The posterior mean is,

$\bar{\theta} = E_{\theta \sim p(\theta|y)}(\theta) = \frac{y+1}{n+2}$

According to Q1, $y = 118$, $n = 129$, so,

$\bar{\theta} = \frac{119}{131} \approx 0.908$

The 95% credible interval for the posterior mean is,

```
p <- c(0.05, 0.95)
qbeta(p = p, shape1 = 119, shape2 = 12)
```

```
[1] 0.8638292 0.9458711
```

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: The prior Beta(10, 10) means that there are 10 successes and 10 failures. Beta(10, 10) assumes that we know more than having a Beta(1, 1) prior, which corresponds to a uniform distribution, as in Q2. Beta(10, 10) has a unimodal shape, which concentrates the probability on its mode.

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 following plot shows the likelihood, the priors, and the corresponding posteriors. The posterior from the Beta(1,1) prior is peaked closer to likelihood peak (the MLE value), while the posterior from the Beta(10, 10) prior is peaked further from the MLE. This also results in a lower posterior mean from the Beta(10, 10) prior than that from the Beta(1, 1) prior.

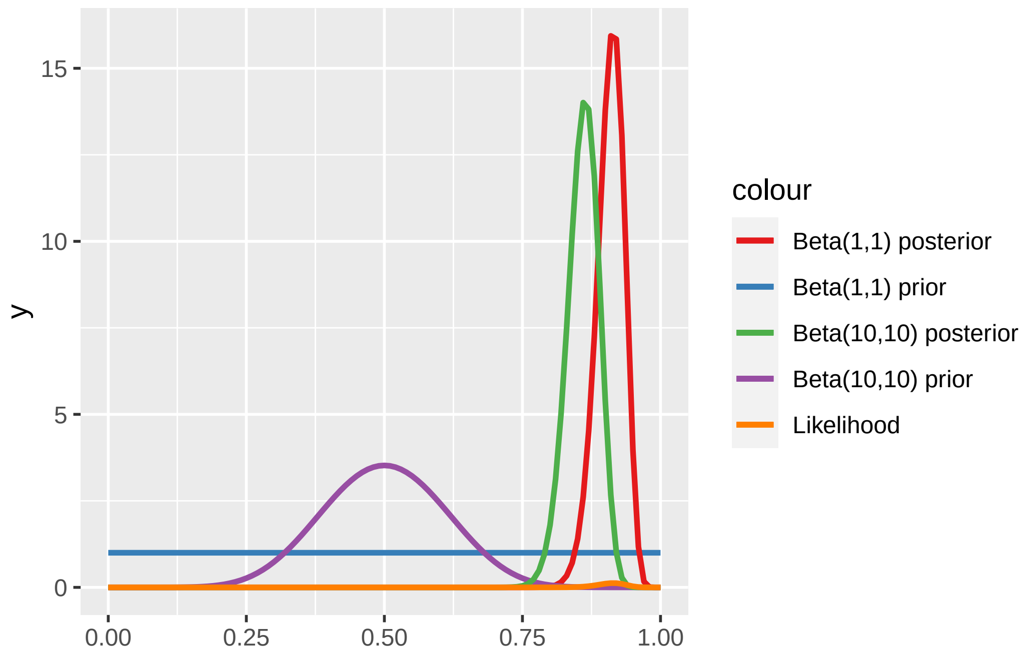
```
library(ggplot2)

n <- 129
y <- 118

data = data.frame(theta=seq(0, 1, by=0.001))
likelihood = function(theta){
```

```
  choose(n,y)*(theta^y)*(1-theta)^(n-y)
}

ggplot(data=data) + scale_colour_brewer(palette = "Set1") +
  stat_function(fun=dbeta, args=c(1,1),
                aes(color="Beta(1,1) prior"), size=1) +
  stat_function(fun=dbeta, args=c(y+1,n-y+1),
                aes(color="Beta(1,1) posterior"), size=1) +
  stat_function(fun=dbeta, args=c(10,10),
                aes(color="Beta(10,10) prior"), size=1) +
  stat_function(fun=dbeta, args=c(y+10,n-y+10),
                aes(color="Beta(10,10) posterior"), size=1) +
  stat_function(fun=likelihood, aes(color="Likelihood"), size=1)
```



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
- A subjective/informative prior based on your best knowledge

Answer: A noninformative prior in this case is a uniform distribution over θ, such as U(0, 1) for positive improvements, or U(-1, 1), if negative progress from the baseline can also occur.

An informative prior assuming a positive progress of the free-throwing score could be a normal distribution with a positive mean in (0, 1), such as N(0.5, 0.5).