Chapter 2 Single-parameter Models

Hendra Bunyamin

12/20/2020

Let's import ggplot2 for plotting.

```
library(ggplot2)
theme_set(theme_minimal())
```

Probability of a girl birth given placenta previa

437 girls and 543 boys have been observed. Calculate and plot the posterior distribution of the proportion of girls θ, using uniform prior on θ.

Analysis using a uniform prior distribution

Under a uniform prior distribution for the probability of a girl birth, the posterior distribution is Beta(438, 544).
`seq` creates evenly spaced values as follows:

```
df1 <- data.frame(theta = seq(0.375, 0.525, 0.001))
a <- 438
b <- 544
head(df1)
```

```
##   theta
## 1 0.375
## 2 0.376
## 3 0.377
## 4 0.378
## 5 0.379
## 6 0.380
```

As we read page 35 from the book, the posterior density for θ is

$$\begin{aligned}\Pr(\theta \mid y) &\propto \theta^y(1-\theta)^{n-y}\theta^{\alpha-1}(1-\theta)^{\beta-1} \\ &= \theta^{y+\alpha-1}(1-\theta)^{n-y+\beta-1} \\ &= \text{Beta}(\theta \mid \alpha+y, \beta+n-y).\end{aligned}$$

Specifically, `dbeta` computes the posterior density.

```
df1$p <- dbeta(df1$theta, a, b)
head(df1)
```

```
##   theta           p
## 1 0.375 0.0008507679
## 2 0.376 0.0011418738
## 3 0.377 0.0015257278
## 4 0.378 0.0020295190
```

```
## 5 0.379 0.0026876416
## 6 0.380 0.0035433725
```

Next, we compute 95% **central interval**.
The `seq` creates evenly spaced values from 2.5% quantile to 97.5% quantile (i.e., 95% central interval).
The `qbeta` computes the value for a given quantile given parameters `a` and `b`.

```
df2 <- data.frame(theta = seq(qbeta(0.025, a, b), qbeta(0.975, a, b), length.out = 100))
```

Subsequently, we compute the *posterior density*.

```
df2$p <- dbeta(df2$theta, a, b)
```

We also plot the posterior (Beta$(438, 544)$) and 48.8% line for population average.