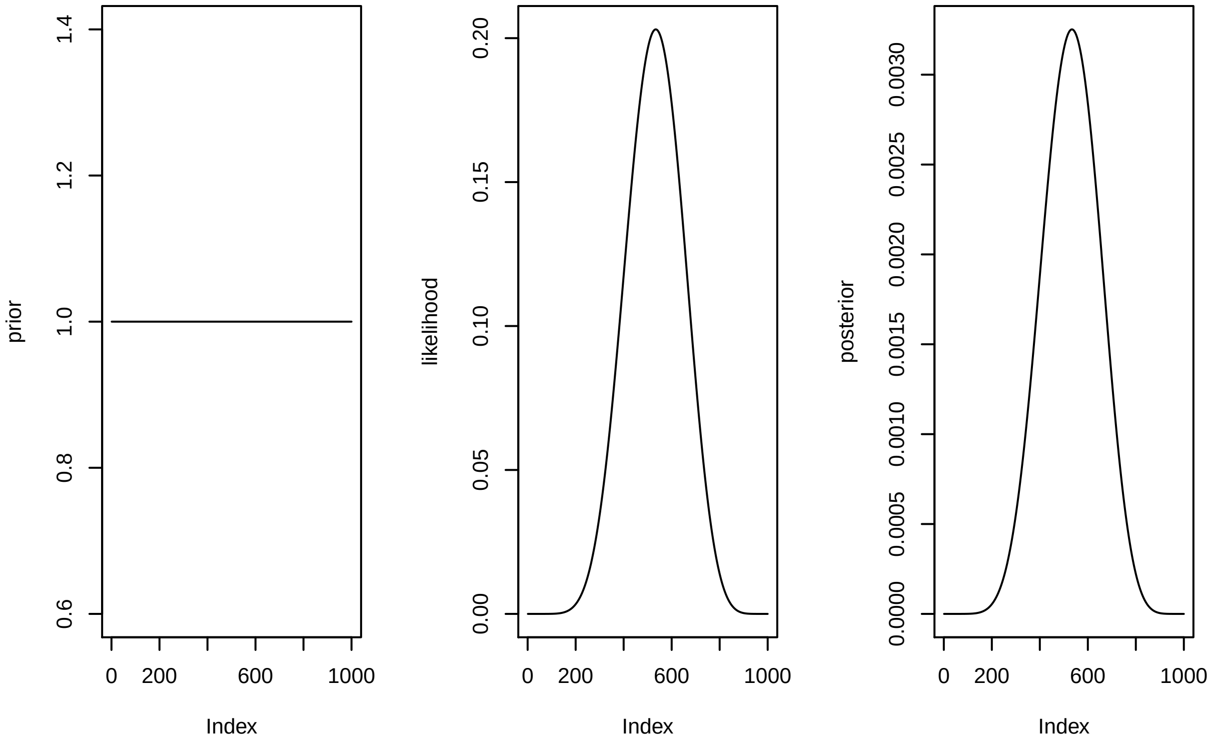
Statistical Rethinking Week 1

Jeff Shen

26 July, 2020

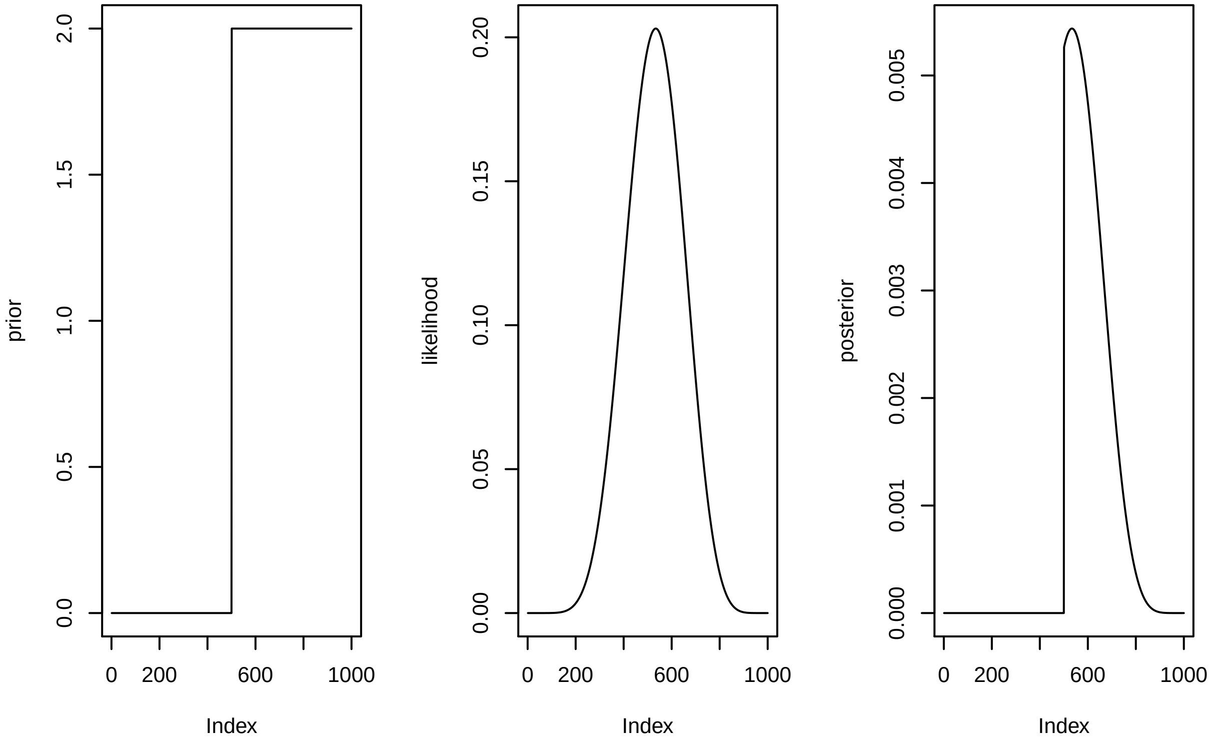
Problem 1

```r
prior <- rep(1, 1000)
grid <- seq(0, 1, length.out=1000)
likelihood <- dbinom(8, size=15, prob=grid)
posterior <- prior * likelihood
posterior <- posterior / sum(posterior)
par(mfrow=c(1, 3))
plot(prior, type = 'l')
plot(likelihood, type = 'l')
plot(posterior, type = 'l')
```



Problem 2

```
prior <- c(rep(0, 500), rep(2, 500))
grid <- seq(0, 1, length.out=1000)
likelihood <- dbinom(8, size=15, prob=grid)
posterior <- prior * likelihood
posterior <- posterior / sum(posterior)
par(mfrow=c(1, 3))
plot(prior, type = 'l')
plot(likelihood, type = 'l')
plot(posterior, type = 'l')
```



Problem 3

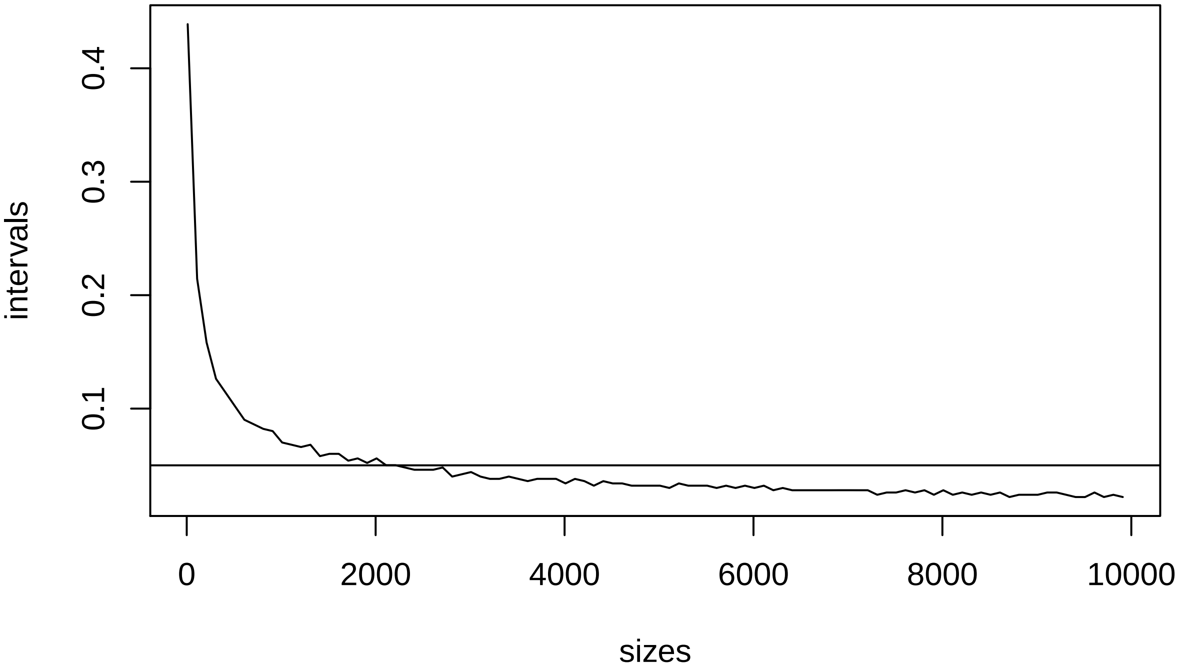
Assume the same prior as in Problem 2. Then we can generate a posterior for all the different sample sizes and calculate the 99% percentile interval for each. We would expect that as you use larger sample sizes, the width of the interval goes down as you are more and more sure of your answer.

```r
library(rethinking);
sizes <- seq(10, 1e4, 100)

gen_intervals <- function(N) {
    prior <- c(rep(0, 250), rep(2, 250))
    grid <- seq(0, 1, length.out=500)
    likelihood <- dbinom(0.7 * N, size=N, prob=grid)
    posterior <- prior * likelihood
    posterior <- posterior / sum(posterior)
    samples <- sample(grid, prob=posterior, size=1e3, replace=TRUE)
    interval <- PI(samples, prob=0.99)
    return(interval[2] - interval[1])
}

intervals <- sapply(sizes, gen_intervals)

plot(sizes, intervals, type='l')
abline(h = 0.05)
```



```r
minidx <- as.numeric(which.min((intervals - 0.05)^2))
print(sizes[minidx])
```

```
## [1] 2110
```