Statistical Rethinking Exercises - Chapter 2

Anthony Chau

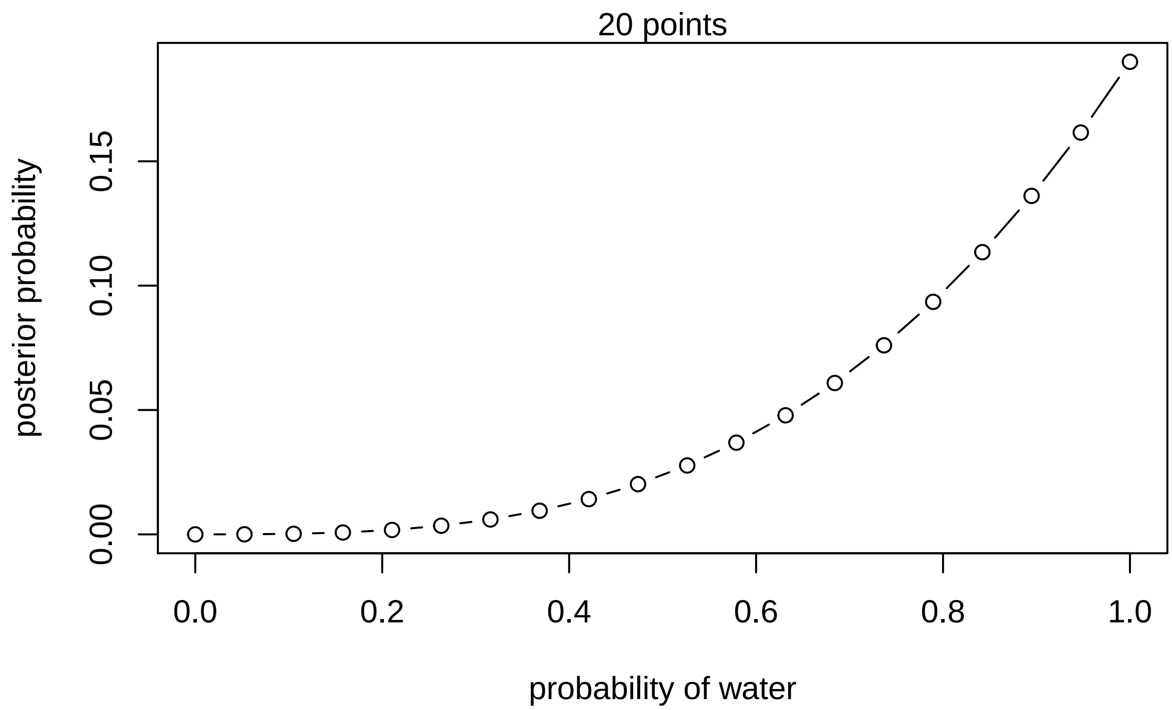
November 25, 2018

Contents

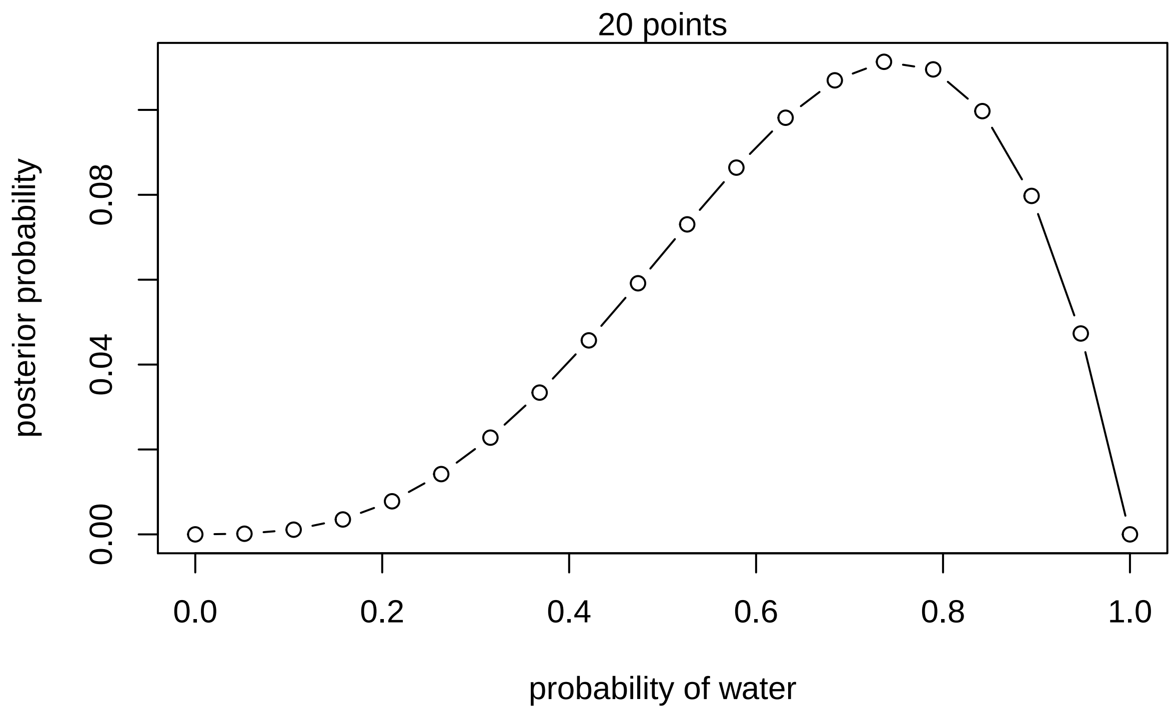
Exercise 2M1

We notice the posterior distribution is only one-tailed and left-skewed when the number of successes equals the number of trials. For the other two cases, the posterior distribution is shifted to the right if the probability of success is higher for the set of observations. For example, in case 2 ($3/4 = 0.75$) and in case 3 ($4/7 = 4/7$). The observed probability of success is higher in case 2, hence the posterior distribution is shifted more to the right.

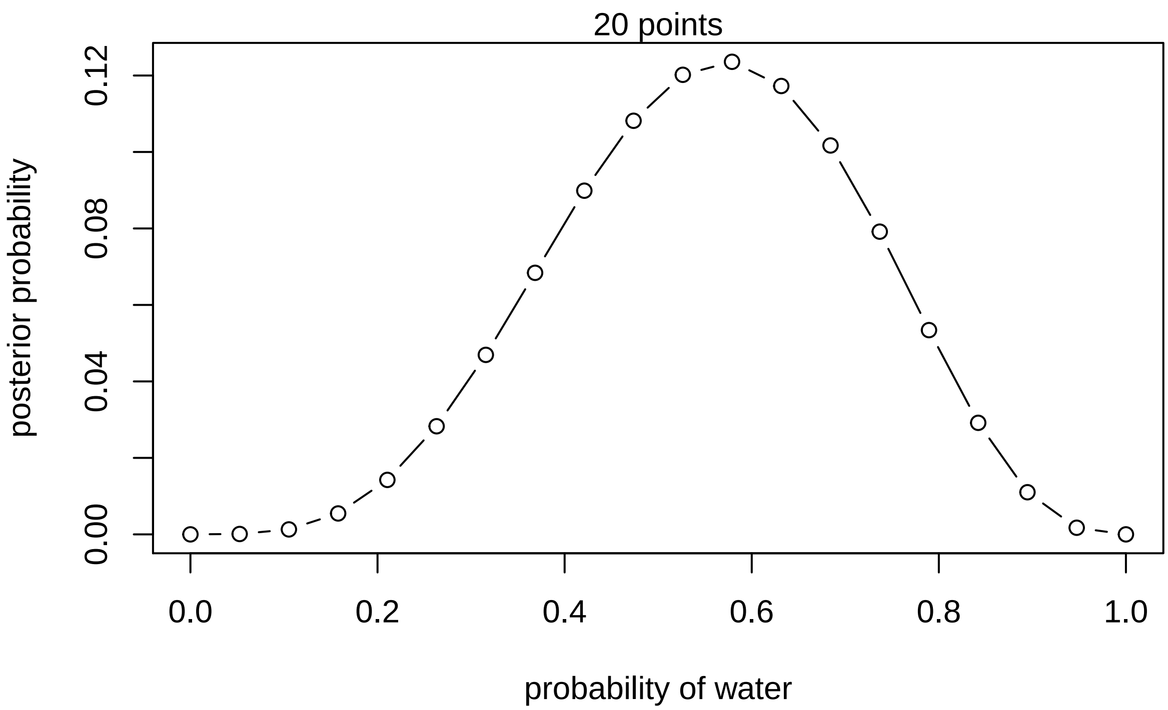
```
# grid approximation with uniform prior

gridApprox <- function(x, n, iterLength){
  # define grid
  p_grid <- seq(from=0 , to=1 , length.out= iterLength)

  # define prior
  prior <- rep(1 , iterLength)

  # compute likelihood at each value in grid
  likelihood <- dbinom(x , size=n , prob=p_grid)

  # compute product of likelihood and prior
  unstd.posterior <- likelihood * prior

  # standardize the posterior, so it sums to 1
  posterior <- unstd.posterior / sum(unstd.posterior)

  plot( p_grid , posterior , type="b" ,
      xlab="probability of water" , ylab="posterior probability" )
  mtext(paste(iterLength, "points"))
}

# Case 1: 3 successes with 3 trials
gridApprox(3, 3, 20)
```




```r
# Case 2: 3 successes with 4 trials
gridApprox(3, 4, 20)
```




```
# Case 3: 3 successes with 7 trials
gridApprox(4, 7, 20)
```

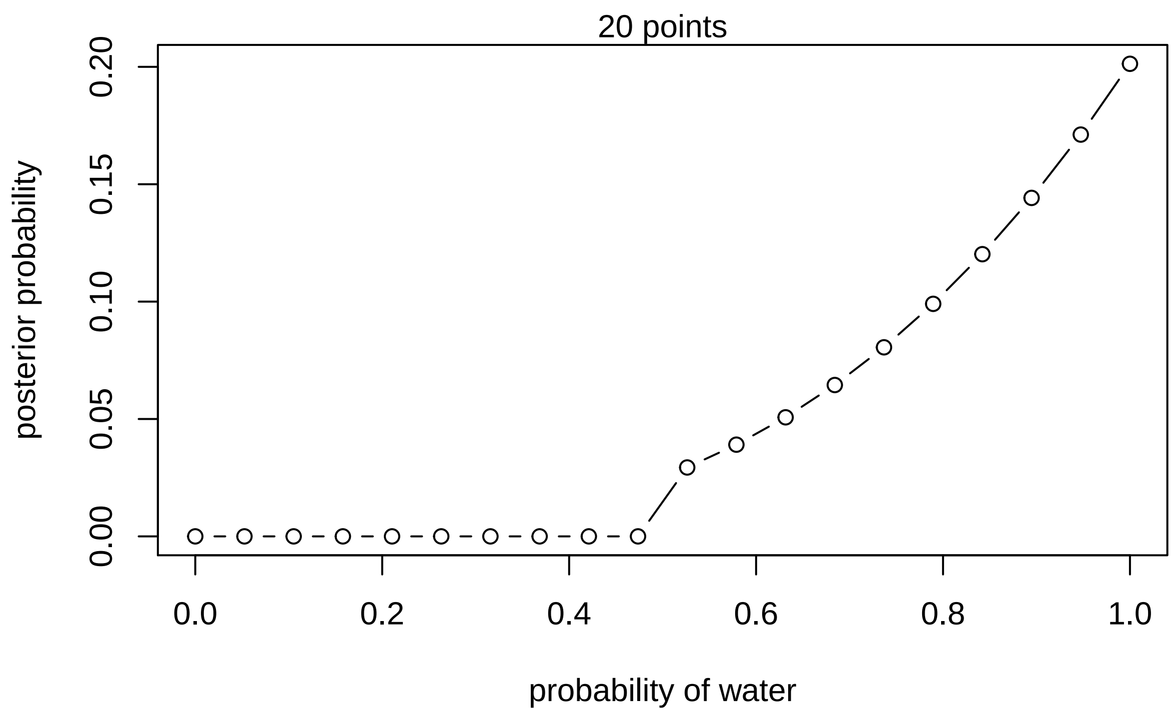


Exercise 2M2

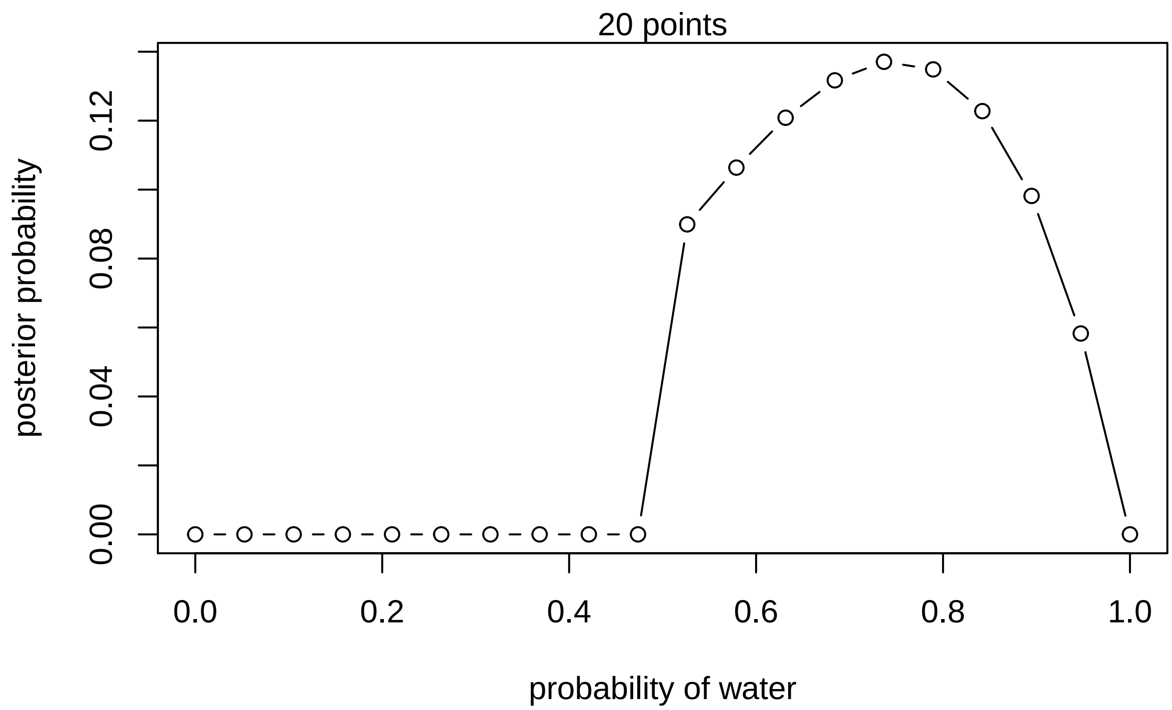
The problem above is modified by changing the specification of the prior. The stepwise prior results in a truncated prior at $p = 0.5$ for all cases.

```
gridApprox2 <- function(x, n, iterLength){
  # define grid
  p_grid <- seq(from=0 , to=1 , length.out= iterLength)

  # prior is 0 when p < 0.5 and equal to a positive constant otherwise.
  prior2 <- ifelse(p_grid < 0.5, 0, 5)

  # compute likelihood at each value in grid
  likelihood <- dbinom(x , size=n , prob=p_grid)

  # compute product of likelihood and prior
  unstd.posterior <- likelihood * prior2

  # standardize the posterior, so it sums to 1
  posterior <- unstd.posterior / sum(unstd.posterior)

  plot( p_grid , posterior , type="b" ,
      xlab="probability of water" , ylab="posterior probability" )
  mtext(paste(iterLength, "points"))
```

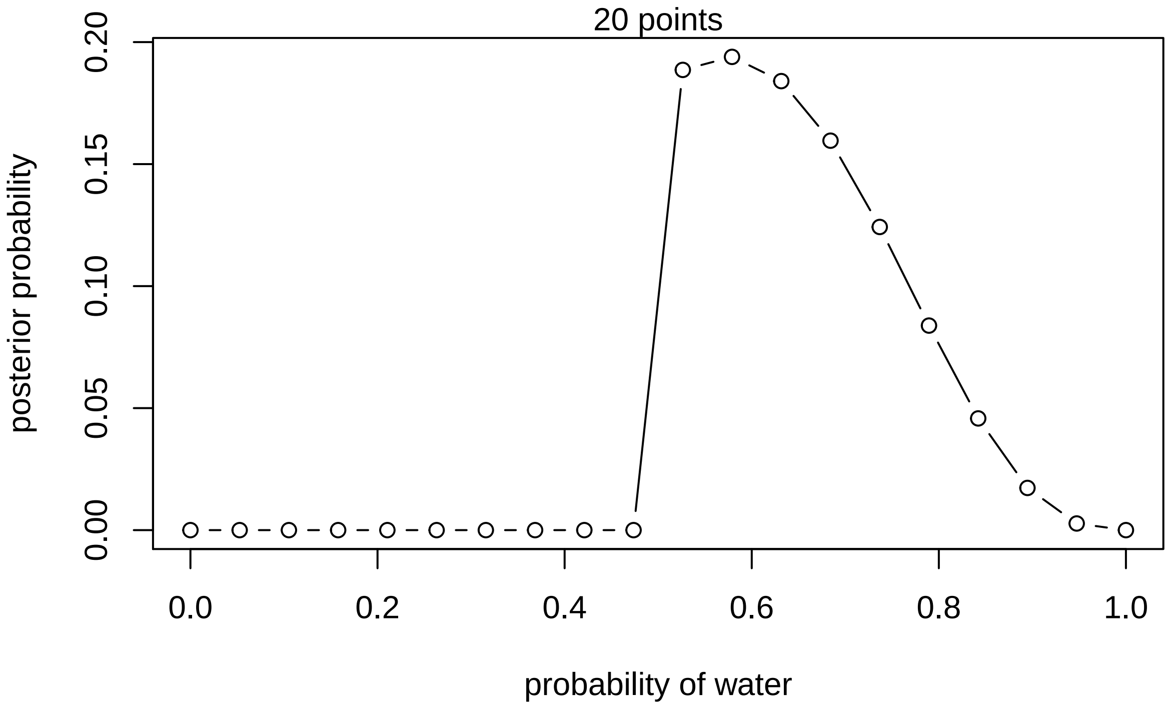
```r
}

# Case 1: 3 successes with 3 trials
gridApprox2(3, 3, 20)
```




```r
# Case 2: 3 successes with 4 trials
gridApprox2(3, 4, 20)
```




```
# Case 3: 3 successes with 7 trials
gridApprox2(4, 7, 20)
```



Exercise 2M3

This problem is an application of Bayes Rule.

$$\begin{aligned}
\Pr(Land|Earth) &= 0.3 && \text{(Likelihood)} \\
\Pr(Earth) &= 0.5 && \text{(Prior)} \\
\Pr(Land) &= Pr(Land|Earth) \cdot Pr(Earth) + Pr(Land|Mars) \cdot Pr(Mars) && \text{(Average Likelihood)} \\
&= 0.3 \cdot 0.5 + 1 \cdot 0.5 \\
&= 0.15 + 0.5 \\
&= 0.65
\end{aligned}$$

Hence, we can apply Bayes Rule to compute $Pr(Earth|Land)$.

$$\begin{aligned}
\Pr(Earth|Land) &= \frac{Pr(Land|Earth) \cdot Pr(Earth)}{Pr(Land)} \\
&= \frac{0.3 \cdot 0.5}{0.65} \\
&= 0.23
\end{aligned}$$