Bayesian Inference

2022-01-05

The Earth problem

- 70% water, 30% land
- A randomized Montecarlo process: impact of celestial bodys, but we could use an inflatable globe and count the finger is pointed.

Samples

Imagine we had this sample: L W L L W W W L W W

- How should we use the sample?
- How to produce a summary?
- How to represent uncertainty?

Bayesian data analysis: (very modest approach)

- For each possible explanation of the data, count all the ways data can happen.
- Explanations with more ways to produce the data are more plausible.

```
samples <- c('L', 'W', 'L', 'L', 'W', 'W', 'W', 'L', 'W', 'W')
```

Garden of Forking Data

- A bag with 4 marbles of color **L** and **W**
- Possible contents:
  1. L L L L
  2. W L L L
  3. W W L L
  4. W W W L
  5. W W W W Observe, SAMPLED WITH REPLACEMENT:
  - W L W

Assume: (2) W L L L, how many ways to observe W L W? (Figure 2.2)

3 ways to see WLW in (2)

Unglamorous basis of applied probability:

> Things that can happen more ways are more plausible.

```
        | W L W | p    | plausibility
L L L L | 0     | 0    |   0
W L L L | 3     | 0.25 |   0.15
W W L L | 8     | 0.5  |   0.40
W W W L | 9     | 0.75 |   0.45
W W W W | 0     | 1    |   0
```

```
ways <- c(0,3,8,9,0)
ways
```

```
## [1] 0 3 8 9 0
```

```
ways / sum(ways)
```

```
## [1] 0.00 0.15 0.40 0.45 0.00
```

With Bayesian updating:

```
           | W L W | p    | plausibility  | W || W L W W
   L L L L | 0     | 0    |   0           | 0 || 0
   W L L L | 3     | 0.25 |   0.15        | 1 || 3
   W W L L | 8     | 0.5  |   0.40        | 2 || 16
   W W W L | 9     | 0.75 |   0.45        | 3 || 27
   W W W W | 0     | 1    |   0           | 4 || 0
```

- Rules:
  1. State a causal model for how the observations arise, given each possible explanation
  2. Count ways data could arise for each explanation
  3. Relative plausibility is relative value of (2)

```
new_ways  <- ways * c(0, 1, 2, 3, 4)
new_ways
```

```
## [1]  0  3 16 27  0
```

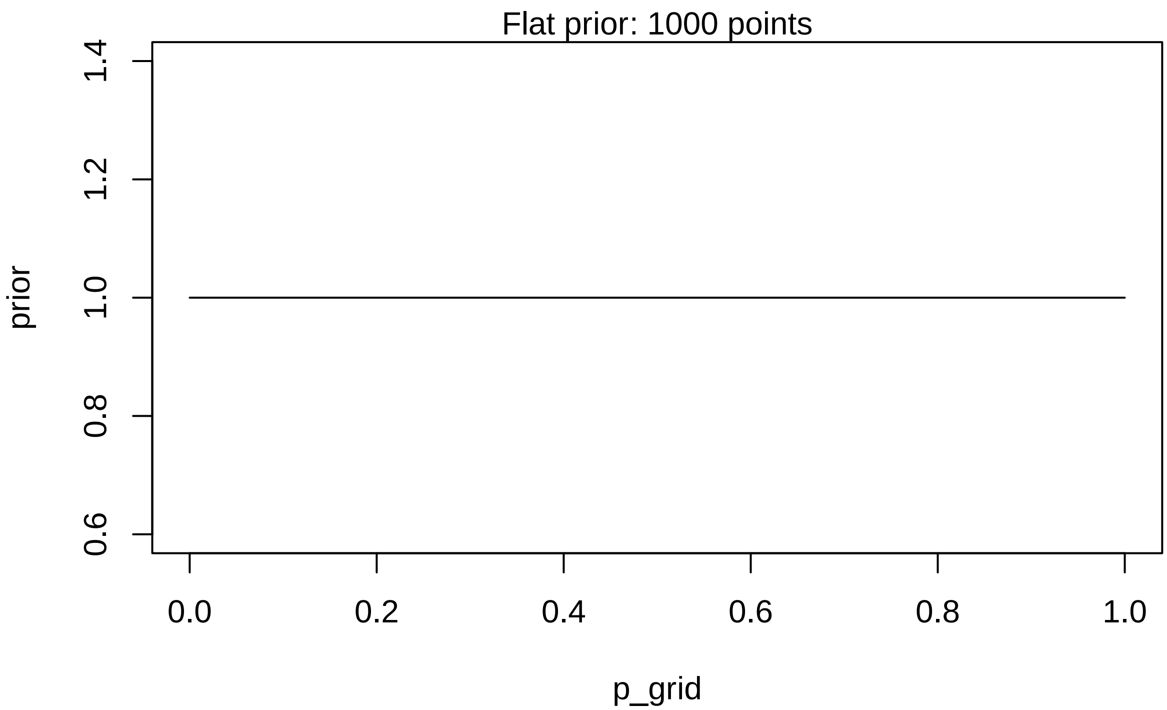
```
new_ways / sum(new_ways)
```

```
## [1] 0.00000000 0.06521739 0.34782609 0.58695652 0.00000000
```

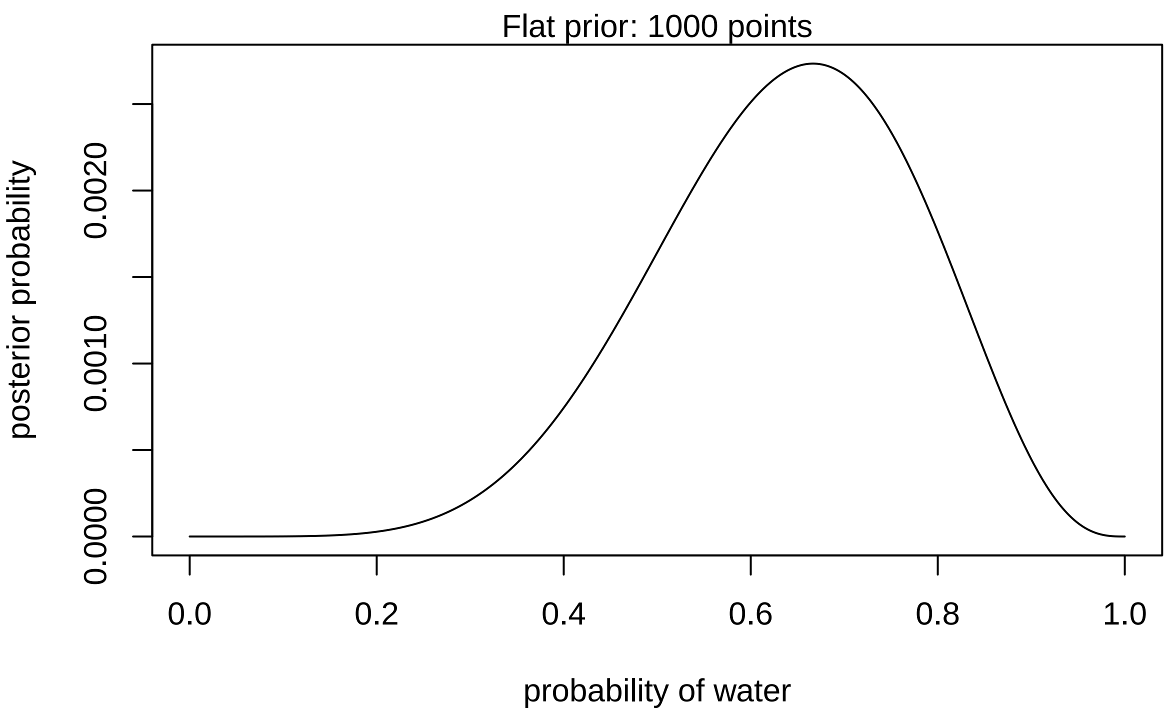
Globe tossing

```
len = 1000
p_grid  <- seq(from=0, to=1, length.out=len)
prior <- rep(1,len)
likelihood <- dbinom(6, size=9, prob=p_grid)
unstd.posterior  <- likelihood * prior
posterior  <- unstd.posterior / sum(unstd.posterior)

plot(p_grid , prior , type="l")
mtext(stringr::str_c("Flat prior: ", len, " points"))
```



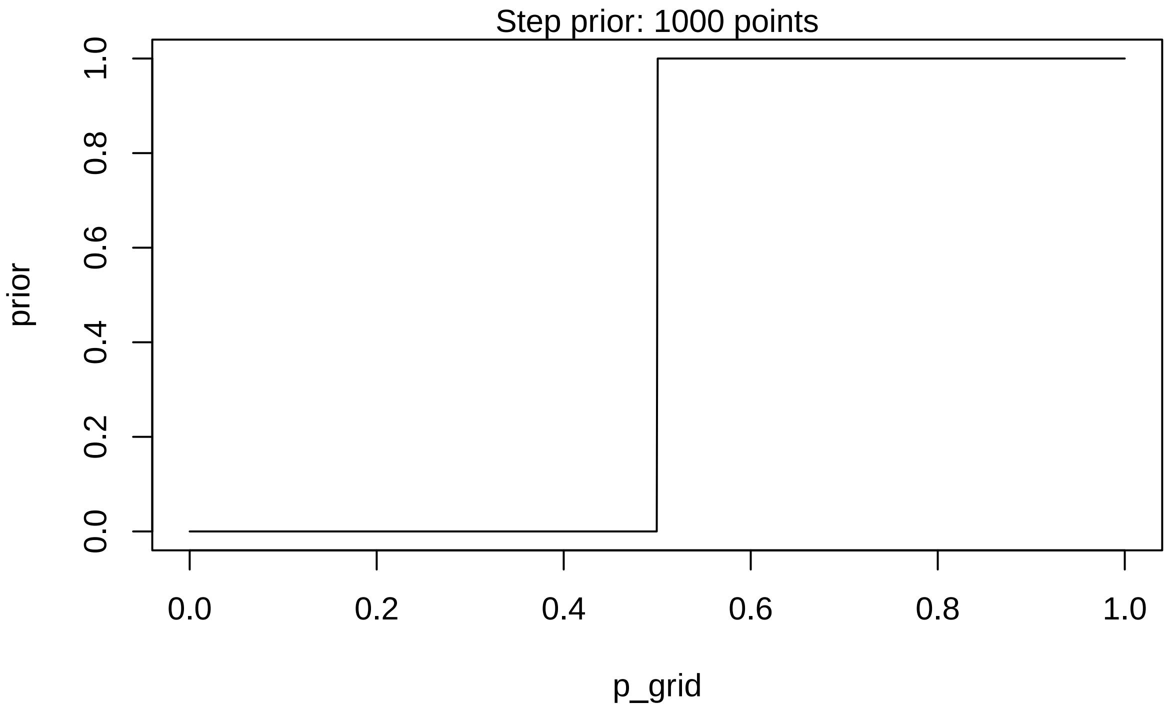

```
plot(p_grid , posterior , type="l" ,
    xlab="probability of water" , ylab="posterior probability" )
mtext(stringr::str_c("Flat prior: ", len, " points"))
```



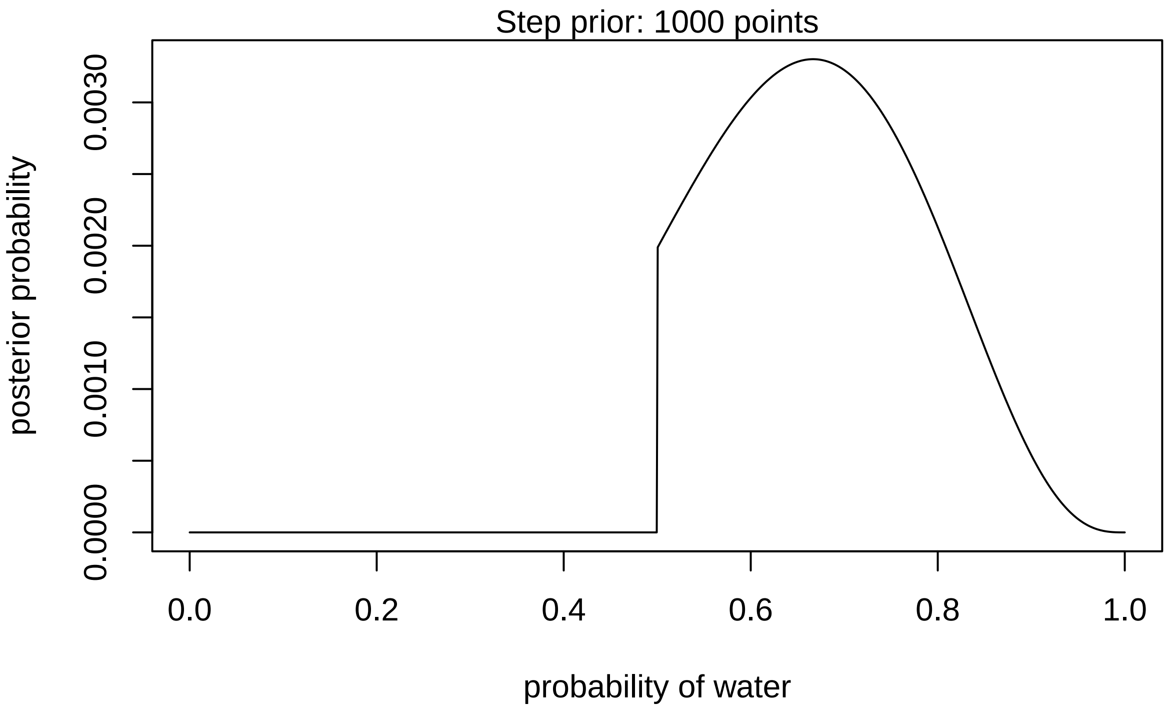

```
prior <- ifelse( p_grid < 0.5 , 0 , 1 )
unstd.posterior  <- likelihood * prior
posterior  <- unstd.posterior / sum(unstd.posterior)

plot(p_grid , prior , type="l")
```

```r
mtext(stringr::str_c("Step prior: ", len, " points"))
```

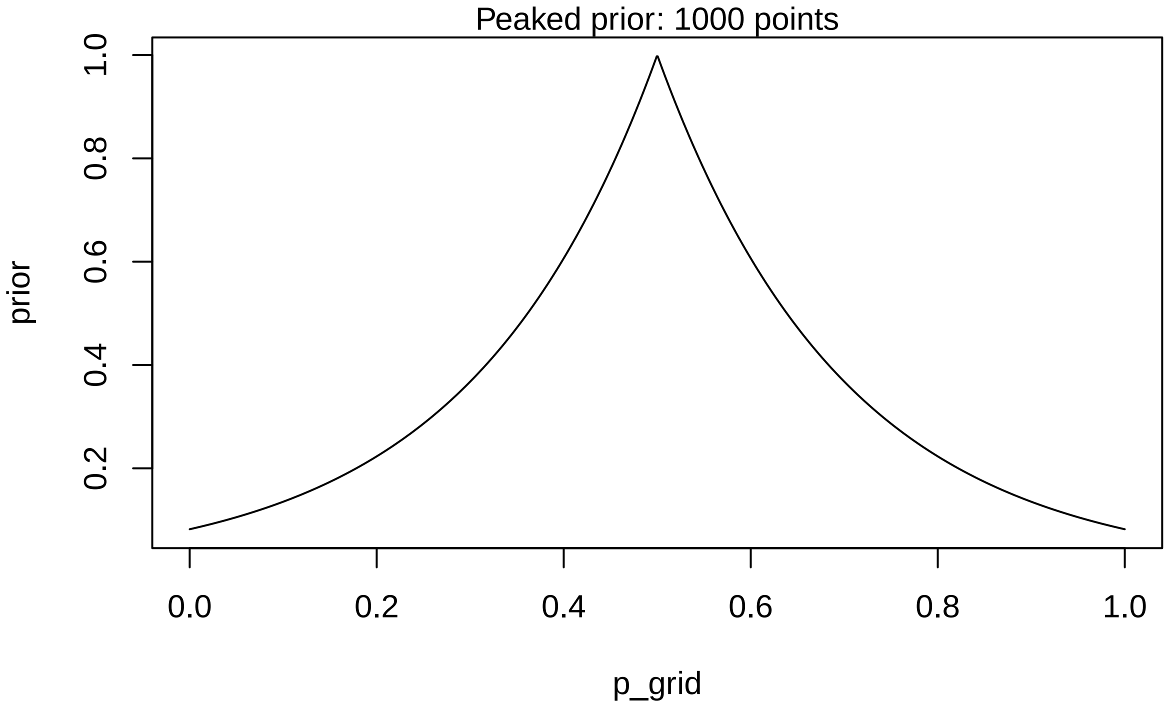



```r
plot(p_grid , posterior , type="l" ,
    xlab="probability of water" , ylab="posterior probability" )
mtext(stringr::str_c("Step prior: ", len, " points"))
```




```r
prior <- exp( -5*abs( p_grid - 0.5 ) )
unstd.posterior  <- likelihood * prior
posterior  <- unstd.posterior / sum(unstd.posterior)
```

```
plot(p_grid , prior , type="l")
mtext(stringr::str_c("Peaked prior: ", len, " points"))
```




```
plot(p_grid , posterior , type="l" ,
    xlab="probability of water" , ylab="posterior probability" )
mtext(stringr::str_c("Peaked prior: ", len, " points"))
```