Week 1

Chris Beeley

11/02/2021

Chapter 2

E1

Probability of rain on Monday would be written as Pr(Rain | Monday)

E2

Pr(Monday | Rain) means the probability of Monday given that it is raining

E3

The probability that it is Monday given that it is raining would be written Pr(Monday | Rain)

E4

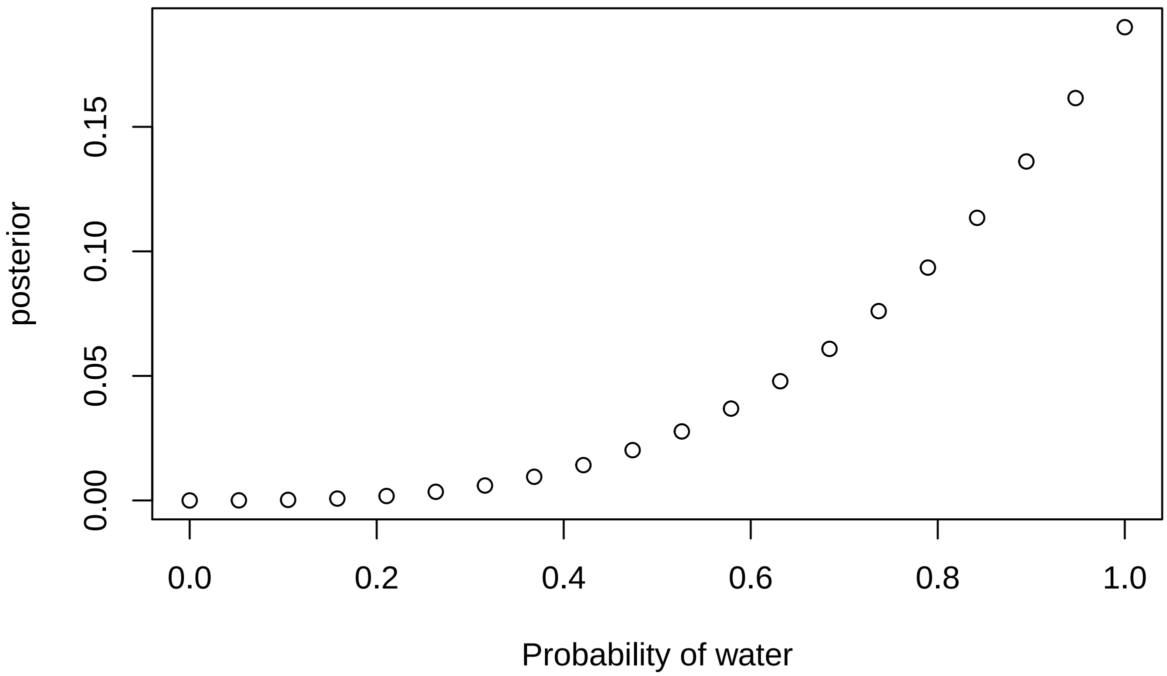
Bruno de Finetti wrote PROBABILITY DOES NOT EXIST which is of course quite true. Everything that happened, happened, and everything that didn't happen, didn't happen. However probability expresses a truth about reality in the sense that if you spin a globe and point to a random part of it, and repeat this observation, the average frequency of water will approach 0.7 at infinity. A similar process describes radioactive decay. So probability is false when it is used to describe events, but true when it describes processes, especially processes that are repeated.

M1

Compute and plot the grid approximation for the globe tossing experiment (randomly sampling water and land)

Plot- W, W, W

```r
p_grid <- seq(0, 1, length.out = 20)

prior <- rep(1, 20)

likelihood <- dbinom(3, size = 3, prob = p_grid)

unstd.posterior <- likelihood * prior

posterior <- unstd.posterior / sum(unstd.posterior)

plot(p_grid, posterior, xlab = "Probability of water")
```

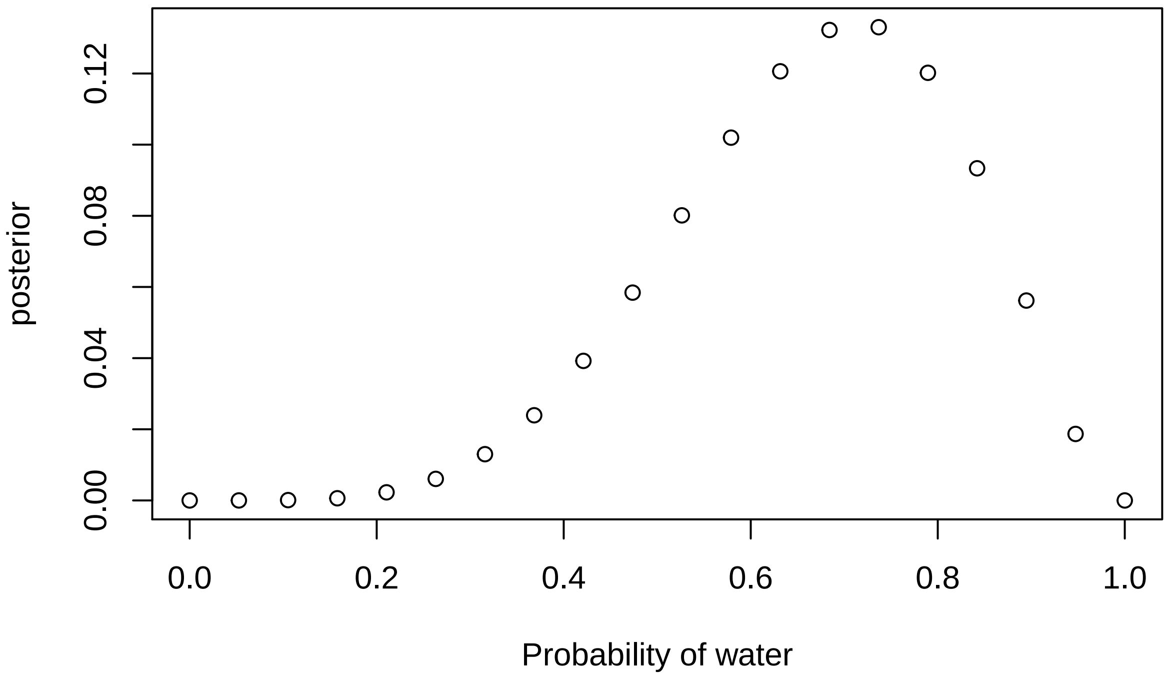


Plot- W, W, W, L

```r
# produce a function

grid_approximation <- function(water, trials){

  p_grid <- seq(0, 1, length.out = 20)

  prior <- rep(1, 20)

  likelihood <- dbinom(water, size = trials, prob = p_grid)

  unstd.posterior <- likelihood * prior

  posterior <- unstd.posterior / sum(unstd.posterior)

  plot(p_grid, posterior, xlab = "Probability of water")
}
```

Plot L, W, W, L, W, W, W

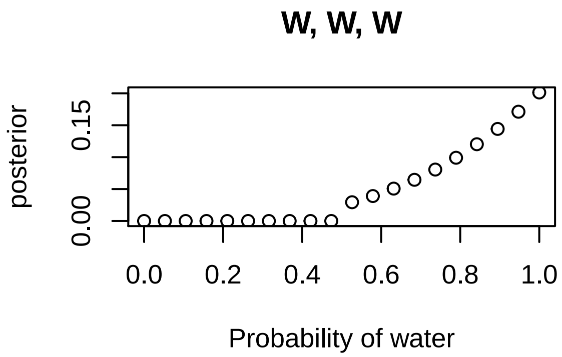
```r
grid_approximation(5, 7)
```

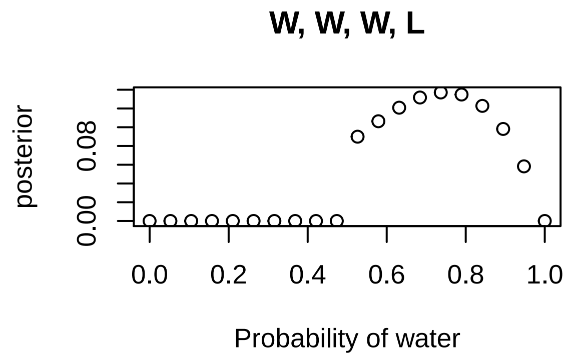


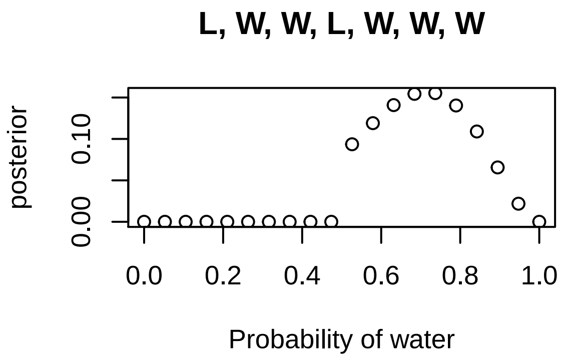
M2

Now assume a prior of zero when p < .5 and a constant when p >= 0.5, and compute the above again.

```r
# produce a function

grid_approximation_prior <- function(water, trials, title){

  p_grid <- seq(0, 1, length.out = 20)

  prior_f <- ifelse(p_grid < .5, 0, 1)

  likelihood <- dbinom(water, size = trials, prob = p_grid)

  unstd.posterior <- likelihood * prior_f

  posterior <- unstd.posterior / sum(unstd.posterior)

  plot(p_grid, posterior, xlab = "Probability of water",
       main = title)
}

par(mfrow = c(2, 2))

grid_approximation_prior(3, 3, "W, W, W")
grid_approximation_prior(3, 4, "W, W, W, L")
grid_approximation_prior(5, 7, "L, W, W, L, W, W, W")

par(mfrow = c(1, 1))
```







M3

Assume there are two globes. Mars is 100% land, Earth is 70% land. One is sampled and produces a "land" observation. Assume they are equally likely to be sampled. Show that Pr(Earth | Land) is 0.23.

$$p(A|B) = \frac{p(B|A).p(A)}{p(B)}$$

$$p(Earth|Land) = \frac{p(Land|Earth).p(Earth)}{p(Land)}$$

$$p(Earth|Land) = \frac{0.7 * 0.5}{0.85}$$

$$p(Earth|Land) = .411$$

This is clearly not the right answer

M4

Suppose you have 3 cards. One is black on one side and white on the other, one is WW, and one is BB. If you turn over a card and one side is black show that p(other side is black) = $\frac{2}{3}$

Counting method:

W1W2 - No W2W1 - No

B1W2 - No

WW	W	|
	W	|
BW	$\underline{\text{B}}$	-> W
	W	|
BB	$\underline{\text{B}}$	-> $\underline{\text{B}}$
	$\underline{\text{B}}$	-> $\underline{\text{B}}$

In the second column we see the possibilities that exist for a black side- there are 3. In the third column we see of these two possibilities only 2 lead to black on the other side. Therefore, the probability is $\frac{2}{3}$

H1

$$p(A|B) = \frac{p(B|A).p(A)}{p(B)}$$

$$p(SpeciesA|Twins) = \frac{p(Twins|SpeciesA).p(SpeciesA)}{p(Twins)}$$

$$p(SpeciesA|Twins) = \frac{p(.2).p(.5)}{p(.15)}$$

$$p(SpeciesA|Twins) = \frac{2}{3}$$

This is also clearly wrong! I'm not getting this I don't think!

Chapter 3

E 1-6

This is using the globe tossing example again, the code below is in the book to help

```
p_grid <- seq(0, 1, length.out = 1000)
prior <- rep(1, 1000)
likelihood <- dbinom(6, 9, prob = p_grid)

posterior <- likelihood * prior
posterior <- posterior / sum(posterior)

set.seed(100)
samples <- sample(p_grid, prob = posterior, size = 1e4, replace = TRUE)
```

E1. How much posterior probability lies below .2?

```
mean(samples < .2)
```

```
## [1] 4e-04
```

E2. How much posterior probability lies above .8?

```
mean(samples > .8)
```

```
## [1] 0.1116
```

E4. 20% of the posterior probability lies below what value of p?

```
sort(samples)[length(samples) * .2]
```

```
## [1] 0.5185185
```

E6. Which values of p contain the narrowest interval equal to 66% of the posterior probability?

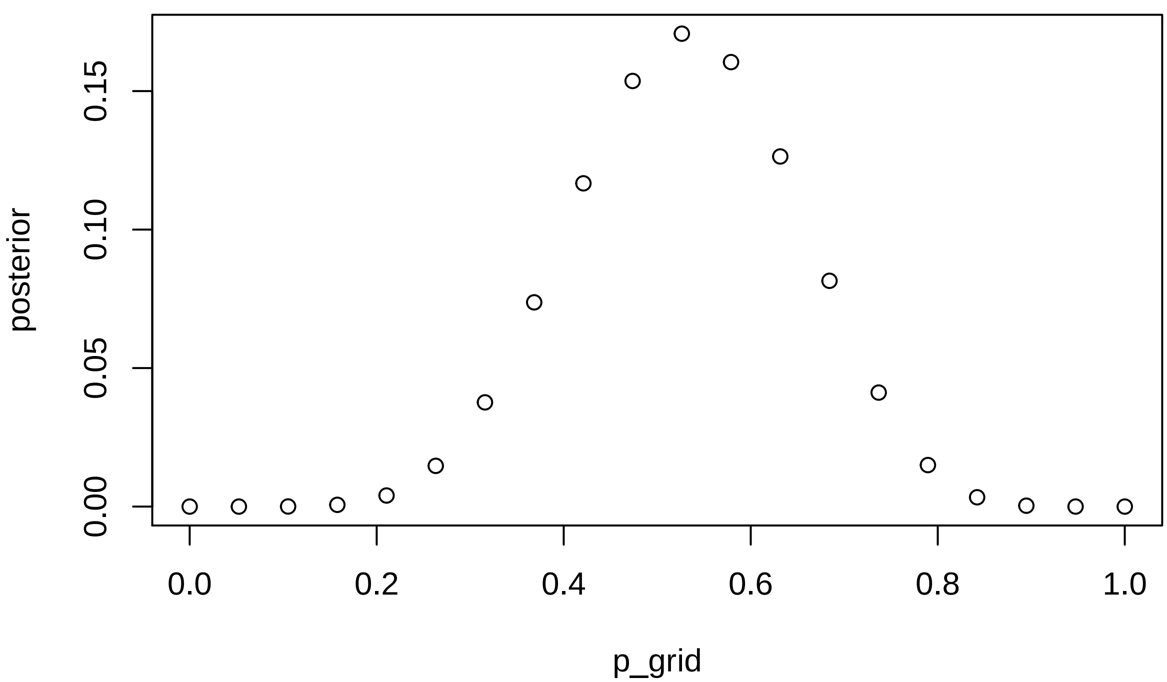
```
rethinking::HPDI(samples, prob = .66)
```

```
##     |0.66     0.66|
## 0.5085085 0.7737738
```

M1

Construct the globe tossing example again, with 8 water in 15 tosses. Construct the posterior distribution with a grid approximation.

```
p_grid <- seq(0, 1, length.out = 20)

prior <- rep(1, 20)

likelihood <- dbinom(8, size = 15, prob = p_grid)

unstd.posterior <- likelihood * prior

posterior <- unstd.posterior / sum(unstd.posterior)

plot(p_grid, posterior)
```



M2

Draw 1e4 samples from the grid approximation. Then calculate the 90% Highest Posterior Density Interval (HPDI).

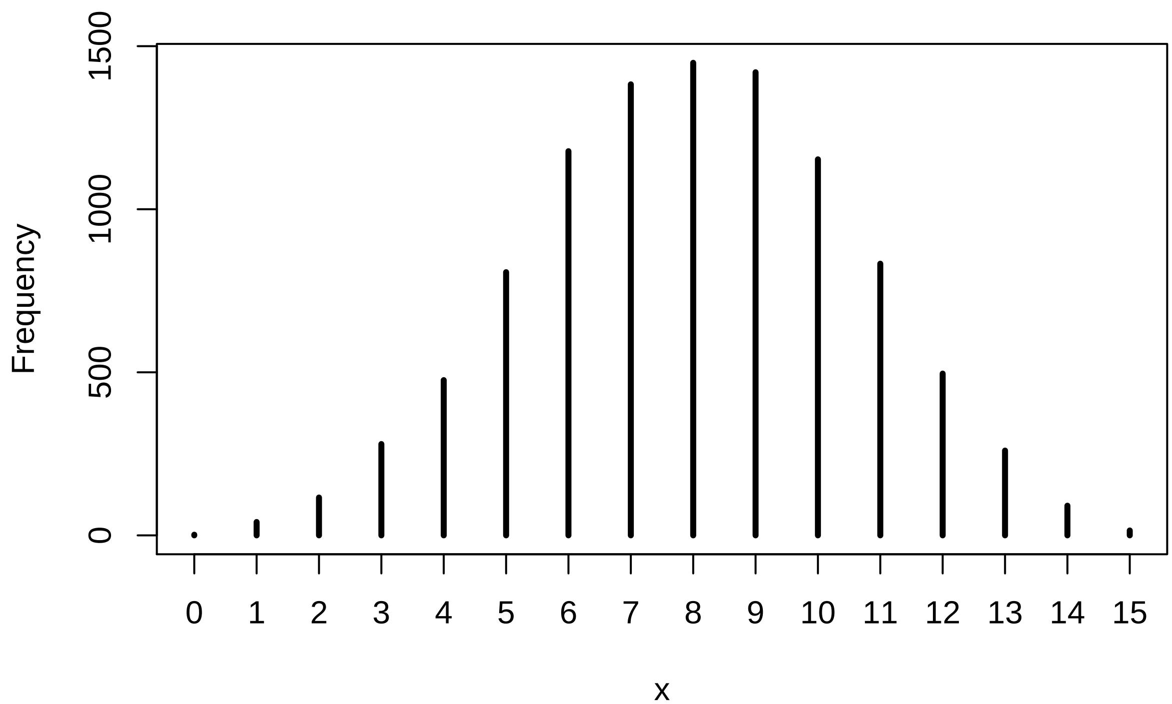
```r
set.seed(100)
samples <- sample(p_grid, prob = posterior, size = 1e4, replace = TRUE)

rethinking::HPDI(samples, .9)
```

```
##      |0.9      0.9|
## 0.3157895 0.6842105
```

M3

Construct a posterior predictive check for this model and data. This means simulate the distribution of samples, averaging over the posterior uncertainty in p. What is probability of observing 8 water in 15 tosses?

```r
ppc <- rbinom(1e4, size = 15, prob = samples)

rethinking::simplehist(ppc)
```




```r
sum(ppc == 8) / sum(ppc %in% c(0 : 7, 9 : 15))
```

```
## [1] 0.1694539
```