

# BST 140.651 Problem Set 3

Kai Aragaki

10/2/2020

## Problem 1

---

*a*

---

The likelihood of a parameter requires observations as well as some underlying probability density function, but we do NOT know the parameters that define that function.

If we are given the results of many coin flips, we do not necessarily know if the coin is fair or not - more formally, we do not know the probability parameter of this pdf that describes coin flips. What we do know is that a pdf for a single coin flip follows a Bernoulli distribution. The trick is to find the probability that describes the Bernoulli distribution.

The general definition of a likelihood function is the multiplication of the individual pdfs of a series of (in this case, at least) coin flips with an unknown defining parameter (in this case, the fairness of the coin (or the probability of getting a head)).

So for a single flip, the likelihood function is

$$\mathcal{L}(\theta|x) = p\theta(x) = \theta^x(1 - \theta)^{1-x}$$

Where  $x \in \{0, 1\}$  (where  $0 = \text{Tails}$  and  $1 = \text{Head}$ ), and  $\theta$  is the probability of getting a head.

For multiple flips, the likelihood function is simply the product of these pdfs. For a certain number of flips (let's say  $n$  flips) the general function will look like this:

$$\mathcal{L}(\theta|x) = \theta^{n_H}(1 - \theta)^{n - n_H}$$

To find the maximum likelihood, we can take the derivative of this function, set it equal to 0, and find the value of theta that makes this true.

$$\begin{aligned} & \left( \theta^{n_H}(1 - \theta)^{n - n_H} \right)' \\ &= n_H \theta^{n_H - 1} (1 - \theta)^{n - n_H} - (n - n_H) \theta^{n_H} (1 - \theta)^{n - n_H - 1} \end{aligned}$$

Setting this to 0 and simplifying yields

$n_H$  = number of heads

$$\begin{aligned}
& n_H \theta^{n_H-1} (1-\theta)^{n-n_H} - (n-n_H) (1-\theta)^{n-n_H-1} \theta^{n_H} = 0 \\
\implies & n_H \theta^{n_H-1} (1-\theta)^{n-n_H} = (n-n_H) (1-\theta)^{n-n_H-1} \theta^{n_H} \\
\implies & n_H (1-\theta) = (n-n_H) \theta \\
\implies & n_H - n_H \theta = (n-n_H) \theta \\
\implies & n_H = (n) \theta \\
\implies & \theta = \frac{n_H}{n}
\end{aligned}$$

Since we know for these trials that the number of heads was 7 and the number of trials was 10,

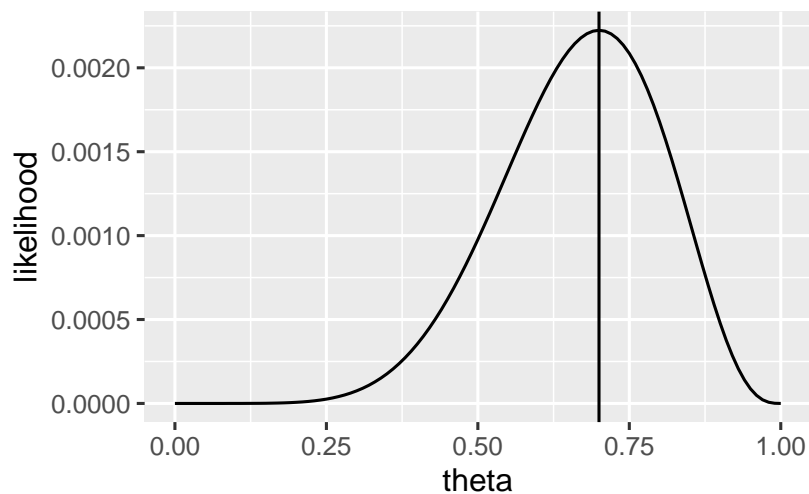
$$\mathcal{L}(\theta|x) = 0.7$$

*b*

---

We can show this by plotting:

```
theta <- seq(0, 1, by = 0.01)
likelihood <- theta^7 * (1 - theta)^(3)
tibble(theta, likelihood) %>% ggplot(aes(theta, likelihood)) + geom_line() + geom_vline(xintercept = 0.7)
```



To give us a better idea of what this means in more comparable terms, we can renormalize it (ie multiply the function by some value) such that the maximum is equal to 1. To do this, we first find the current maximum likelihood at the given theta:

```
m1 <- 0.7^7 * (1 - 0.7)^(3)
m1

## [1] 0.002223566
```

And then determine what number we need to multiply that value by to get 1:

```
scalar <- 1/ml  
scalar  
  
## [1] 449.728
```

So now we can replot with this scalar out front:

```
theta <- seq(0, 1, by = 0.01)  
likelihood <- scalar * theta^7 * (1 - theta)^(3)  
dat <- tibble(theta, likelihood)  
ggplot(dat, aes(theta, likelihood)) + geom_line() + geom_vline(xintercept = 0.7)
```

