Maximum Likelihood Estimation

Week 2 - PSYC 5316

September 4, 2017

Recall

Last time, we gave a formal definition for a probability function. An example was the *binomial* distribution for N independent Bernoulli trials (e.g., coin flips):

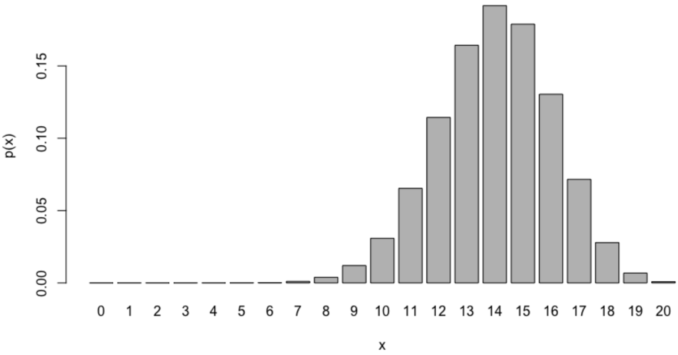
$$f(x \mid \theta) = \binom{N}{x} \theta^x (1-\theta)^{N-x}$$

where $x = \#$ of successes, and $\theta =$ probability of success.

Probability function

Suppose $N = 20$ and $\theta = 0.7$.

```
barplot(dbinom(0:20,size=20,prob=0.7),
        names.arg=0:20,
        ylab="p(x)",
        xlab="x")
```



Data and parameters

$$f(x \mid \theta) = \binom{N}{x} \theta^x (1-\theta)^{N-x}$$

This function gives us the probability of data, *given* a specific parameter

Data and parameters

What if we switched these?

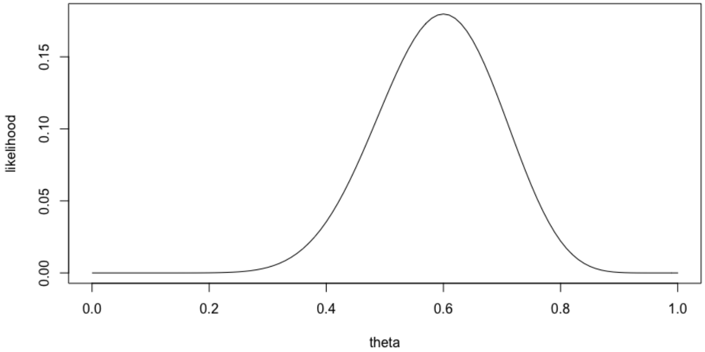
$$f(\theta \mid x) = \binom{N}{x} \theta^x (1-\theta)^{N-x}$$

This function then gives us the likelihood of a range of parameters, *given* a specific data point

Likelihood function

Suppose we observed 12 successes in 20 trials:

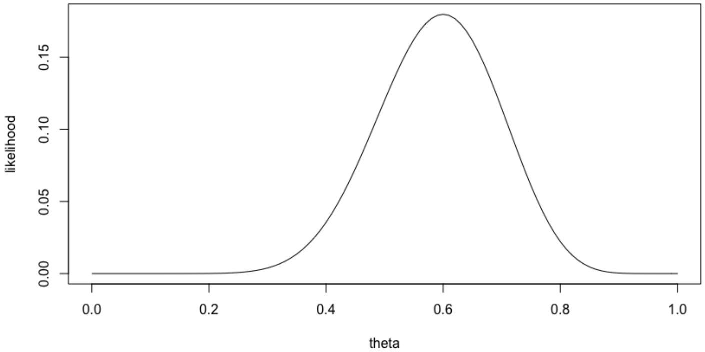
```
theta=seq(from=0, to=1, by=0.01)
plot(theta, dbinom(x=12, size=20, prob=theta),
     type="l",ylab="likelihood")
```



Likelihood function

Suppose we observed 12 successes in 20 trials:

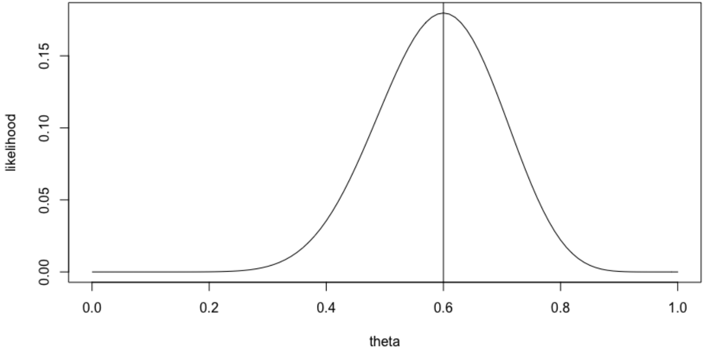
Natural question – what value of θ is most likely, given the data?



Likelihood function

Suppose we observed 12 successes in 20 trials:

Natural question – what value of θ is most likely, given the data?
Answer: $\theta = 0.6$



Maximum likelihood estimation

A key problem in statistical inference is how to infer from sample data to population parameters.

Maximum likelihood estimation is one solution to this problem

Maximum likelihood estimation

Basic workflow:

1. collect data
2. decide on a "model" for the data (e.g., binomial, normal, etc.)
3. define a likelihood function based on the underlying model
4. find the parameter value(s) that maximize the likelihood function