Likelihood

Bayesian Modeling for Socio-Environmental Data

Mary B. Collins

August 2016

Likelihood forms the fundamental link between models and data in a Bayesian framework. In addition, ***maximum likelihood*** *is a widely used alternative to Bayesian methods for estimating parameters in socio-ecological models.*

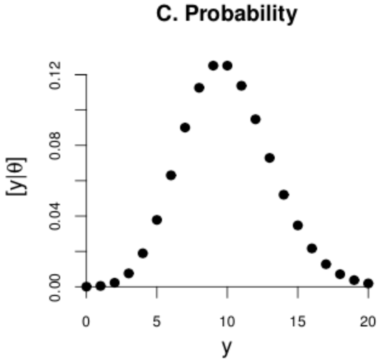
-Hobbs and Hooten 2015

Outline

- Probability functions
- Likelihood concepts
- Maximum likelihood

Probability functions

For a discrete random variable, Y, the probabiliy that the random variable Y takes on a specific value y is a probability function.



Tadpole Example

- You collect data on the number of tadpoles per volume of water in a pond. You observe 14 tadpoles in a one liter sample.
- You **know** the true average number of tadpoles per liter of water to be 23.
- The probability of your data is

$[y_i|\lambda] =$

What is the probability of your data?

$[y_i|\lambda] = \text{Poisson}(y_i = 14|\lambda = 23)$

```
lambda = 23
y = 14
dpois(y, lambda)
```

```
## [1] 0.01364609
```

In this example, what did we treat as fixed and what did we treat as random?

In this example, what did we treat as fixed and what did we treat as random?

Parameter values (λ or θ) are fixed and the data (y) are random.

What if, instead, you want to know the likelihood of the parameter given the observed data?

This evaluation can be accomplished using a *likelihood function* $\mathrm{L}(\theta|y)$

What is a likelihood function?

$$\mathrm{L}(\theta|y) = [y|\theta]$$

$$\mathrm{L}(\theta|\boldsymbol{y}) = \prod_{i=1}^{n} [y_i|\theta]$$

What is a likelihood function?

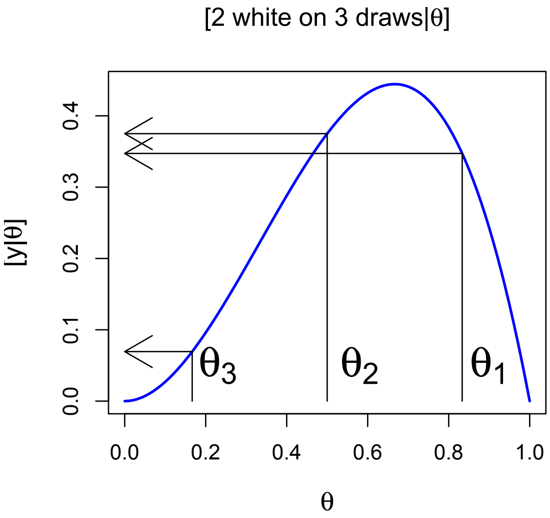
$$\underbrace{\mathsf{L}(\theta|\boldsymbol{y})}_{\text{“Likelihood Function"}} = \overbrace{\prod_{i=1}^{n}[y_i|\theta]}^{\text{“Likelihood Model" or “Data Model"}}$$

Cut to Tom's cans of beans

Parameter	Likelihood $[y \mid \theta_i]$
θ_1	.347
θ_2	.375
θ_3	.069
$\Sigma_{i=1}^{3}$	.791

Table 1: Probability of two whites on three draws conditional on θ_i

Likelihood profile



Likelihood as a concept

- Likelihood is the $[y|\theta]$.

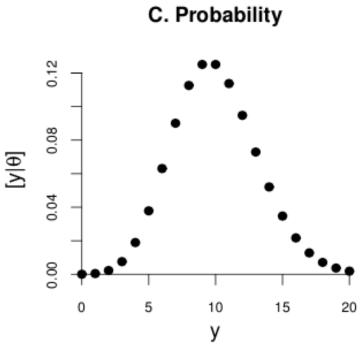
Likelihood as a concept

- Likelihood is the $[y|\theta]$.
- Likelihood is the chance of observing your data given theta.

Likelihood as a concept

- Likelihood is the $[y|\theta]$.
- Likelihood is the chance of observing your data given theta.
- Likelihood is the probability of observing your data conditional on your hypothesis θ.

What are the main differences between a PMF (or PDF) and a likelihood profile?



What are the main differences between a PMF (or PDF) and a likelihood profile?

Probability Density/Mass

- Data are treated as random variables.

Likelihood

What are the main differences between a PMF (or PDF) and a likelihood profile?

Probability Density/Mass

- Data are treated as random variables.
- Parameters are fixed.

Likelihood

What are the main differences between a PMF (or PDF) and a likelihood profile?

Probability Density/Mass

- Data are treated as random variables.
- Parameters are fixed.
- Area under the curve $= 1$.

Likelihood

What are the main differences between a PMF (or PDF) and a likelihood profile?

Probability Density/Mass

- Data are treated as random variables.
- Parameters are fixed.
- Area under the curve $= 1$.

Likelihood

- Data are fixed.

What are the main differences between a PMF (or PDF) and a likelihood profile?

Probability Density/Mass

- Data are treated as random variables.
- Parameters are fixed.
- Area under the curve $= 1$.

Likelihood

- Data are fixed.
- Parameters are varried.

What are the main differences between a PMF (or PDF) and a likelihood profile?

Probability Density/Mass

- Data are treated as random variables.
- Parameters are fixed.
- Area under the curve $= 1$.

Likelihood

- Data are fixed.
- Parameters are varried.
- Area under the curve $\neq 1$.

What are the main differences between a PMF (or PDF) and a likelihood profile?

Probability Density/Mass

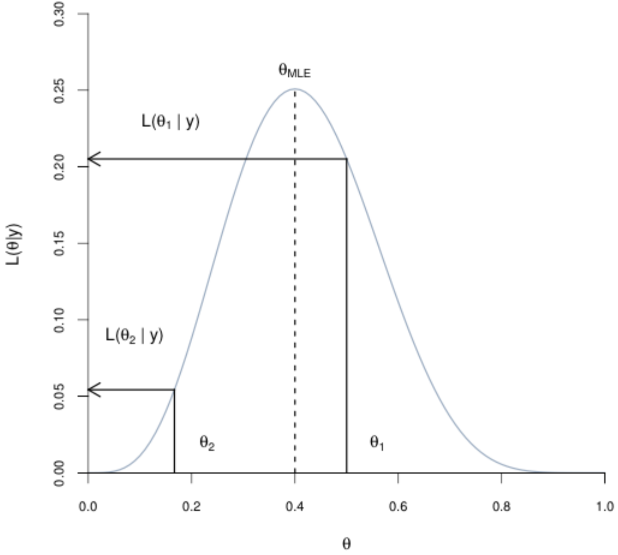
- Data are treated as random variables.
- Parameters are fixed.
- Area under the curve $= 1$.

Likelihood

- Data are fixed.
- Parameters are varried.
- Area under the curve $\neq 1$.
- Y-axis values are arbitrary and scalable.

Understanding the likelihood profile

What is the meaning of any one point on the likelihood profile curve?



Maximum Likelihood

Knowing the likelihood of a specific parameter value doesn't tell us anything useful in the absence of a comparison value. Therefore the evidence provided by data is expressed as the likelihood ratio,

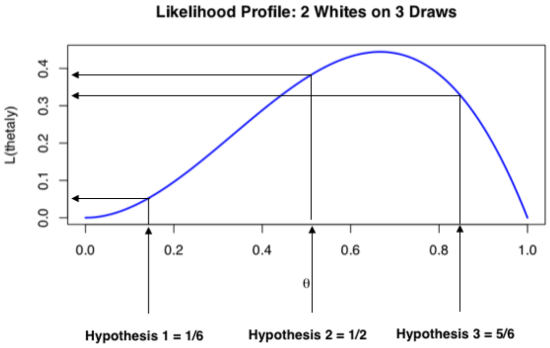
$$\frac{\mathsf{L}(\theta_1|y)}{\mathsf{L}(\theta_2|y)} = \frac{[y|\theta_1]}{[y|\theta_2]}$$

Practically, we often want to know the value of parameter θ that has the maximum *support* in the data, which is the peak of the likelihood profile. This is the value of θ that maximizes the likelihood function.

Likelihood example

Consider we have a jar of white and black beans and want to estimate the probabiliy, *p*, of choosing a white bean. We draw 3 beans and 2 are white. Plot the probability of the data conditional on θ as a function of all possible θ.

```r
p <- seq(0,1,.01)
w <- 2 #num whites
n <- 3 #num draws
y <- dbinom(x=w, size=n, prob=p)
```


Likelihood Profile: 2 Whites on 3 Draws
L(theta|y)
0.0
0.1
0.2
0.3
0.4
0.0
0.2
0.4
0.6
0.8
1.0
θ
Hypothesis 1 = 1/6
Hypothesis 2 = 1/2
Hypothesis 3 = 5/6