Likelihood inference

McElreath Ch 2

Figs 2.2 - 2.4

The likelihood principle

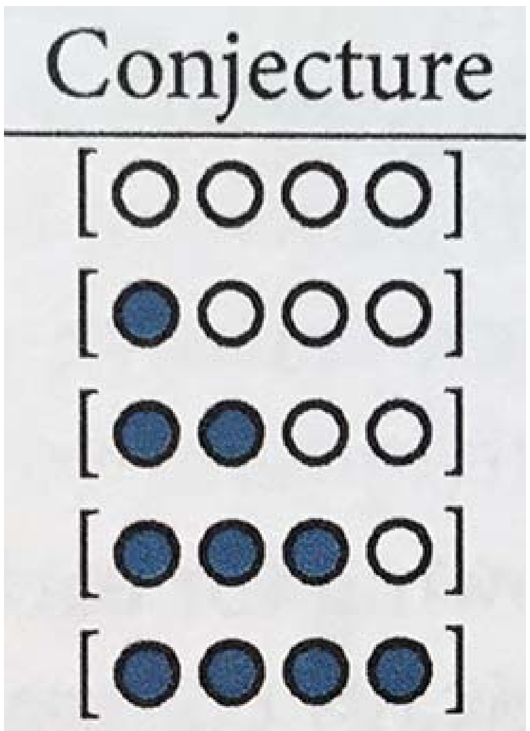
All the evidence in an observation (data) about the parameters (model) is in the likelihood function

The likelihood function

Counts all the ways the data could have happened for a given model or hypothesis

Marbles in a bag

We know: 4 marbles, 2 colors, marbles drawn randomly with replacement
Goal: what is in the bag?



hypothesis, model

H_1, M_1

H_2, M_2

H_3, M_3

H_4, M_4

H_5, M_5

Data

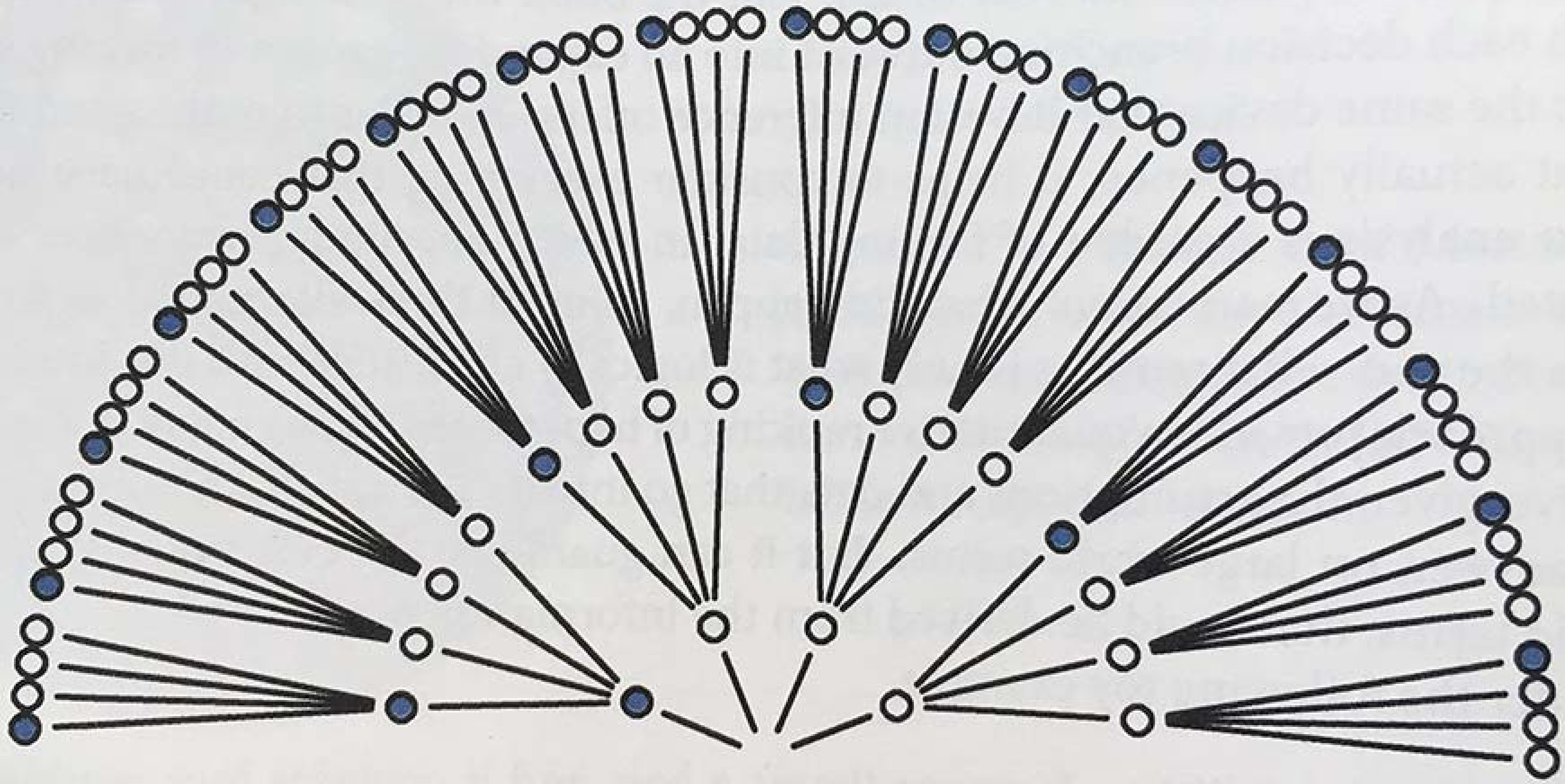
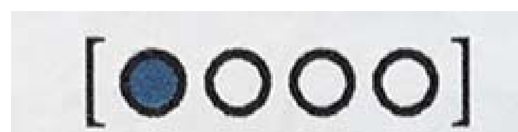


FIGURE 2.2. The 64 possible paths generated by assuming the bag contains one blue and three white marbles.

i.e. assuming we have H_2, M_2



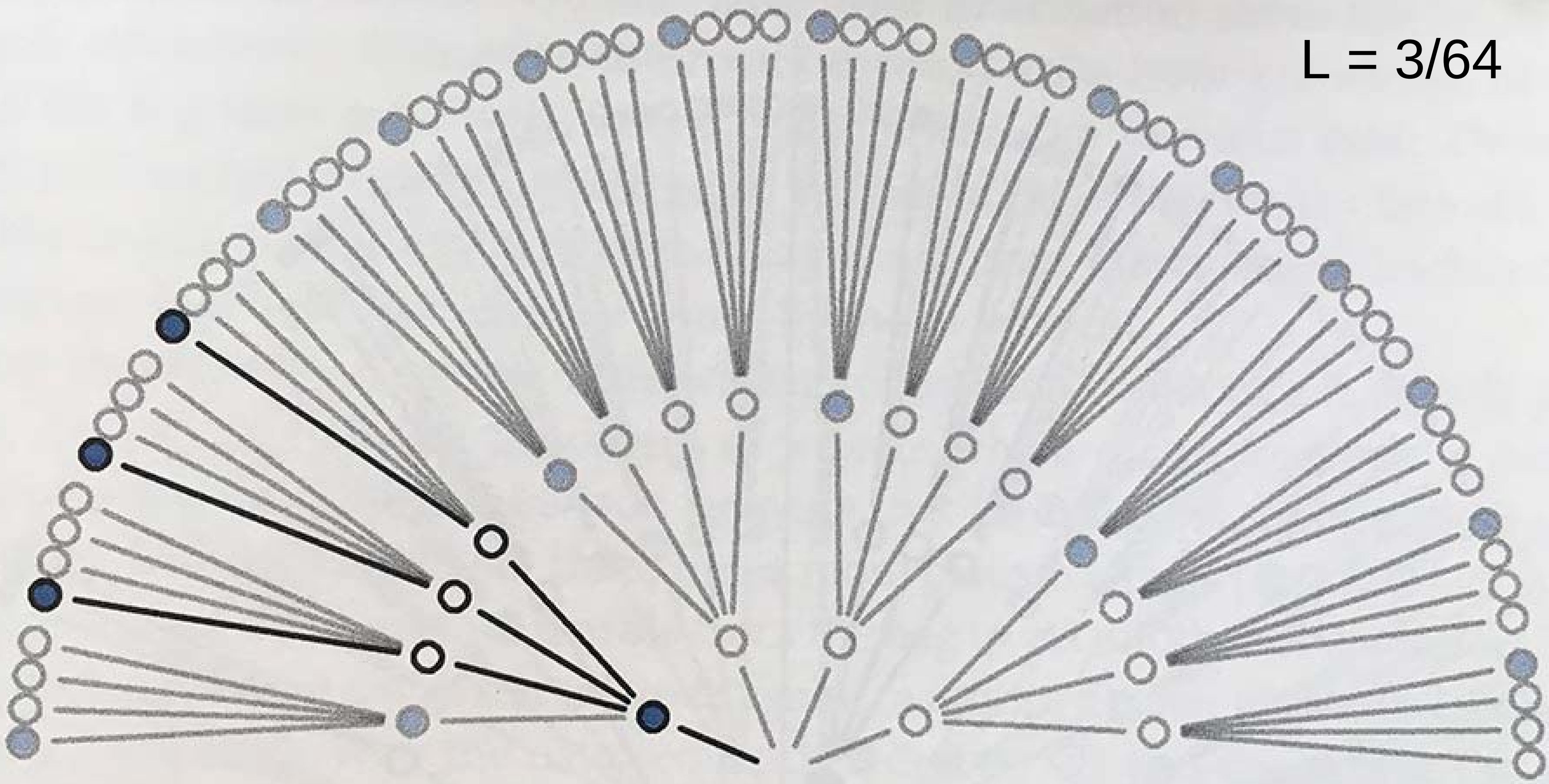
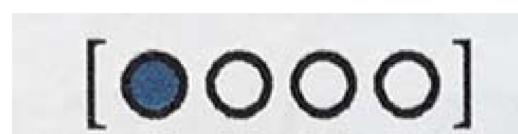
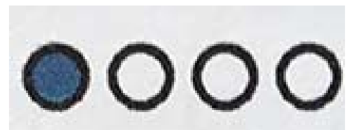


FIGURE 2.3. After eliminating paths inconsistent with the observed sequence, only 3 of the 64 paths remain.

Paths for data  given M_2 

The likelihood

- Probability of the data given a model

L=P( | )

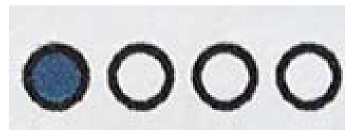
Data — Model 2

"given"

A conditional probability

The likelihood

- Probability of the data given a model

L=P( | )

Data "given" Model 2

A conditional probability

$$P(y|M_2) = P(y|\theta_2) \qquad y = [\text{"b"},\text{"w"},\text{"b"}]$$

could be a vector

θ indicates parameters
(number of blue & white)

Likelihood of model or H

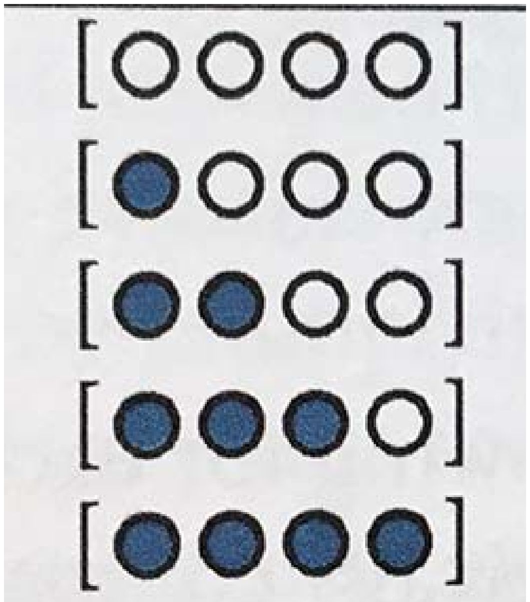
A model is more likely than another if it is the model for which the data are more probable

Notice that this doesn't mention the *probability* of the model, only the probability of the data.

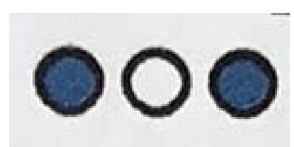
Inference: likelihood ratio

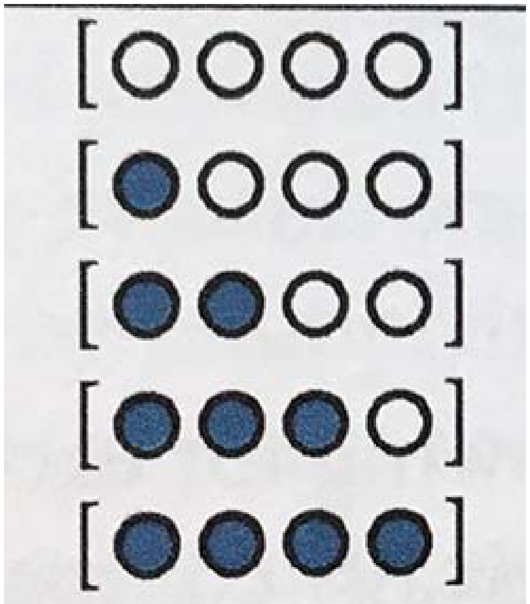
$$\frac{P(y \mid \theta_2)}{P(y \mid \theta_1)}$$

Strength of evidence for model 2 compared to model 1



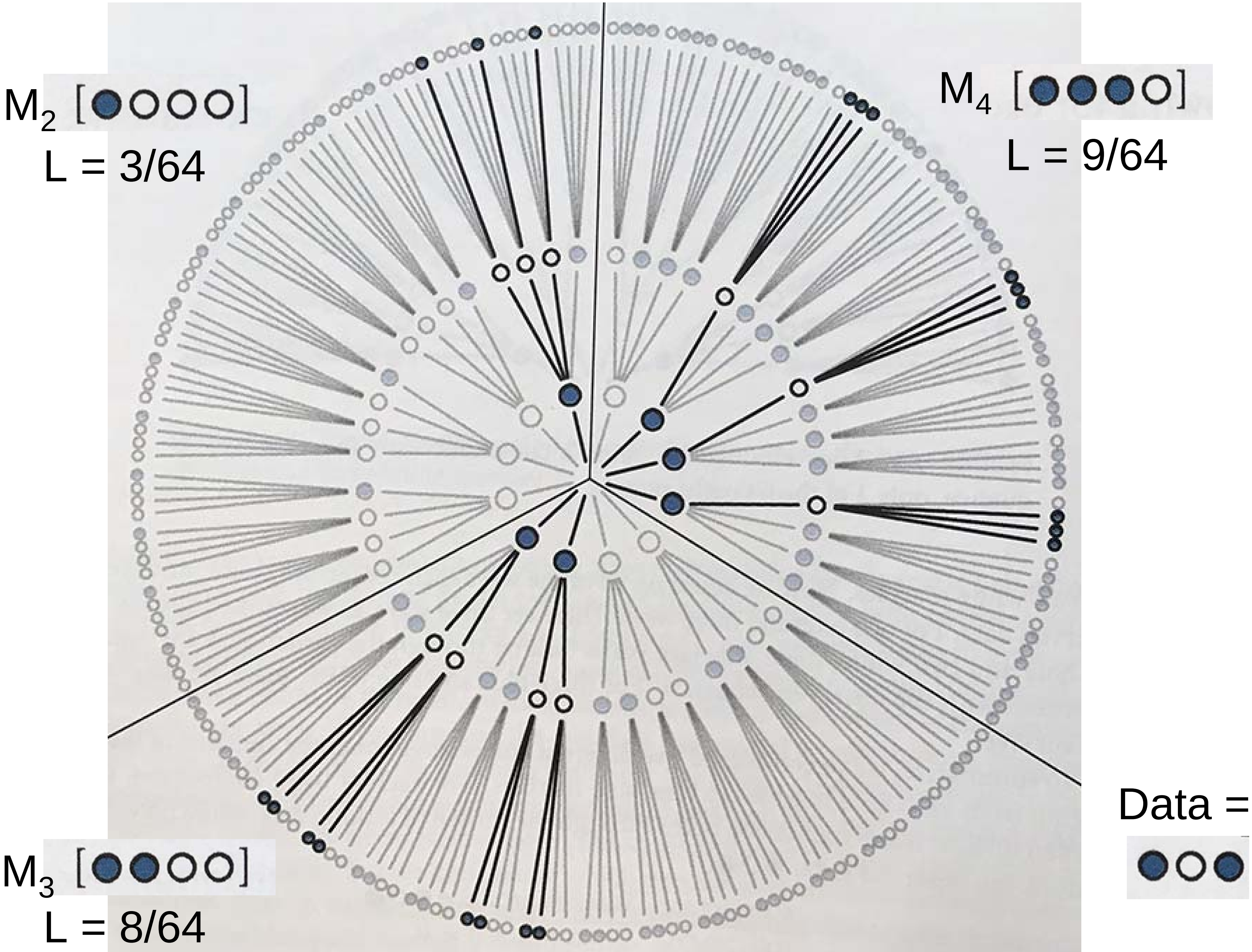
Models	Pathways	Likelihood
M_1	?	
M_2	3	3/64
M_3	?	
M_4	?	
M_5	?	

Data = 

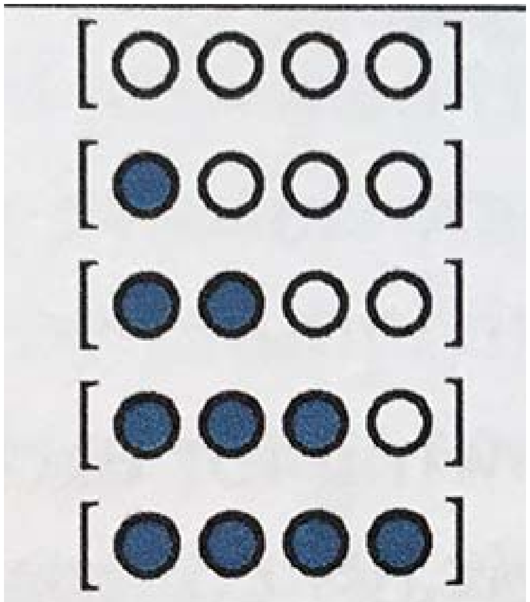


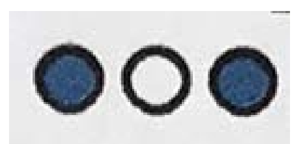
Models	Pathways	Likelihood
M_1	0	0
M_2	3	3/64
M_3	?	
M_4	?	
M_5	0	0

Data = 



M_2 [●○○○]
L = 3/64
M_4 [●●●○]
L = 9/64
M_3 [●●○○]
L = 8/64
Data =
●○●

	Models	Pathways	Likelihood
M_1		0	0
M_2		3	3/64
M_3		8	8/64
M_4		9	9/64
M_5		0	0

Data = 

$$\frac{P(y \mid \theta_4)}{P(y \mid \theta_2)} = \frac{9}{3} = 3 \qquad \frac{P(y \mid \theta_4)}{P(y \mid \theta_3)} = \frac{9}{8} = 1.125 \qquad \frac{P(y \mid \theta_3)}{P(y \mid \theta_2)} = \frac{8}{3} = 2.\dot{6}$$

Notes

- Not frequentist
- Not the same or even similar to a sampling distribution
  - we have not invoked multiple repeated samples
  - probability of the data, not probability of a sample statistic