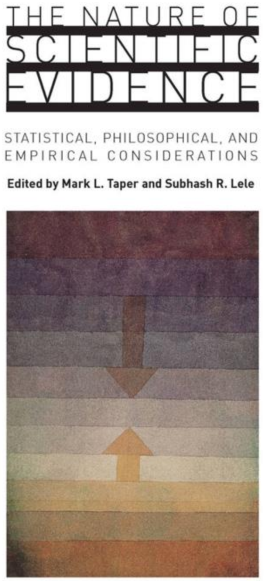
Have you learned how to drive an F1 car by watching Checo Pérez race on TV?

- If you only listen to lectures and are not actively reading, engaging with content, and making notes, you will not learn
- Practice many times the computer tutorials
- Find others in the class who can discuss ideas with you
- Seek for help early!

Introduction to the likelihood function

Rosana Zenil-Ferguson
Assistant Professor. University of Kentucky
MOLE at Woods Hole workshop 2026

Purely likelihood-based inference is also known as Evidential statistics



A branch of statistics interested in understanding how to infer processes through evidence (data)

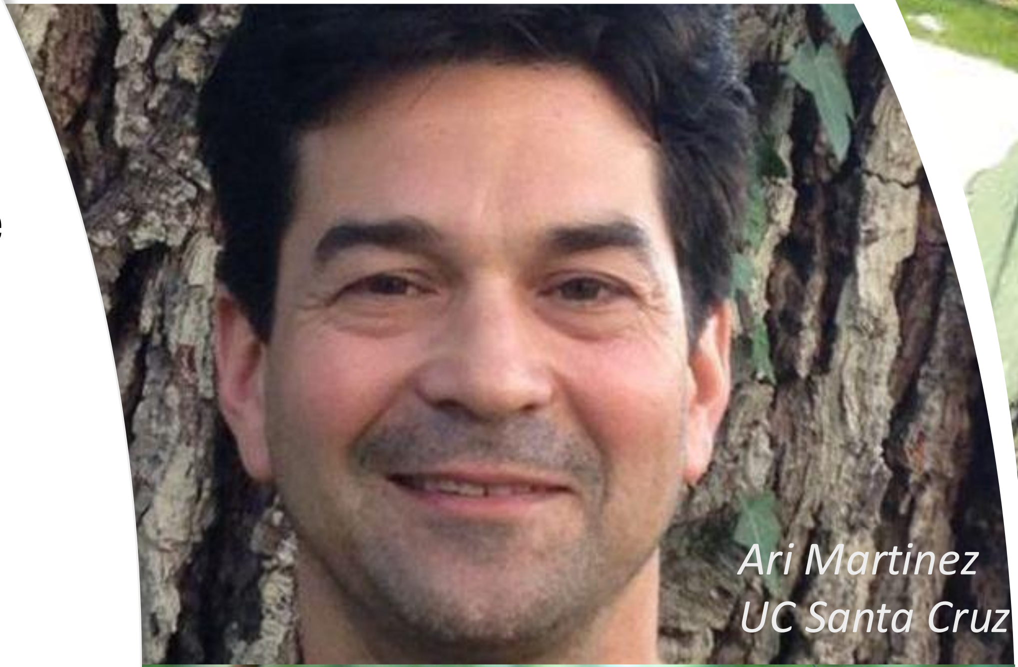
Goal: To propose a way to measure evidence connecting it to hypotheses

Our example today: Bird alarm calls in the Peruvian Amazonia

Ari goes to the jungle and observes the behavior of two mixed flocks.

Dead-leaf gleaners feeding on the ground would freeze after an alarm call.

Flycatchers feeding on the top of the trees would resume eating almost instantly after an alarm call.



Ari Martinez
UC Santa Cruz

Thamnomanes ardesiacus

The experiment and gathering the data

What do you notice about these data?

Response is amount of time a bird freezes in seconds

DL - dead-leaf gleaners

F - flycatchers

Response	Forage
18	DL
11	DL
23	DL
0	F
0	F
6	F

Definition of a random variable

- A random variable is a function that takes us from the space of events (Ω) to a number, in this case a continuous number ($\mathbb{R}$)

$$X : \Omega \longrightarrow \mathbb{R}$$

omega is the space of events

real numbers

random variables always written using capital letters

Event: The amount of time a bird freezes $\longrightarrow$ 6

(six seconds is an observation)

Probability spaces and measure

Random variables are random because they need a probability measure

$$(\Omega, \underbrace{\sigma - algebra}_{\text{space with combination of events}}, \mathbb{P}) \rightarrow (\mathbb{R}, B, P)$$

space of events

Space with combination of events

probability meausure

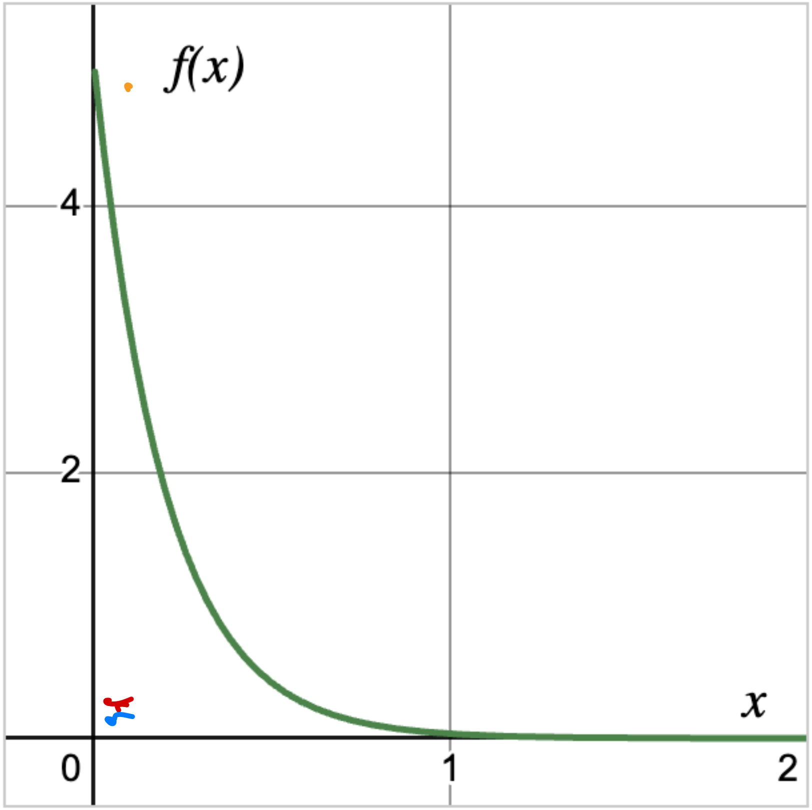
real numbers

Borel set : combinations of numbers

probabilty distribution function $f_X(x)$

$$\mathbb{P}(\text{of a bird freezing for six seconds}) \longrightarrow P(X=6) = f_X(6)$$

Probability model for freezing (waiting) times



$f_X(x)$

X values of the random variable

$X \sim Exp(\beta)$

(X has a distribution exponential with parameter β)

$$f_X(x) = \frac{1}{\beta} e^{-x/\beta}$$

what is β?

$$\mathbb{E}[X] = \beta$$

it is the expected value "average" of the random variable

Likelihood function (flycatchers only)

data

$$\mathcal{L}(\beta; \mathbf{X}) = P(\mathbf{X}|\beta)$$

the model is $Exp(\beta)$

The probability of the sample given the model or hypothesis represented by the symbol β

1. The model: $f_X(x) = \frac{1}{\beta} e^{-\frac{x}{\beta}}$

$\mathbf{X} = \{0, 0, 6\}$ observed values

$\mathbf{X} = \{x_1, x_2, x_3\}$ (lower case means observed values)

2. Data are assumed independent:

given

$$P(\mathbf{X} | \beta) = P(x_1 | \beta) \times P(x_2 | \beta) \times P(x_3 | \beta) = f_X(x_1) \times f_X(x_2) \times f_X(x_3)$$

independent means I can multiply the probabilities of each bird.

Calculating the likelihood function

$$f_X(x) = \frac{1}{\beta} e^{-\frac{x}{\beta}} \qquad x_1 = 0 \qquad x_2 = 0 \qquad x_3 = 6$$

$\boxed{\beta = 1}$ (assuming expected freezing time is 1 second)

$$f_X(0) = \frac{1}{1} e^{-\frac{0}{1}} = 1$$

for bird 1 and 2

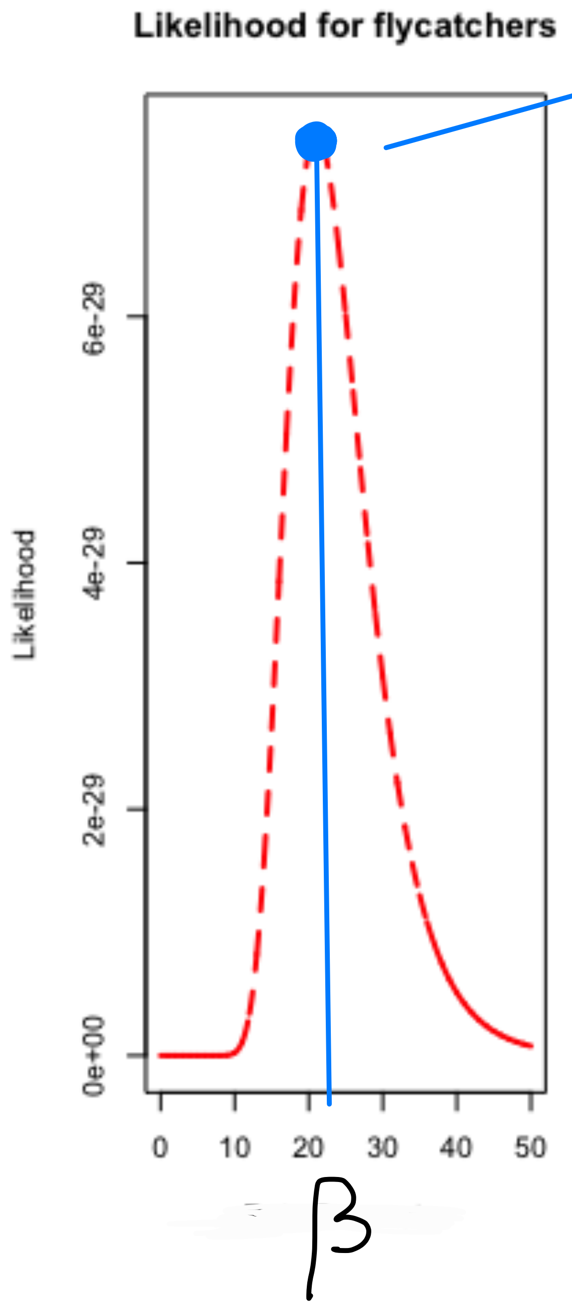
$$f_X(6) = \frac{1}{1} e^{-\frac{6}{1}} = e^{-6}$$

$$= 0.002$$

$$\boxed{\mathcal{L}(1;\cancel{X}) = \underbrace{1}_{\text{Bird 1}} \times \underbrace{1}_{\text{Bird 2}} \times \underbrace{e^{-6}}_{\text{Bird 3}} = 0.002}$$

Maximum likelihood estimate (MLE)

$$\mathcal{L}(\beta; \mathbf{x}) = P(\mathbf{x} \mid \beta)$$

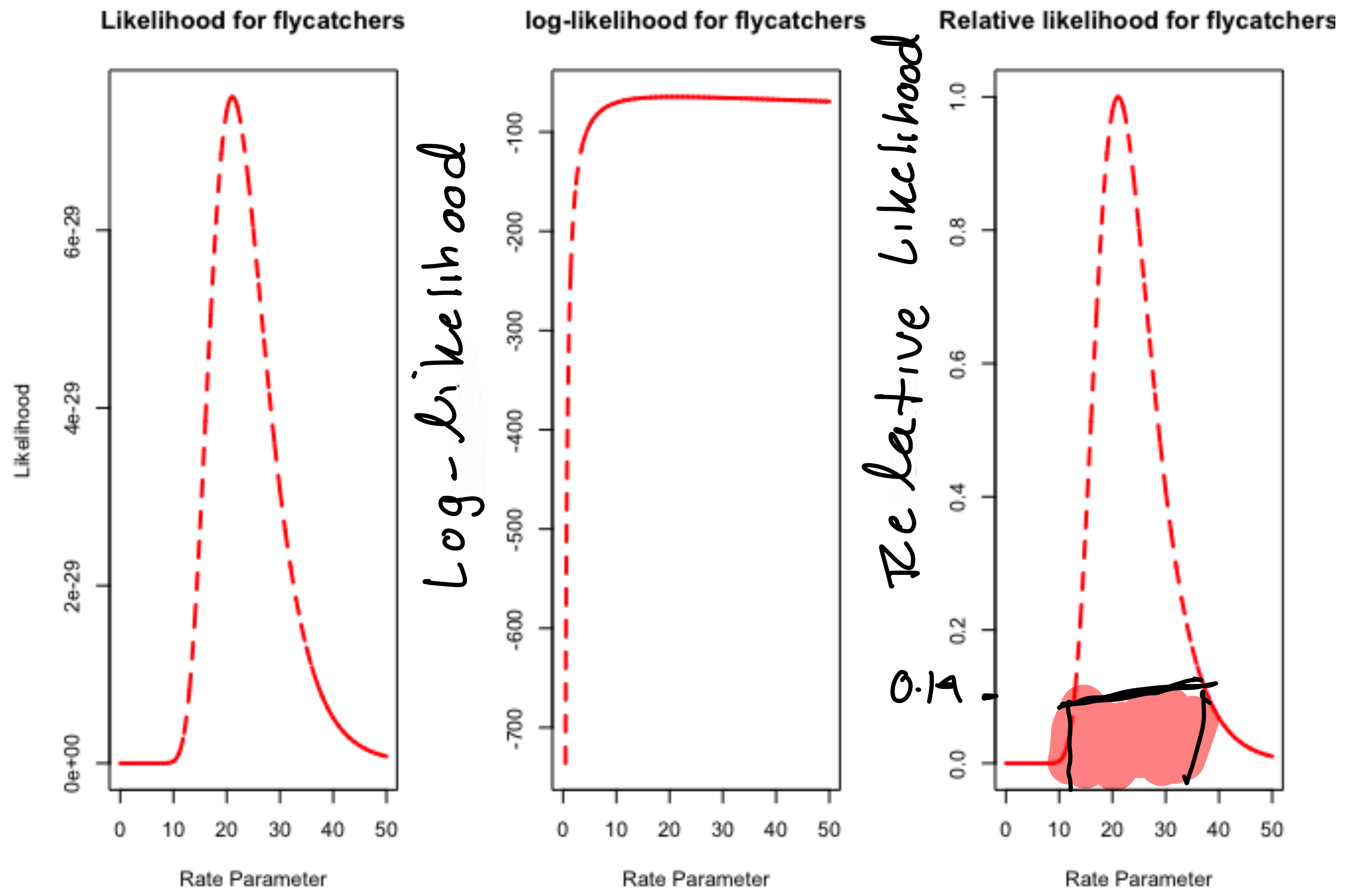


Maximum likelihood Estimate

$$\hat{\beta} = \beta_{MLE}$$

The value of β that maximizes the probability of observing our data.

Different ways of visualizing the likelihood



Read y-axis for scale

95%
(10-35) = Confidence Likelihood Interval

Exercise (in class if we have time)

- Calculate the likelihood for the dead-leaf gleaners:

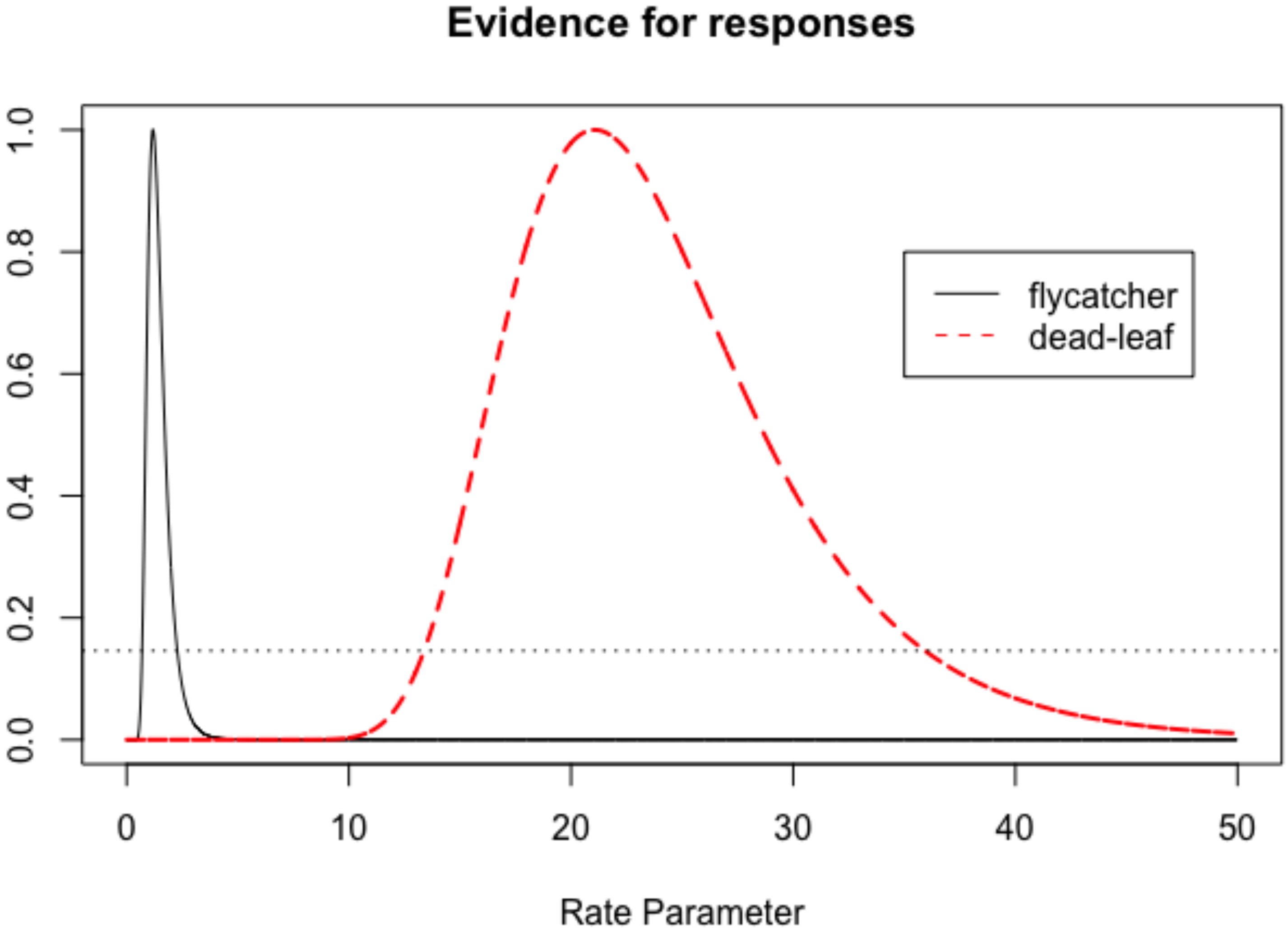
Data: $X = \{18, 11, 23\}$ for parameter values $\beta = 15, 16, 17, 18$

$$\hat{\beta} = \beta_{MLE} = 17.3$$

for the exponential distribution the maximum likelihood estimate of the parameter is exactly the average.

Differences between the groups

when two groups have no over-lapping CI (95%) this is significant evidence that the DL and F behaviors are different.



Martínez and Zenil, 2012. *Behavioral Ecology*

Advantages of using a likelihood function

(Sprott and Kalbfleisch, 1965)

1. It represents a concise summary of all of the information in the data, and when graphed, it gives a pictorial summary of what the data have to say.

2. Inferences based on the likelihood are exact for samples of all sizes; no assumption or mathematical investigation is required to ascertain if the sample is large enough to justify the method.

3. The rather complicated mathematics of asymptotic or large sample theory is not necessary, giving a relatively simple theoretical approach to inference.

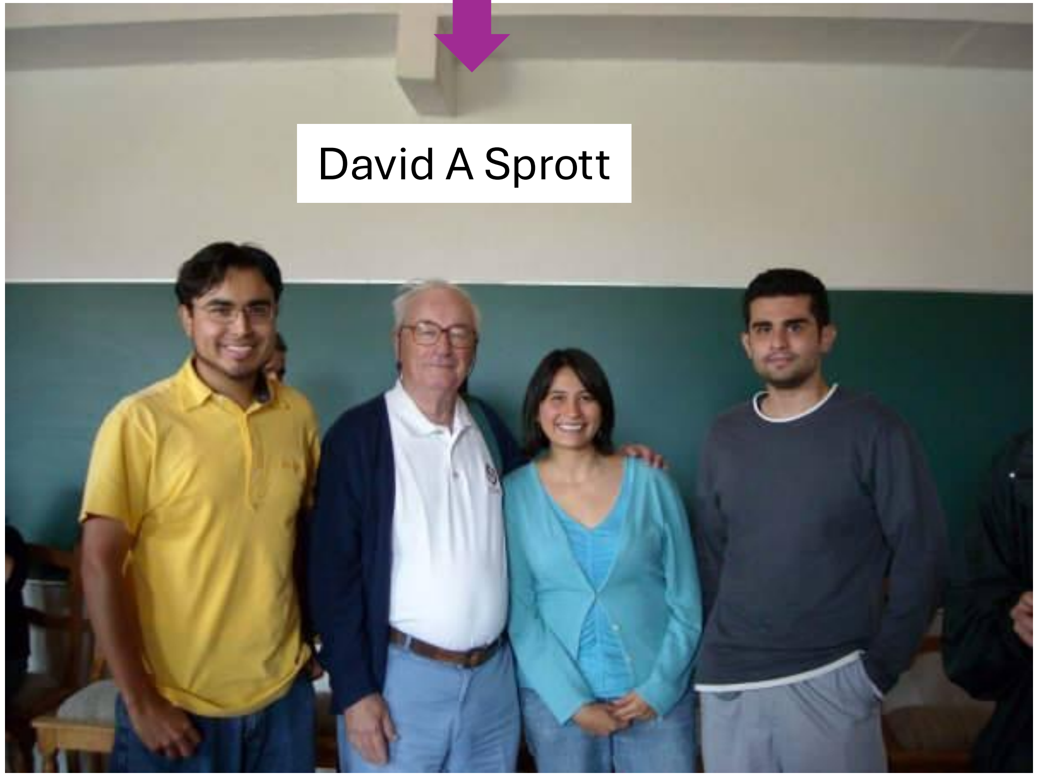
Disadvantages

(Sprott and Kalbfleisch, 1965)

1. Increased numerical computation in addition to the uncertainty being expressed (exactly) in terms of likelihood rather than (approximately in large samples) in terms of probability.

2. The likelihood method is not a test of significance and cannot be used to test the validity of a single hypothesis without reference to alternatives

Fisher



Collective dynamics of ‘small-world’ networks

Duncan J. Watts* & Steven H. Strogatz