# Statistics in Cosmology

## Day 1 – Interpretation

*11th TRR33 Winter School in Cosmology*
*10-16 December 2017, Passo del Tonale - Italy*


Emille E. O. Ishida

*Laboratoire de Physique de Clermont - Université Clermont-Auvergne*
*Clermont Ferrand, France*

# Emille E. O. Ishida

## Researcher at the Laboratoire de Physique de Clermont

*Clermont Ferrand, France*

Type Ia Supernova + Machine Learning + Bayesian Statistics

Interdisciplinary science development

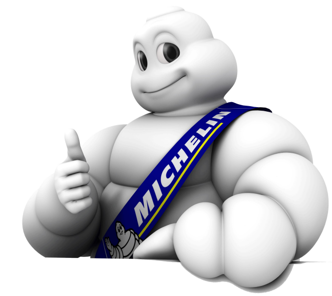
Co-chair of the Cosmostatistics Initiative (COIN)

# Clermont Ferrand, France

*Cheese*

*Truffade*









*Puy-de-Dome*

# Overview

1. Interpretation

   *Bayesian x Frequentists*

2. Parameter inference

   *Markov Chain Monte Carlo*

3. Model Selection

   *The art of making choices*

4. Approximate Bayesian Computation

   *Forward modeling for likelihood-free inference*

   4b. A note on interdisciplinarity

   *The whole is more than the sum of its parts*

# Acknowledgments

Alan Heavens
David Hogg
Eric Feigelson
Ewan Cameron
Jake VanderPlas
Jessi Cisewski
Licia Verde
Martin Kilbinger
Rafael S. de Souza
Roberto Trotta
The Cosmostatistics Initiative

*In memory of Joseph M. Hilbe (1944-2017)*

*"Begin at the beginning,"*
*the King said, very gravely,*
*"and go on till you come to the end:*
*then stop".*

Alice in Wonderland, Lewis Carroll (1865)

# In the beginning ...

Statistics ⟷ Status

Mean

Variance

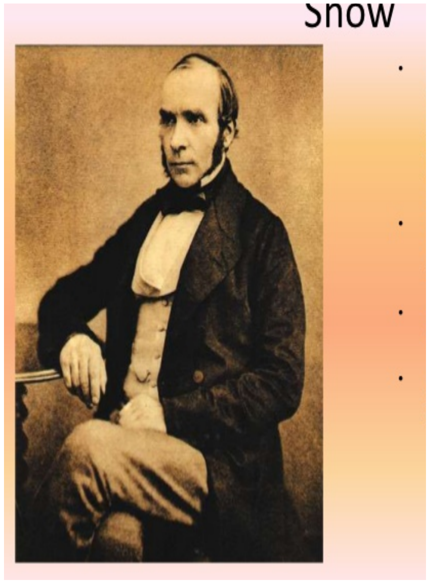
Frequency

Taxes

Medicine

# In the beginning ...

## Statistics ⟷ Status




John Snow (1813 – 1858)
*The father of epidemiology*

Used data to prove that cholera was transmitted via contaminated water

Taxes

Medicine

# ... then, gambling!

Blaise Pascal

1654

Pierre de Fermat

Christian Huygens

*On Reasoning in Games of Chance*, 1657

Chance

*http://homepages.wmich.edu/~mackey/Teaching/145/probHist.html*

# ... then, gambling!

Blaise Pascal

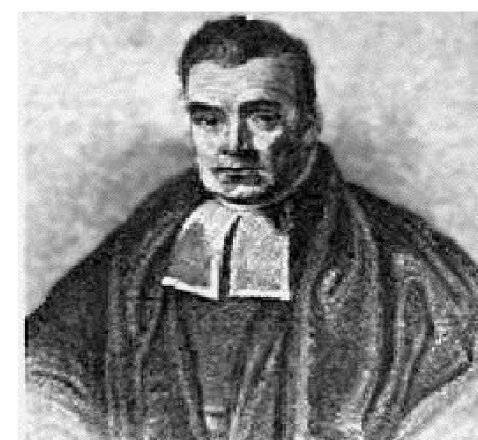
1654

Pierre de Fermat

Christian Huygens

*On Reasoning in Games of Chance*, 1657

*An Essay towards solving a Problem in the Doctrine of Chances*, 1763



Thomas Bayes

*http://homepages.wmich.edu/~mackey/Teaching/145/probHist.html*

# What is a probability?

Frequentist:

Frequency

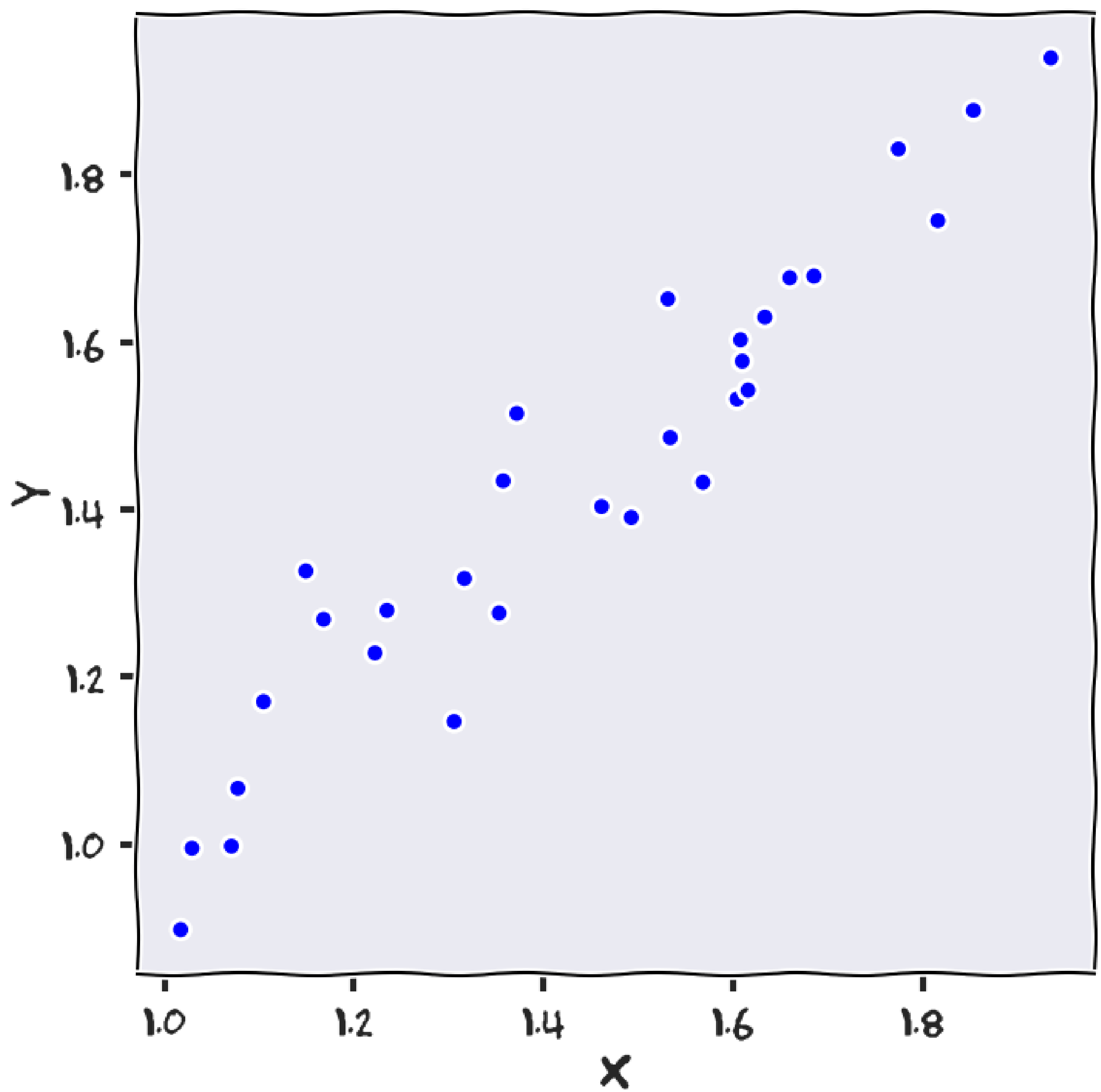
Bayesian:

State of knowledge

*Or*

Degree of belief

# A first example:



$$y = a\,x + b$$

*"Physical" model*

Least Squares:

$$\{a, b\} \leftarrow \min \left[ \sum_{i=1}^{N} \frac{(y_i - (ax_i + b))^2}{\sigma^2} \right]$$

Main assumptions:

- independent observations (iid)
  *no correlations*
- all the points are equally valid
  *no outliers*
- uncertainties are Gaussian distributed
  *homoscedasticity*
- no exterior information
  *no priors*

Statistical model

*https://www.albert.io/blog/key-assumptions-of-ols-econometrics-review/*

$$y = a\,x + b$$

*Model of the mean*

Least Squares: $$\{a, b\} \leftarrow \min \left[ \sum_{i=1}^{N} \frac{(y_i - (ax_i + b))^2}{\sigma^2} \right]$$

Main assumptions:

- independent observations (iid)
  *no correlations*
- all the points are equally valid
  *no outliers*
- uncertainties are Gaussian distributed
  *homoscedasticity*
- no exterior information
  *no priors*

Model for the uncertainties

*https://www.albert.io/blog/key-assumptions-of-ols-econometrics-review/*

# Frequentist approach

If {X, Y} are random variables:

X ~ *Uniform*(1.0, 2.0)

Y = aX + b + ε

ε ~ *Normal*(0, σ)

or

Y ~ *Normal*(aX + b, σ)

Likelihood for 1 point:

$$\mathcal{L}(\boxed{x, y|a, b, \sigma}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(y - (ax+b))^2}{2\sigma^2}\right]$$

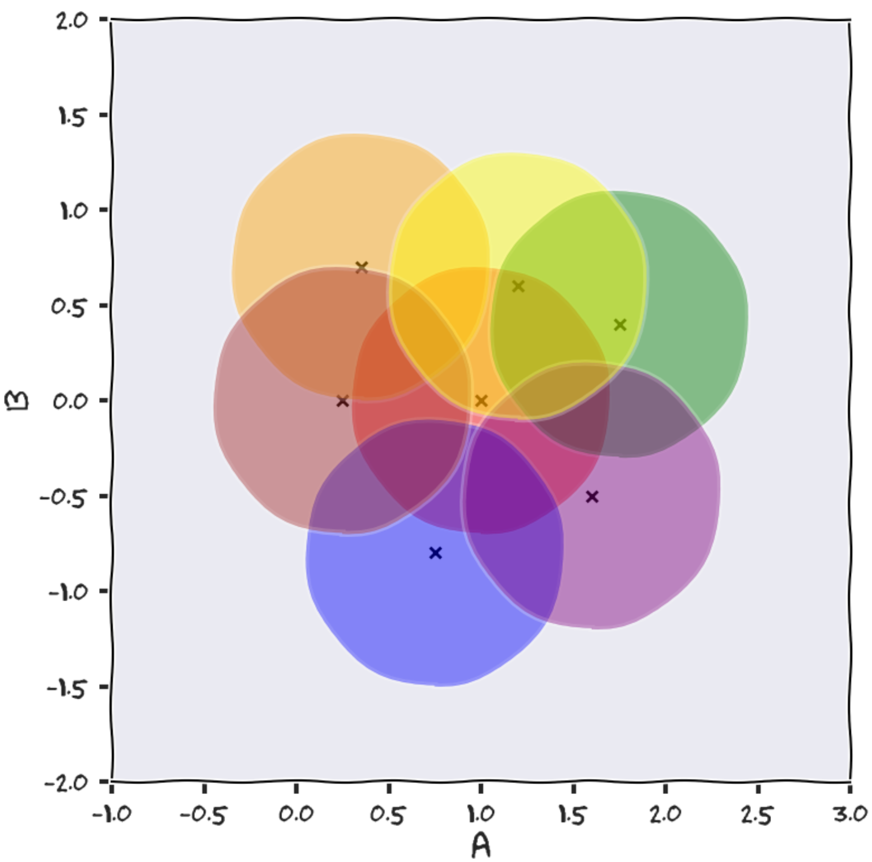
Likelihood for the complete data set:

$$\begin{aligned}\mathcal{L}(\boxed{\vec{x}, \vec{y}|a, b, \sigma}) &= \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(y_i - (ax_i+b))^2}{2\sigma^2}\right] \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\sum_{i=1}^{N} \frac{(y_i - (ax_i+b))^2}{\sigma^2}\right]\end{aligned}$$

# $y = a\,x + b$

Frequentist:

$$\mathcal{L}(\vec{x},\vec{y}|a,b,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\sum_{i=1}^{N}\frac{(y_i-(ax_i+b))^2}{\sigma^2}\right]$$




Frequentist:

95% confidence interval →
95% of the intervals derived from possible data
will contain the true value

# Bayesian approach

If {X, Y} are random variables:

$X \sim \textit{Uniform}(1.0, 2.0)$

$Y = aX + b + \varepsilon$

$\varepsilon \sim \textit{Normal}(0, \sigma)$

or

$Y \sim \textit{Normal}(aX + b, \sigma)$

+

$a \sim \textit{Normal}(0, 5.0)$

$b \sim \textit{Normal}(0, 5.0)$

Bayes theorem

$$
\begin{aligned}
P(A \cap B) &= P(A|B)P(B) \\
P(B \cap A) &= P(B|A)P(A) \\
P(A|B)P(B) &= P(B|A)P(A) \\
P(A|B) &= \frac{P(B|A)P(A)}{P(B)}
\end{aligned}
$$

# Bayesian approach

If $\{X, Y\}$ are random variables:

Bayes theorem

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

$X \sim \textit{Uniform}(1.0, 2.0)$

$Y = aX + b + \varepsilon$

$\varepsilon \sim \textit{Normal}(0, \sigma)$

or

$Y \sim \textit{Normal}(aX + b, \sigma)$

+

$a \sim \textit{Normal}(0, 5.0)$

$b \sim \textit{Normal}(0, 5.0)$

# Bayesian approach

If {X, Y} are random variables:

$X \sim \mathit{Uniform}(1.0, 2.0)$

$Y = aX + b + \varepsilon$

$\varepsilon \sim \mathit{Normal}(0, \sigma)$

or

$Y \sim \mathit{Normal}(aX + b, \sigma)$

+

$a \sim \mathit{Normal}(0, 5.0)$

$b \sim \mathit{Normal}(0, 5.0)$

Bayes theorem

$$P(\vec{\theta}|\mathcal{D}) = \frac{P(\mathcal{D}|\vec{\theta})P(\vec{\theta})}{P(\mathcal{D})}$$

# Bayesian approach

If {X, Y} are random variables:

X ~ *Uniform*(1.0, 2.0)

Y = aX + b + ε

ε ~ *Normal*(0, σ)

or

Y ~ *Normal*(aX + b, σ)

\+

a ~ *Normal*(0, 5.0)

b ~ *Normal*(0, 5.0)

Bayes theorem

$$P(a, b|\vec{x}, \vec{y}, \sigma) = \frac{\mathcal{L}(\vec{x}, \vec{y}|a, b, \sigma)p(a)p(b)}{E(\vec{x}, \vec{y})}$$

# Bayesian approach

If {X, Y} are random variables:

Bayes theorem

$$P(a,b|\vec{x},\vec{y},\sigma) = \frac{\mathcal{L}(\vec{x},\vec{y}|a,b,\sigma)p(a)p(b)}{\int_a \int_b \mathcal{L}(\vec{x},\vec{y}|a,b,\sigma)p(a)p(b)dadb}$$

X ~ *Uniform*(1.0, 2.0)

Y = aX + b + ε

ε ~ *Normal*(0, σ)

or

Y ~ *Normal*(aX + b, σ)

\+

a ~ *Normal*(0, 5.0)

b ~ *Normal*(0, 5.0)

# Bayesian approach

If {X, Y} are random variables:

X ~ *Uniform*(1.0, 2.0)

Y = aX + b + ε

ε ~ *Normal*(0, σ)

or

Y ~ *Normal*(aX + b, σ)

+

a ~ *Normal*(0, 5.0)

b ~ *Normal*(0, 5.0)

Bayes theorem

$$\boxed{P(a,b|\vec{x},\vec{y},\sigma) = \frac{\mathcal{L}(\vec{x},\vec{y}|a,b,\sigma)p(a)p(b)}{\int_a \int_b \mathcal{L}(\vec{x},\vec{y}|a,b,\sigma)p(a)p(b)dadb}}$$

$$\mathcal{L}(\vec{x},\vec{y}|a,b,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\sum_{i=1}^{N}\frac{(y_i-(ax_i+b))^2}{\sigma^2}\right]$$

$$p(a) = \frac{1}{\sqrt{2\pi(5)^2}} \exp\left[-\frac{(a-0)^2}{2\times(5^2)}\right]$$

$$p(b) = \frac{1}{\sqrt{2\pi(5)^2}} \exp\left[-\frac{(b-0)^2}{2\times(5^2)}\right]$$

# Bayesian approach

Likelihood

Posterior

Priors

$$P(a,b|\vec{x},\vec{y},\sigma) = \frac{\mathcal{L}(\vec{x},\vec{y}|a,b,\sigma)p(a)p(b)}{\int_a \int_b \mathcal{L}(\vec{x},\vec{y}|a,b,\sigma)p(a)p(b)\,da\,db}$$

Evidence

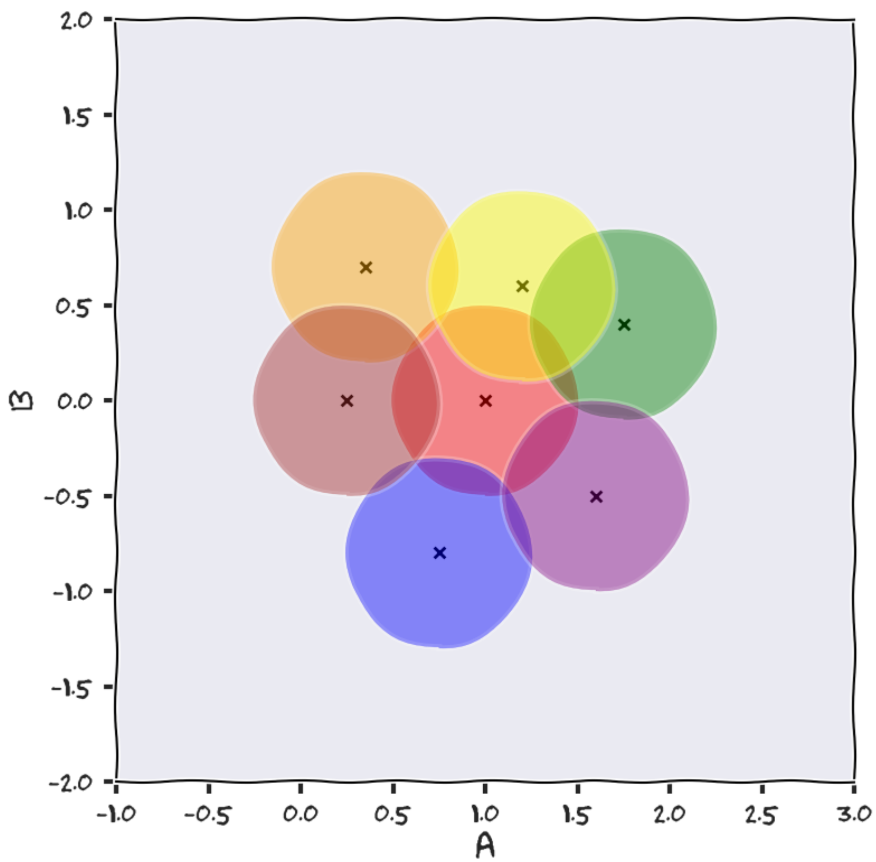
# Bayesian approach

$$P(\boxed{a, b | \vec{x}, \vec{y}, \sigma}) = \frac{\mathcal{L}(\vec{x}, \vec{y} | a, b, \sigma) p(a) p(b)}{\int_a \int_b \mathcal{L}(\vec{x}, \vec{y} | a, b, \sigma) p(a) p(b) da db}$$

# $y = a\,x + b$

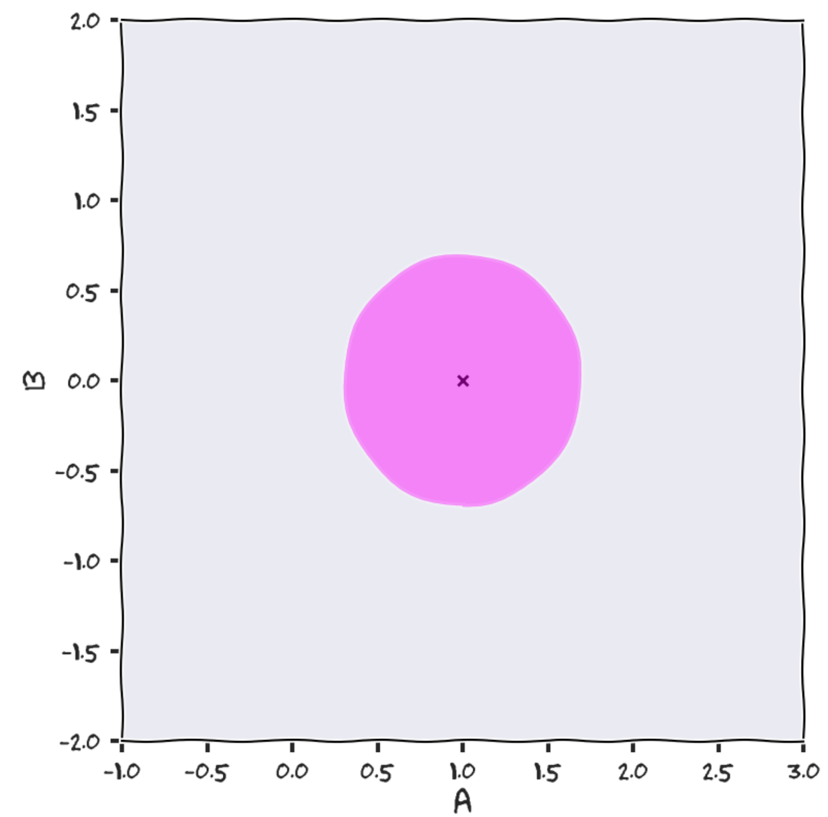
## Frequentist:

$$\mathcal{L}(\vec{x},\vec{y}|a,b,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{1}{2}\sum_{i=1}^{N}\frac{(y_i-(ax_i+b))^2}{\sigma^2}\right]$$




## Bayesian:

$$P(a,b|\vec{x},\vec{y},\sigma) = \frac{\mathcal{L}(\vec{x},\vec{y}|a,b,\sigma)p(a)p(b)}{\int_a\int_b \mathcal{L}(\vec{x},\vec{y}|a,b,\sigma)p(a)p(b)dadb}$$




<u>Frequentist:</u>

95% confidence interval →
*95% of the intervals derived from possible data will contain the true value*

belief

<u>Bayesian:</u>

95% credible interval →
*There is a 95% chance that the true value is within this interval*

# The Bayesian approach is a process…

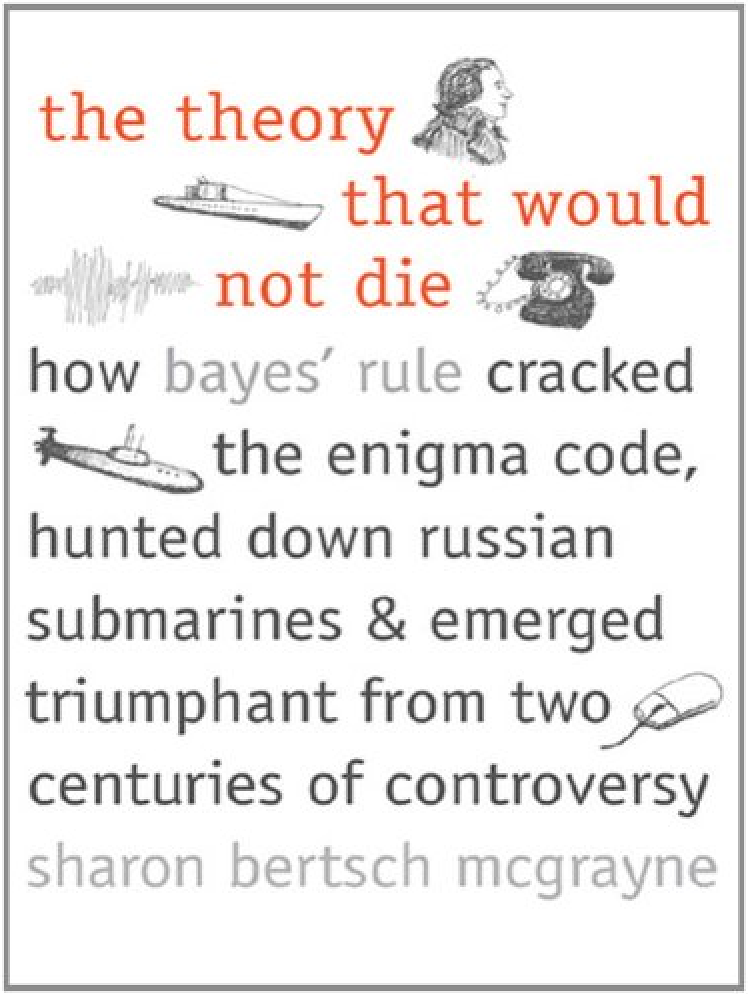
$$P(a,b|\vec{x},\vec{y},\sigma) = \frac{\mathcal{L}(\vec{x},\vec{y}|a,b,\sigma)p(a)p(b)}{\int_a \int_b \mathcal{L}(\vec{x},\vec{y}|a,b,\sigma)p(a)p(b)dadb}$$

# Bayes in the Human Brain


*https://www.ted.com/talks/laura_schulz_the_surprisingly_logical_minds_of_babies*

Can you see the connection between Bayes and this TED talk?

# Bayes in History




*Yale University Press, 2012*

# Bayes in Astronomy

1991 – I Statistical Challenges in Modern Astronomy – Penn State University

*by Eric Feigelson (astronomer) and Jogesh Babu (statistician)*

Contributions: 2 Bayesian/22 total – Pedagogical

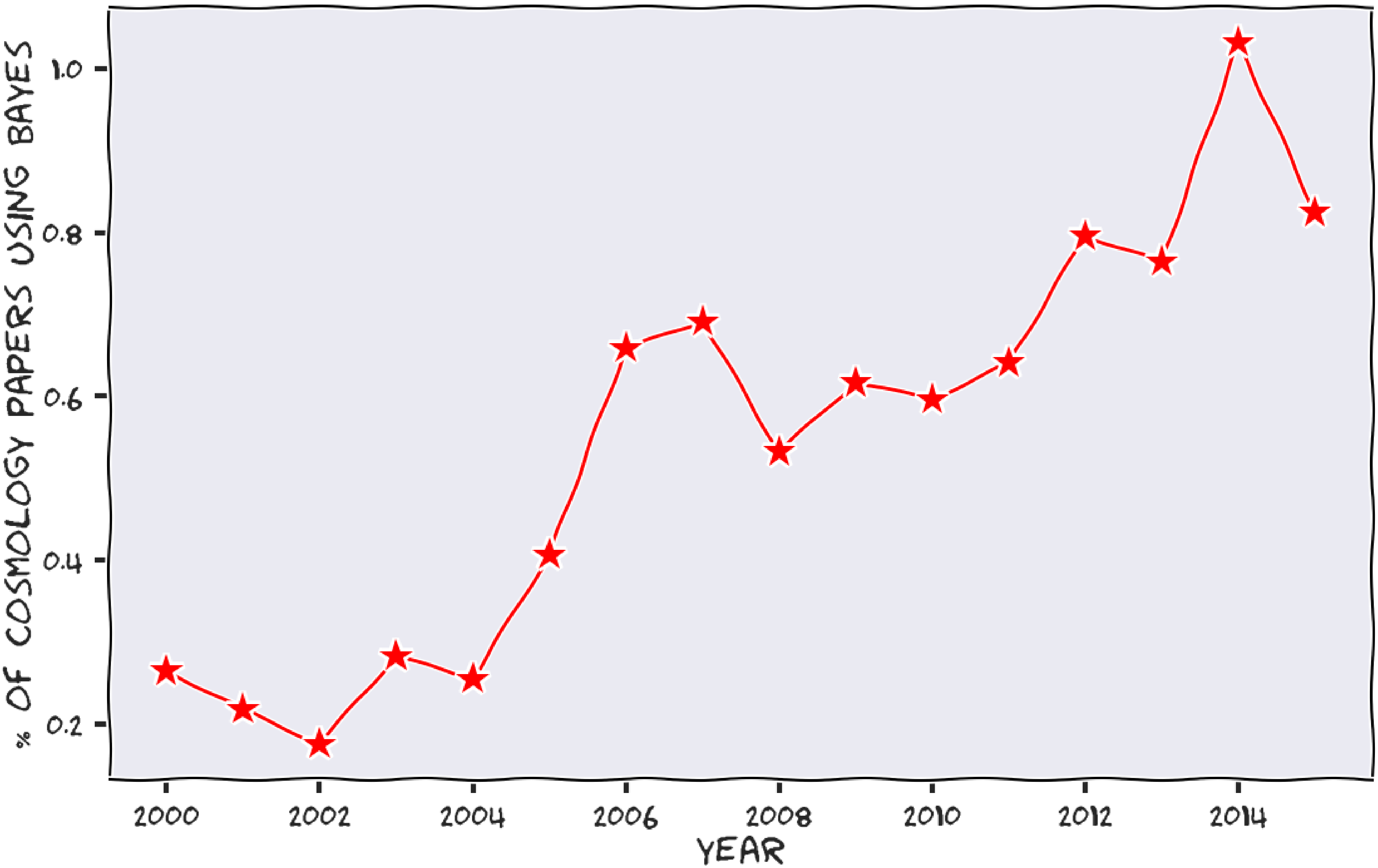
2011 – V Statistical Challenges in Modern Astronomy – Penn State University

*by Eric Feigelson (astronomer) and Jogesh Babu (statistician)*

Contributions: 14 Bayesian/32 total – Results

Tom Loredo, *Bayesian astrostatistics: a backward look to the future*, 2010, arXiv:astro-ph/1208.3036

# Bayes in Cosmology



Frequentist

Bayesian

How to extract information about the parameters of your model?

*To be continued …*

Check reference list and additional material at

*https://github.com/emilleishida/StatisticsInCosmology*