ONE RULE TO RULE THEM ALL

BAYESIAN ANALYSIS IN THE SCIENCES

DR. BRIAN BLAIS

CROSS-DISCIPLINARY SYMPOSIUM

- Fall Cross-Disciplinary Research, Teaching and Technology Symposium
- Date: Wednesday November 9th, 2:00PM – 5:00PM, AIC 118

This event is for faculty who are interested in using tools from data science in their research and teaching, or who are already involved in such work.

Organizers: Dave Louton, Gao Niu, Hakan Saraoglu, Kristin Scaplen, M.L. Tlachac, Chen Zhang, Tingting Zhao

ABSTRACT

Bayesian methods are a mainstay in the sciences, especially in high energy physics and astrophysics but many still see these methods as an arcane, challenging, and overly technical. Bayesian analysis is especially perceived as inappropriate for undergraduate teaching to the point where there are essentially no undergraduate Bayesian textbooks. In this presentation I am going to challenge these notions and provide concrete techniques for doing Bayesian inference that I have used with students for both within the classroom and for research. I will also provide the rationale for why everyone should be using Bayesian techniques, both practically and philosophically.

INTRODUCTION

- I am not a data scientist
- I am a scientist - computational neuroscience, paleoclimate, epidemiology (zombies!), chemical kinetics, anything that interests me
- Messages in this presentation
  - Positive message —
    - Bayesian methods give you a uniform approach to all problems
    - Bayesian methods are easy to interpret
    - Bayesian methods get you to think more deeply
  - Negative message —
    - Standard tools fail
- My teaching and research goal — make technical topics approachable

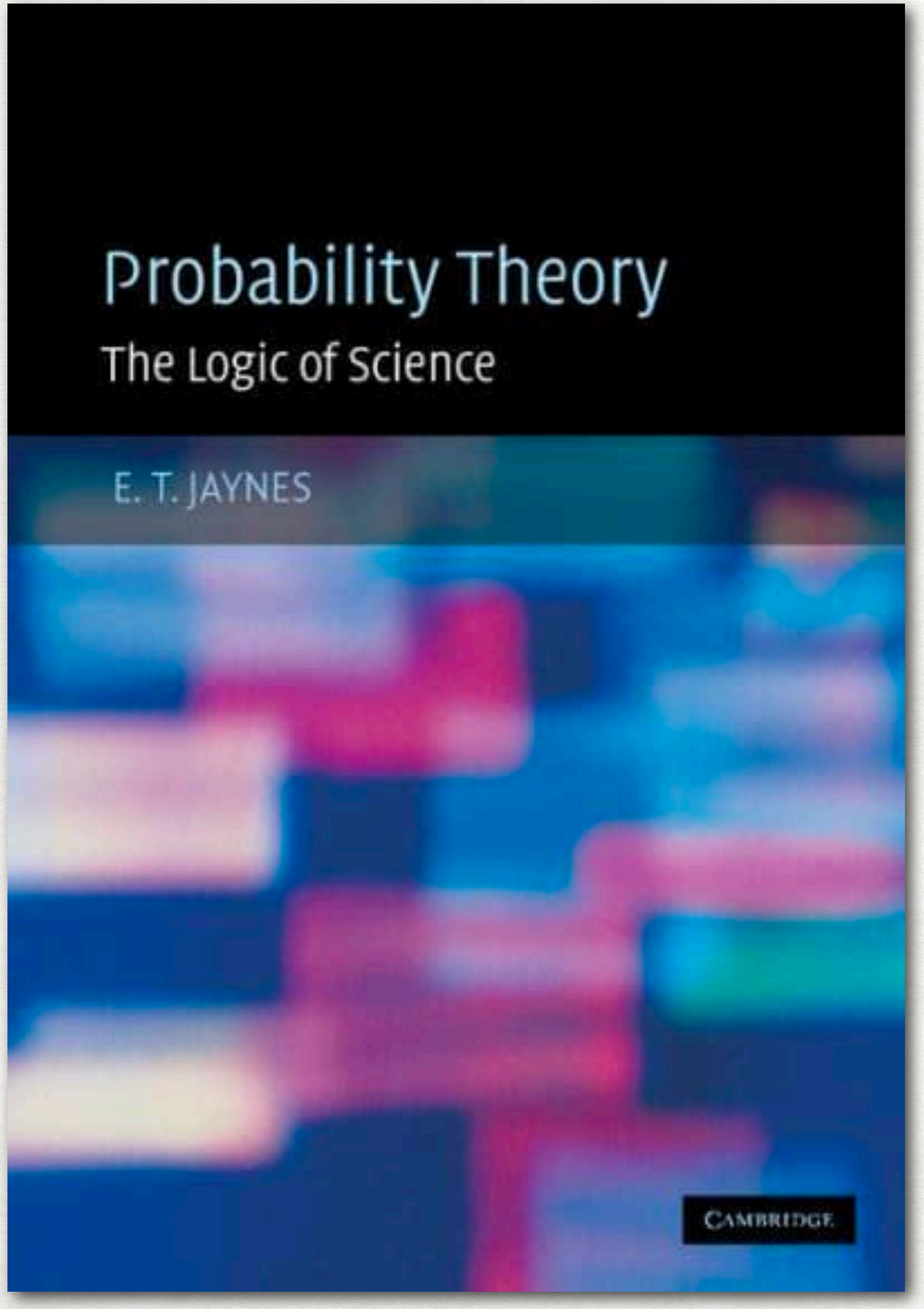
THERE IS STILL RESISTANCE TO BAYESIAN METHODS

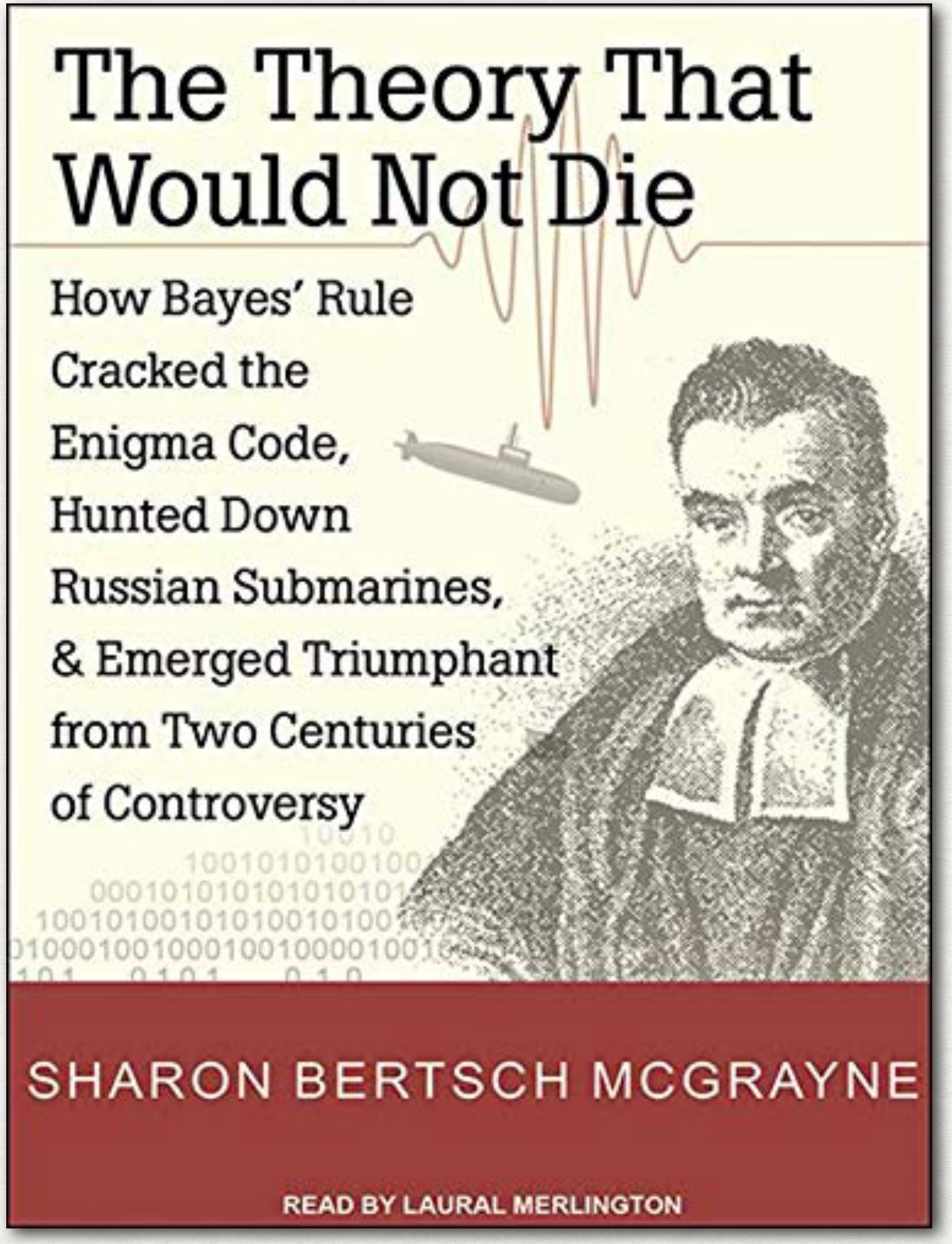
ESPECIALLY IN THE CLASSROOM

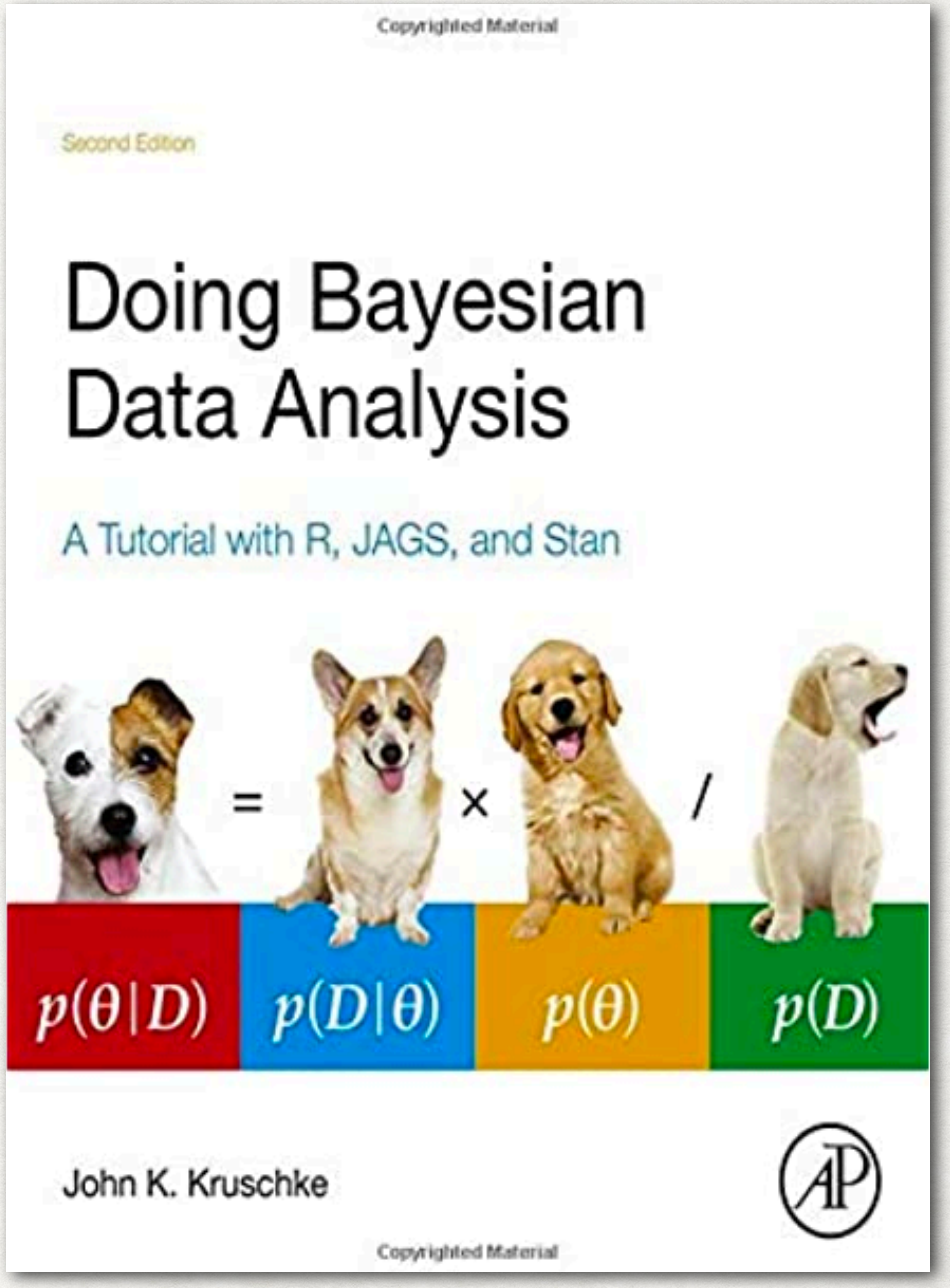
- Reasons
  - “Subjective” priors vs “Objective” frequencies
  - Math is hard
  - Inertia (we’ve always done things this way…)


ONE DOES NOT SIMPLY
ASSUME BAYESIAN METHODS ARE HARD

BOOK RECOMMENDATIONS







WHAT ARE BAYESIAN METHODS?

- Application of probability theory as an extension of logic
- "Probability theory is nothing but common sense reduced to calculation." - Laplace
- Bayes' Rule, Bayes Theorem, etc… is just an algebraic step from the multiplication rule of probability

RULES OF PROBABILITY

- $p(A) = 0$ certain that A is false
- $p(A) = 1$ certain that A is true
- Limited Sum Rule $p(A) + p(\overline{A}) = 1$
- Full Sum Rule ("or") $p(A + B) \quad = \quad p(A) + p(B) - p(AB)$
- Product Rule ("and")

$$\begin{aligned} p(AB) &= p(A|B)p(B) \\ &= p(B|A)p(A) \end{aligned}$$

- Bayes Rule

$$\underbrace{p(A|B)}_{\text{posterior}} = \frac{\overbrace{p(B|A)}^{\text{likelihood}}\ \overbrace{p(A)}^{\text{prior}}}{\underbrace{p(B)}_{\text{normalization}}}$$

PLAUSIBILITY AND AXIOMS

E. T. JAYNES, 2003

- (I) Degrees of plausibility are represented by real numbers
- (II) Qualitative corresp
  - (a) direction of value
  - (b) consistent with t
- (IIIa) If a conclusion ca … must lead to the same result.
- (IIIb) The robot always … It does not arbitrarily ignore some of the inf … ot is non-ideological
- (IIIc) The robot always represents equivalent states if knowledge by equivalent plausibility assignments. That is, if in two problems the robot's state of knowledge is the same (except perhaps for the labeling of the propositions), then it must assign the same plausibilities in both

Amazingly, these few axioms are enough to specify a completely consistent, mathematical framework for plausibilities…

…and this mathematical framework is exactly the same as the rules developed by Laplace for probabilities

FORMS OF BAYES

This isn't just a method to do quantitative analysis, it is a general way to structure rational thought. In other words, to not think this way will violate one of Jaynes' axioms… and one will, by definition, be irrational.

evidence

old knowledge

$$p(\text{model}|\text{data}) = \frac{p(\text{data}|\text{model}) \cdot p(\text{model})}{p(\text{data})}$$

alternatives

updated knowledge

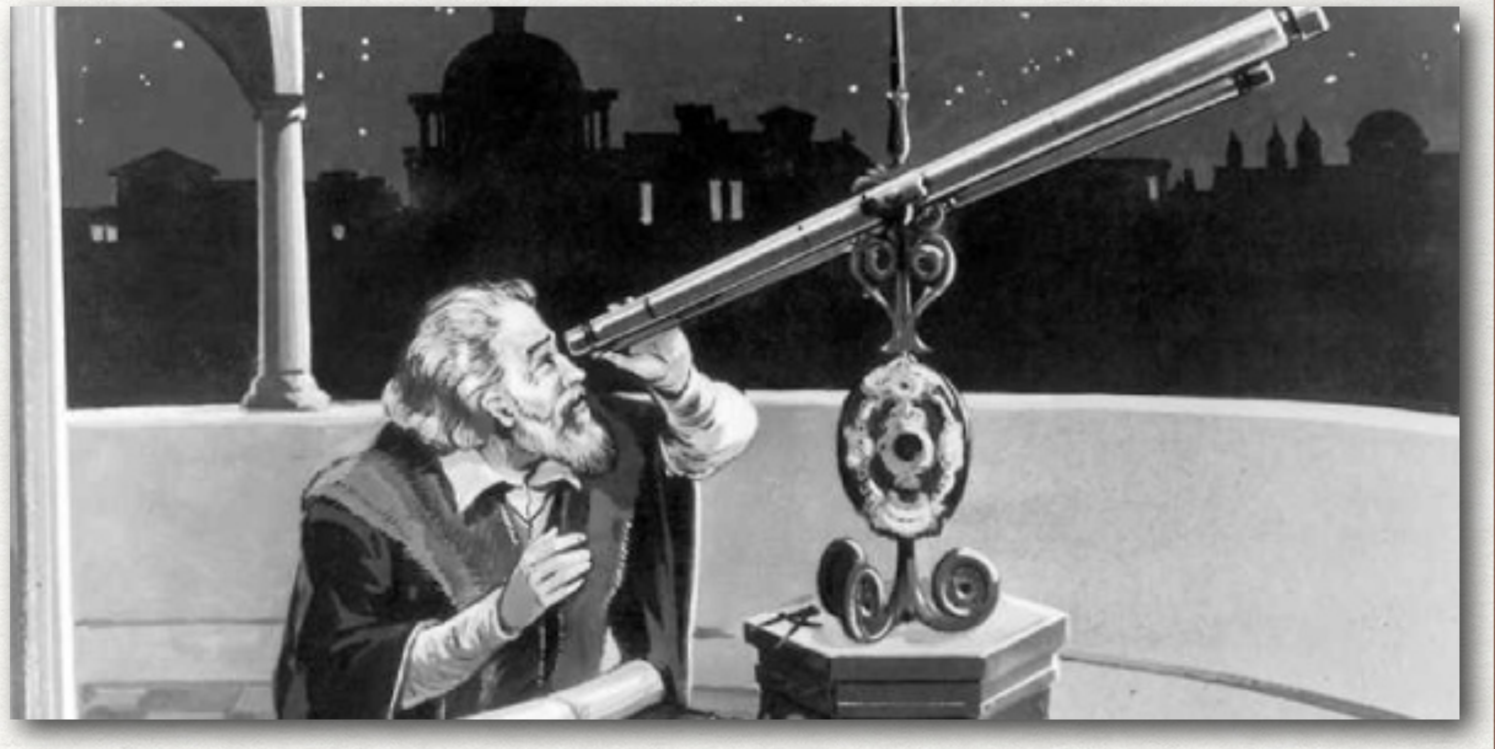
COMPARISON WITH ORTHODOX STATISTICS

AKA FREQUENTIST STATISTICS

Frequentist aka Orthodox Statistics aka Standard Methods	**Bayesian aka Probability Theory as Logic**
P(A) = long-run relative frequency of occurrence of A in a sequence of "identical" repetitions	P(A) = real-number value of the plausibility of A with incomplete information
Testing a hypothesis, e.g. true (or population) mean $\mu=0$, from a sample, e.g. x_1 , x_2 , x_3 , …, one *imagines* many repetitions of the sample and compares the sample mean in these repetitions to the hypothesis.	Testing a hypothesis, e.g. true (or population) mean $\mu=0$, from a sample, e.g. x_1 , x_2 , x_3 , …, one looks at the probability of that hypothesis given the data, $P(\mu \mid x_1, x_2, x_3, \ldots)$

GALILEAN PROBLEMS

- What Galileo did with his telescope was to take something that was invisible and magnify it so that one can easily see the truth. He used it to compare different approaches to explaining the cosmos.



- A Galilean problem is one which is small enough that ones intuition is enough to determine its truth or falsity.

EASY PROBLEMS

TRUE VALUE WITH KNOWN NOISE

- Data: $\{x_i\} = \{12, 14, 16\}, \sigma = 1$
- Question: Is the true (population) value , μ, less than 13?

STANDARD (NON-BAYESIAN) SETUP

Z-TEST

- Choose a statistic (i.e. some function of the data) with certain properties you'd like (sufficiency, unbiased, etc…)
- In this case we use the sample mean, $\bar{x}$
- The sample mean has a distribution (in the long run, or over a large population) that is Normal with mean of the true value, μ, and standard deviation $\sigma/\sqrt{N}$.
- In the distribution of the (imagined) population, where does our hypothesis fall?
- Distribution always over *data where the hypothesis (i.e. parameter) is always* constant

STANDARD (NON-BAYESIAN) COMPUTATION

Z-TEST

- Data: $\{x_i\} = \{12,14,16\}, \sigma = 1$
- True value (μ) less than 13?

$$\bar{x} = 14, \hat{\sigma} = 1/\sqrt{3}$$

$$z = \frac{14 - 13}{1/\sqrt{3}} = 1.732$$

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-1.2	.11507	.11314	.11123	.10935	.10749	.10565	.10383	.10204	.10027	.09853
-1.3	.09680	.09510	.09342	.09176	.09012	.08851	.08692	.08534	.08379	.08226
-1.4	.08076	.07927	.07780	.07636	.07493	.07353	.07215	.07078	.06944	.06811
-1.5	.06681	.06552	.06426	.06301	.06178	.06057	.05938	.05821	.05705	.05592
-1.6	.05480	.05370	.05262	.05155	.05050	.04947	.04846	.04746	.04648	.04551
-1.7	.04457	.04363	.04272	.04182	.04093	.04006	.03920	.03836	.03754	.03673
-1.8	.03593	.03515	.03438	.03362	.03288	.03216	.03144	.03074	.03005	.02938
-1.9	.02872	.02807	.02743	.02680	.02619	.02559	.02500	.02442	.02385	.02330
-2	.02275	.02222	.02169	.02118	.02068	.02018	.01970	.01923	.01876	.01831
-2.1	.01786	.01743	.01700	.01659	.01618	.01578	.01539	.01500	.01463	.01426
-2.2	.01390	.01355	.01321	.01287	.01255	.01222	.01191	.01160	.01130	.01101
-2.3	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

A STRATEGIC CHOICE FOR TEACHING

- Use the computer for calculations
- For Bayesian calculations — use MCMC
- Otherwise, one can get lost in analysis

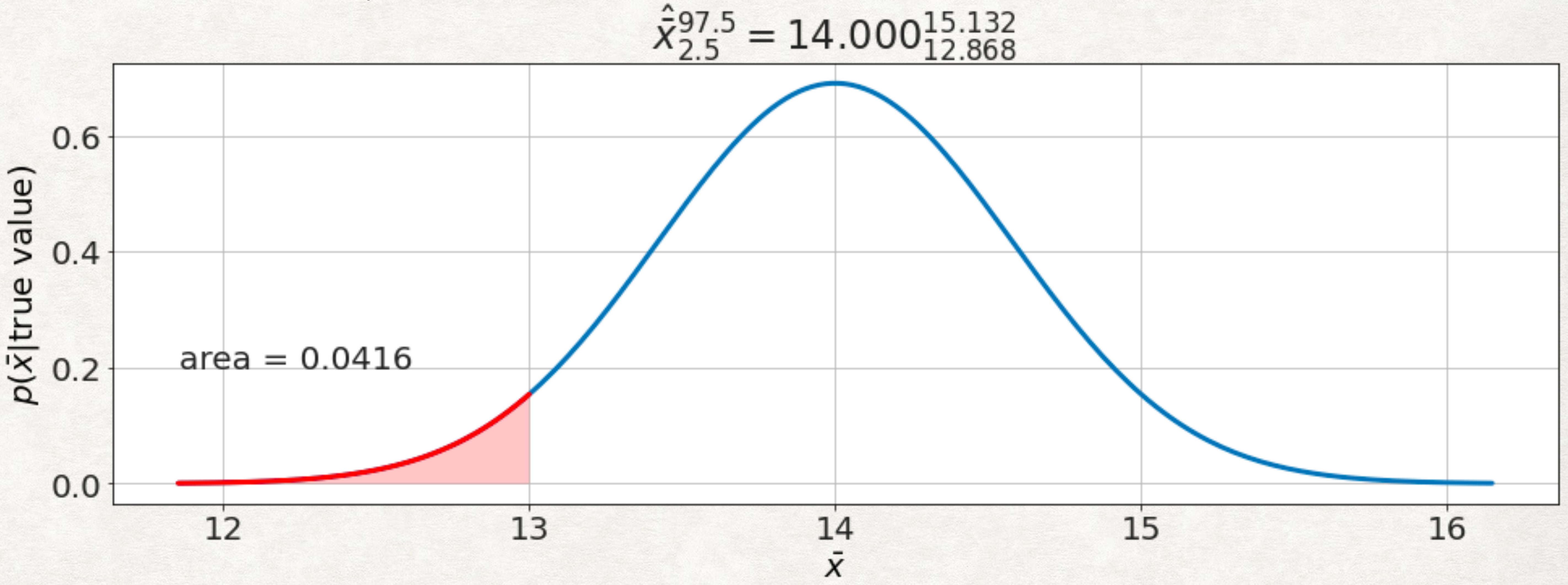
STANDARD (NON-BAYESIAN) COMPUTATION

Z-TEST WITH PYTHON

- Data: $\{x_i\} = \{12,14,16\}, \sigma = 1$

$$\bar{x} = 14, \hat{\sigma} = 1/\sqrt{3}$$

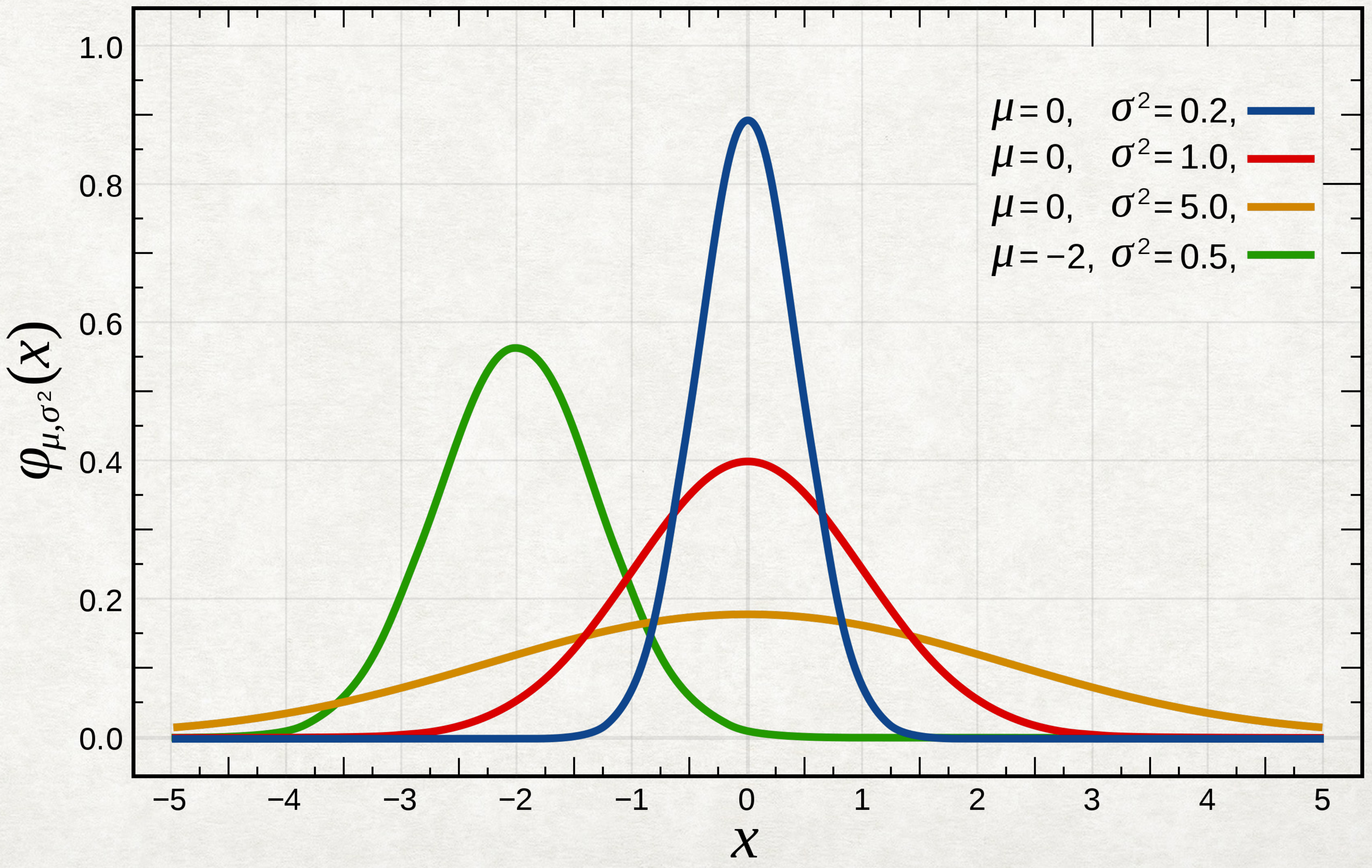
```
x=[12,14,16]
σ=1
N=len(x)
dist=Normal(mean=mean(x),std=σ/sqrt(N))

plot_distribution(dist,fill_left=13)
```



BAYESIAN SETUP

NORMAL NOISE, ESTIMATE LOCATION PARAMETER

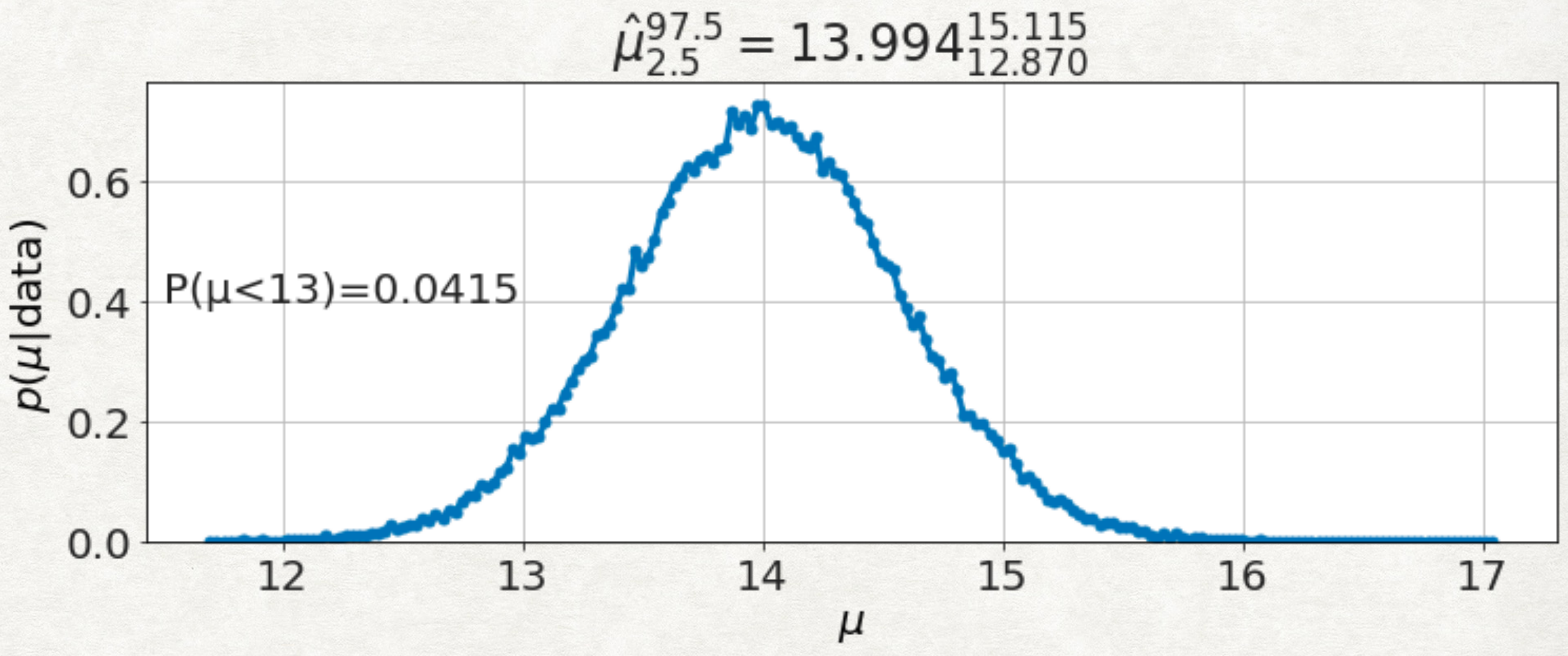
- Data: $\{x_i\} = \{12,14,16\}, \sigma = 1$
- Bayes $P(\mu \,|\, \{x_i\}, \sigma) \sim \prod_i$ likelihood$(x_i \,|\, \mu, \sigma) \times$ prior(μ)
- Likelihood: $P(x_i \,|\, \mu, \sigma) \sim$ Normal$(x_i - \mu, \sigma)$
    - $P(\{x_i\} \,|\, \mu, \sigma) \sim \prod_i$ Normal$(x_i - \mu, \sigma)$
- Prior: $P(\mu) \sim$ Uniform(μ)
- Distribution over *parameter* not *data*



BAYESIAN COMPUTATION

NORMAL NOISE, ESTIMATE LOCATION PARAMETER

```
def lnlike(data,μ):
    x=data
    return lognormalpdf(x,μ,σ)

data=array([12.0,14,16])
σ=1
model=MCMCModel(data,lnlike,
                μ=Uniform(-50,50),
                )
model.run_mcmc(2000,repeat=2)
model.plot_distributions()
```



EASY PROBLEMS

TRUE VALUE WITH UNKNOWN NOISE

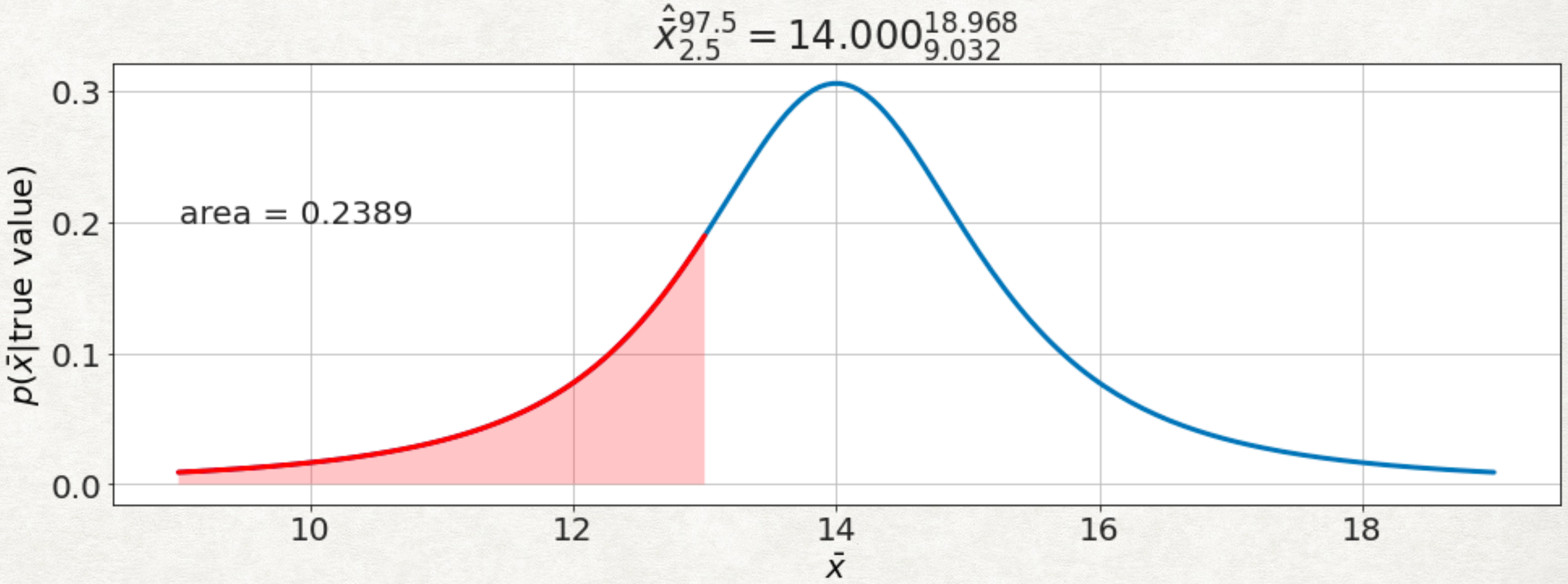
- Basis for the Student-T test
- Data: $\{x_i\} = \{12,14,16\}$

STANDARD (NON-BAYESIAN) COMPUTATION

T-TEST

- Basis for the Student-T test
- Data: $\{x_i\} = \{12,14,16\}$

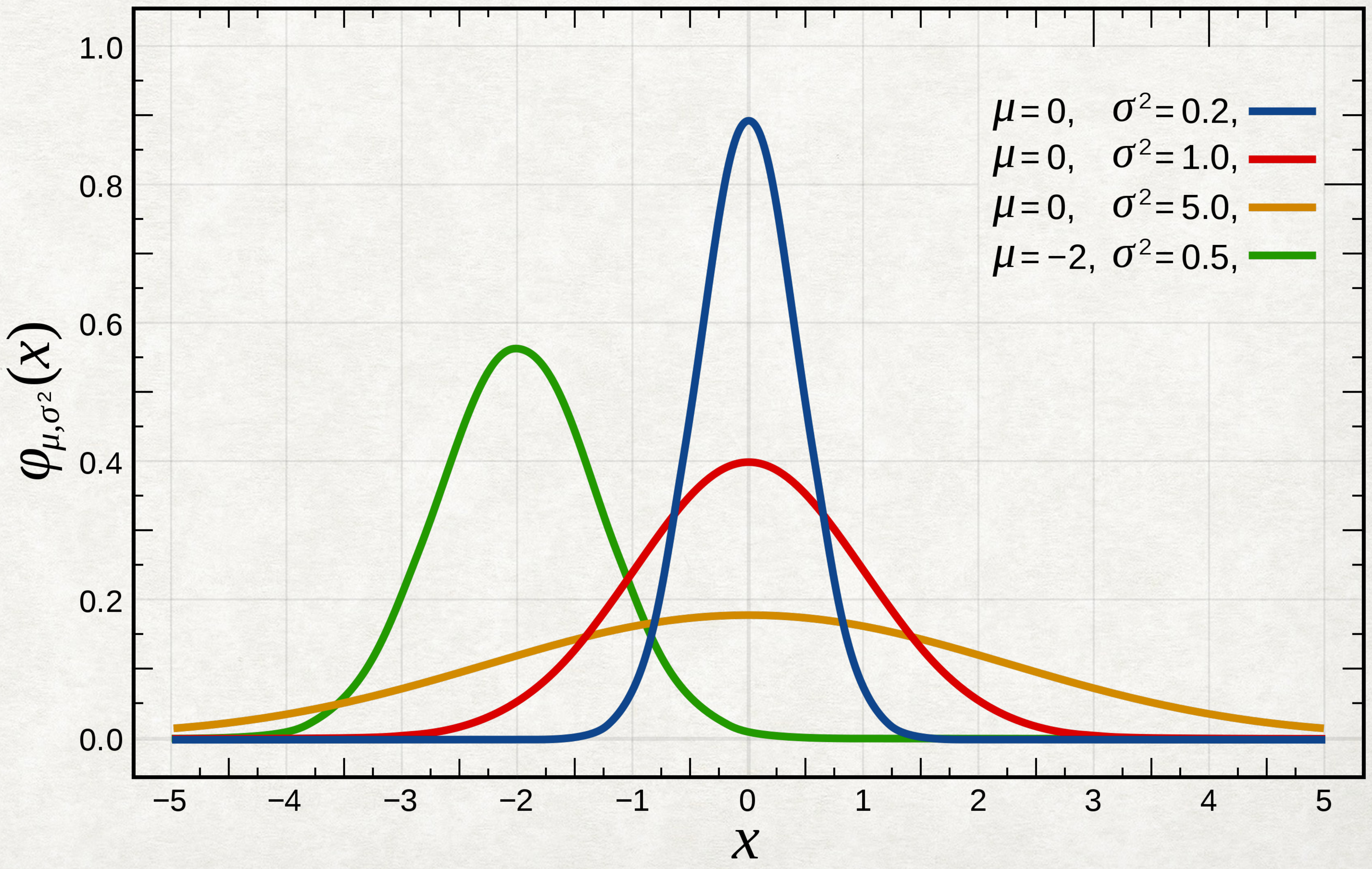
```
x=[12,14,16]
dof=len(x)

dist=StudentT(mean=mean(x),
              std=std(x)/sqrt(N-1),
              dof=N-1)

plot_distribution(dist,fill_left=13,xlim=[9,19])
```



BAYESIAN SETUP

NORMAL NOISE, ESTIMATE LOCATION AND SCALE PARAMETER

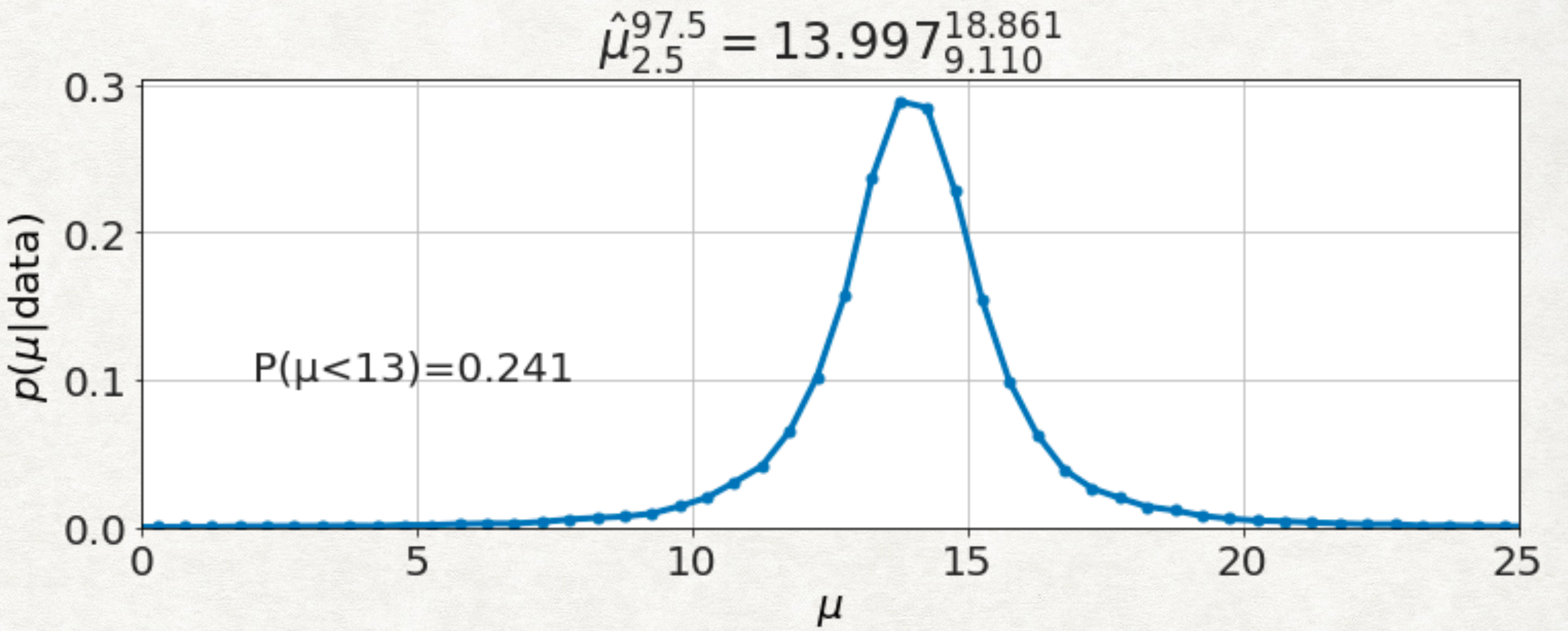
- Data: $\{x_i\} = \{12{,}14{,}16\}$
- Bayes: $P(\mu, \sigma \,|\, \{x_i\}) \sim \prod_i \text{likelihood}(x_i \,|\, \mu, \sigma) \times \text{prior}(\mu, \sigma)$
- Likelihood: $P(x_i \,|\, \mu, \sigma) \sim \text{Normal}(x_i - \mu, \sigma)$
    - $P(\{x_i\} \,|\, \mu, \sigma) \sim \prod_i \text{Normal}(x_i - \mu, \sigma)$
- Prior: $P(\mu, \sigma) \sim \text{Uniform}(\mu) \times \text{Uniform}(\log(\sigma))$

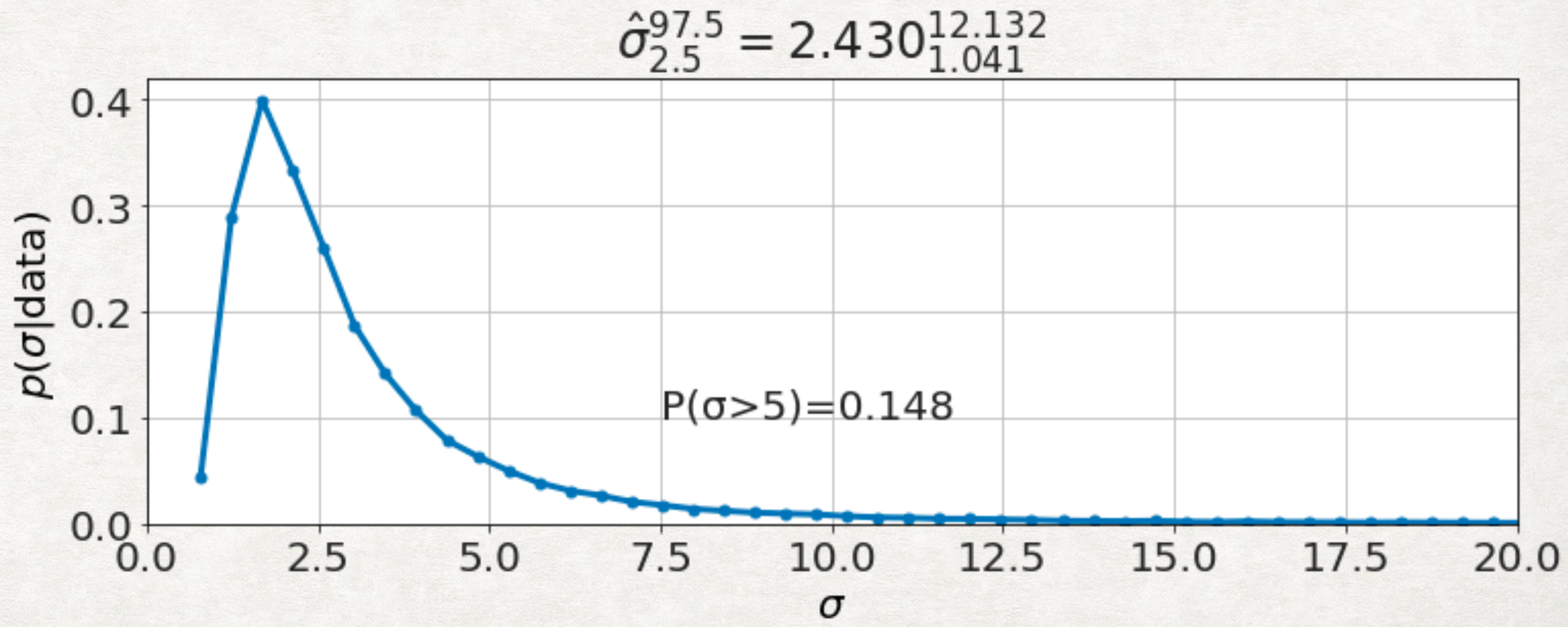


BAYESIAN COMPUTATION

NORMAL NOISE, ESTIMATE LOCATION AND SCALE PARAMETER

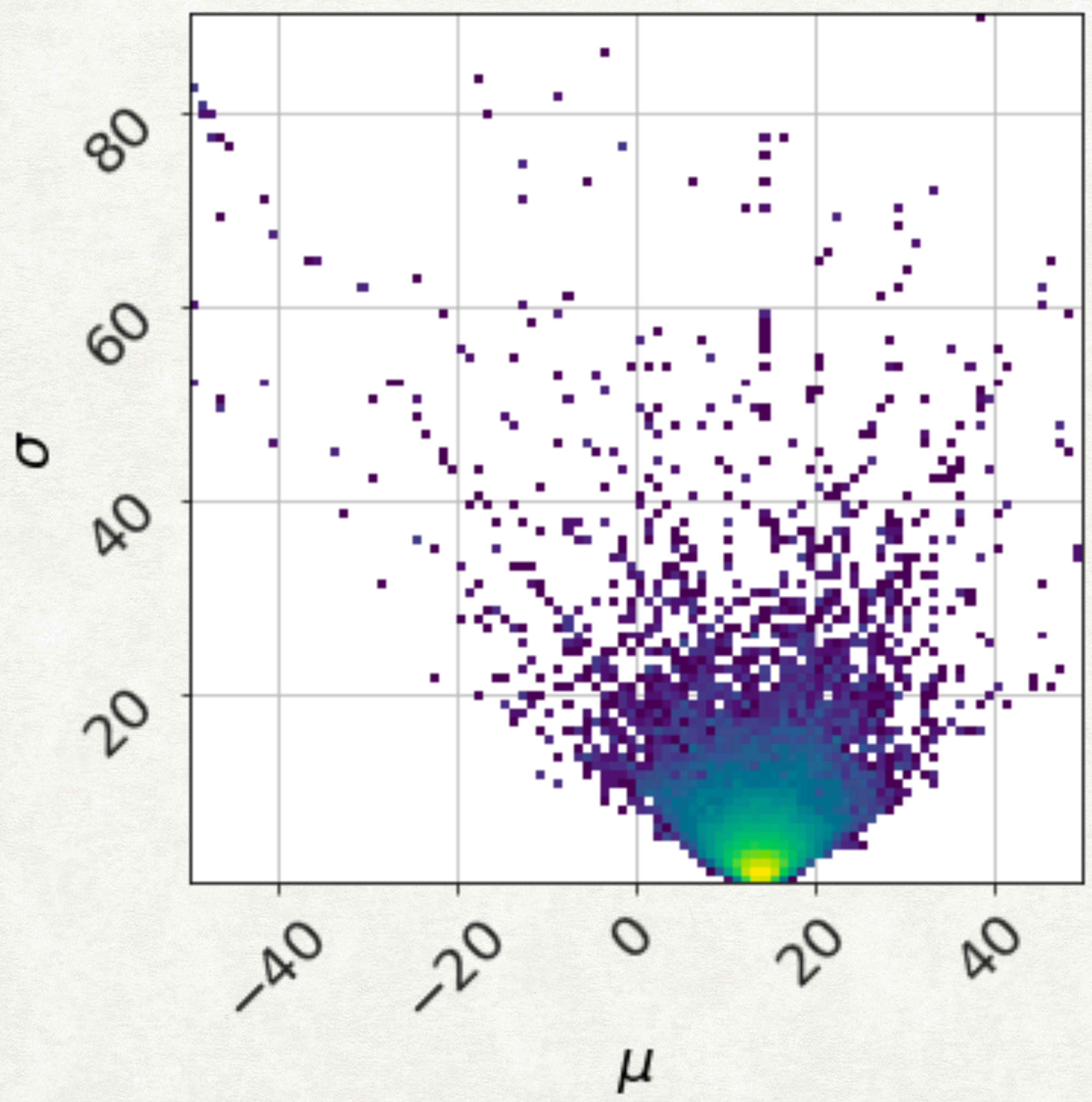
```python
def lnlike(data,μ,σ):
    x=data
    return lognormalpdf(x,μ,σ)

data=array([12.0,14,16])

model=MCMCModel(data,lnlike,
                μ=Uniform(-50,50),
                σ=Jeffreys(),
                )
model.run_mcmc(2000,repeat=2)
```

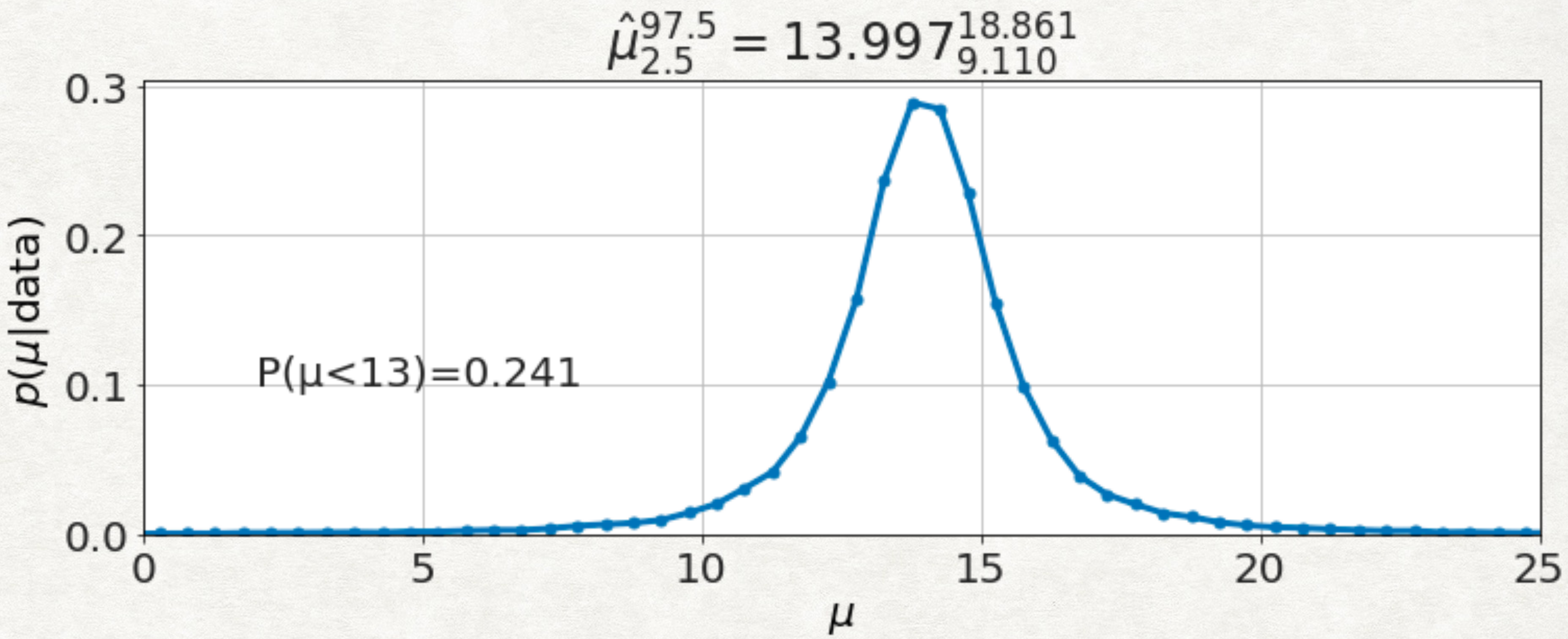


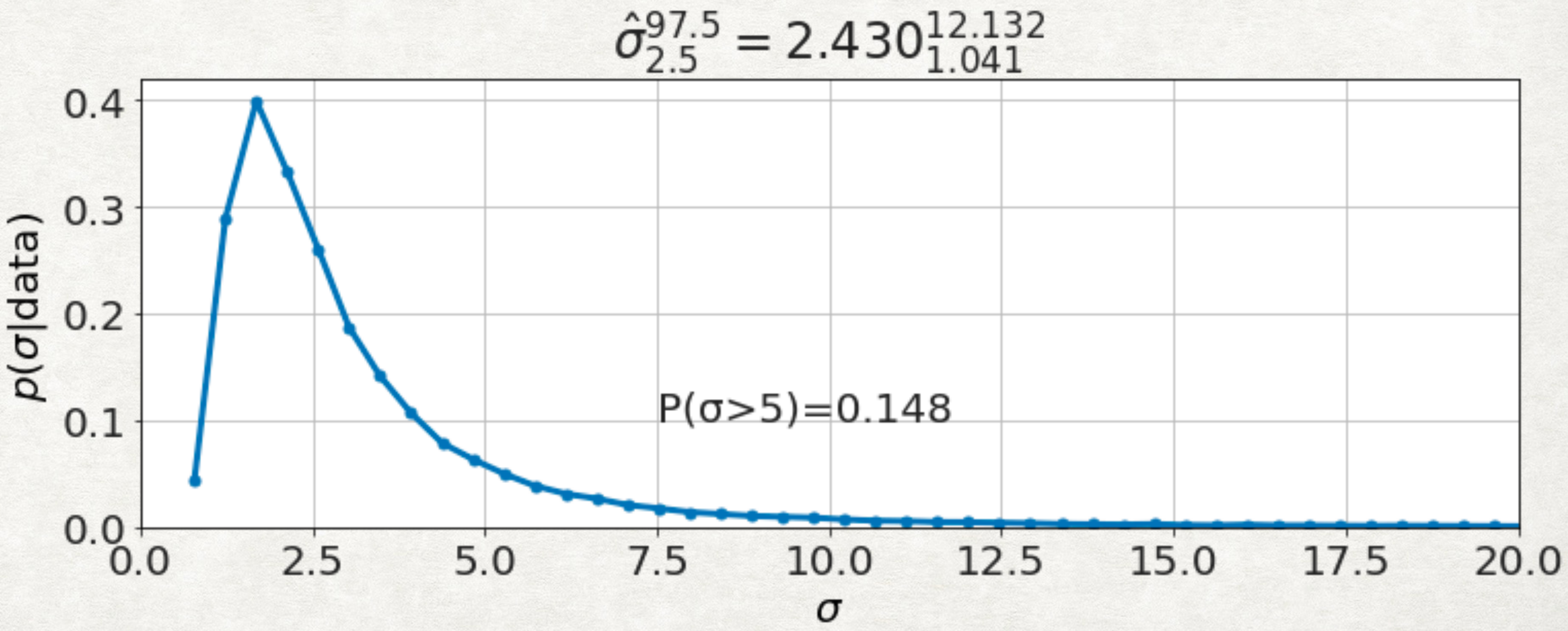


BAYESIAN COMPUTATION

NORMAL NOISE, ESTIMATE LOCATION AND SCALE PARAMETER







THE LIGHTHOUSE PROBLEM

THE LIGHTHOUSE PROBLEM

EASY PROBLEMS

TRUE VALUE WITH UNKNOWN CAUCHY DISTRIBUTED NOISE

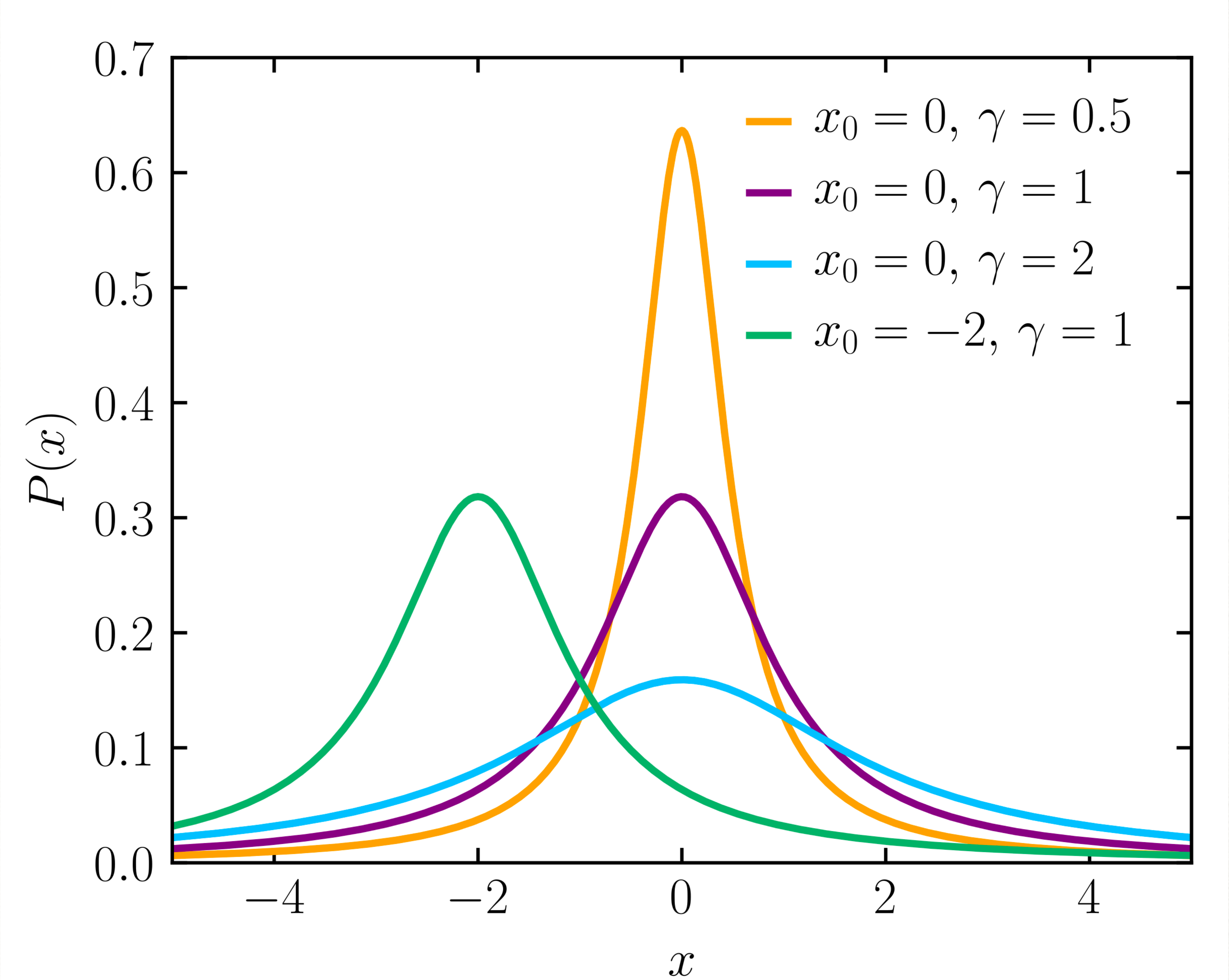
- Data: $\{x_i\} = \{12,14,16\}$

$$P(x_i \,|\, \mu, \sigma) \sim \text{Cauchy}(x_i - \mu, \gamma)$$

- Likelihood: $$\sim \frac{1}{\pi\gamma}\left(\frac{\gamma^2}{(x-\mu)^2+\gamma^2}\right)$$

- Prior: $P(\mu, \gamma) \sim \text{Uniform}(\mu) \times \text{Uniform}(\log(\gamma))$

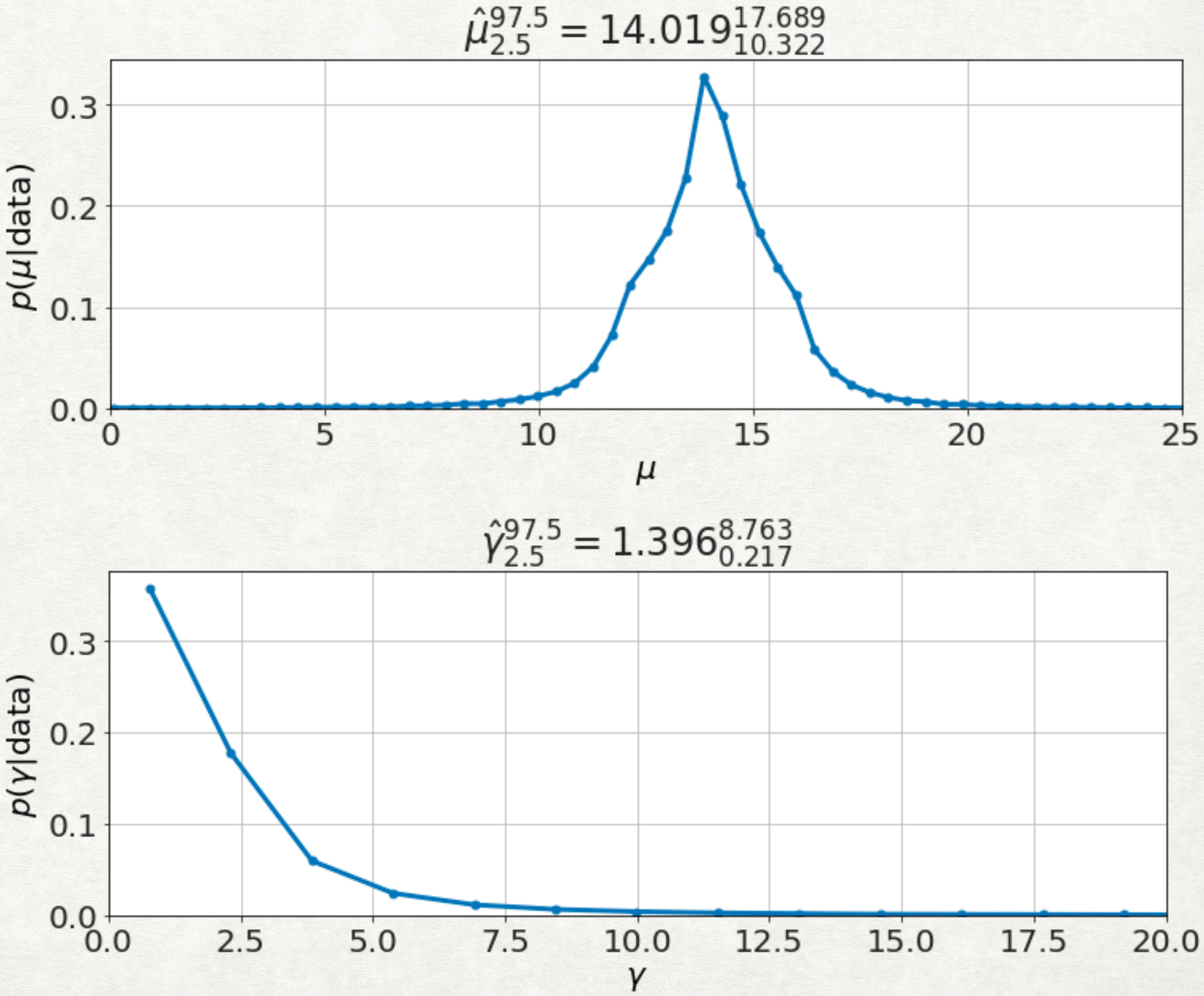
- Bayes: $$P(\mu, \gamma \,|\, \{x_i\}) = \frac{\prod_i \text{Cauchy}(x_i - \mu, \gamma) \times P(\mu, \gamma)}{P(\{x_i\})}$$



BAYESIAN COMPUTATION

CAUCHY NOISE, ESTIMATE LOCATION AND SCALE PARAMETER

```
def lnlike(data,μ,γ):
    x=data
    return logcauchypdf(x,μ,γ)

data=array([12.0,14,16])

model=MCMCModel(data,lnlike,
                μ=Uniform(-50,50),
                γ=Jeffreys(),
                )
model.run_mcmc(2000,repeat=2)
```



THERE IS NO ORTHODOX/FREQUENTIST SOLUTION

- Here we have a trivial problem, only slightly more complex than the easiest problem addressed in an introductory Statistics course and *none of the tools presented in undergraduate Statistics are able to address it*.

Parameters	x_0 location (real) $\gamma > 0$ scale (real)
Support	$x \in (-\infty, +\infty)$
PDF	$\frac{1}{\pi\gamma\left[1+\left(\frac{x-x_0}{\gamma}\right)^2\right]}$
CDF	$\frac{1}{\pi}\arctan\left(\frac{x-x_0}{\gamma}\right)+\frac{1}{2}$
Quantile	$x_0+\gamma\,\tan[\pi(p-\frac{1}{2})]$
Mean	undefined
Median	x_0
Mode	x_0
Variance	undefined
MAD	γ
Skewness	undefined
Ex. kurtosis	undefined
Entropy	$\log(4\pi\gamma)$
MGF	does not exist
CF	$\exp(x_0\,i\,t-\gamma\,\|t\|)$
Fisher information	$\frac{1}{2\gamma^2}$

THERE IS NO ORTHODOX/FREQUENTIST SOLUTION

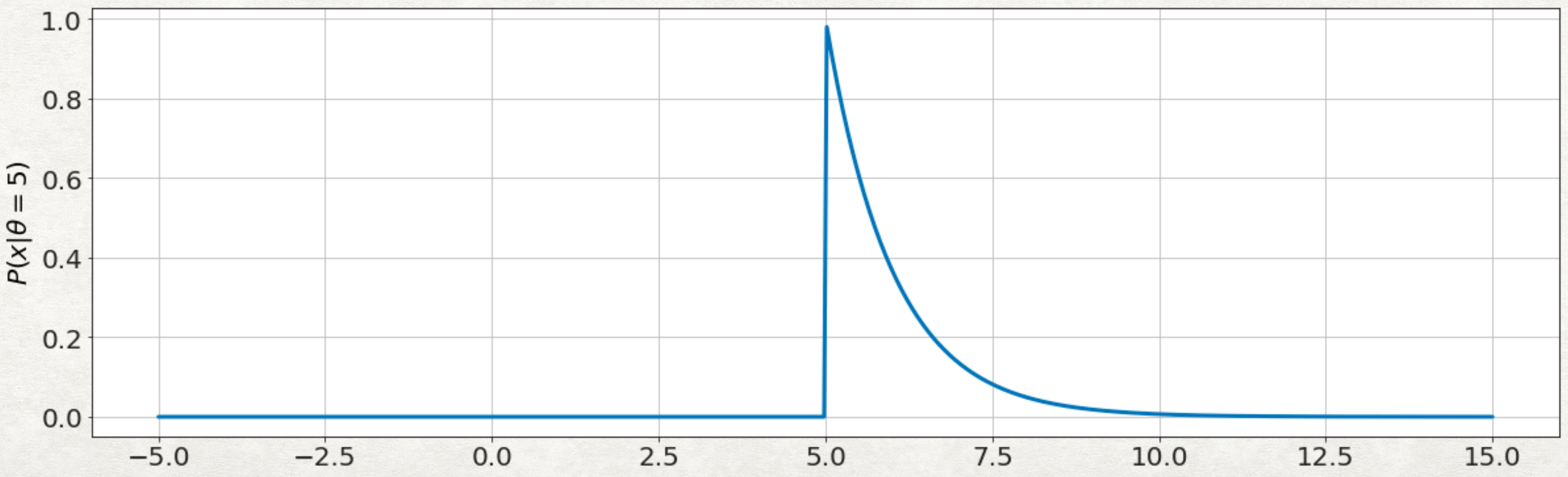
- Here we have a trivial problem, only slightly more complex than the easiest problem addressed in an introductory Statistics course and *none of the tools presented in undergraduate Statistics are able to address it*. (Because the sampling distribution is pathological)

- The Bayesian solution is a straightforward process and can easily be seen as a slight generalization from the easiest problem.

- This isn't the worst problem with the standard statistical tools!

A WORSE ORTHODOX/FREQUENTIST SOLUTION

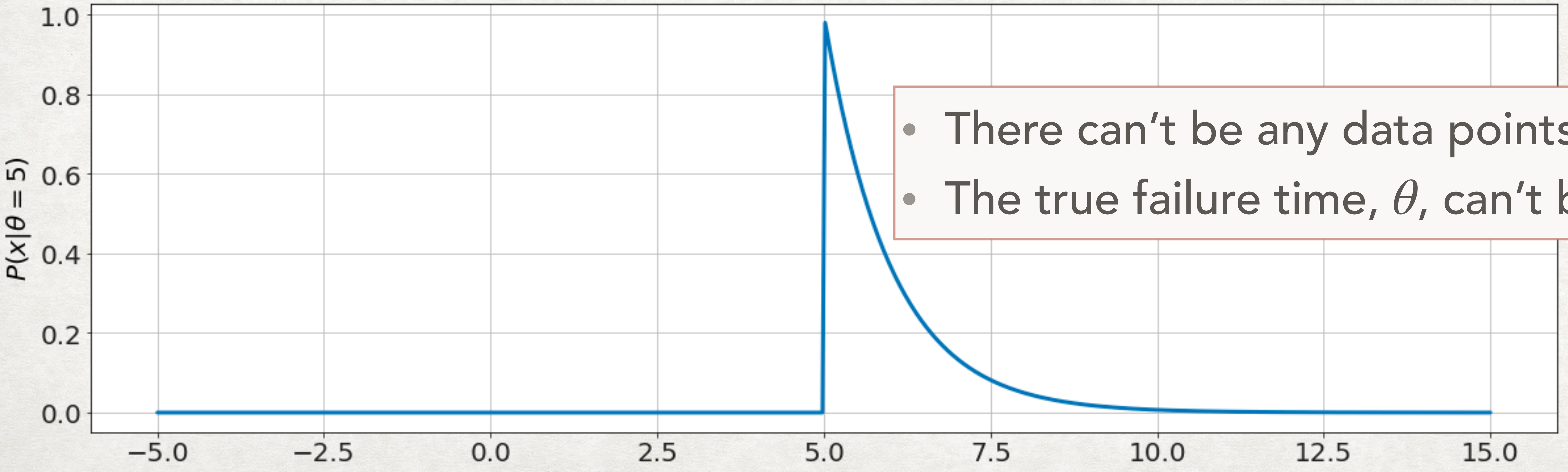
- (E T Jaynes) A device will operate without failure for a time, θ, because of a protective chemical inhibitor injected into it; but at time, θ, the supply of this chemical is exhausted, and failures then commence, following the exponential failure law. It is not feasible to observe the depletion of this inhibitor directly; one can observe only the resulting failures. From data on actual failure times, estimate the time θ of guaranteed safe operation by a confidence interval. Here we have a continuous sample space, and we are to estimate a location parameter, θ, from the sample values $\{x_i\} = \{12,14,16\}$.



EASY PROBLEMS

TRUE VALUE EXPONENTIAL DISTRIBUTION

- Data: $\{x_i\} = \{12, 14, 16\}$
- Likelihood: $p(x \mid \theta)dx = \begin{Bmatrix} \exp(\theta - x)\mathrm{d}x, & x > \theta \\ 0 & x < \theta \end{Bmatrix}$
- Prior: $P(\theta) \sim$ Uniform(θ)
- Bayes: $P(\theta \mid \{x_i\}) = \dfrac{\prod_i p(x_i - \theta) \times P(\theta)}{P(\{x_i\})}$



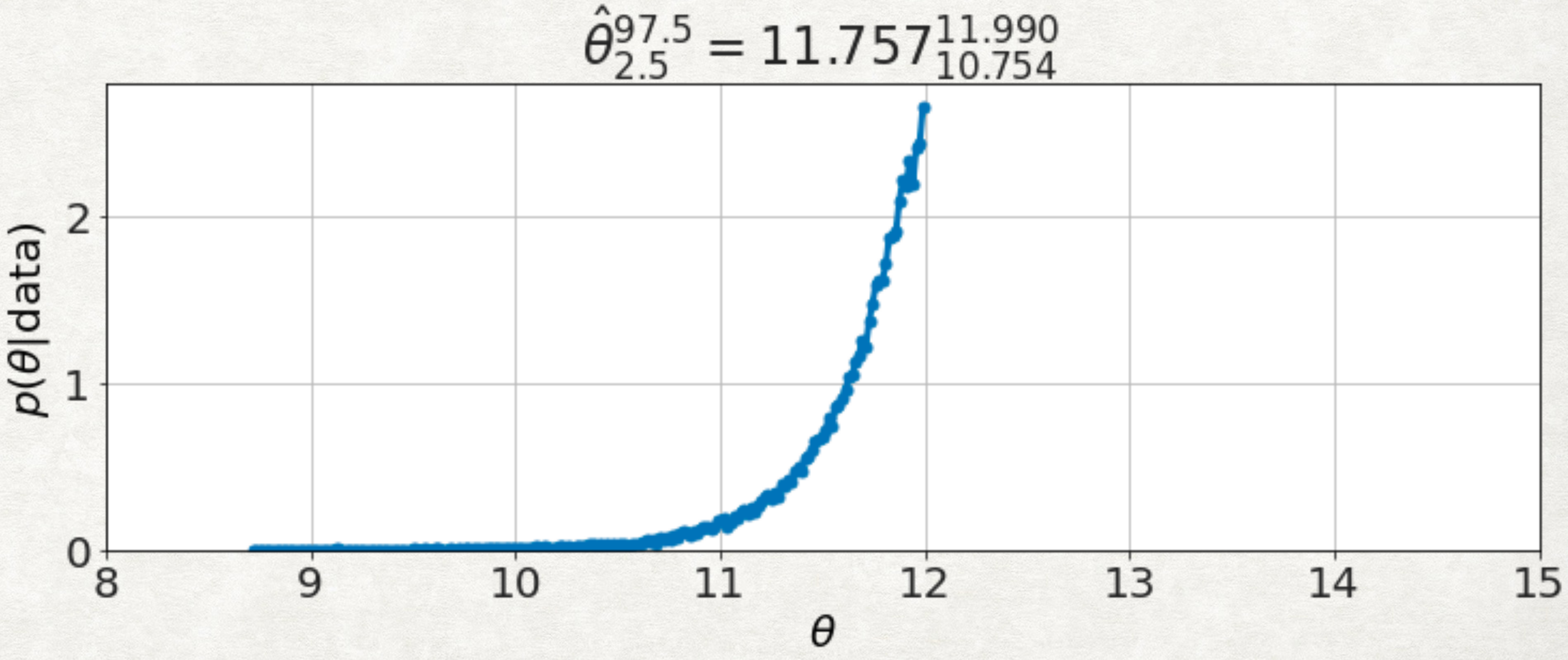
- There can't be any data points less than the failure time, $x_i < \theta$
- The true failure time, θ, can't be greater than any of the data points, x_i

BAYESIAN COMPUTATION

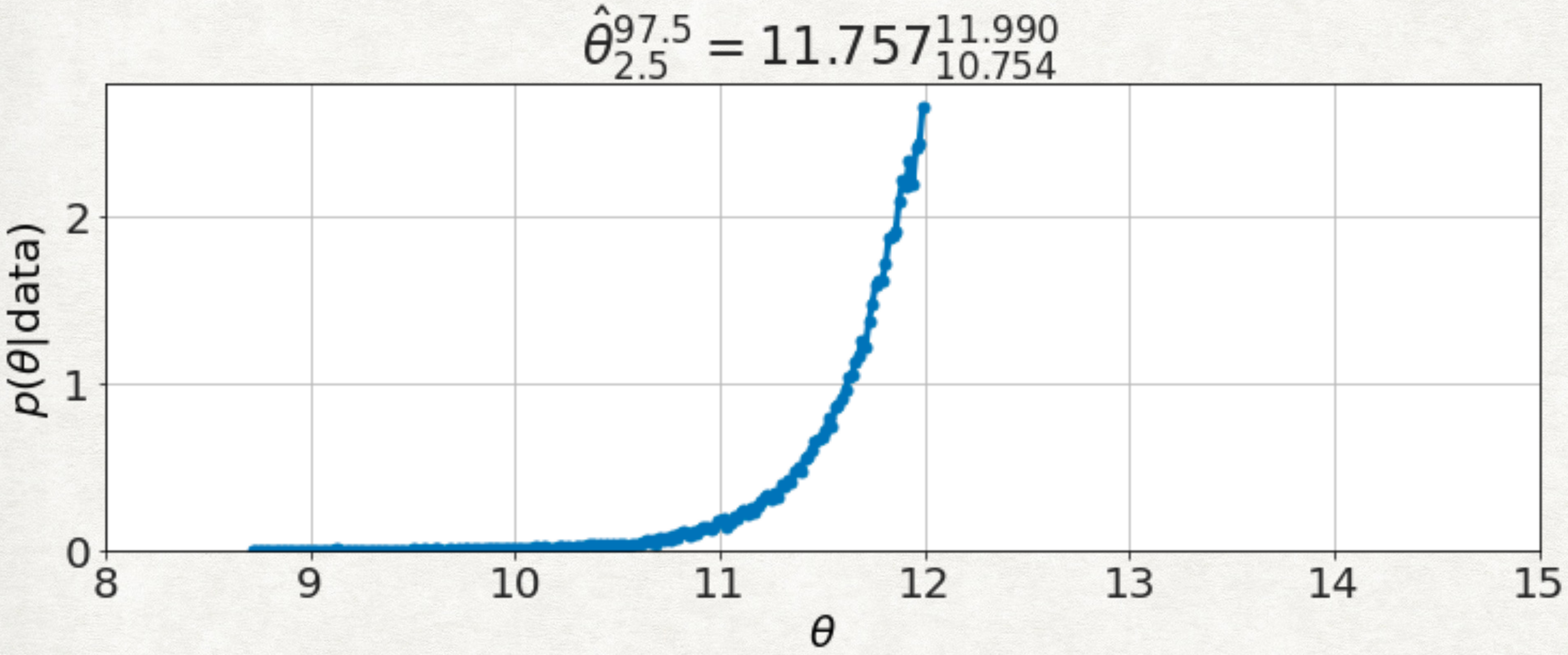
EXPONENTIAL DISTRIBUTION ESTIMATE LOCATION PARAMETER

- Data: $\{x_i\} = \{12,14,16\}$
- Likelihood: $p(x \mid \theta)dx = \begin{cases} \exp(\theta - x)\mathrm{d}x, & x > \theta \\ 0 & x < \theta \end{cases}$

```python
def lnlike(data,θ):
    x=data
    if np.any(x<θ):
        return -np.inf

    return np.sum(θ-x)

data=array([12.0,14,16])

model=MCMCModel(data,lnlike,
                θ=Uniform(-50,50),
                )
model.run_mcmc(1000,repeat=3)
```



BAYESIAN VS FREQUENTIST



- Data: $\{x_i\} = \{12, 14, 16\}$

Let us compare the confidence intervals obtained from two different estimators with the Bayesian intervals. The population mean is $E(x) = \theta + 1$, and so

$$(16) \qquad \theta^*(x_1 \ldots x_N) \equiv \frac{1}{N} \sum_{i=1}^{N} (x_i - 1)$$

is an unbiased estimator of θ. By a well-known theorem, it has variance $\sigma^2 = N^{-1}$, as we are accustomed to find. We must first find the sampling distribution of θ^*; by the method of characteristic functions we find that it is proportional to $y^{N-1} \exp(-Ny)$ for $y > 0$, where $y \equiv (\theta^* - \theta + 1)$.

$$(17) \qquad \theta^* - 0.8529 < \theta < \theta^* + 0.8264$$

or, with the above sample values, the shortest 90% confidence interval is

$$(18) \qquad 12.1471 < \theta < 13.8264 .$$

BAYESIAN VS FREQUENTIST

$$(15) \qquad p(\mathrm{d}x \mid \theta) = \begin{cases} \exp(\theta - x)\,\mathrm{d}x\,, & x > \theta \\ 0\,, & x < \theta \end{cases}.$$

(c) WHAT WENT WRONG?

Let us try to understand what is happening here. It is perfectly true that, *if* the distribution (15) is indeed identical with the limiting frequencies of various sample values, and *if* we could repeat all this an indefinitely large number of times, then use of the confidence interval (17) *would* lead us, in the long run, to a correct statement 90% of the time. But it would lead us to a wrong answer 100% of the time in the subclass of cases where $\theta^* > x_1 + 0.85$; and *we know from the sample whether we are in that subclass.*

Let us compare the confidence intervals obtained from two different estimators with the Bayesian intervals. The population mean is $E(x) = \theta + 1$, and so

$$(16) \qquad \theta^*(x_1 \ldots x_N) \equiv \frac{1}{N} \sum_{i=1}^{N} (x_i - 1)$$

is an unbiased estimator of θ. By a well-known theorem, it has variance $\sigma^2 = N^{-1}$, as we are accustomed to find. We must first find the sampling distribution of θ^*; by the method of characteristic functions we find that it is proportional to $y^{N-1} \exp(-Ny)$ for $y > 0$, where $y \equiv (\theta^* - \theta + 1)$.

$$(17) \qquad \theta^* - 0.8529 < \theta < \theta^* + 0.8264$$

or, with the above sample values, the shortest 90% confidence interval is

$$(18) \qquad 12.1471 < \theta < 13.8264\,.$$

- Here we have a trivial problem, only slightly more complex than the easiest problem addressed in an introductory Statistics course and *the standard tools give the wrong answer.*
- The Bayesian solution is a straightforward process and can easily be seen as a slight generalization from the easiest problem.

PAUSE TO REFLECT

- Shown some simple problems where the standard tools either cannot give a correct answer or give wrong answers
- Also shown that the Bayesian methods are straightforward to write and implement, and can handle each of these problems with minimal changes
- The Bayesian methods also provide more information — you get the individual distributions for all parameters along with correlations between parameters with no extra work

A LITTLE OF MY FRUSTRATION

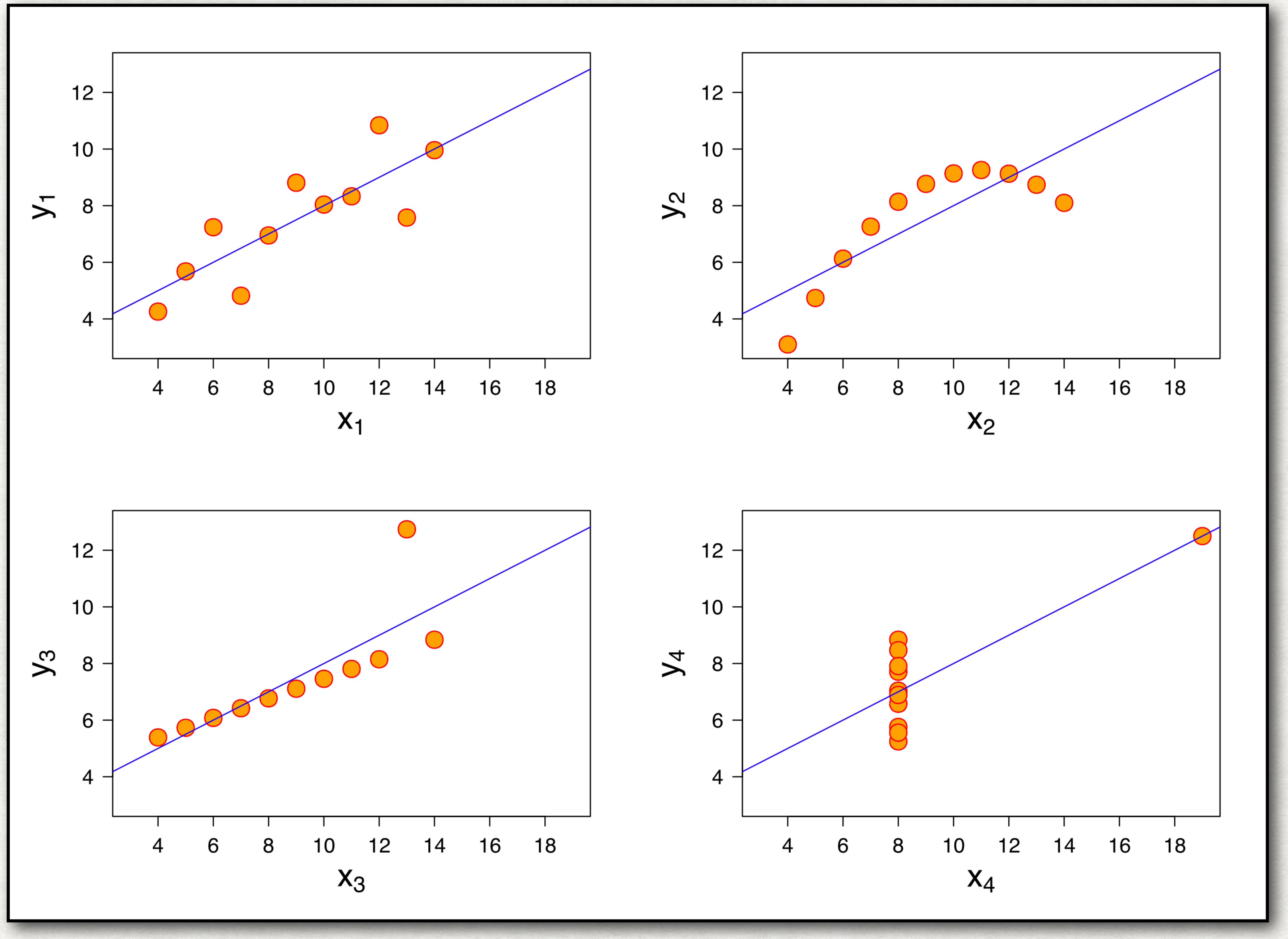
- If I was using a method to solve problems and was shown:

1. The method I was using broke in some simple cases
2. There is an alternative that is only marginally more complex but….
    1. …gives reasonable results on all well-posed problems
    2. …gives the *exact same* results I get on easy problems
    3. …is easier to interpret
    4. …generally gives more information than the methods I had been using

I know what I would do, but what would you do?

ANSCOMBE'S QUARTET

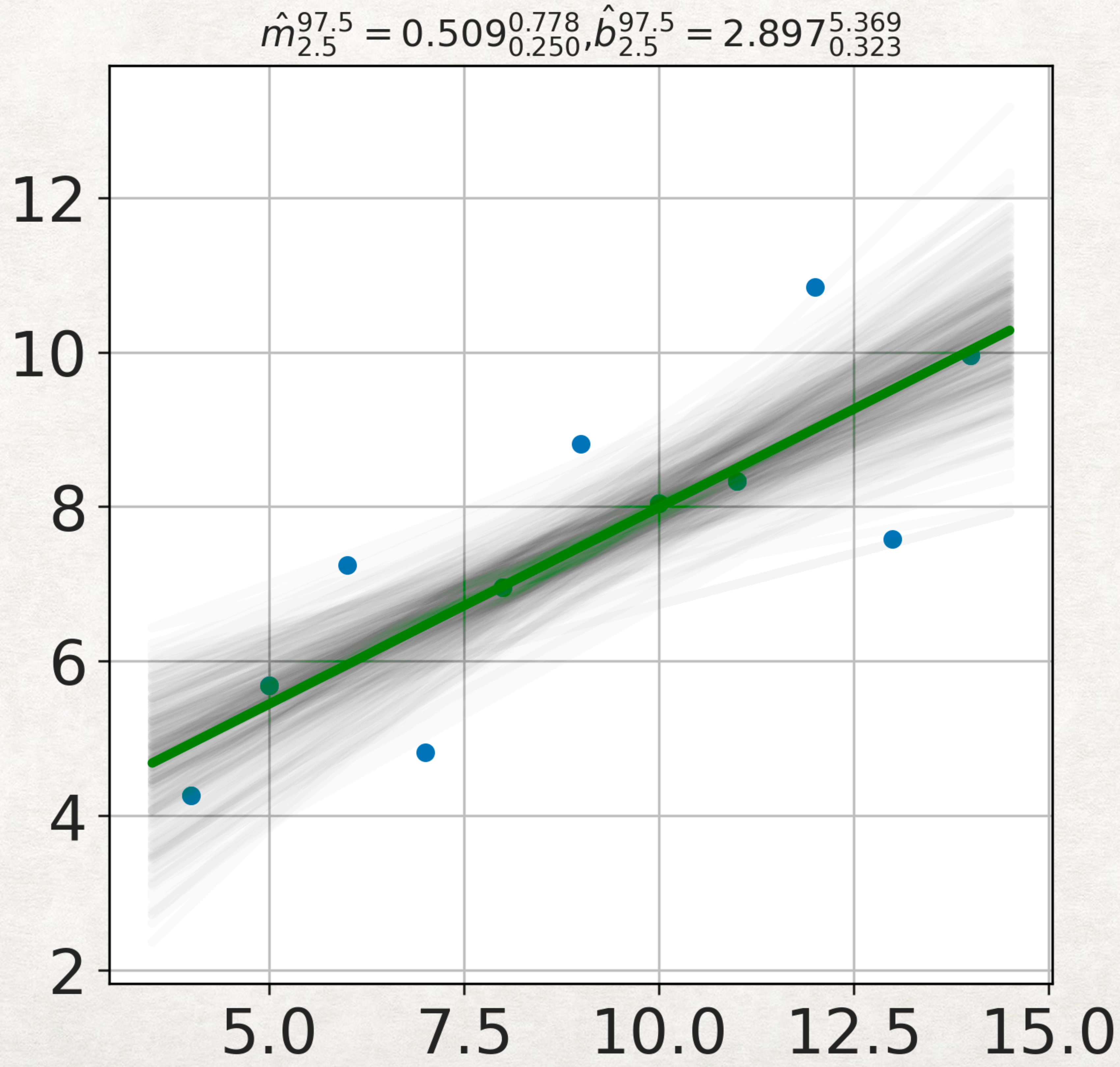
Property	Value	Accuracy
Mean of *x*	9	exact
Sample variance of *x* : s_x^2	11	exact
Mean of *y*	7.50	to 2 decimal places
Sample variance of *y* : s_y^2	4.125	±0.003
Correlation between *x* and *y*	0.816	to 3 decimal places
Linear regression line	$y = 3.00 + 0.500x$	to 2 and 3 decimal places, respectively
Coefficient of determination of the linear regression : R^2	0.67	to 2 decimal places



https://en.wikipedia.org/wiki/Anscombe%27s_quartet

LINEAR REGRESSION

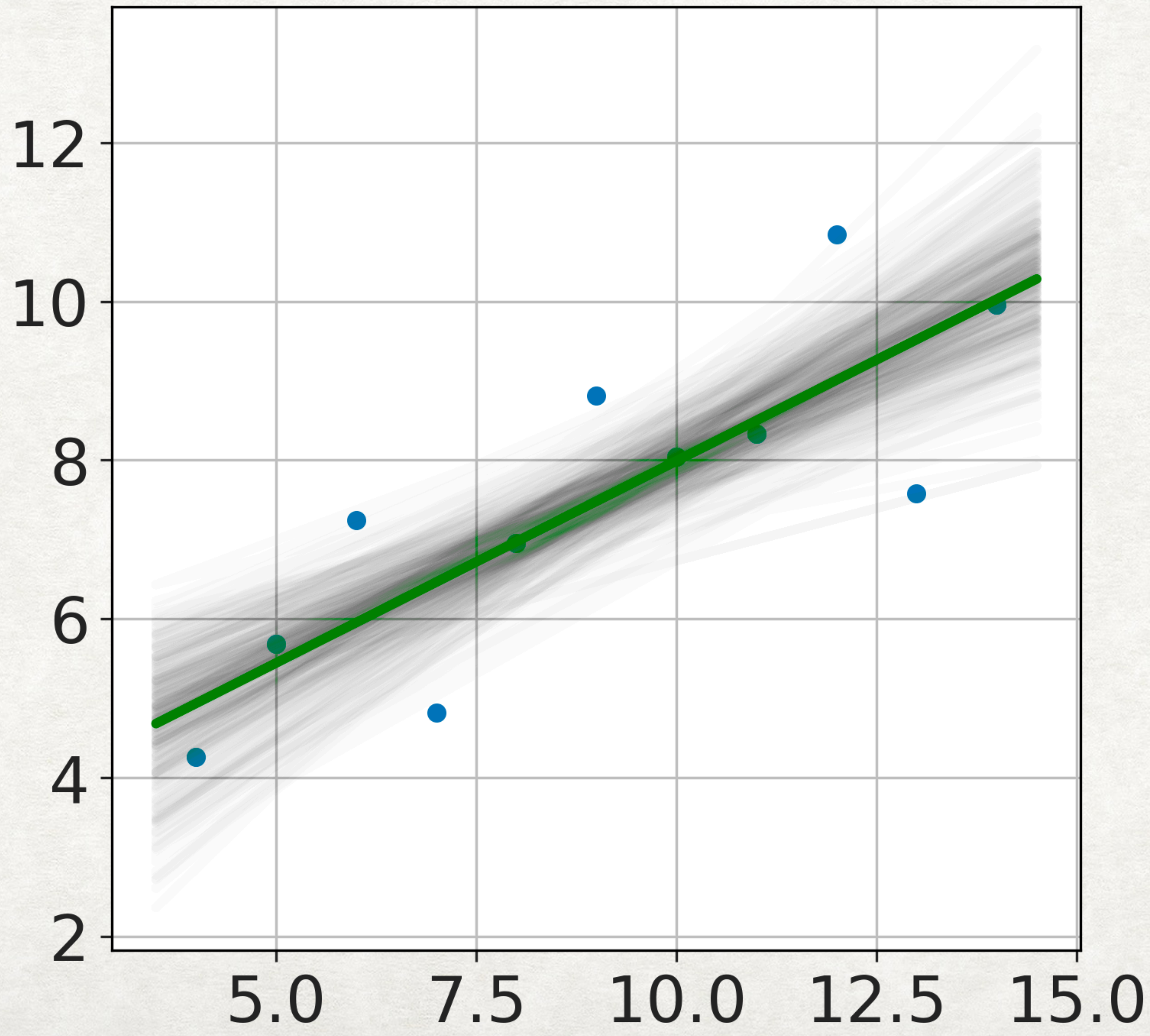
- Bayes: $P(m, b, \sigma \,|\, \{x_i\}) \sim$ likelihood$(x_i, y_i \,|\, m, b, \sigma) \times$ prior(m, b, σ)
- Likelihood: $P(y_i, x_i \,|\, m, b, \sigma) \sim$ Normal$(y_i - m \cdot x_i + b, \sigma)$
- Prior: $P(m, b, \sigma) \sim$ Uniform$(m, b) \times$ Uniform$(\log(\sigma))$



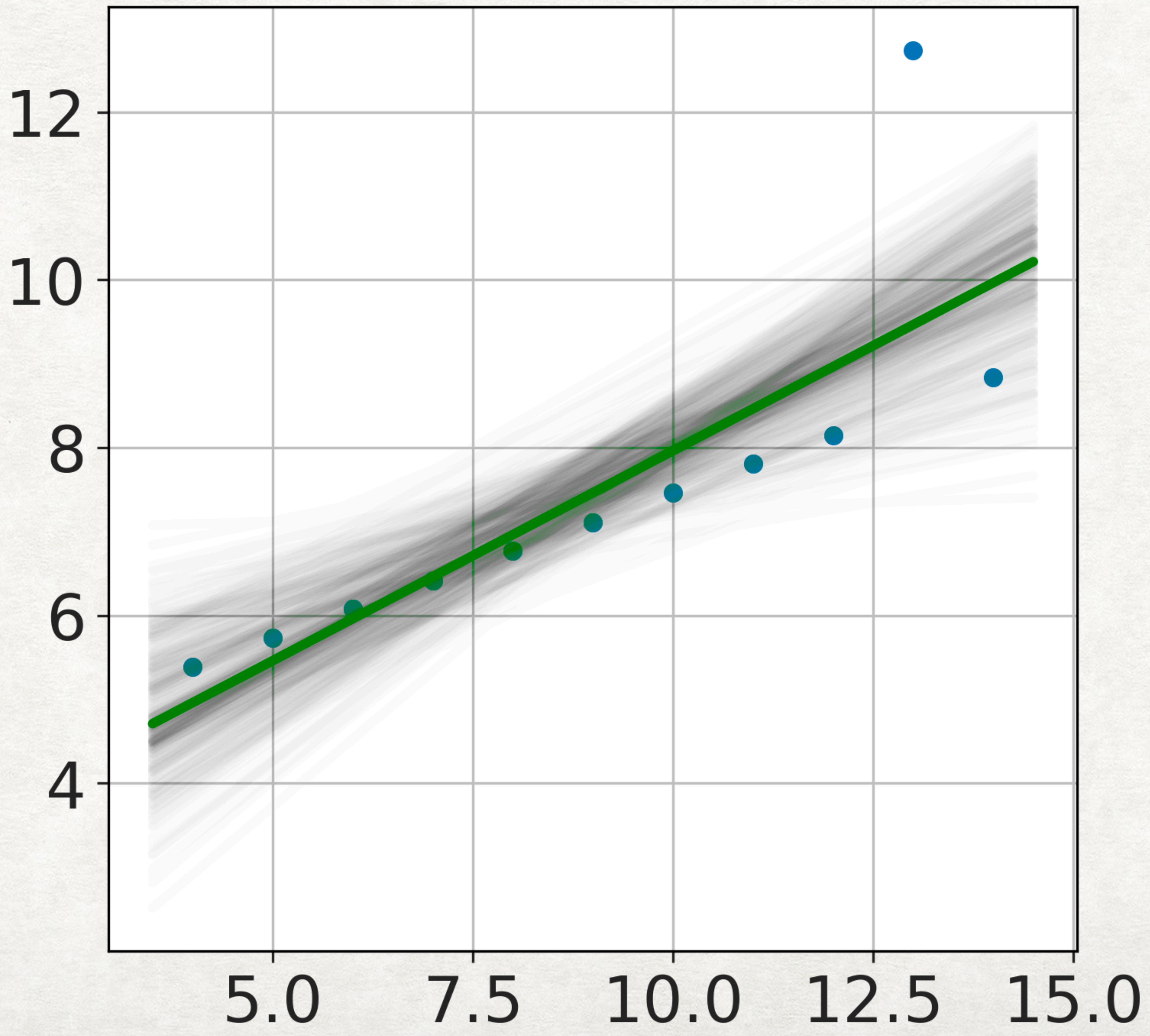
LINEAR REGRESSION

- Bayes: $P(m, b, \sigma \,|\, \{x_i\}) \sim \text{likelihood}(x_i, y_i \,|\, m, b, \sigma) \times \text{prior}(m, b, \sigma)$
- Likelihood: $P(y_i, x_i \,|\, m, b, \sigma) \sim \text{Normal}(y_i - m \cdot x_i + b, \sigma)$
- Prior: $P(m, b, \sigma) \sim \text{Uniform}(m, b) \times \text{Uniform}(\log(\sigma))$

$\hat{m}_{2.5}^{97.5} = 0.509_{0.250}^{0.778}, \hat{b}_{2.5}^{97.5} = 2.897_{0.323}^{5.369}$



$\hat{m}_{2.5}^{97.5} = 0.501_{0.246}^{0.776}, \hat{b}_{2.5}^{97.5} = 2.958_{0.374}^{5.378}$



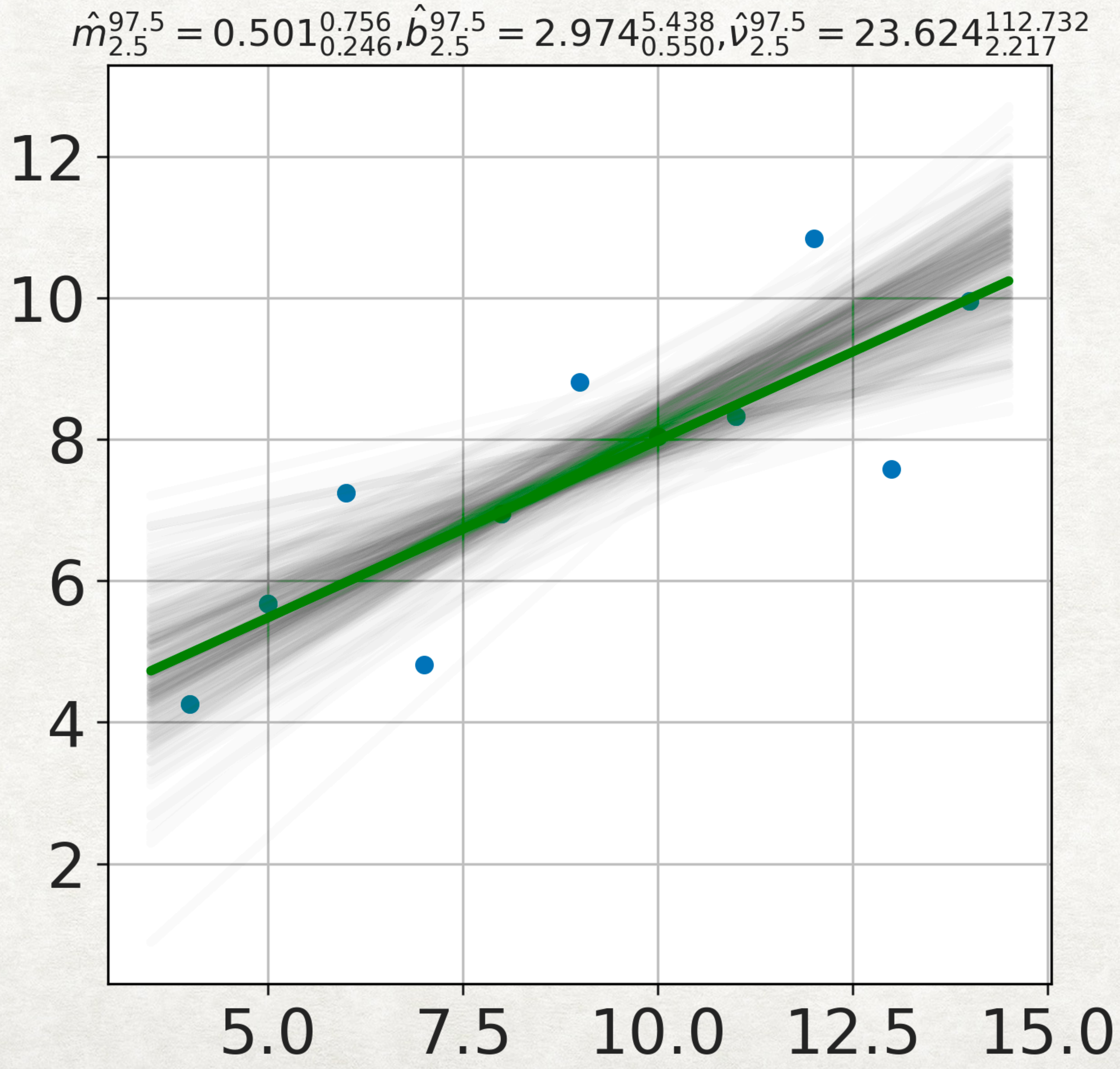
LINEAR REGRESSION WITH OUTLIERS

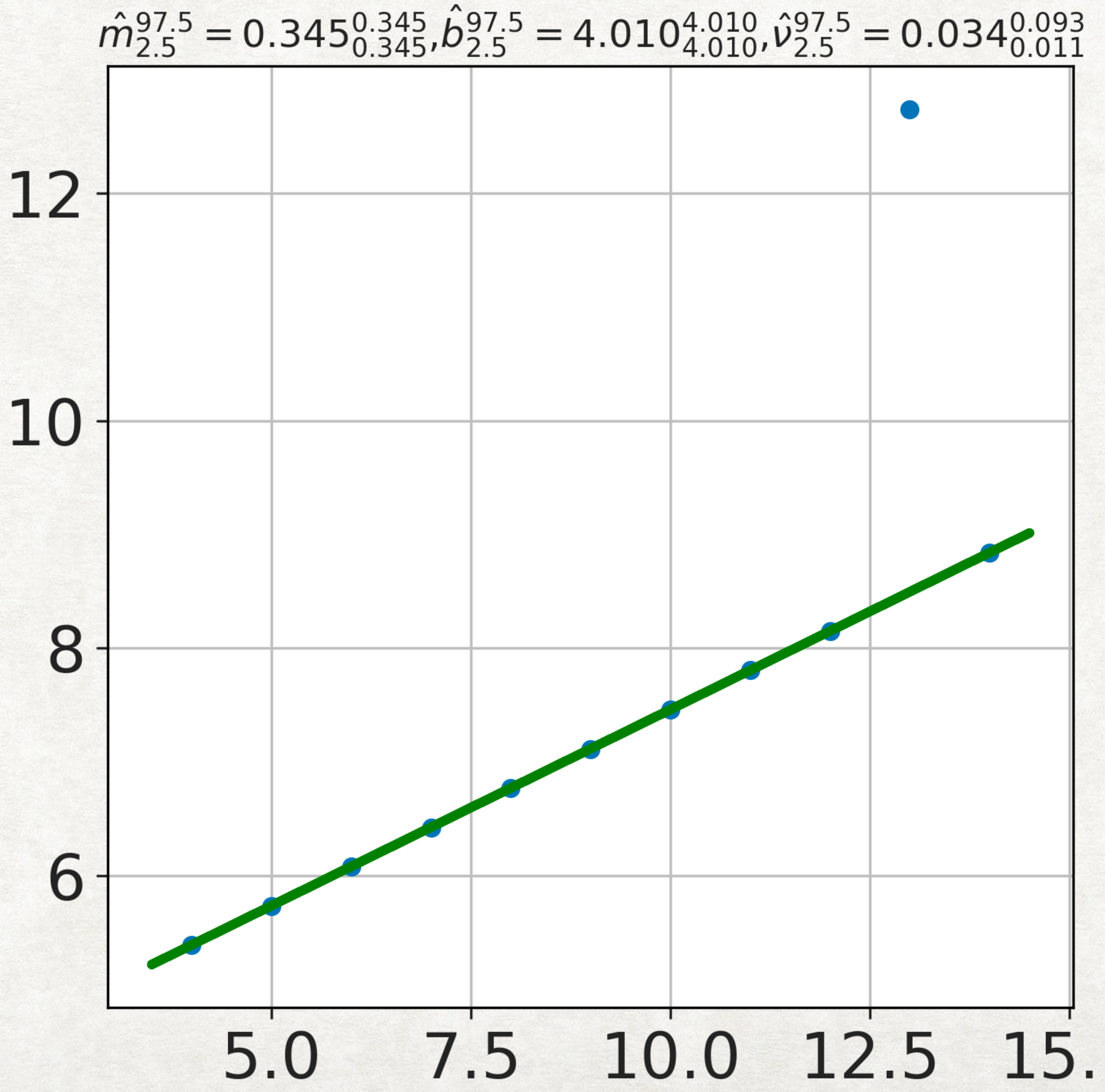
(J. Kruschke)

- Bayes: $P(m, b, \sigma, \nu \,|\, \{x_i\}) \sim \text{likelihood}(x_i, y_i \,|\, m, b, \sigma, \nu) \times \text{prior}(m, b, \sigma, \nu)$
- Likelihood: $P(y_i, x_i \,|\, m, b, \sigma, \nu) \sim \text{Student-T}(y_i - m \cdot x_i + b; \sigma, \nu)$
- Prior: $P(m, b, \sigma) \sim \text{Uniform}(m, b) \times \text{Uniform}(\log(\sigma)) \times \text{Exponential}(\nu)$

$\nu \to \infty \Rightarrow$ Normal

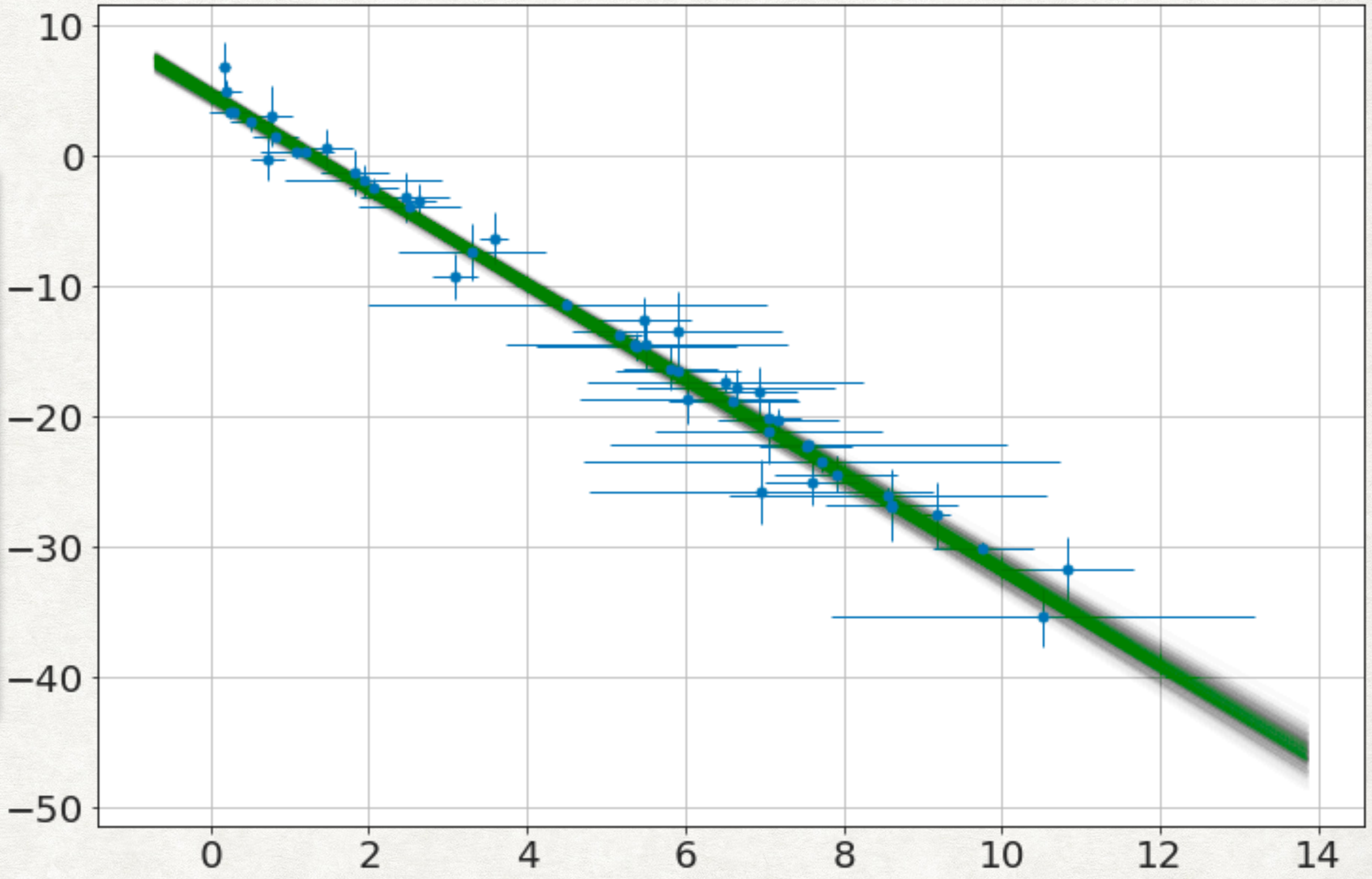
$\nu \to 0 \Rightarrow$ Outliery





LINEAR REGRESSION WITH ERRORS IN BOTH VARIABLES

```
def lnlike(data,m,b):
    x,y,xerr,yerr=data
    model = m * x + b

    d=(-x*m+y-b)**2/(m**2+1)
    sigma2=(yerr**2+m**2*xerr**2)/(1+m**2)  # projection of error along the line
    inv_sigma2 = 1.0/sigma2
    return -0.5*(np.sum(d**2*inv_sigma2 - np.log(inv_sigma2)))

data=x,y,xerr,yerr
model=MCMCModel2(data,lnlike,
                m=Uniform(-15,15),
                b=Uniform(-20,20),
                )
```



NON-LINEAR FITTING

Bayesian Analysis of Epidemics - Zombies, Influenza, and other Diseases

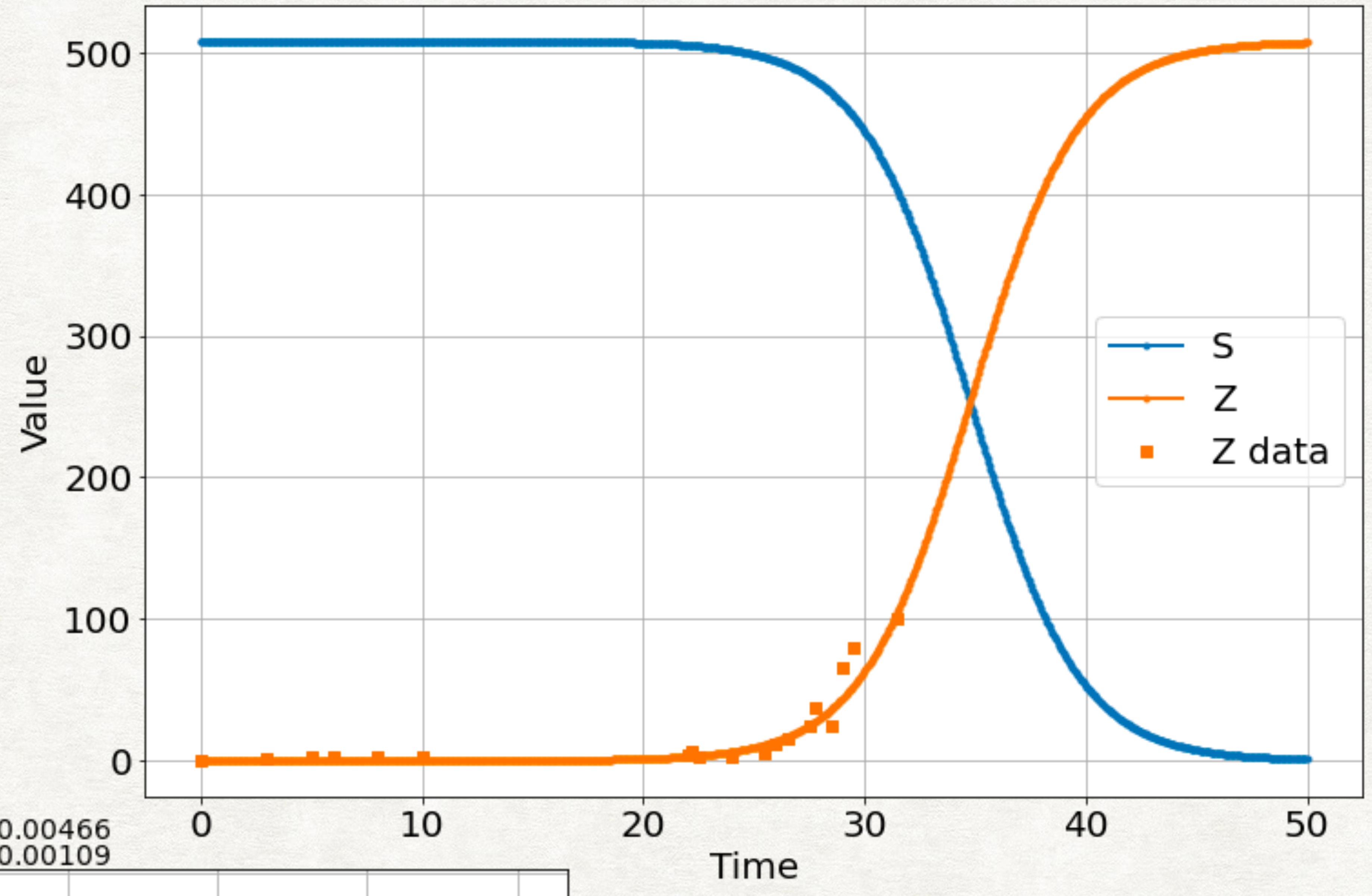
Caitlyn Witkowski[1,*], Brian Blais[1,2]
1 Science and Technology Department, Bryant University, Smithfield RI 02917
2 Institute for Brain and Neural Systems, Brown University, Providence RI
* Email: cwitkows@bryant.edu

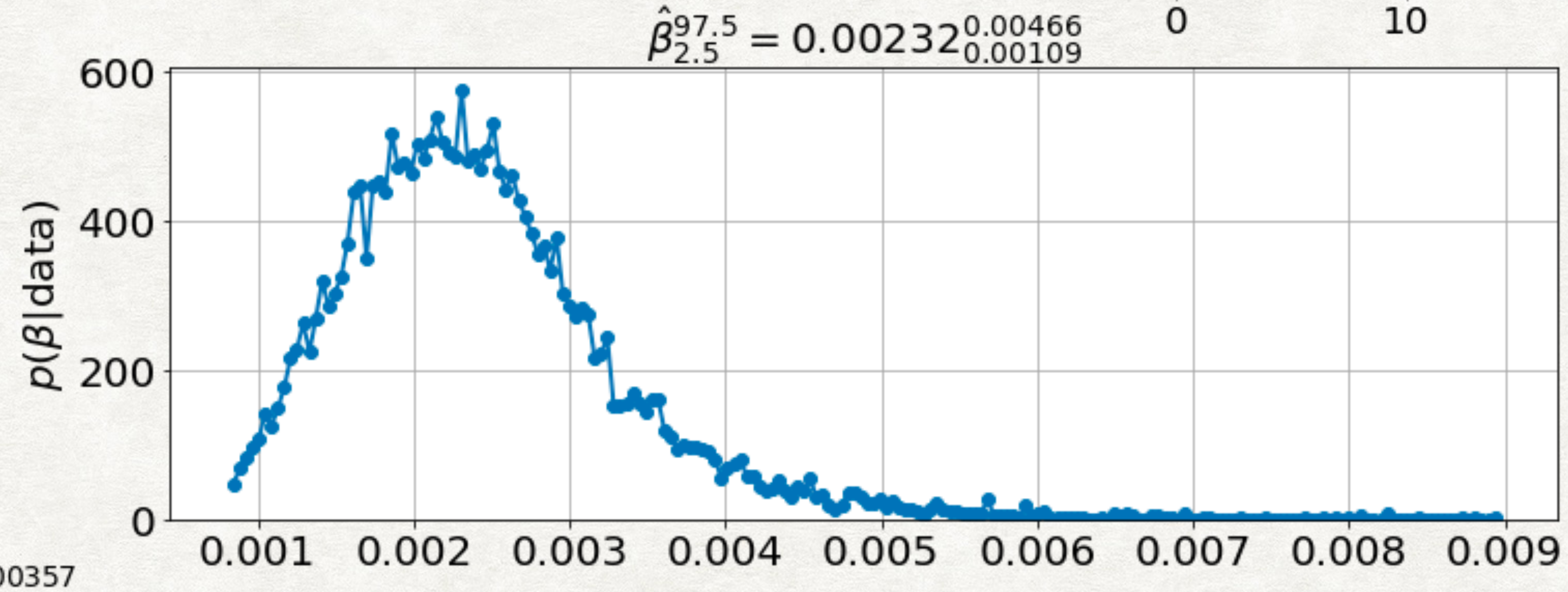
$$\frac{dS}{dt} = -\beta S \cdot Z$$

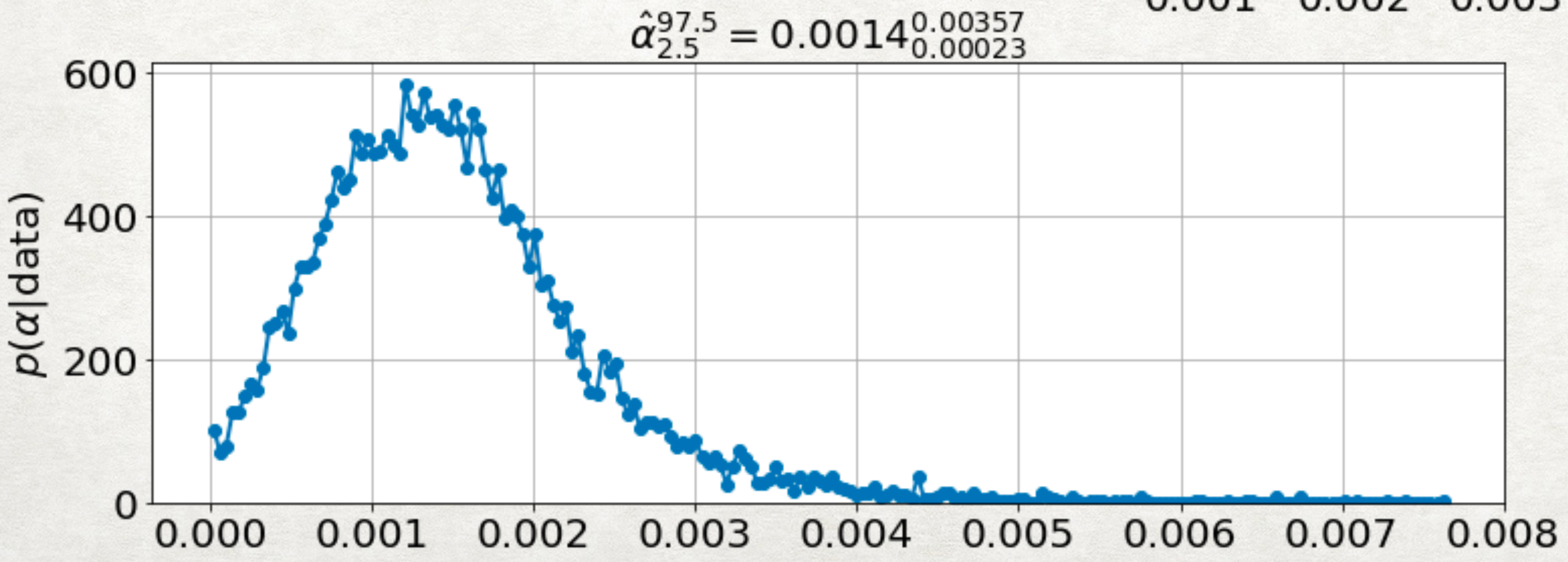
$$\frac{dE}{dt} = +\beta S \cdot Z - \zeta E$$

$$\frac{dZ}{dt} = +\zeta E - \alpha S \cdot Z$$

$$\frac{dR}{dt} = +\alpha S \cdot Z$$





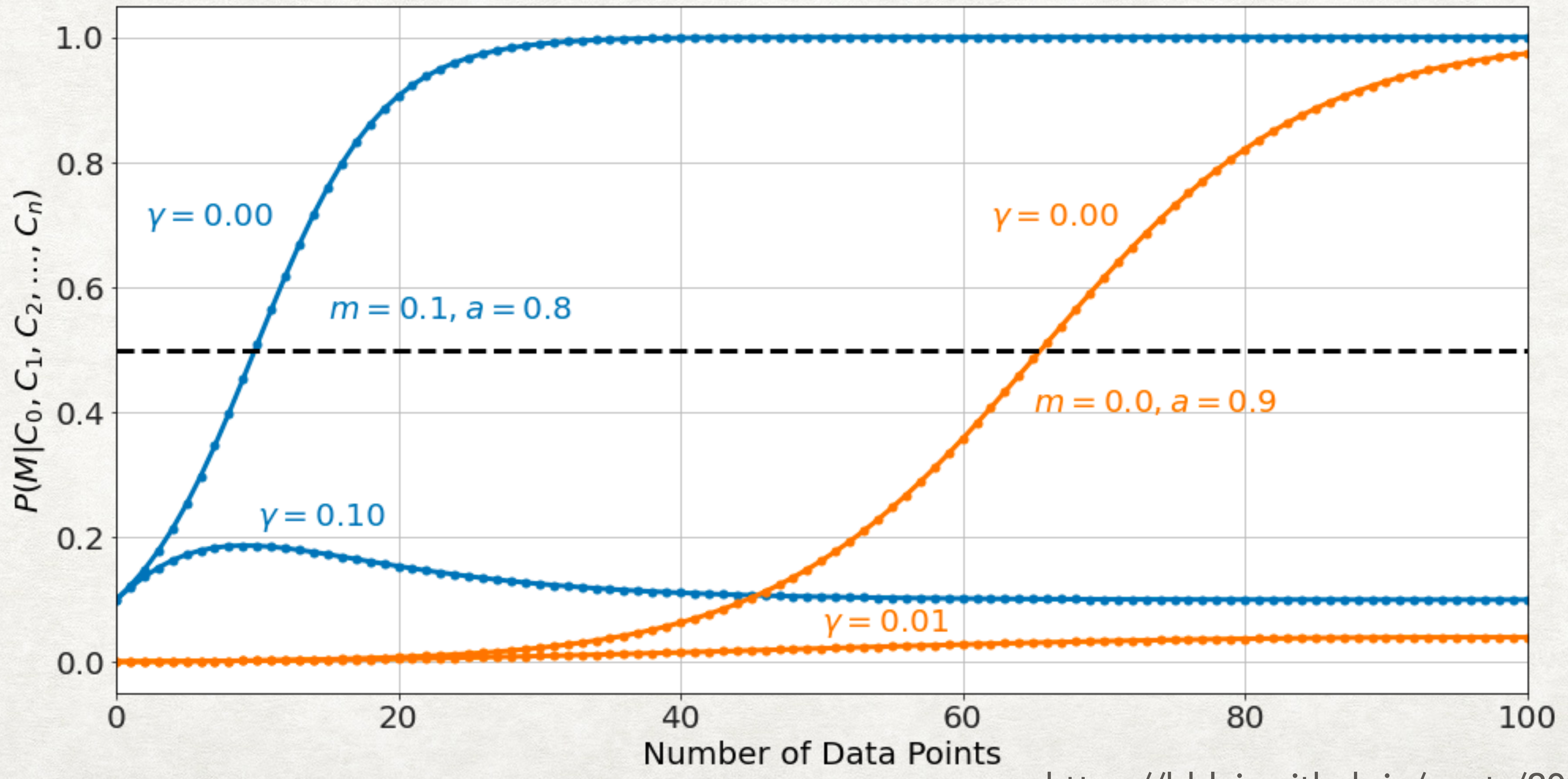


MATHEMATICS OF TESTIMONY

ANALYSIS OF THE PHILOSOPHER DAVID HUME'S ESSAY ON MIRACLES

- M = a miracle happened, m = prior probability of a miracle
- C_0, C_1, C_2, … = claims of a miracle
- a = reliability (a=1 unreliable, a=0 reliable)
- γ = amount reliability changes after each claim is debunked

$$P(M \mid C_0, C_1, C_2, \ldots, C_n) = \frac{m}{m + \left((a_0 + \gamma - 1) \cdot (1-\gamma)^{n-1} + 1\right)^n \cdot (1-m)}$$



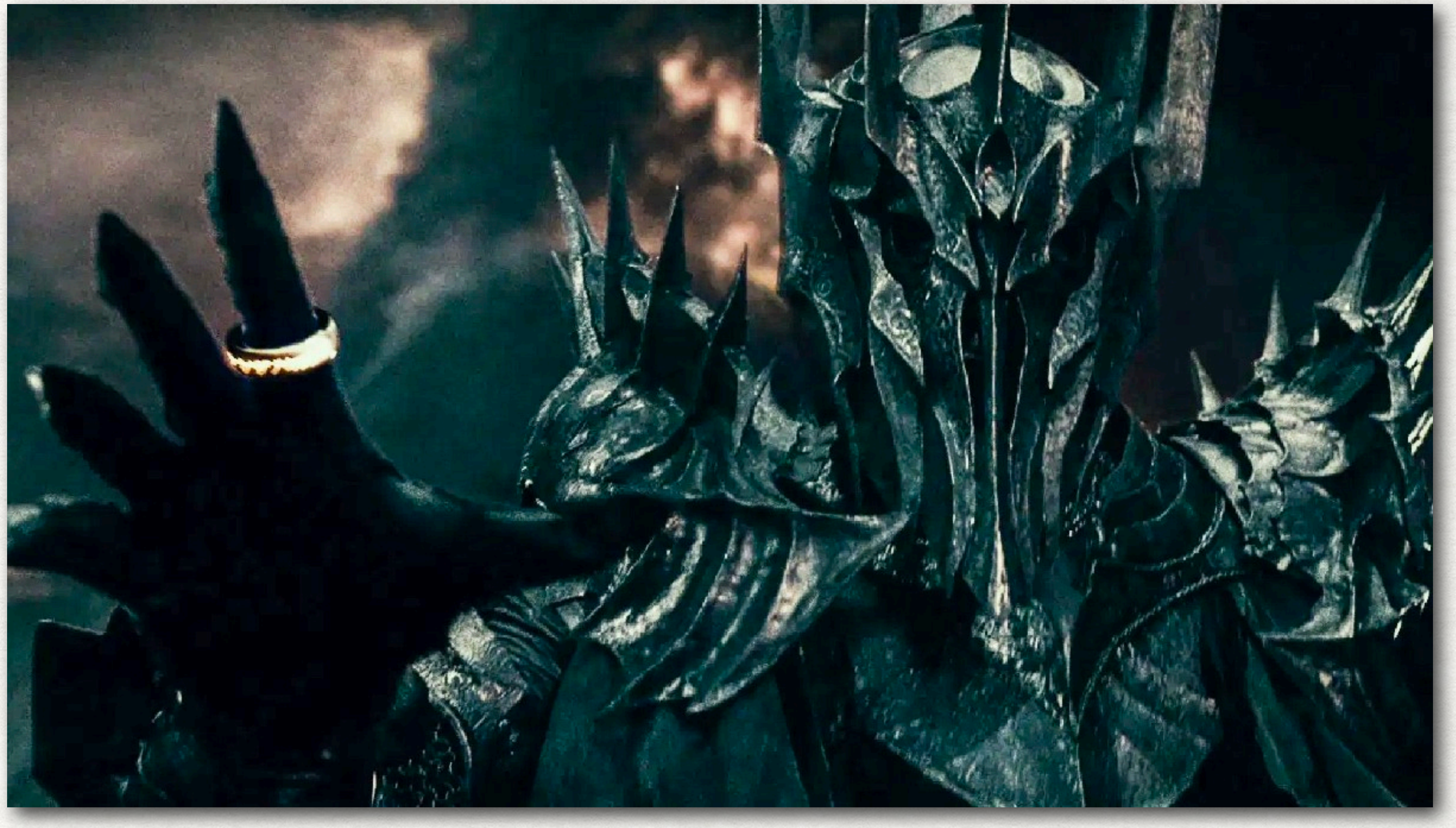
https://bblais.github.io/posts/2022/Jun/14/sometimes-more-testimony-is-worse/

CONCLUSIONS

- Messages in this presentation
  - Positive message —
    - Bayesian methods give you a uniform approach to all problems
    - Bayesian methods are easy to interpret
    - Bayesian methods give you more information about your problem
    - Bayesian methods get you to think more deeply about your data rather than reaching for a grab-bag of tools
  - Negative message —
    - Standard tools fail on some simple, well-defined problems
- My teaching and research goal — make technical topics approachable — can I help you with your projects?

GANDALF VS SAURON

BAYESIAN



NOT BAYESIAN

THANK YOU!