STSCI 4780
The big picture

Tom Loredo, CCAPS, Cornell University

2018-01-25

Course info

https://github.com/CU-BDA-2018/CourseInfo

The meteorologist

Joint and marginal frequencies of actual
and predicted weather

		Actual Weather		
		Snow	Shine	
Prediction	Snow	**1/4**	**1/2**	3/4
	Shine	**0**	**1/4**	1/4
		1/4	3/4	**Marginals**

Forecaster is right only 50% of the time

Observer notes a prediction of 'Sun' *every day* would be right 75% of the time, and applies for the forecaster's job

Should he or she get the job?

	Actual	
Prediction	Snow	Sun
Snow	1/4	1/2
Sun	0	1/4

Forecaster: You'll never be in an unpredicted snow

Observer: You'll be in an unpredicted snow 1 day out of 4

Bayesian viewpoint

The value of an inference lies in its usefulness in the individual case

Long run performance is not an adequate criterion for assessing the usefulness of inferences

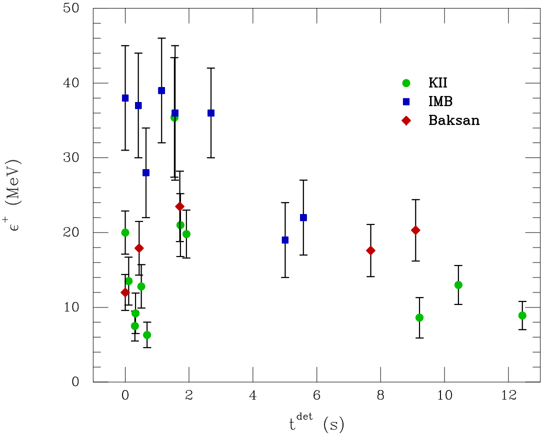
When long run performance is important, it needs to be separately evaluated

Modeling SN 1987A Neutrino Data

Before 1987

Feb 1987

Arrival times and energies for SN 1987A νs



- Do we understand the basics of core collapse?
- Can we discriminate between prompt and delayed shock models?
- How does the nascent NS compare with predictions?
- What is the $\bar{\nu}_e$ rest mass?

So far a *unique event*. . .

The big picture

1. **Statistics and scientific method**
2. **Models**
3. **Confidence intervals vs. credible intervals**

Agenda

1. **Statistics and scientific method**

2. **Models**

3. **Confidence intervals vs. credible intervals**

Scientific method

Science is more than a body of knowledge; it is a way of thinking.
The method of science, as stodgy and grumpy as it may seem,
is far more important than the findings of science.
—Carl Sagan

Scientists *argue!*

Argument $\equiv$ Collection of statements comprising an act of reasoning from *premises* to a *conclusion*

A key goal of science: Explain or predict *quantitative measurements* (data!)

Data analysis: Constructing and appraising arguments that reason from data to interesting scientific conclusions (explanations, predictions)

The role of data

Data do not speak for themselves!

"No body of data tells us all we need to know about its own analysis."
— John Tukey, *EDA*

We don't just *tabulate* data, we *analyze* data

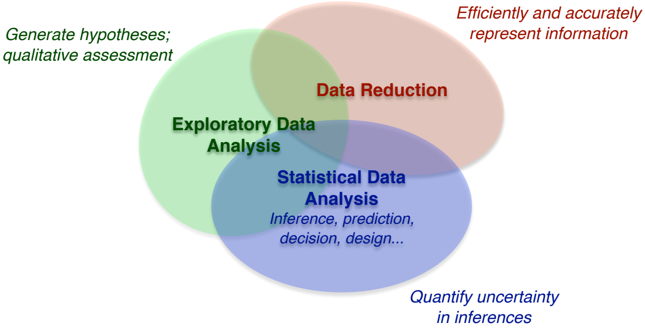
We gather data so they may speak for or against existing hypotheses, and guide the formation of new hypotheses

A key role of data in science is to be among the premises in scientific arguments

Data analysis

Building & Appraising Arguments Using Data

Modes of Data Analysis



Inference: Learning models (populations, signals. . .) for the data generating process from observed data—just one of several interacting modes of analyzing data

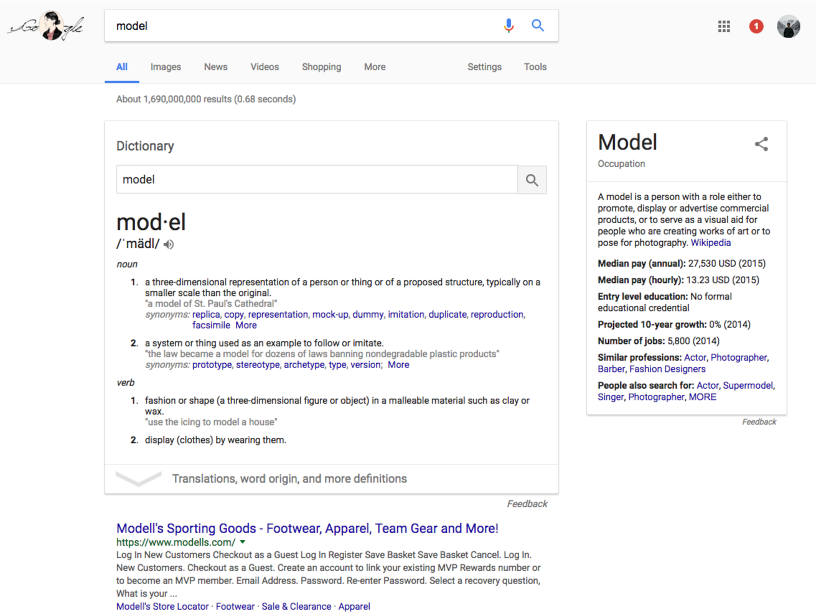
Agenda

1. Statistics and scientific method

2. **Models**

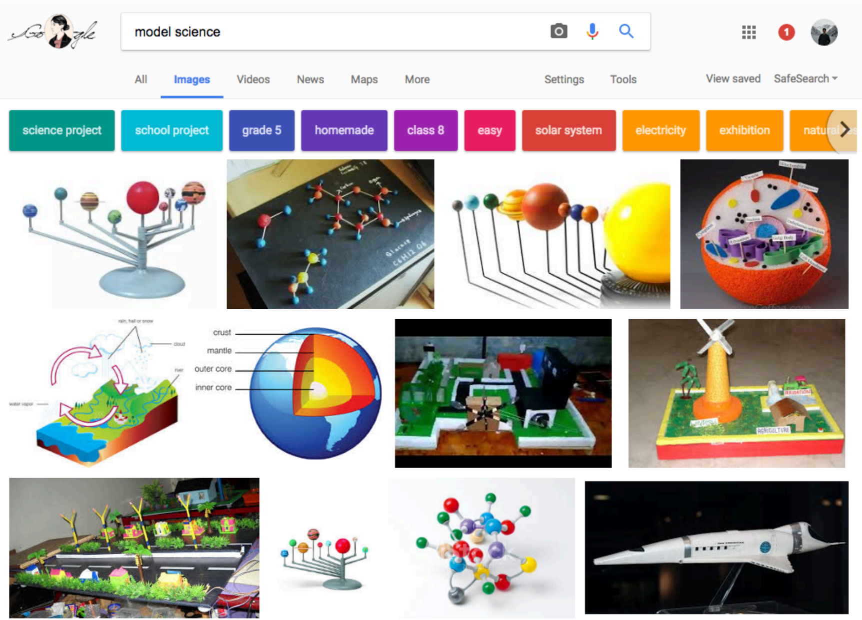
3. Confidence intervals vs. credible intervals

Models


model
All
Images
News
Videos
Shopping
More
Settings
Tools
About 1,690,000,000 results (0.68 seconds)
Dictionary
model
mod·el
/ˈmädl/
noun
1. a three-dimensional representation of a person or thing or of a proposed structure, typically on a smaller scale than the original.
"a model of St. Paul's Cathedral"
synonyms: replica, copy, representation, mock-up, dummy, imitation, duplicate, reproduction, facsimile More
2. a system or thing used as an example to follow or imitate.
"the law became a model for dozens of laws banning nondegradable plastic products"
synonyms: prototype, stereotype, archetype, type, version; More
verb
1. fashion or shape (a three-dimensional figure or object) in a malleable material such as clay or wax.
"use the icing to model a house"
2. display (clothes) by wearing them.
Translations, word origin, and more definitions
Feedback
Modell's Sporting Goods - Footwear, Apparel, Team Gear and More!
https://www.modells.com/
Log In New Customers Checkout as a Guest Log In Register Save Basket Save Basket Cancel. Log In. New Customers. Checkout as a Guest. Create an account to link your existing MVP Rewards number or to become an MVP member. Email Address. Password. Re-enter Password. Select a recovery question, What is your ...
Modell's Store Locator · Footwear · Sale & Clearance · Apparel
Model
Occupation
A model is a person with a role either to promote, display or advertise commercial products, or to serve as a visual aid for people who are creating works of art or to pose for photography. Wikipedia
Median pay (annual): 27,530 USD (2015)
Median pay (hourly): 13.23 USD (2015)
Entry level education: No formal educational credential
Projected 10-year growth: 0% (2014)
Number of jobs: 5,800 (2014)
Similar professions: Actor, Photographer, Barber, Fashion Designers
People also search for: Actor, Supermodel, Singer, Photographer, MORE
Feedback

Classes of models

- **Descriptive:** Aims to describe how something *is*
- **Normative:** Aims to describe how something *should be* (by some criteria)


model science
All
Images
Videos
News
Maps
More
Settings
Tools
View saved
SafeSearch
science project
school project
grade 5
homemade
class 8
easy
solar system
electricity
exhibition
natural
crust
mantle
outer core
inner core

model science

All Images Videos News Maps More Settings Tools

About 71,300,000 results (0.73 seconds)

Scientific modelling - Wikipedia
https://en.wikipedia.org/wiki/Scientific_modelling
Jump to **Model**-based learning in education - **Model**–based learning in education, particularly in relation to learning **science** involves students creating **models** for **scientific** concepts in order to: Gain insight of the **scientific** idea(s); Acquire deeper understanding of the subject through visualization of the **model**; Improve ...
Overview · Basics of scientific ... · Types of scientific modelling · Applications

People also ask

What is the model of science?

What is an example of a model in science?

What is a mathematical model in science?

What are the different types of models used in science?

Feedback

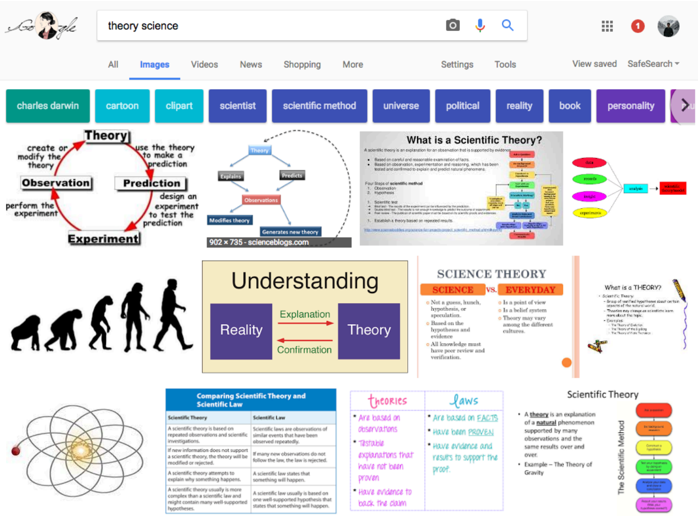
Images for model science

More images for model science

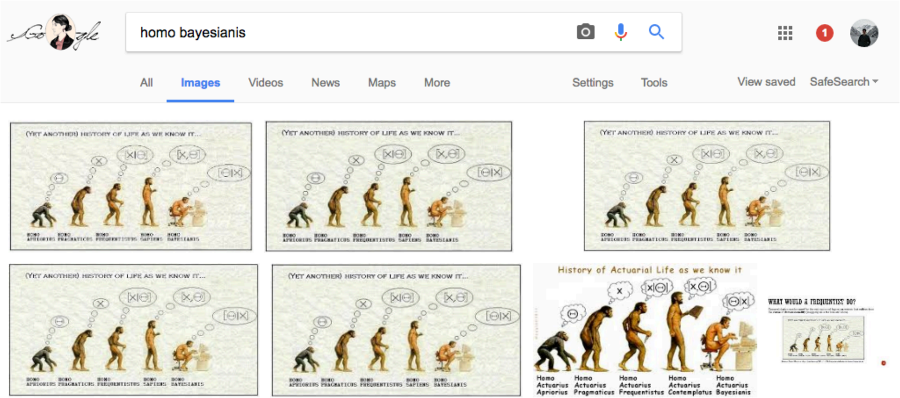
Report images

Making a Solar System Model : Science School Project - YouTube

https://www.youtube.com/watch?v=gubsXJXqAWc
Jul 26, 2017 - Uploaded by ExamFear Education
Making a Solar System **Model** : **Science** School Project This is a simple DIY science experiment that can be ...



Theory (scientific): Principles/laws/desiderata guiding construction of scientific models


homo bayesianis
All
Images
Videos
News
Maps
More
Settings
Tools
View saved
SafeSearch
History of Actuarial Life as we know it
Homo Actuarius Aprioruis
Homo Actuarius Pragmaticus
Homo Actuarius Frequentistus
Homo Actuarius Contemplatus
Homo Actuarius Bayesianis

[zoom!]

modeling science instruction

All Images Videos Shopping News More Settings Tools

About 7,570,000 results (0.59 seconds)

Scholarly articles for modeling science instruction

... framework and implications for **science instruction** - Chinn - Cited by 1853
Inquiry, **modeling**, and metacognition: Making **science** ... - White - Cited by 1731
Modeling instruction: An effective model for **science** ... - Jackson - Cited by 176

American Modeling Teachers Association – Fostering brilliant ...

https://modelinginstruction.org/

AMTA is a professional organization of teachers, by teachers and for teachers who utilize **Modeling Instruction**(TM) in their **Science**, Technology, Engineering and Mathematics (STEM) teaching practice. Our mission is to provide professional development for teachers in the **Modeling** Method of **Instruction**, to provide ...

Synopsis of Modeling ... · Modeling Instruction ... · Middle School Storylines · News

Modeling Instruction in High School Physics

modeling.asu.edu/modeling-HS.html

Dec 14, 2015 - The program cultivates physics teachers as school experts on effective use of guided inquiry in **science** teaching, thereby providing schools and school districts with a valuable resource for broader reform. Program goals are fully aligned with National **Science** Education Standards. The **Modeling** Method ...

Modeling Instruction Program

modeling.asu.edu/

The ASU **Modeling Instruction** program and MNS degree program meet the following needs. The National **Science** Education Standards (NRC, 1996) emphasize that "coherent and integrated programs" supporting "lifelong professional development" of **science** teachers are essential for significant reform. "The conventional ...

[PDF] **Modeling Instruction: An Effective Model for Science Education - Eric**

https://files.eric.ed.gov/fulltext/EJ851867.pdf

by J Jackson - Cited by 175 - Related articles

Introduction. **Modeling Instruction** is an evolving, research-based program for high school **science** education reform that was supported by the National. **Science** Foundation (NSF) from. 989 to 2005. The name **Modeling**. **Instruction** expresses an emphasis on the construction and application of conceptual **models** of physical.

“Modeling is the main activity of scientists. . . ” — David Hestenes

Modeling

Modeling Theory aims at elucidating the roles of **models** and **modeling** in scientific knowledge and practice. It provides a foundation for a realist philosophy of science and a constructivist teaching methodology. Modeling pedagogy and its implementation in physics teaching is addressed at the Modeling Instruction Program site.

Conceptual Modeling in physics, mathematics and cognitive science

Abstract. Scientific thinking is grounded in the evolved human ability to freely create and manipulate mental models in the imagination. This modeling ability enabled early humans to navigate the natural world and cope with challenges to survival. Then it drove the design and use of tools to shape and control the environment. Spoken language facilitated the sharing of mental models in cooperative activities like hunting and in maintaining tribal memory through storytelling. The evolution of culture accelerated with the invention of written language, which enabled creation of powerful symbolic systems and tools to think with. That includes deliberate design of mathematical tools that are essential for physics and engineering . A mental model coordinated with a symbolic representation is called a conceptual model. Conceptual models provide symbolic expressions with meaning. This essay proposes a Modeling Theory of cognitive structure and process. Basic definitions, principles and conclusions are offered. Supporting evidence from the various cognitive sciences is sampled. The theory provides the foundation for a science pedagogy called Modeling Instruction, which has been widely applied with documented success and recognized most recently with an Excellence in Physics Education award from the American Physical Society.

D. Hestenes, *SemiotiX*, November 2015.
http://semioticon.com/semiotix/2015/11/.

Modeling Theory for Math and Science Education

Abstract. Mathematics has been described as the science of patterns. Natural science can be characterized as the investigation ofpatterns in nature. Central to both domains is the notion of model as a unit of coherently structured knowledge. Modeling Theory is concerned with models as basic structures in cognition as well as scientific knowledge. It maintains a sharp distinction between mental models that people think with and conceptual models that are publicly shared. This supports a view that cognition in science, math, and everyday life is basically about making and using mental models. We review and extend elements of Modeling Theory as a foundation for R&D in math and science education.

D. Hestenes, Modeling Theory for Math and Science Education, In R. Lesh, P. Galbraith, C. Hines, A. Hurford (eds.) *Modeling Students' Mathematical Competencies* (New York: Springer, 2010).

Scientific models

What is a scientific model?

- A model is a surrogate
- A model is an idealization, a partial representation
- A model is often an abstraction
- A model is provisional—a consequence of idealiziation
- Many models are quantitative—*mathematical models*
- . . .

Bayesian inference

- Bayesian inference uses probability theory to *quantify the strength of data-based arguments about mathematical models* (i.e., a more abstract view than restricting PT to describe variability in repeated "random" experiments)
- Uses probability theory as a *normative model* for reasoning in the presence of uncertainty
- A different approach to *all* statistical inference problems (i.e., not just another method in the list: BLUE, linear regression, least squares/χ^2 minimization, maximum likelihood, ANOVA, survival analysis . . .)
- Focuses on *deriving consequences of modeling assumptions* rather than *devising and calibrating procedures*

Bayesian data analysis (BDA): Using Bayesian ideas across various data analysis tasks—not just inference, but also prediction, decision, design, EDA, data reduction. . .

Frequentist vs. Bayesian hypothesis appraisal

"The data $D_{\rm obs}$ support conclusion C . . . "

Frequentist appraisal

"C was selected with a procedure that's right 95% of the time over a set $\{D_{hyp}\}$ that includes D_{obs}."

Probabilities are properties of *procedures*, not of particular results

Bayesian appraisal

"The strength of the chain of reasoning from the model and D_{obs} to C is 0.95, on a scale where 1= certainty."

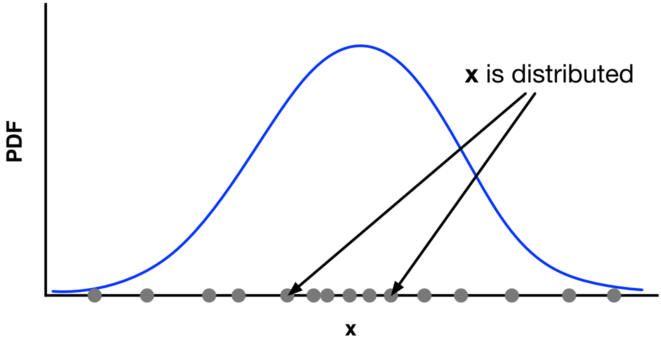
Probabilities are associated with *specific, observed data*

Long-run performance must be separately evaluated (and is often good by frequentist criteria)

Interpreting PDFs

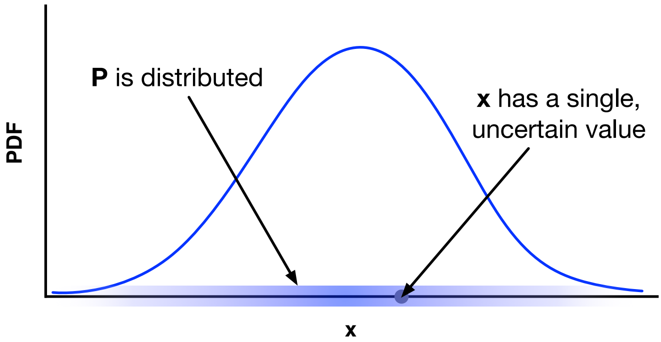
Frequentist

Probabilities are always (limiting) rates/proportions/frequencies that *quantify variability* in a sequence of trials. $p(x)$ describes how the *values of x* would be distributed among infinitely many trials:



Bayesian

Probability *quantifies uncertainty* in an inductive inference. $p(x)$ describes how *probability* is distributed over the possible values x might have taken in the single case before us:



Agenda

1. Statistics and scientific method
2. Models
3. **Confidence intervals vs. credible intervals**

A Simple (?) confidence region

Problem

Estimate the location (mean) of a Gaussian distribution from a set of samples $D = \{x_i\}$, $i = 1$ to N

Report a *point estimate*, and a *region* summarizing the uncertainty

Model

$$p(x_i|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x_i-\mu)^2}{2\sigma^2}\right]$$

$$\text{Equivalently,} \quad x_i \sim \mathcal{N}(\mu,\sigma^2)$$

Here assume σ is *known*; we are uncertain about μ

Classes of variables

- μ is the unknown we seek to estimate—the *parameter*. The *parameter space* is the space of possible values of μ—here the real line (perhaps bounded). *Hypothesis space* is a more general term.
- A particular set of N data values $D = \{x_i\}$ is a *sample*. The *sample space* is the N-dimensional space of possible samples.

Standard inferences

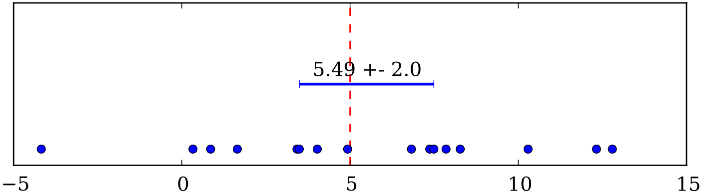
Let $\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$.

- "Standard error" (rms error) is $\sigma/\sqrt{N}$
- "1σ" interval: $\bar{x} \pm \sigma/\sqrt{N}$ with conf. level CL $= 68.3\%$
- "2σ" interval: $\bar{x} \pm 2\sigma/\sqrt{N}$ with CL $= 95.4\%$

Some simulated data

Take $\mu = 5$ and $\sigma = 4$ and $N = 16$, so $\sigma/\sqrt{N} = 1$

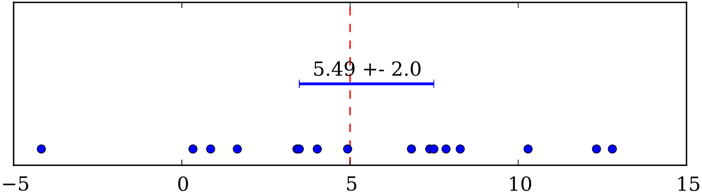
What is the CL associated with this interval?



Some simulated data

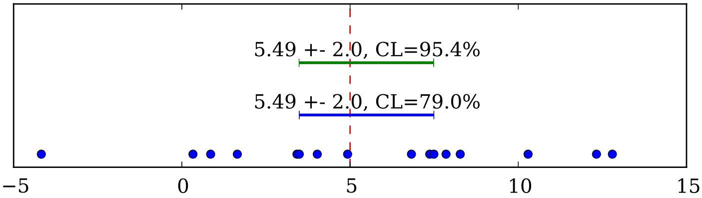
Take $\mu = 5$ and $\sigma = 4$ and $N = 16$, so $\sigma/\sqrt{N} = 1$

What is the CL associated with this interval?



The (frequentist) confidence level for this interval is 79.0%

Two intervals

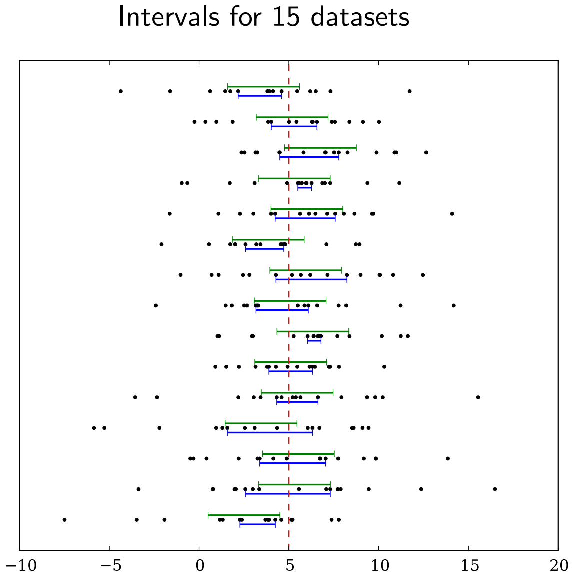


- Green interval: $\bar{x} \pm 2\sigma/\sqrt{N}$
- Blue interval: Let $x_{(k)} \equiv k$'th order statistic
  Report $[x_{(6)}, x_{(11)}]$ (i.e., leave out 5 outermost each side)

Moral

The confidence level is a property of the **procedure**, *not of the particular interval reported for a given dataset*

Performance of intervals



Confidence interval for a normal mean

Suppose we have a sample of $N = 5$ values x_i, with

$$x_i \sim N(\mu, 1)$$

We want to estimate μ, including some *quantification of uncertainty* in the estimate: an interval *with a probability attached*

Frequentist approaches: method of moments, BLUE, least-squares/χ^2, maximum likelihood

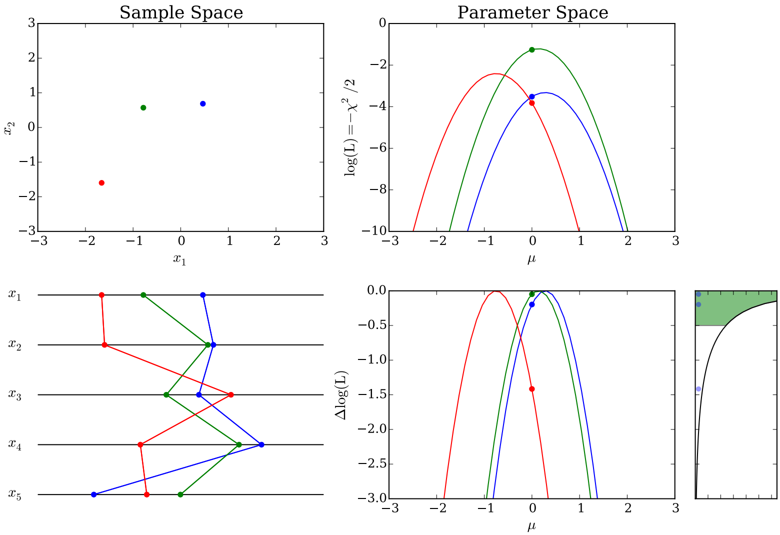
Focus on likelihood (equivalent to χ^2 here); this is closest to Bayes:

$$\begin{aligned}
\mathcal{L}(\mu) &= p(\{x_i\}|\mu) \\
&= \prod_i \frac{1}{\sigma\sqrt{2\pi}} e^{-(x_i-\mu)^2/2\sigma^2}; \qquad \sigma = 1 \\
&\propto e^{-\chi^2(\mu)/2}
\end{aligned}$$

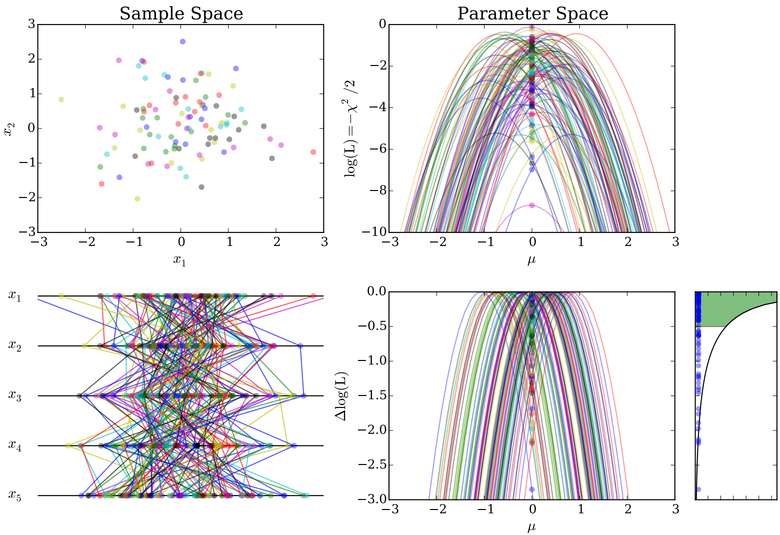
Estimate μ from maximum likelihood (minimum χ^2)
Define an interval and its coverage frequency from the $\mathcal{L}(\mu)$ curve

Construct an interval procedure for known μ

Likelihoods for 3 simulated data sets, $\mu = 0$

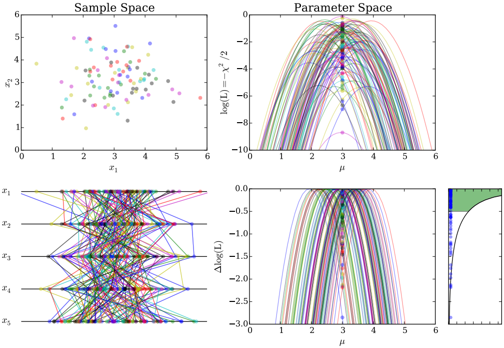


Likelihoods for 100 simulated data sets, $\mu = 0$



Explore dependence on μ

Likelihoods for 100 simulated data sets, $\mu = 3$

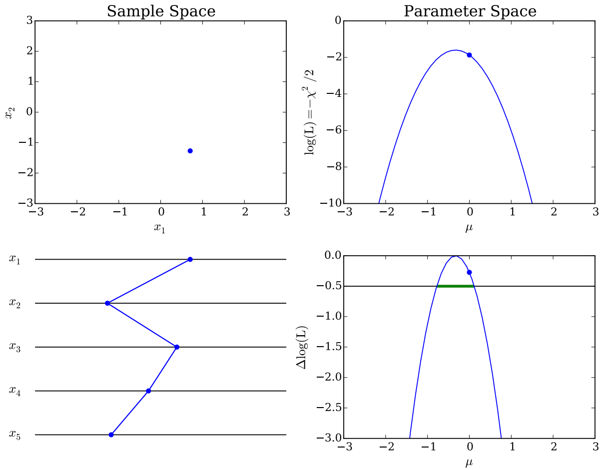


Luckily the $\Delta \log \mathcal{L}$ distribution is the same!
($\Delta \log \mathcal{L}$ is a *pivotal quantity*)

If it weren't, define *confidence level* $=$ maximum coverage over all μ (confidence level $=$ conservative guarantee of coverage)

Parametric bootstrap: Skip this step; just report the coverage based on $\mu = \hat{\mu}(\{x_i\})$ for the observed data. Theory shows the error in the coverage falls faster than $\sqrt{N}$.

Apply to observed sample



Report the green region, with coverage as calculated for ensemble of hypothetical data (red region, *previous slide*)

Likelihood to probability via Bayes's theorem

Recall the likelihood, $\mathcal{L}(\mu) \equiv p(D_{\text{obs}}|\mu)$, is a probability for the observed data, but *not* for the parameter μ

Convert likelihood to a probability distribution over μ via *Bayes's theorem*:

$$
\begin{aligned}
p(A, B) &= p(A)p(B|A) \\
&= p(B)p(A|B) \\
\rightarrow p(A|B) &= p(A)\frac{p(B|A)}{p(B)}, \quad \text{Bayes's th.}
\end{aligned}
$$

$$\Rightarrow p(\mu|D_{\text{obs}}) \quad \propto \quad \pi(\mu)\mathcal{L}(\mu)$$

$p(\mu|D_{\text{obs}})$ is called the *posterior probability distribution*

Requires a prior probability density, $\pi(\mu)$, often taken to be constant over the allowed region if there is no significant information available (or sometimes constant wrt some reparameterization motivated by a symmetry in the problem)

Gaussian problem posterior distribution

For the Gaussian example, a bit of algebra ("complete the square") gives:

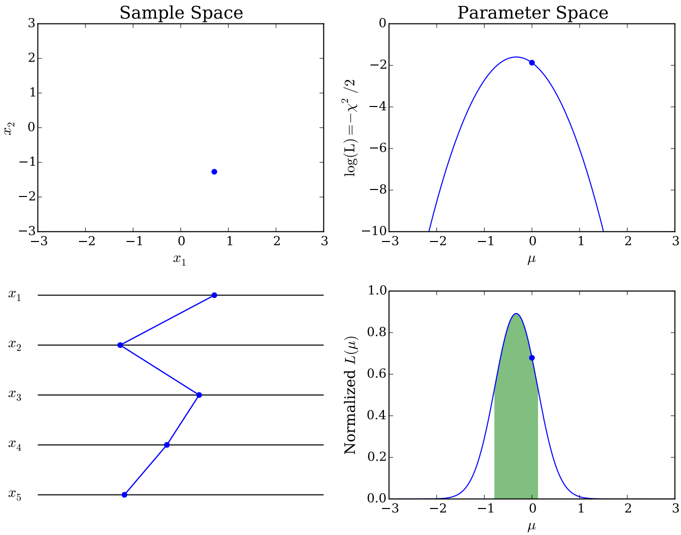
$$\begin{aligned}\mathcal{L}(\mu) &\propto \prod_i \exp\left[-\frac{(x_i-\mu)^2}{2\sigma^2}\right] \\ &\propto \exp\left[-\frac{(\mu-\bar{x})^2}{2(\sigma/\sqrt{N})^2}\right]\end{aligned}$$

The likelihood is Gaussian in μ

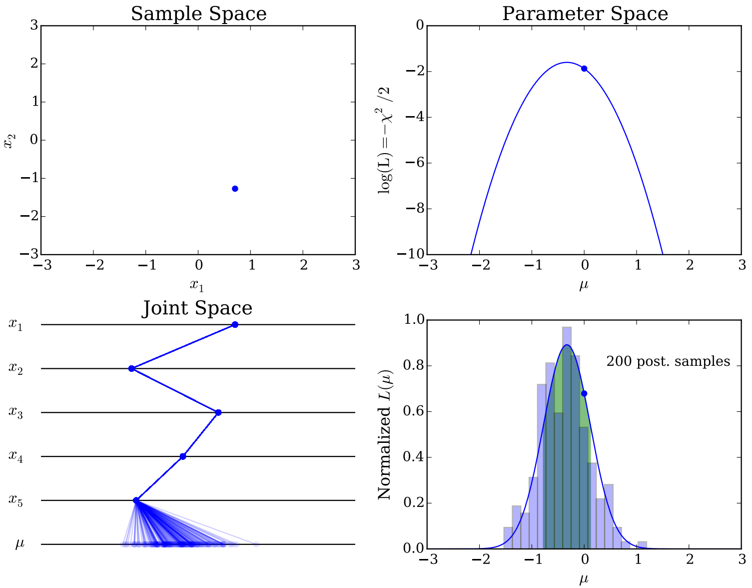
Flat prior $\rightarrow$ posterior density for μ is $\mathcal{N}(\bar{x}, \sigma^2/N)$

Bayesian credible region

Normalize the likelihood for the observed sample; report the region that includes 68.3% of the normalized likelihood



Credible region via Monte Carlo: *posterior sampling*



Posterior summaries

- Posterior mean is $\langle\mu\rangle \equiv \int d\mu\, \mu\, p(\mu|D_{\rm obs}) = \bar{x}$
- Posterior mode is $\hat{\mu} = \bar{x}$
- Posterior std dev'n is $\sigma/\sqrt{N}$
- $\bar{x} \pm \sigma/\sqrt{N}$ is a 68.3% *credible region*:

$$\int_{\bar{x}-\sigma/\sqrt{N}}^{\bar{x}+\sigma/\sqrt{N}} d\mu\ p(\mu|D_{\rm obs}) \approx 0.683$$

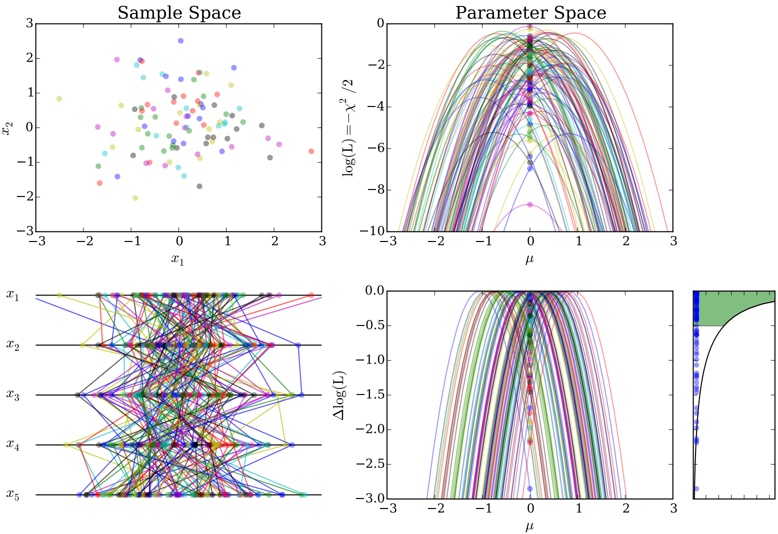
- $\bar{x} \pm 2\sigma/\sqrt{N}$ is a 95.4% credible region

The credible regions above are *highest posterior density* credible regions (*HPD regions*); these are the smallest regions with a specified probability content

These reproduce familiar frequentist results, but this is a *coincidence* due to special properties of Gaussians

Confidence region calculation (recap)

Likelihoods for 100 simulated data sets, $\mu = 0$



When They'll Differ

Both approaches report $\mu \in [\bar{x} - \sigma/\sqrt{N}, \bar{x} + \sigma/\sqrt{N}]$, and assign 68.3% to this interval (with different meanings)

This matching is a *coincidence*!

When might results differ? ($\mathcal{F}$ = frequentist, $\mathcal{B}$ = Bayes)

- If $\mathcal{F}$ procedure doesn't use likelihood directly
- If $\mathcal{F}$ procedure properties depend on params (nonlinear models, need to find pivotal quantities)
- If likelihood shape varies strongly between datasets (conditional inference, ancillary statistics, recognizable subsets)
- If there are extra uninteresting parameters (nuisance parameters, corrected profile likelihood, conditional inference)
- If $\mathcal{B}$ uses important prior information

Also, for a different task—comparison of parametric models—the approaches are qualitatively different (significance tests & info criteria vs. Bayes factors)

Bayesian and Frequentist inference

Brad Efron, ASA President (2005)

> The 250-year debate between Bayesians and frequentists is unusual among philosophical arguments in actually having *important practical consequences*. . . . The physicists I talked with were really bothered by our 250 year old Bayesian-frequentist argument. Basically there's only one way of doing physics but there seems to be at least two ways to do statistics, and *they don't always give the same answers*. . . .
>
> Broadly speaking, Bayesian statistics dominated 19th Century statistical practice while the 20th Century was more frequentist. What's going to happen in the 21st Century?. . . I strongly suspect that statistics is in for a burst of new theory and methodology, and that this burst will feature a combination of Bayesian and frequentist reasoning. . . .

Roderick Little, ASA President's Address (2005)

Pragmatists might argue that good statisticians can get sensible answers under Bayes or frequentist paradigms; indeed maybe two philosophies are better than one, since they provide more tools for the statistician's toolkit. . . . I am discomforted by this "inferential schizophrenia." Since *the Bayesian (B) and frequentist (F) philosophies can differ even on simple problems*, at some point decisions seem needed as to which is right. I believe our credibility as statisticians is undermined when we cannot agree on the fundamentals of our subject. . . .

An assessment of strengths and weaknesses of the frequentist and Bayes systems of inference suggests that *calibrated Bayes*. . . captures the strengths of both approaches and provides a roadmap for future advances.

[*Calibrated Bayes* = Bayesian inference within a specified space of models + frequentist ideas for model checking; Andrew Gelman uses *"Bayesian data analysis"* similarly]

(see arXiv:1208.3035 [by TL] for discussion/references)