Bayesians and Frequentists

Models, Assumptions, and Inference

George Casella
Department of Statistics
University of Florida

Approaches to Statistics

- Frequentists: From Neymann/Pearson/Wald setup. An orthodox view that sampling is infinite and decision rules can be sharp.
- Bayesians: From Bayes/Laplace/de Finetti tradition. Unknown quantities are treated probabilistically and the state of the world can always be updated.
- Likelihoodists: From Fisher. Single sample inference based on maximizing the likelihood function and relying on the Birnbaum (1962) Theorem. *Bayesians - But they don't know it.*

- So let's look at some critical differences between Frequentists and Bayesians...

Differences Between Bayesians and Non-Bayesians
According to my friend Jeff Gill

Typical Bayesian

Typical Non-Bayesian

Differences Between Bayesians and Non-Bayesians
What is Fixed?

Frequentist:

- Data are a repeatable random sample - there is a frequency
- Underlying parameters remain constant during this repeatable process
- Parameters are fixed

Bayesian:

- Data are observed from the realized sample.
- Parameters are unknown and described probabilistically
- Data are fixed

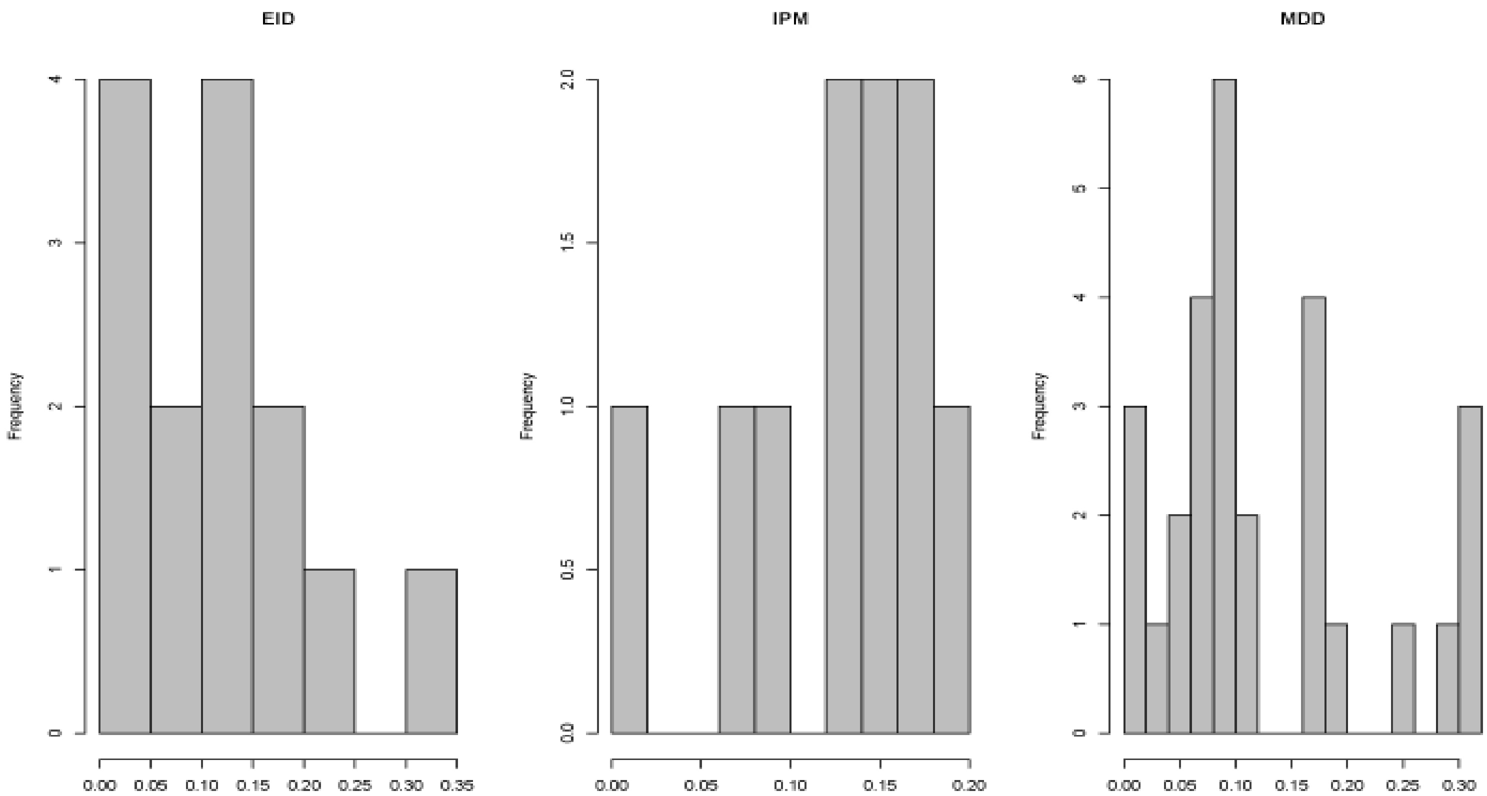
Example: Application of Bayes Theorem to Aminoglycoside-Associated Nephrotoxicity (AAN)

- Kim *et al.* (2004 *Journal of Clinical Pharmacology*)
- Examine the incidence of AAN related to
  - Extended-interval dosing (EID)
  - Individualized pharmacokinetic monitoring (IPM)
  - Multiple-daily dosing (MDD)
- Meta-analysis of published results
- Bayesian methods used

Example: Application of Bayes Theorem to AAN
-The Data-

	EID		IPM		MDD	
	Incidence of Nephrotoxicity	AAN Related	Incidence of Nephrotoxicity	AAN Related	Incidence of Nephrotoxicity	AAN Related
	34	8	80	13	66	11
	179	25	62	0	1756	129
	141	15	36	5	272	48
	187	14	98	12	151	18
Studies	⋮	⋮	⋮	⋮	⋮	⋮
	71	11	95	14	146	28
	40	2	78	7	140	14
	35	0			113	11
	34	2			108	10
	61	9				

Example: Application of Bayes Theorem to AAN -Histograms-



- Histograms of Relative Frequencies of AAN
- Protocols have similar means but different patterns

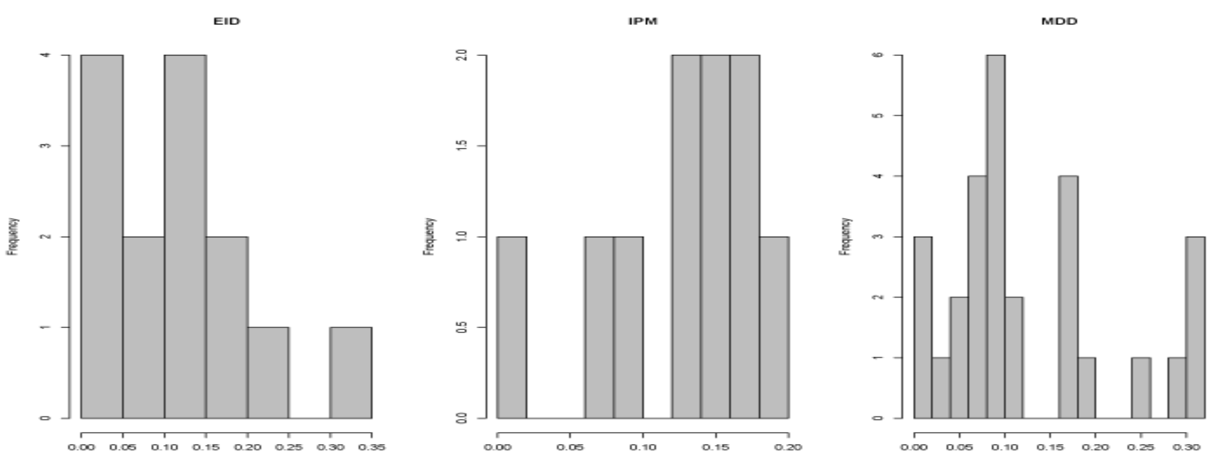
Example: Application of Bayes Theorem to AAN -Sampling Models-

Frequentist:

- For Protocol i, $= 1, 2, 3$, X=AAN frequency
- For Study j in Protocol i
  - $X_j \sim \text{Binomial}(n_j, p_i)$
- p_i is the same for each study

Bayesian:

- For Protocol i, $= 1, 2, 3$, X=AAN frequency
- For Study j in Protocol i
  - $X_j \sim \text{Binomial}(n_j, p_i)$
- p_i can vary from study to study



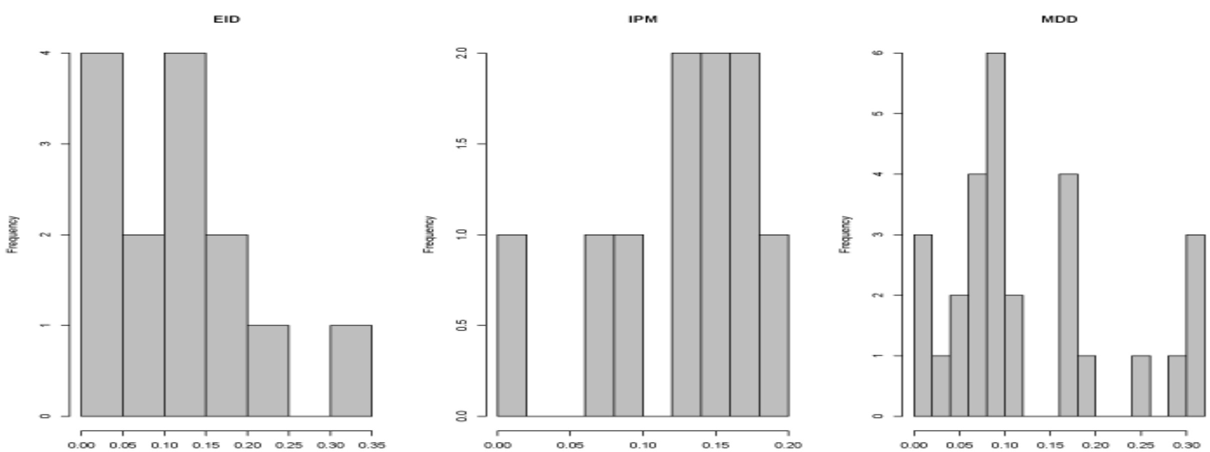
Example: Application of Bayes Theorem to AAN -Inference-

Frequentist:

- The true AAN rates p_1, p_2, p_3 are fixed
- The data are repeated
- Determine if p_1, p_2, p_3 are different

Bayesian:

- The data from the studies are fixed
- The true AAN rates p_1, p_2, p_3 can vary
- Determine if p_1, p_2, p_3 are different



Differences Between Bayesians and Non-Bayesians
What is Fixed?

Frequentist:

- Data are a repeatable random sample - there is a frequency
  -The studies are repeatable
- Underlying parameters remain constant during this repeatable process
  -The studies (in protocol) have same AAN rate
- Parameters are fixed

Bayesian:

- Data are observed from the realized sample
  -The studies are fixed
- Parameters are unknown and described probabilistically
  -The studies (in protocol) have varying AAN rates
- Data are fixed

- We see why Kim *et al.* used Bayesian Inference
- Difficult to assume that this "experiment" is repeatable
- The collection of studies is a one-time phenomenon

Differences Between Bayesians and Non-Bayesians
General Inference

Frequentist:

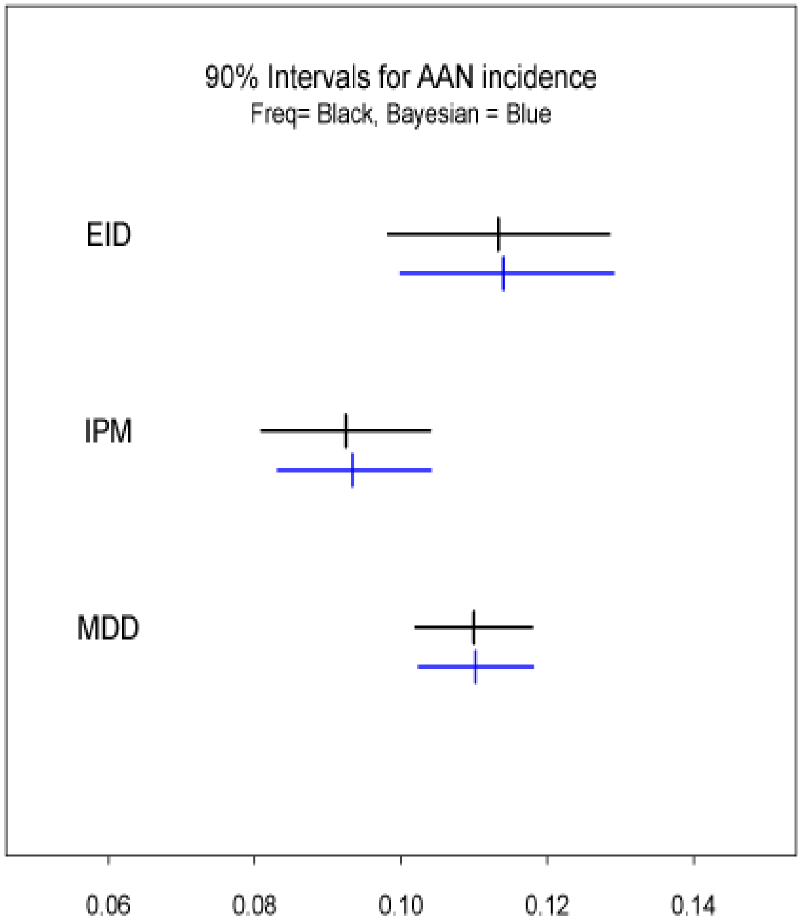
- Point estimates and standard errors or 95% *confidence* intervals.
- Deduction from $P(data|H_0)$, by setting α in advance.
- Accept H_1 if $P(data|H_0) < \alpha$.
- Accept H_0 if $P(data|H_0) \geq \alpha$.

Bayesian:

- Induction from $P(\theta|data)$, starting with $P(\theta)$.
- Broad descriptions of the posterior distribution such as means and quantiles.
- Highest posterior density intervals indicating region of highest posterior probability, regardless of contiguity.

- Frequentist: $P(data|H_0)$ is the sampling distribution of the data given the parameter
- Bayesian: $P(\theta)$ is the prior distribution of the parameter (before the data are seen)
  - $P(\theta|data)$ is the posterior distribution of the parameter
  - Update of the prior with the data (more later)

Differences Between Bayesians and Non-Bayesians
90% Intervals



Frequentist:

- In repeated sampling 90% of realized intervals cover the true parameter

Bayesian:

- For these data, with probability 90% the parameter is in the interval

- These are different probabilities

Example: Application of Bayes Theorem to AAN
-Construction of Confidence Intervals-

For Protocol $i, = 1, 2, 3$, X=AAN frequency

Frequentist:

- For Study j in Protocol i
  - $X_j \sim \text{Binomial}(n_j, p_i)$
- p_i is the same for each study
- Describe variability in X_j for fixed p_i

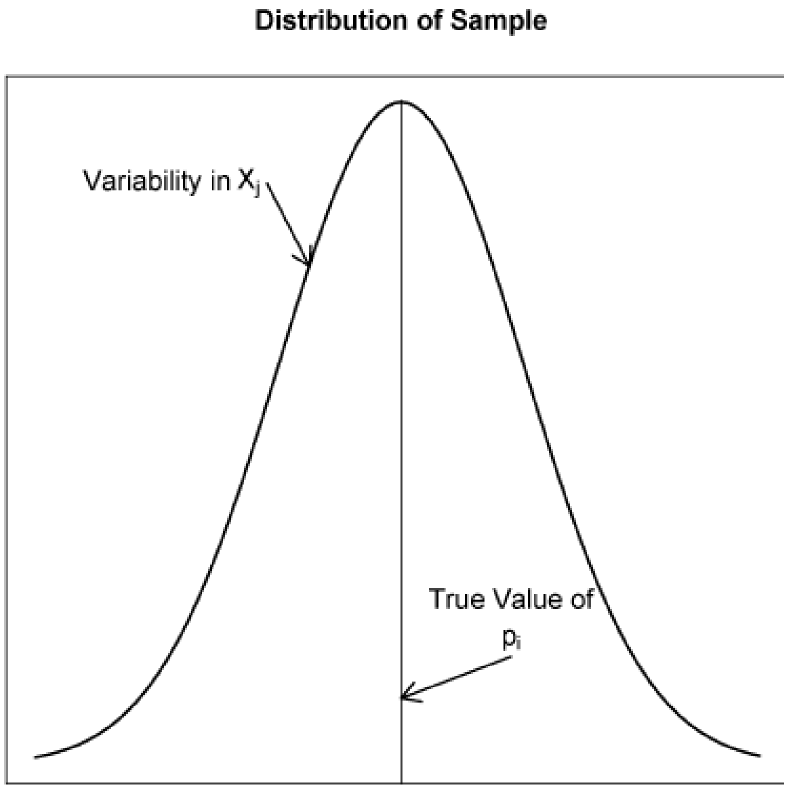
Bayesian:

- For Study j in Protocol i
  - $X_j \sim \text{Binomial}(n_j, p_i)$
- p_i has a prior distribution
- Describe variability in p_i for fixed X_j

-Construction of Confidence Intervals-

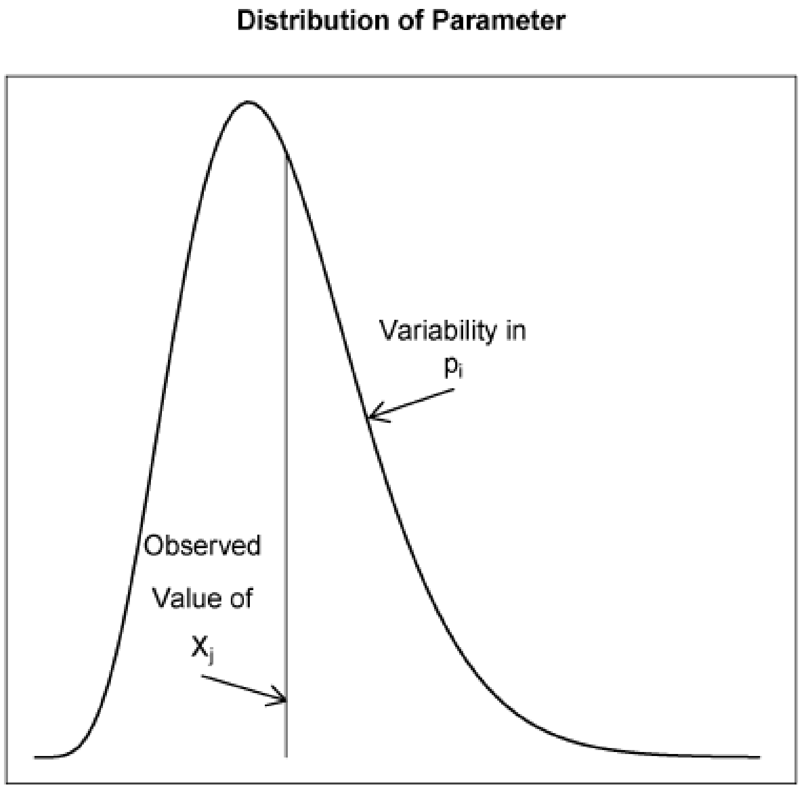
Frequentist:

- Describe variability in X_j for fixed p_i



Bayesian:

- Describe variability in p_i for fixed X_j



Three General Steps for Bayesian Modeling

I. Specify a probability model for unknown parameter values that includes some prior knowledge about the parameters if available.

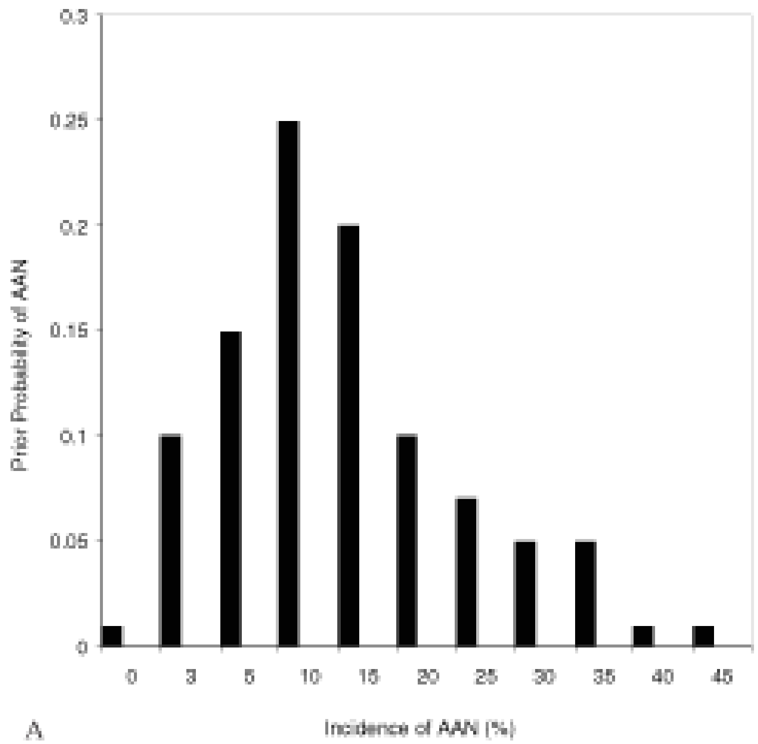
II. Update knowledge about the unknown parameters by conditioning this probability model on observed data.

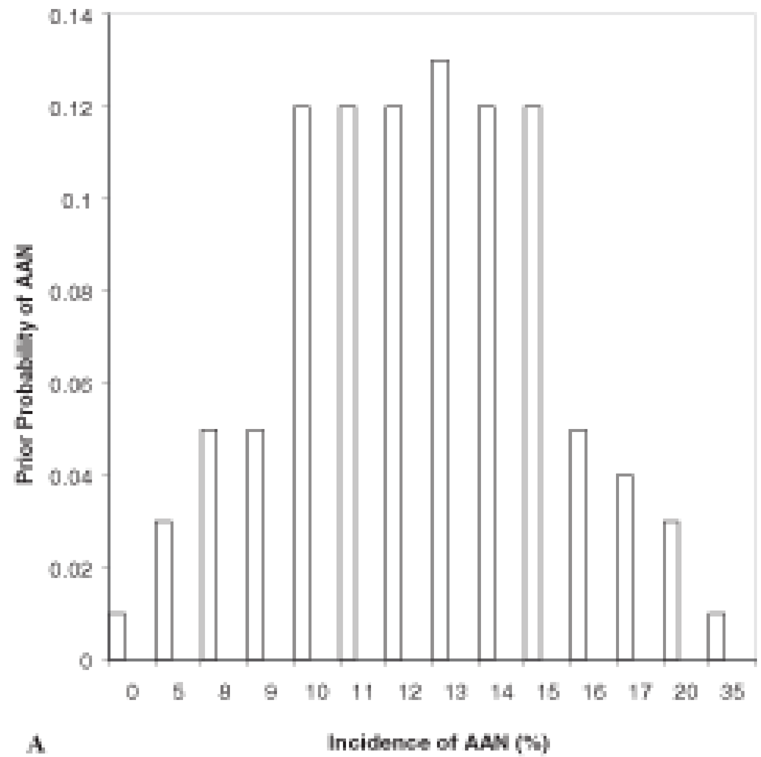
III. Evaluate the fit of the model to the data and the sensitivity of the conclusions to the assumptions. (Another time)

Where Do Priors Come From?

- Previous studies, published work.
- Researcher intuition.
- Substantive Experts
- Convenience (conjugacy, vagueness).
- Nonparametrics and other data sources.

Kim *et al.* 2004 Priors

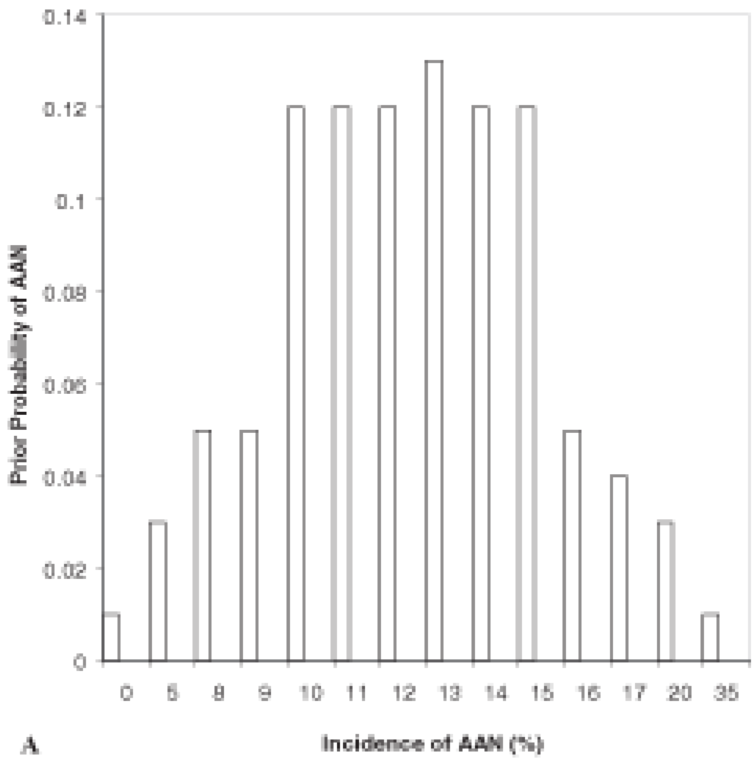


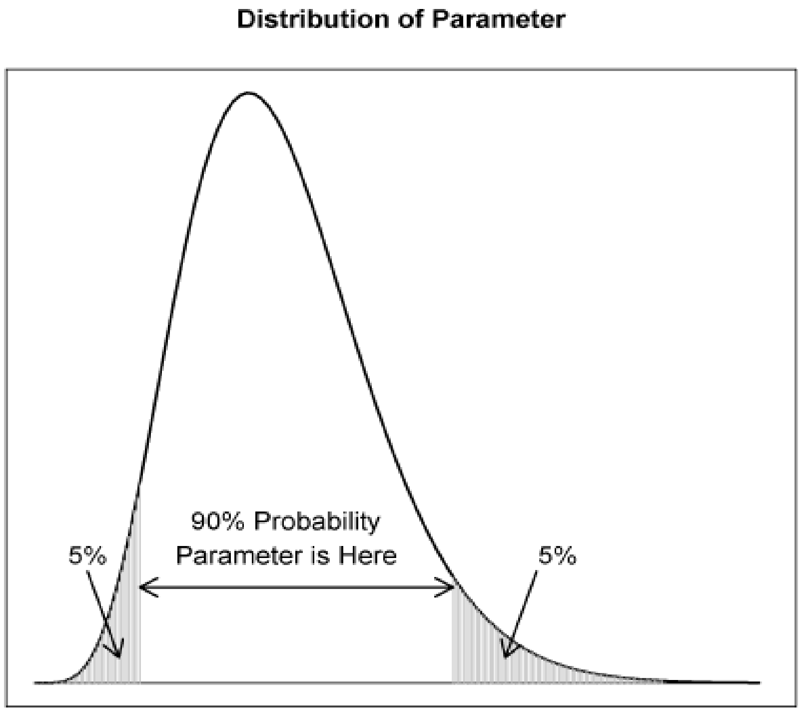


- First Prior
- From Review of Literature
- And Expert Judgement

- Second Prior
- Eliminate Extremes
- AAN > 35% Unlikely

Priors and Posteriors -Posterior Interval-

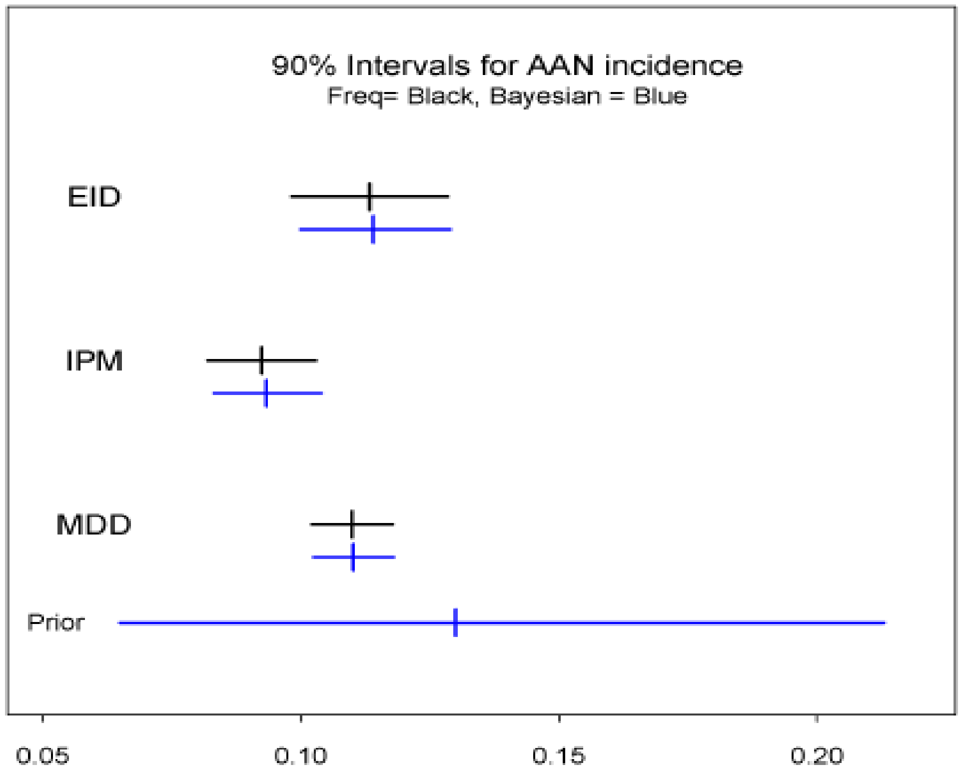




- Prior Distribution
- Before Data are Seen

- Posterior Distribution
- Prior Updated with Data
- 90% “Credible” Interval

Priors and Posteriors -Effect of the Prior-

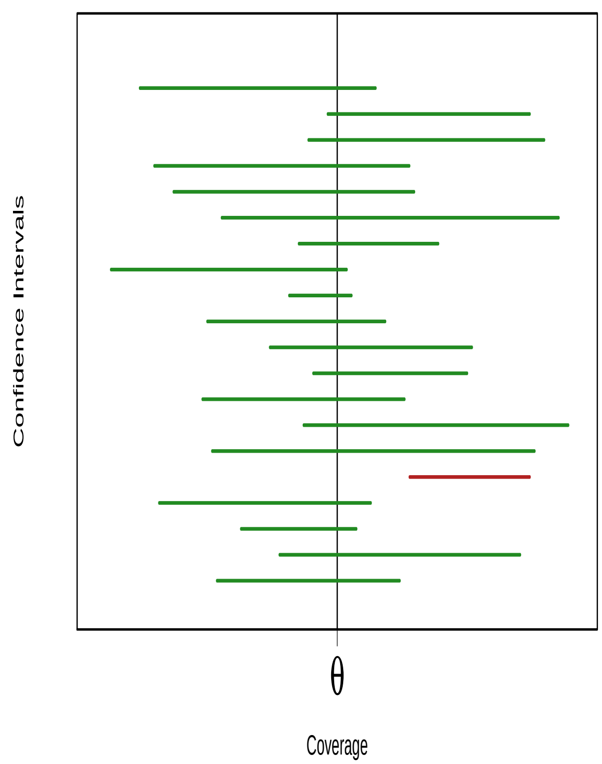


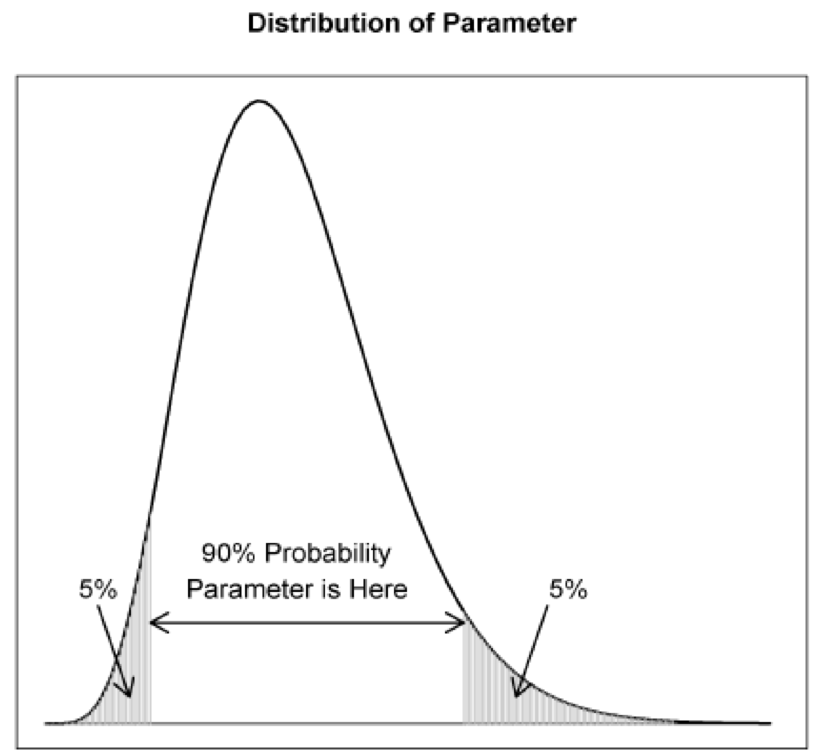
- Prior has high variability
- Data information is very strong
- Intervals Similar
- But Inference Remains Different

Interpretations of Confidence

- Frequentist: A collection of intervals with 90% of them containing the true parameter

- Bayesian: An interval that has a 90% chance of containing the true parameter.





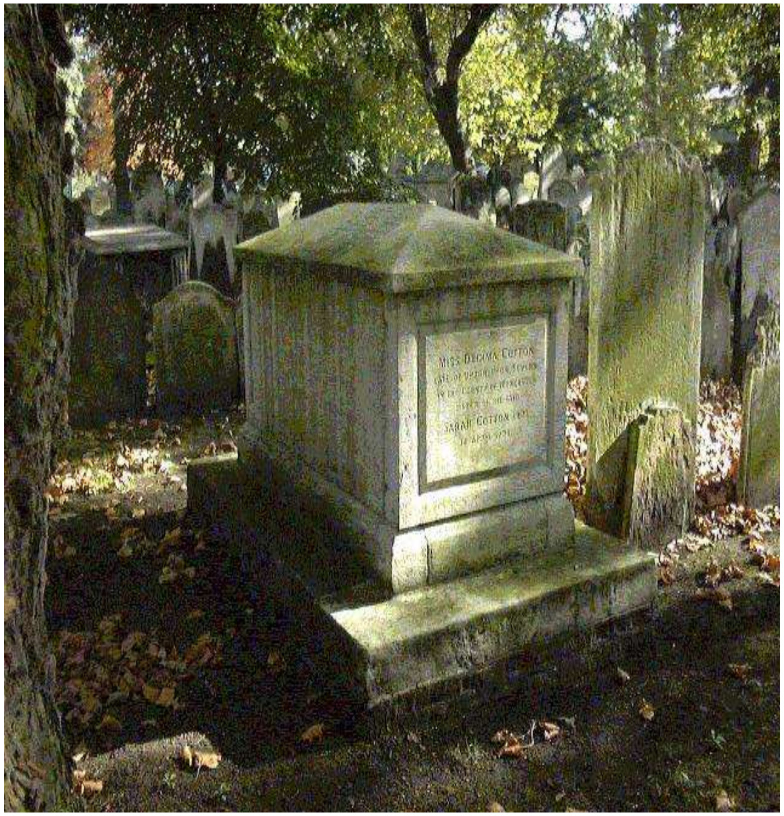
- Which interpretation *preferred*?.

So Why Did Frequency Win?

- 1950 – 1990 Nobody did Bayesian Analysis
  - Well Some, but on the fringe
- We want very automated, “cookbook” type procedures - or that is what we sold.
- Computers were slow and relatively unavailable.
- Bayesian Statistics need Lots of computation

And the everything changed....

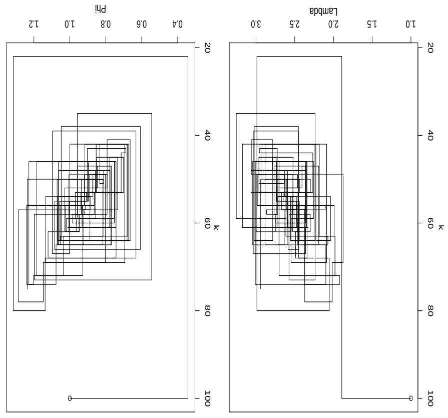
The History of Bayesian Statistics–Milestones



- Reverend Thomas Bayes (1702-1761).
- Pierre Simon Laplace.
- Pearson (Karl), Fisher, Neyman and Pearson (Egon), Wald.
- Jeffreys, de Finetti, Good, Savage, Lindley, Zellner.
- A world divided (mainly over practicality).
- The revolution: Gelfand and Smith (1990).
- Today…

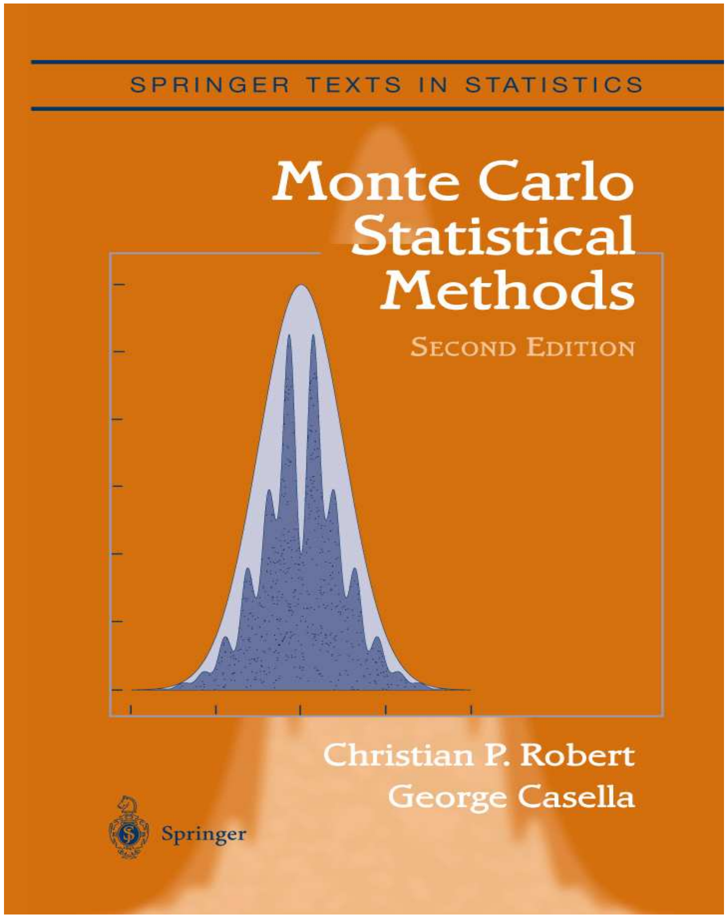
Technologies that have changed my life:

- microwave ovens
- ATM machines
- pay-at-the-pump
- Gibbs sampling



Markov Chain Monte Carlo
Gibbs Sampling and Variations

- Computational algorithms
- Cracked open countless problems
- Research explosion 1990 – 2005
- Allowed solutions of
  - Practical Problems
  - Complex Models



PKPD Medical Models - An Example from the Book

- Pharmacokinetics is the modeling of the relationship between the dosage of a drug and the resulting concentration in the blood.
- Estimate pharmacokinetic parameters using mixed-effects model and
- For a given dose d_i administered at time 0 to patient i, the measured log concentration in the blood at time t_{ij}, X_{ij}, is assumed to follow a normal distribution

$$X_{ij} \sim N(\log g_{ij}, \sigma^2),$$
$$g_{ij}(\lambda_i) = \frac{d_i}{V_i} \exp\left(-\frac{C_i \, t_{ij}}{V_i}\right).$$

- C_i represents *clearance*
- V_i represents *volume* for patient i.

PKPD Medical Models - Cadralazine Concentration

- Wakefield *et al.* 1994 *applied Statistics*
- Data on 10 Cardiac Failure Patients
- Plasma Concentration after 30mg Dose

	Hours After Administration					
Patient	2	4	6	8	10	24
1	1.09	0.7	0.53	0.34	0.23	0.02
2	2.03	1.28	1.2	1.02	0.83	0.28
3	1.44	1.3	0.95	0.68	0.52	0.06
4	1.55	0.96	0.8	0.62	0.46	0.08
5	1.35	0.78	0.5	0.33	0.18	0.02
6	1.08	0.59	0.37	0.23	0.17	0
7	1.32	0.74	0.46	0.28	0.27	0.03
8	1.63	1.01	0.73	0.55	0.41	0.01
9	1.26	0.73	0.4	0.3	0.21	0
10	1.3	0.7	0.4	0.257	0.14	0

PKPD Medical Models

Estimates of Log Clearance and Log Volume

Patient	Log Clearance	Log Volume
1	1.269	2.854
2	0.2877	2.624
3	0.6723	2.721
4	0.8287	2.71
5	1.219	2.642
6	1.472	2.763
7	1.257	2.666
8	0.8884	2.599
9	1.309	2.682
10	1.328	2.624

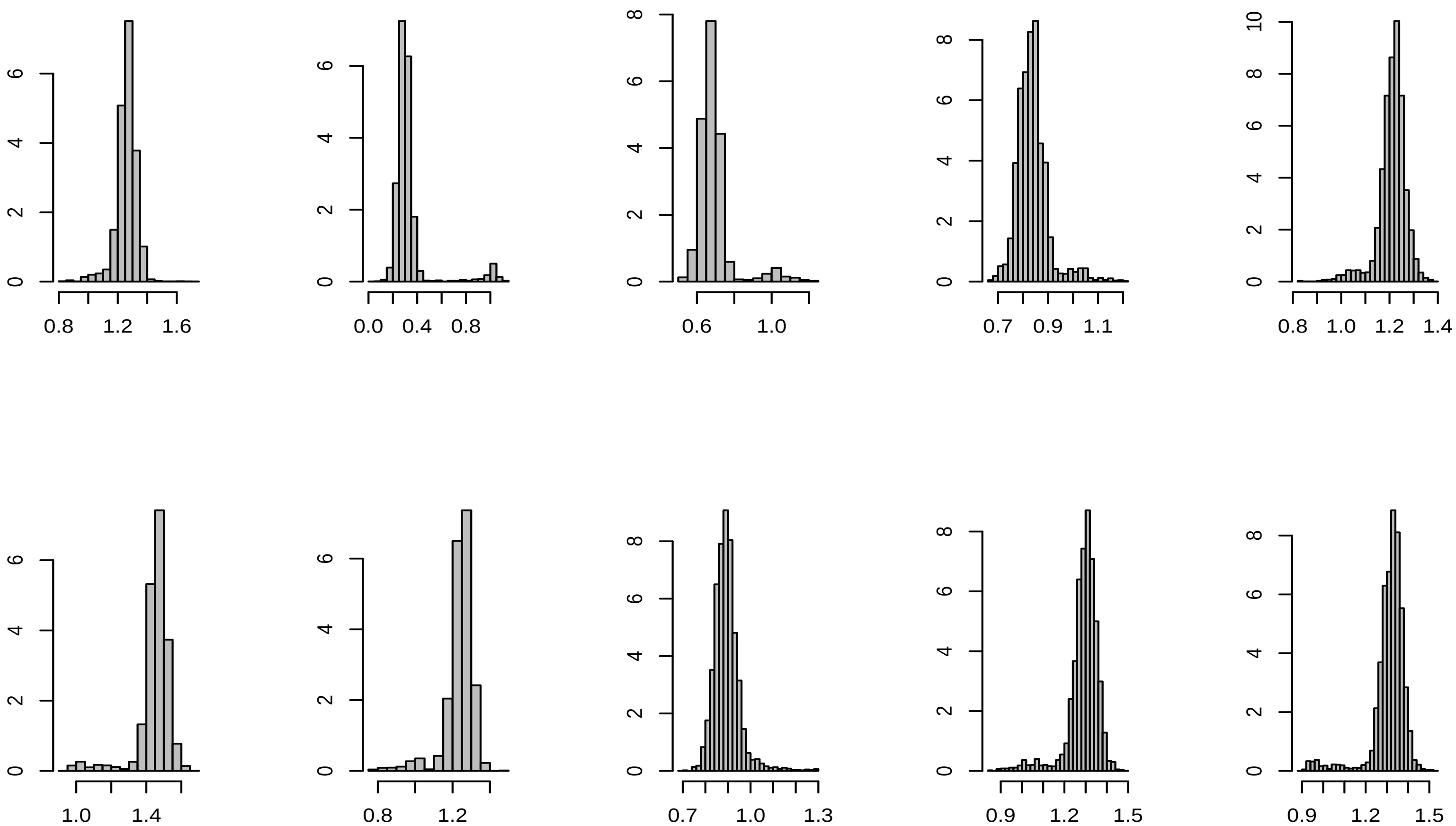
- Std Error $\approx .07$

- Variability in Clearance
- Lesser variability in Volume?

- With MCMC we can see more
- Have the entire posterior distribution for C and V

PKPD Medical Models

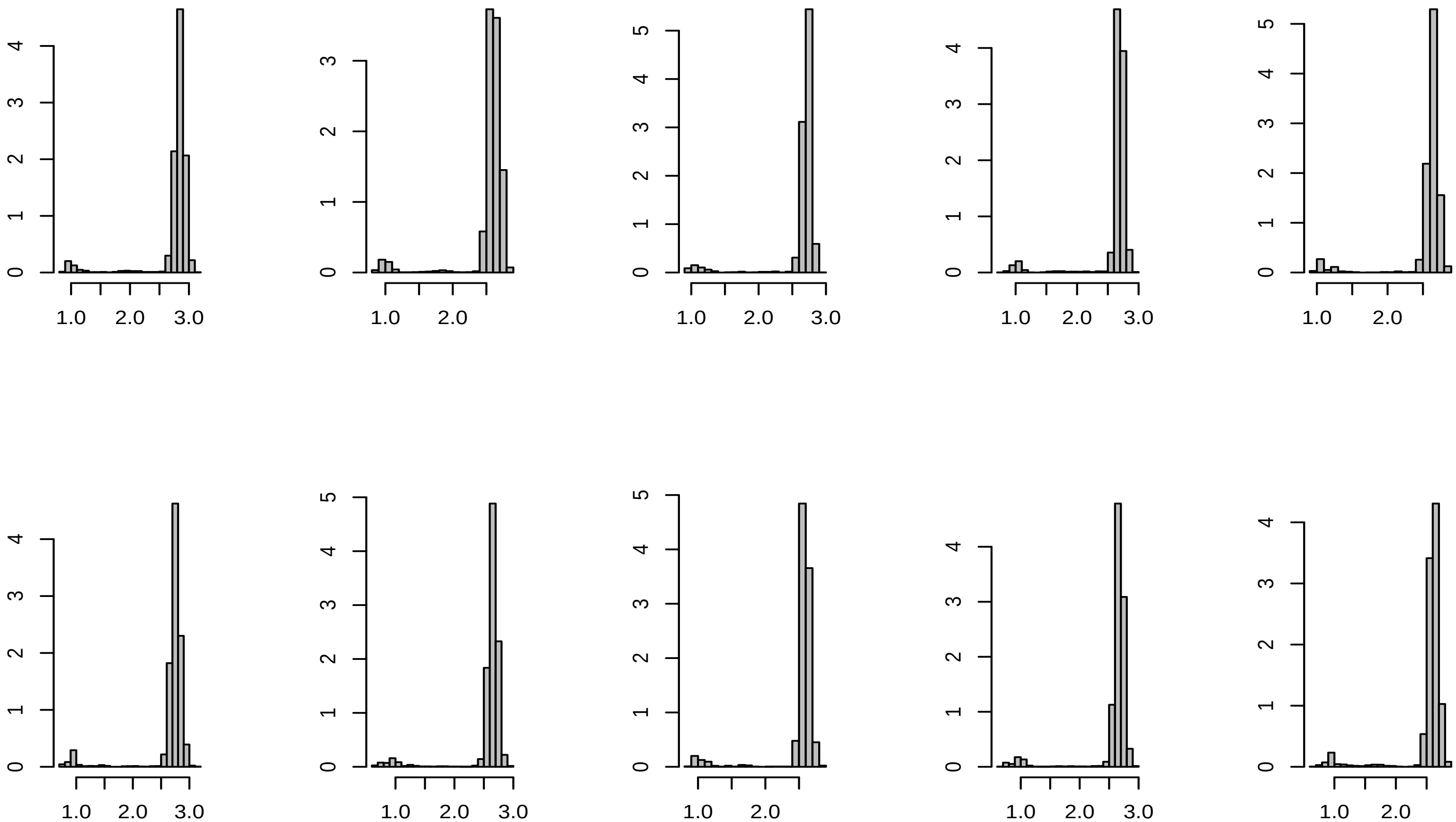
Posterior Distribution of Log Clearance For All Patients



- Large Mean Differences
- Evidence of Bimodality

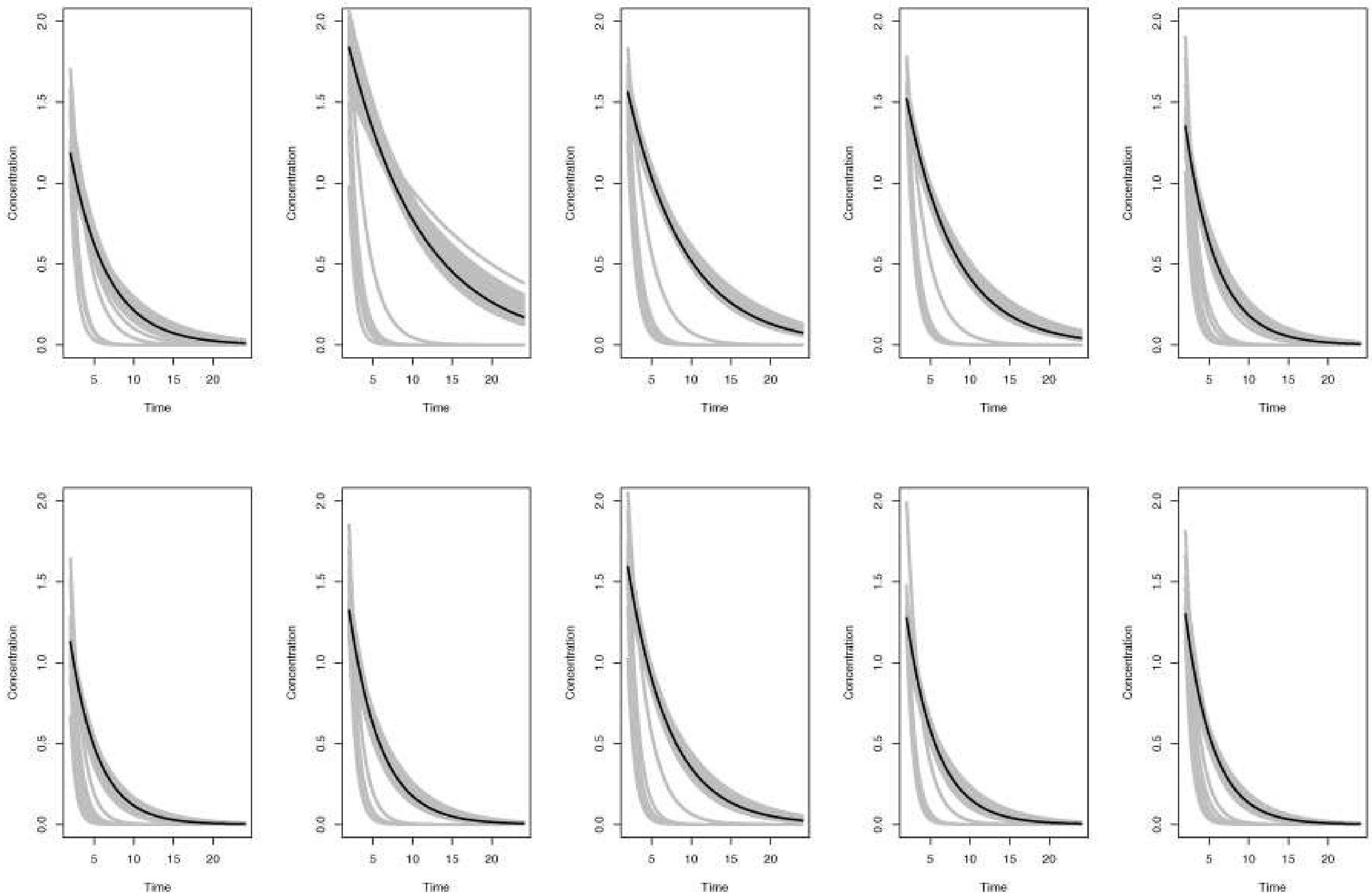
PKPD Medical Models

Posterior Distribution of Log Volume For All Patients



- Small Mean Differences
- Evidence of Bimodality

PKPD Medical Models - Shape of the Curves



- ▶ Multimodal parameter distribution ⇒ distinct curves

Differences Between Bayesians and Frequentists

Frequentist:

- The parameters of interest are fixed and unchanging under all realistic circumstances.
- No information prior to the model specification.

Bayesian:

- View the world probabilistically, rather than as a set of fixed phenomena that are either known or unknown.
- Prior information abounds and it is important and helpful to use it.

Differences Between Bayesians and Frequentists

Frequentist:

- Statistical results assume that data were from a controlled experiment.
- Nothing is more important than repeatability, no matter what we pay for it.

Bayesian:

- Very careful about stipulating assumptions and are willing to defend them.
- Every statistical model ever created in the history of the human race is subjective; we are willing to admit it.

- Berger and Berry "Statistical Analysis and the Illusion of Objectivity" *American Scientist* 1988

But in the End

- We are Statisticians
- We should use all of our tools

Frequentist:

- Evaluative Paradigm
- Repeatability can be Important

Bayesian:

- Modeling Paradigm
- Inference can be appropriate

- Bring what is needed to Solve the Problem

Thanks for your Attention

Thank You and Go Gators

casella@ufl.edu