Why We Should All Be Bayesians

The Problem of Inferential Statistics in Organizational Science

David Krackhardt ...

Test Results of Pregnancy Test in Maternity Clinic*

θ

y

	Actually Pregnant	Actually Not Pregnant	
Test says Pregnant	197	0	1.0
Test says Not Pregnant	1	2	.667
Total	198	2	

A different kind of being right

$$p(\theta \mid y)$$

Probability of being right: .995 1.0

$$p(y \mid \theta)$$

*Barnett, 1994

Critical Insight #1:

What we care about as scientists is $p(\theta \mid y)$

... not $p(y \mid \theta)$

And, they are almost always NOT the same thing.

Test Results of Pregnancy Test in Maternity Clinic

	Actually Pregnant	Actually Not Pregnant
Test Says Pregnant	197	0
Test Says Not Pregnant	1	2
Total	198	2

Probability of being right: .995 1.0

Barnett: Problem is based on tests of 2 women who were not pregnant (i.e., small sample size).

The pharma collects more data...

	Actually Pregnant	Actually Not Pregnant		
Test Says Pregnant	197,000	0	**1.0**	**Probability of Truth given the test results (observed data)**
Test Says Not Pregnant	1,000	2,000	**.667**	
Total	**198,000**	**2,000**		
Probability of test being right given the Truth:	**.995**	**1.0**		

Critical Insight #2:

Having more data does not solve the problem if you are answering the wrong question.

What Traditionalists do...

- **Assume** the population looks like θ, ask: What is the probability that the sample we observe looks like y (hypothesis testing)?

$$p(y \mid \theta)$$

This is **not an inference** – It is a mathematical deduction. Traditional statistics gives a great answer to the wrong question.

What Bayesians do is...

$$p(y \mid \theta) \Rightarrow p(\theta \mid y)$$

$$p(\theta \mid y) = p(y \mid \theta) \; \boldsymbol{x} \; \textbf{\textit{Bayesian Magic}}$$

$$p(\theta \mid y) = p(y \mid \theta) \bullet \frac{p(\theta)}{p(y)}$$

Test Results of Pregnancy Test in Maternity Clinic

	ACTUALLY PREGNANT	ACTUALLY NOT PREGNANT
TEST SAYS PREGNANT	197	0
TEST SAYS NOT PREGNANT	1	2
TOTAL	198	2

$$p(y\,|\,\theta) = \frac{2}{2} = 1.0$$

$$p(\theta) = \frac{2}{200} = .01$$

$$p(y) = \frac{3}{200} = 0.015$$

$$p(\theta\,|\,y) = p(y\,|\,\theta) \bullet \frac{p(\theta)}{p(y)}$$

$$p(\theta\,|\,y) = 1.0 \bullet \frac{.01}{.015} = .667$$

The General Framework in Science

Truth

Data

	Actually Pregnant	Actually Not Pregnant
Test Says Pregnant	?	?
Test Says Not Pregnant	?	?

Truth

	Actually Pregnant	Actually Not Pregnant	
Test Says Pregnant	?	?	198
Test Says Not Pregnant	?	Likelihood: 100% of “Not Pregnant” people will test negative.	3
	Prior = 99.5%	Prior = .5%	200

Data

Normalizing Constant (P(*y*)) = 3/200 =.015%

Conclusion: P(Not pregnant | test says “not pregnant”)

= Prior * Likelihood/Normalizing Constant

= .005 * 1.0/.015 = .333

College Campus

Truth

Data

	Actually Pregnant	Actually Not Pregnant	
Test Says Pregnant	?	0	9,900
Test Says Not Pregnant	?	Likelihood: 100% of “Not Pregnant” people will test negative.	90,100
	Prior = 10%	Prior = 90%	100,000

Normalizing Constant (P(y)) = 90100/100,000 =90.1%

Conclusion: P(Not pregnant | test says “not pregnant”)

= Prior * Likelihood/Normalizing Constant

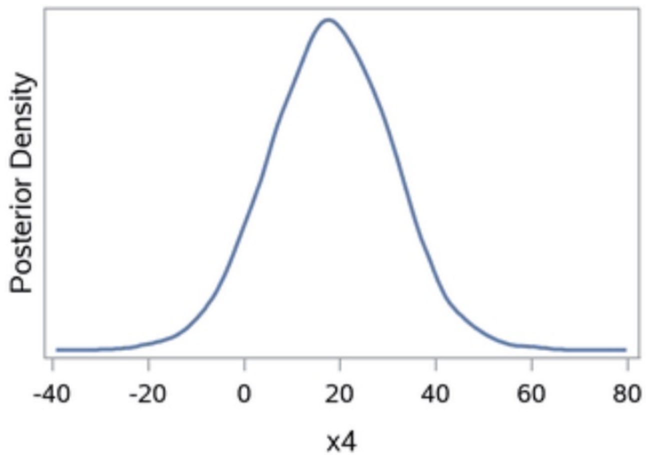
= .9 * 1.0/.901 = .9989

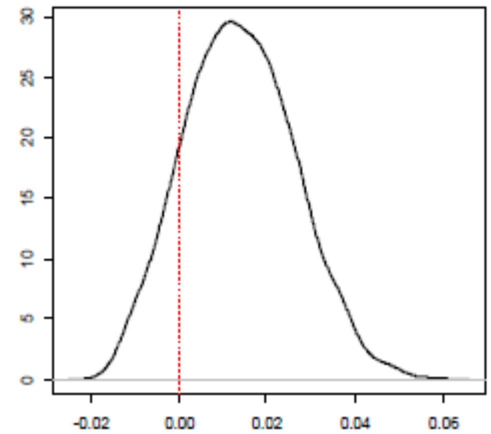
Critical Insight #3

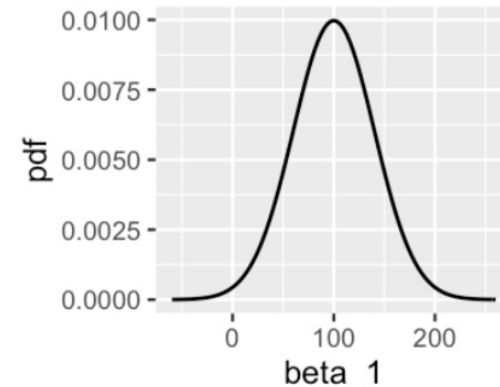
- Bayesians test their assumptions of their prior distributions by looking at how sensitive their results are to other alternative but reasonable priors.
  - They can use hyperparameters to describe and explore this sensitivity.
  - These hyperparameters can be used to look at meta analyses, also.

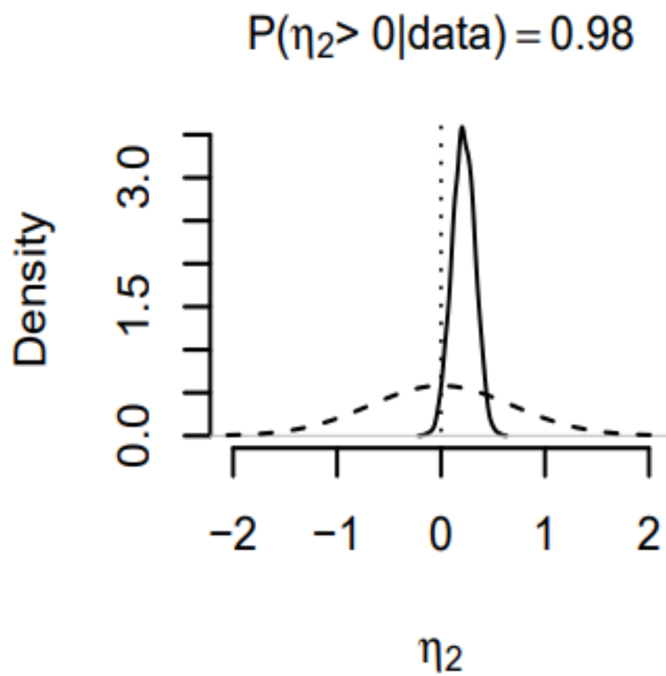
Critical Insight #4

- Bayesians consider population parameters to be random variables, not fixed points. The key, then, is to describe the likelihood distribution of that population parameter.
  - This is often done graphically, not just with numbers.


Posterior Density
-40
-20
0
20
40
60
80
x4


0
5
10
15
20
25
30
-0.02
0.00
0.02
0.04
0.06


pdf
0.0100
0.0075
0.0050
0.0025
0.0000
0
100
200
beta_1


$P(\eta_2 > 0|\text{data}) = 0.98$
Density
0.0
1.5
3.0
−2
−1
0
1
2
η_2

Bayesian Analysis

1. Forces us to confront our assumptions about the population (although can resort to "uninformative priors" if have no idea).
2. Gives us the power to test the **sensitivity** of our results based on our assumptions.
3. **Answers the right question: What can I infer about the population given the data I have?**