Case Western Reserve U
Dept Statistics
Oct 1 2004 Yost 300

Is there statistical inference?
The Bayesian-frequentist divergence!

D A S Fraser
U Toronto
Dept Statistics

A	Preamble	5
B	What is Bayesian? ... frequentist?	8
C	Does prior give Answer	3
D	Summary	2
		18

A Preamble

1) High energy physics

$y \sim \text{Poisson}(\theta)$ $\quad \theta \geq 6.7$ $\quad y^0 = 27$

Is there a Top quark ?

$\theta = 6.7 + \mu$ $\quad$ μ = Mean Top quark count

Is $\mu > 0$?

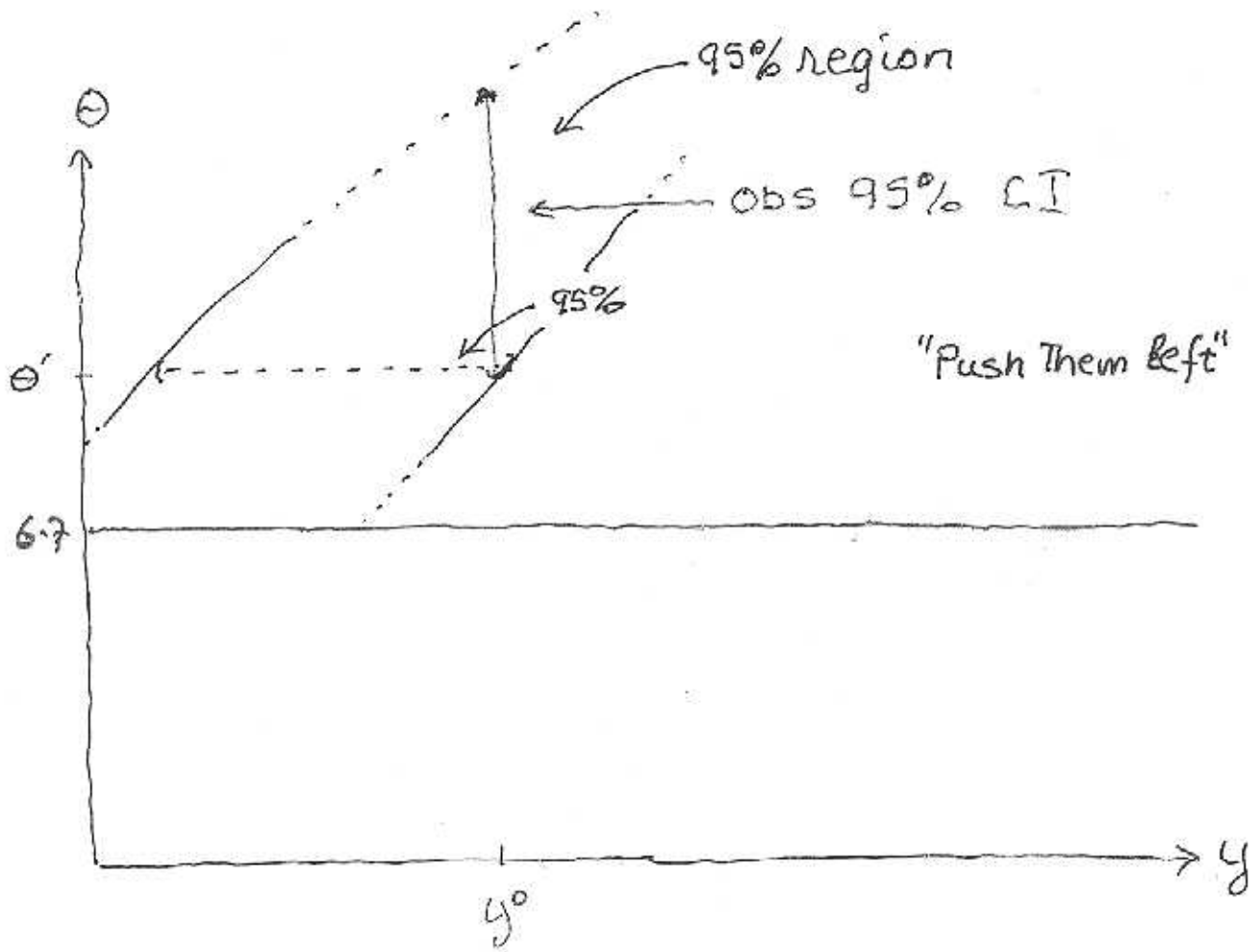
F. Abe et al., *Phys. Rev. Lett.* **73(2)** 225 (1994).

Fraser, D.A.S., Reid, N. and Wong, A. (2004). Setting confidence intervals for bounded parameters: a different perspective. *Physics Review D*, **69**, 033002.

2) What was done?

- Focus on confidence intervals
- Seek procedures that avoid $(0, 6.7]$:

....ingenuity, inventiveness, optimality, ...



- Push acceptance sets to left ... high principle!
- Push CI's to right ... automated!

"Deliberately" distort info re: $\boxed{\text{Is } \theta > 6.7?}$

- Dark days of statistics!
- See Stat Sc 17 Mandelkern

Not addressed in Discussion to M.

3) What can be done?

Record / Plot:

Probability at data $L(\theta)$ *

Probability left of data $p(\theta)$

"Tell whole story!"

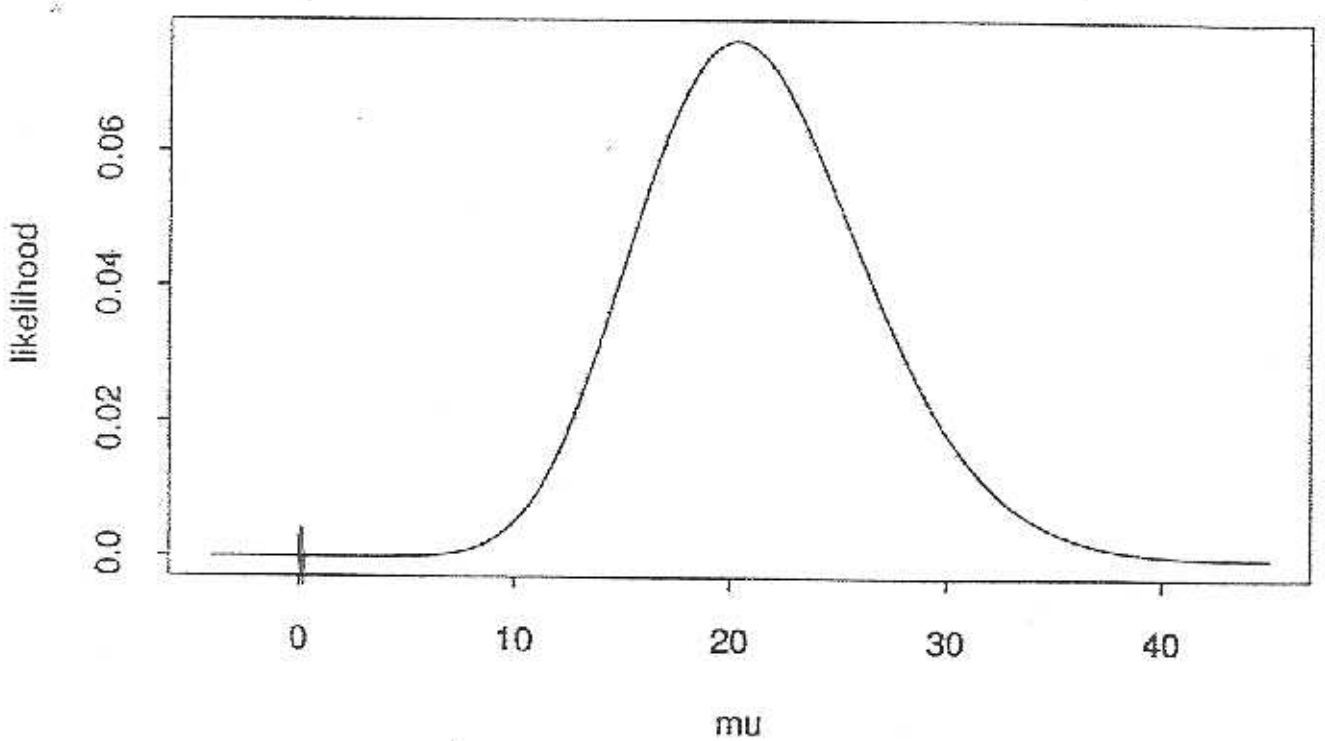
"Forget the decision theory for the problem

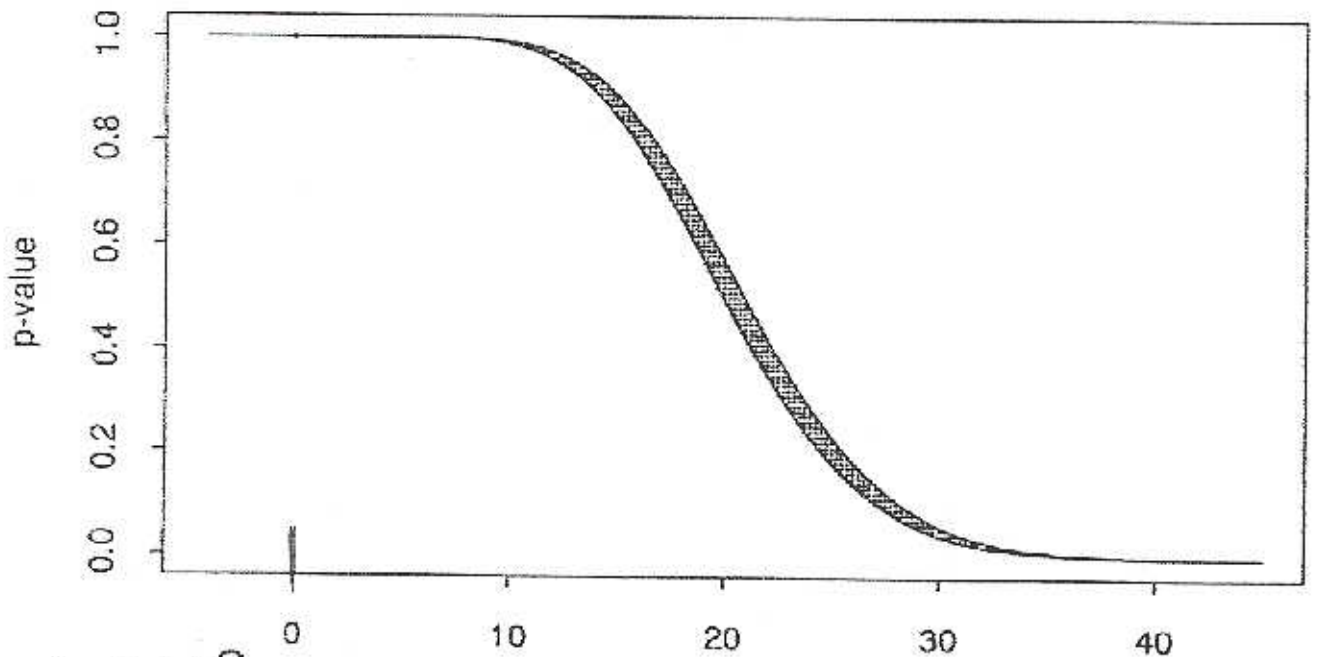
See: Fraser, Reid, Wong Phys Rev D 69 033002

BUT ..... in wide generality ... Computer *

Can do in wide generality, uniquess, ... !

References later...





$1 - 3 \cdot 10^{-8}$
$1 - 10^{-8}$

$\mu \approx \theta - 6.7$

FIG. 2. The likelihood function (top) and p-value function (bottom) for the Poisson model, with $b = 6.7$ and $y^0 = 27$. For $\mu = 0$ the upper and lower p-values are essentially 1.

/www.utstat.toronto.edu/dfraser/documents/224.pdf

4) <u>More on "Our Dean"</u> & "High energy physics"

Publication re "Top Quark" was <u>delayed</u>...

because <u>Bayesian, frequentist disagreement</u>

Next collider <u>delayed</u> .... cost many M$s

.... physicists <u>couldn't</u> find: "Statistical inference"

Astonishing?
Appalling?
Blight on our profession?

Complacency?

Stat Sc 17 "sounds OK"
but it tells a <u>sad story</u> Bankrupt!

B What is Bayesian? frequentist?

(1) Context:

Model: $f(y;\theta)$ $\qquad \boxed{\theta} \to y$

Data: y° $\qquad \boxed{} \to y^\circ$

other contexts { Exploratory analyses, EDA writ-BIG } EDA, algorithmics } other...

Model based statistics of core interest!

Next issue: Statistical Science Vol 19 "May"

(2) Common B-f split

$f(y;\theta) \quad + \quad y^\circ$

now add "prior" $\pi(\theta)$ ...

$$\pi(\cdot) \longrightarrow \boxed{\theta} \longrightarrow y \qquad \text{Observe } y^\circ$$

*	Joint (θ, y)	$\pi(\theta) f(y;\theta)$
	Cond'l $\theta \mid y^\circ$	$c\,\pi(\theta) L(\theta)$

Beyond reproof? .... or?

* Bayesian paradigm,
Information processing,
Cultural split,
Persuasions, belief
Thomas Kuhn

3) Add a prior $\pi(\theta)$??

Where did it come from?

1) Objective (truly!) "corresponds to..." → Prob model

2) "Objective" (default)* → Statistical

3) Subjective → Subjective

Q: Does one want to add it?
No imperative!

and Bayes himself... appealed to invariance*

4) f-B split

Operational definition here....
frequentist inference = Based on Model + Data only

Split

frequentist	Bayesian
Model + Data (Add ~~optimality... Decision Theory~~)	Model + Data + Subjective prior
Model + Data + Objective prior or + Default prior	

These are just frequentist short cuts:
- Weight the likelihood (Fisher, Gauss)
- Welch & Peers (1963) and earlier
- Its not Bayesian; "Territory"

Take as "frequentist"

Take as "Bayesian"

Reclaim the frequentist territory!

5) <u>Lots of subjective priors in Toronto!</u> USA?

- Instance being investigated by a police officer
- Subjective prior

otherwise known as <u>Profiling</u>

This is . . . . . .

Pure subjective Bayes

Detective could be disciplined under Act, where dismissal is the ultimate sanction

Globe and Mail 1/04

POLICE *from page A1*

Toronto Chief Julian Fantino, who always has disputed that racial profiling is widespread but pledged "zero tolerance" where it is found, moved quickly yesterday to order an investigation by the professional standards branch, saying, "I take the judge's findings and comments very seriously."

Potentially, Det. Asselin could be disciplined under the Police Services Act, where dismissal is the ultimate sanction, or prosecuted criminally for such offences as perjury or obstructing justice.

Mr. Khan is a York University graduate and at the time of his arrest in October of 2001, was on medical leave from his job as a teacher with the Toronto District School Board. He is also a successful real-estate broker.

He was stopped on Oct. 22, 2001, by the two officers ostensibly for driving erratically — a claim the judge disbelieved — and during a subsequent search of his car, a kilogram of cocaine worth $100,000 was discovered in a sealed plastic baggie, which was inside another sealed baggie, which was in turn inside a large garbage bag.

Mr. Khan's lawyer, John Struthers, argued that the search was illegal because it was motivated by racial profiling, and that the drug evidence therefore should be excluded.

Judge Molloy agreed, and acquitted Mr. Khan last June.

But in her lengthy reasons, she also ensured that no one could ever claim the charge was dismissed on a technicality.

That technicality — a breach of Mr. Khan's Section 9 Charter-guaranteed right against arbitrary detention — was serious enough.

But the judge also found that Mr. Khan was innocent, didn't know the cocaine was in his car (he had lent the vehicle to his brother for the previous weekend) and that Det. Asselin was lying when he said he had found it under the driver's front seat.

She went to extraordinary lengths to test the veracity of Det. Asselin's claim that the drugs were found under the seat and that the second he opened Mr. Khan's door, the telltale smell of cocaine (like cat urine with a chemical hint, apparently) "came flying out of the car."

"I was curious about that fact," Judge Molloy wrote, so she had the seized cocaine brought into the courtroom and placed on the dais beside her during the trial.

"I could smell nothing," she said, so she had the drug removed from its plastic exhibits bag, and ordered that the bag be fully opened. "I could smell nothing."

Then she had the open bag placed on the floor at her feet "to see if the smell would rise up. It did not. I could smell nothing."

Finally, the judge stuck her nose to the bag, and sure enough, there it was — a detectable smell that "exactly matched" Det. Asselin's description.

"However," the judge said, "it certainly did not come 'flying out' and it is difficult to believe that even the most sensitive of human noses would have smelled it to the extent described by Officer Asselin."

She concluded the drugs were secreted somewhere else in the car where they had gone unnoticed by Mr. Khan.

She also found that Mr. Khan's version of events was bolstered by the officers' own notes and by cellphone records demonstrating that after being stopped, he was allowed for almost 10 minutes to call friends as he tried to arrange for someone to come and take his car home.

"Why would an experienced officer who had just placed a believed drug dealer under arrest . . . think it would make sense for the drug dealer's friends to come and pick up the car? . . . What was he [Det. Asselin] doing?" the judge wondered.

"In my view," she said flatly, "the evidence is overpowering that the testimony of Officer Asselin and Officer James is untrue."

Judge Molloy's findings have their roots in a controversial 1999 case involving the former Toronto Raptors player Decovan (Dee) Brown.

On Nov. 1 that year, Mr. Brown was stopped for speeding and then charged with impaired driving.

At trial, his lawyer argued that Mr. Brown had been stopped only because he was a young black man driving a brand new Ford Expedition.

Eventually, Mr. Brown pleaded guilty, but not before the case went to the Ontario Court of Appeal, which formally recognized that racial profiling exists, and laid down rules for determining when it is at play.

Since police are unlikely to admit to racial basis, the court said, first the subject has to fit the profile of the profiled — be a young black male in a posh car — and second, the evidence has to allow the judge to infer that the police are lying.

Mr. Khan's lawyer, Mr. Struthers, was blunt.

"This is the first 'driving while black' case in Canadian history I'm aware of," he said yesterday at a press conference at his downtown office.

Originally, Mr. Khan was to attend, but had a change of heart. Mr. Struthers attributed it to his client's fear of becoming a police target, but Mr. Khan was not so sure.

'In my view, the evidence is overpowering that the testimony of Officer Asselin and Officer James is untrue.'

Madam Justice Anne Molloy of the Ontario Superior Court

"Mostly," he told reporters by phone, "it's about my comfort level . . . being on the news, like an actor, it's not what I do."

Had he been convicted, he likely would have received a penitentiary sentence of between three and five years, Mr. Struthers said, adding, "It would have killed his life."

As it was, Mr. Khan's bail conditions of the past three years precluded him from marrying his fiancée, because he was required to live with his mother, and robbed him of precious firsts with the couple's baby daughter.

But he was unembittered and pronounced himself "grateful to the process" that vindicated him.

Ironically, according to Det. Asselin, Mr. Khan had first come to his attention because, when their cars met at a four-way stop, and Mr. Khan had the right-of-way, he had waited for the police to proceed first.

He would do it again, Mr. Khan said yesterday. "Out of courtesy, I would give the police a chance to go. It seems like something you should do."

He was just as Judge Molloy found him: A palpably fine young man, who just happens to be black and to own a nice car.

cblatch@globeandmail.ca

P. 9

Judge lashes police for racial profiling

BY CHRISTIE BLATCHFORD, TORONTO

Two Toronto police officers have been branded liars and a 29-year-old man pronounced a victim of racial profiling in what is believed to be the first "driving-while-black" acquittal in Canadian history.

In a remarkable decision released yesterday, Madam Justice Anne Molloy of the Ontario Superior Court concluded unequivocally that constables Glenn Asselin and Craig James "fabricated significant aspects of their evidence," and then asked herself the rhetorical question that legions of young black men have asked aloud for years.

Why had the police singled out Kevin Khan one October Monday about noon and decided to search his car?

"Because," Judge Molloy wrote, "he was a young black male driving an expensive Mercedes."

Constable Asselin has since been promoted to detective, while Mr. James has left the force.

See POLICE on page A11

Globe and Mail
Sep 2004

- "driving while black" acquittal
- A set-back for subjective Bayes...

.....

Let's: Omit the "profiling" / subjective Bayes!

Find: ..... "What model with data says"

You can <u>always</u> add a <u>prior</u> ... <u>later</u> ... if you must!
(for personal, individual reasons)

to get a posterior for personal needs
to make a decision, if needed

but now let's consider
Statistical Inference

C frequentist inference: With or Without Bayes

Have: Model with Data

Want: "What does it say about theta?"

Q? Do the analysis ... !

(~~or add~~ a Prior ("objective")

~~& conditional~~ analysis (B paradigm)?

(1) Recent Likelihood theory (RLT)

In wide generality..., Model with Data gives:

- For any scalar interest $\psi(\theta)$ gives

  $L(\psi)$ $p(\psi)$ 3rd HA

- For any vector interest $\psi(\theta)$ gives

  $L(\psi)$ $p(\psi)$ 3rd 2nd

- For any departure measure $t(y)$ gives

  p-value 3rd

That is: Just stuff it in the computer!

Q: Do we need the "prior" route?

www.utstat.toronto.edu/dfraser/documents/231.pdf

(2) Does prior ("objective"/default) give Answer?

(a) RLT gives conditional results .... given key characteristics of data$^+$

(b) Default B: conditions on data ... uses $L(\theta)$ but prior depends on full model & It 'should' depend on identified model!

ex $\begin{cases} \frac{1}{2} & N \quad \theta \quad 1 \\ \frac{1}{2} & N \quad \theta^3 \quad 1 \end{cases}$ Cox 1958

(Why average over measurements that haven't been taken?)
(Why be marginal?)

(c) With vector parameter and the 'right' prior, one still has marginalization paradoxes (Dawid, Stone Zidek)

Need to target prior! Desparate search!

(d) But then, RLT is available! .... WIDELY
.... "stuff" it in the computer ...

218 page.pdf
211.pdf

D Big Picture

By 1950's:

- Big claims of decision theory
  - but few answers
- Subjective pushed by Jimmy Savage
  - resistance to subjective
  - Profiling

Recent:

Bayesian chase of 'default' 'objective'

- Try to short-circuit the frequentist route

Now:

Recent Likelihood Theory

gives Likelihood $L(\psi)$
p-value $p(\psi)$

for almost any interest $\psi(\theta)$

- Just compute...

and if you must (use subjective / profiling)

- can do it afterwards (Just keep your inference clean)

And the police offer....

- for his personal use ... subjective prior?
- for his official, legal duty?

He'd better rethink...

and the scientist..... subjective

- Not appropriate for him!
- Should make available inference material

likelihood $L(\cdot)$

p-value $p(\cdot)$

and let others use subjective if they must