Why a prior?

df
Dept Statistics
U of Toronto
28 April '05
Munk Centre

Why a prior?

df
Dept Statistics
U of Toronto
28 April '05
Munk Centre

Or Why not?

Why a prior?

df
Dept Statistics
U of Toronto
28 April '05
Munk Centre

Or Why not?

Or Why the question? .... but first...

Nancy Reid
Judith Rousseau
Augustine Wong
Jianrong Wu
Grace Yun Yi
and...

responsibility...
df

Issues Classifying Examples Ex's say
Likelihood says For inference
Where To Enigmas

Many areas in statistics....

but ..... Two persuasions, cultures

$f \quad B \qquad f \cap B = \emptyset$

... Two persuasions, cultures

f $\boxed{f(y;\theta)} \longrightarrow y$ Stat model

Two persuasions, cultures

F $\quad \boxed{f(y;\theta)} \longrightarrow y$ $\qquad$ Stat model

B $\quad \boxed{\pi(\theta)} \longrightarrow \boxed{f(y;\theta)} \longrightarrow y$ $\qquad$ Prob model

f $\quad \boxed{f(y;\theta)} \longrightarrow y$ $\qquad$ Stat model

Hard

B $\quad \boxed{\pi(\theta)} \longrightarrow \boxed{f(y;\theta)} \longrightarrow y$ $\qquad$ Prob model

Easy

but often different results !

Does it matter ?

HEP

Connections, overlap ... with earlier seminars ...
of HEP

1) $y \sim \text{Poisson}(\theta) \qquad \theta \geq \theta_0$

1) F Abe et al Phys Rev Lett 7(2) 1994

Connections, overlap ... with earlier seminars ...
of HEP

1) $y \sim \text{Poisson}(\theta) \qquad \theta \geq \theta_0$

Search for Top Quark ....!
f and B couldn't agree !
Delay in publication ... ~2 years
Cost: Delay in construction ... next generation (Collider Detector
Many mega \$'s

1) F Abe et al Phys Rev Lett 7(2) 1994

Connections, overlap ... with earlier seminars ...

df HEP

1) $y \sim \text{Poisson}(\theta) \qquad \theta \geq \theta_0$

Search for Top Quark ....!

f and B couldn't agree !

Delay in publication ... ~2years

Cost: Delay in construction ... next generation (Collider Detector

Many mega \$'s

Does it matter ? Scientifically; \$M's ; Locally ?

Connections, overlap ... with earlier seminars ...

df HEP

1) $y \sim \text{Poisson}(\theta) \qquad \theta \geq \theta_0$

Search for Top Quark ....!

f and B couldn't agree !

Delay in publication ... ~2 years

Cost: Delay in construction ... next generation (Collider Detector

Many mega \$'s

Does it matter ? Scientifically; \$M's ; Locally ?

Our discipline cost Big \$s ! ?

Connections, overlap ... with earlier seminars ...

df HEP

1) $y \sim \text{Poisson}(\theta) \qquad \theta \geq \theta_0$

Search for Top Quark ....!

f and B couldn't agree !

Delay in publication ... ~2 years

Cost: Delay in construction ... next generation (Collider Detector

Many mega \$'s

Does it matter ? Scientifically; \$m's ; Locally ?

Our discipline cost Big \$s !

That HEP is <u>now our Faculty Dean</u> !

Does it matter? Scientifically; $m's; Locally?

Our discipline cost Big $s!

That HEP is <u>now</u> our Faculty Dean!

Some further:

2) Mandelkern (2002) Stat Sc

3) Fraser, Reid, Wong Phys Rev D (2004)

<u>Classify</u>: What kind of priors $\pi(\theta)$?

1) There is an identified <u>random source</u>

... sampling, genetic theory, related investigation, experience, theory, model

Classify: What kind of priors $\pi(\theta)$?

Called by f B

1) There is an identified random source

objective —

Classify: What kind of priors $\pi(\theta)$?

Called by f B

1) There is an identified random source

objective —

2) Based on invariance re model

... Original Thomas B, Laplace, flat (re model), insufficent r.

Classify: What kind of priors $\pi(\theta)$?

Called by f B

1) There is an identified random source

objective —

2) Based on invariance re model

invariant Objective

Objective ... ?
Based on "model & data only" ... sounds like f territory!
... currently a dynamic area ...

Classify: What kind of priors $\pi(\theta)$?

Called by	f	B
1) There is an identified random source	objective	—
2) Based on invariance re model	invariant	Objective
3) Based on an individual's views ... beliefs, elicitations, ...		

Classify: What kind of priors $\pi(\theta)$?

Called by	f	B
1) There is an identified random source	objective	—
2) Based on invariance re model	invariant	Objective
3) Based on an individual's views	subjective	Subjective

Classify: What kind of priors $\pi(\theta)$?

Called

1) There is an identified random source

objective

2) Based on invariance re model

invariant

3) Based on an individual's views

subjective

Q's Should usage be affected ?

Forbidden Knowledge Science 11 Feb 2005

Examples: Invariance / neutral re model

<u>Ex 1</u> $y \sim f(y-\theta)$ ... the nice case ... should work here

$\longleftrightarrow$ translation / location

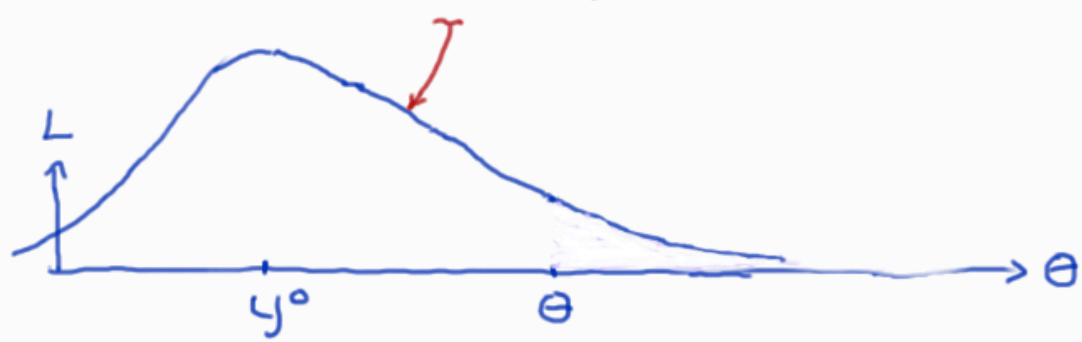
f: Report <u>percentile position</u> of data

$$p(\theta) = F(y^0-\theta) = \int_{-\infty}^{y^0} f(y-\theta)\,dy$$

"as it is!"

and <u>likelihood</u>

$$L(\theta) = f(y^0-\theta)$$



Examples: Invariance / neutral re model

<u>Ex 1</u> $y \sim f(y-\theta)$ ... the nice case ... should work here

$\longleftrightarrow$ translation / location

f: Report <u>percentile position</u> of data

$$p(\theta) = F(y^0-\theta) = \int_{-\infty}^{y^0} f(y-\theta)\,dy$$

"as it is!"

and <u>likelihood</u>

$$L(\theta) = f(y^0-\theta)$$

B: Invariance

$$s(\theta) = \int_{\theta}^{\infty} c\,L(t)\,dt$$

Examples: Invariance / neutral re model

<u>Ex 1</u> $y \sim f(y-\theta)$ ... the nice case ... should work here

$\longleftrightarrow$ translation / location

f: Report <u>percentile position</u> of data

$$p(\theta) = F(y^0-\theta) = \int_{-\infty}^{y^0} f(y-\theta)\,dy$$

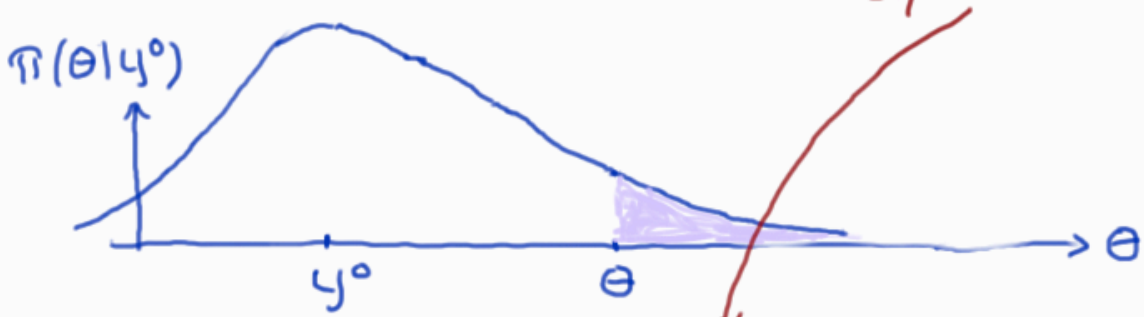
"as it is!"

$$p(\theta) = s(\theta) = \int_{-\infty}^{y^0-\theta} f(t)\,dt$$

Equal

Strong matching

All harmony!



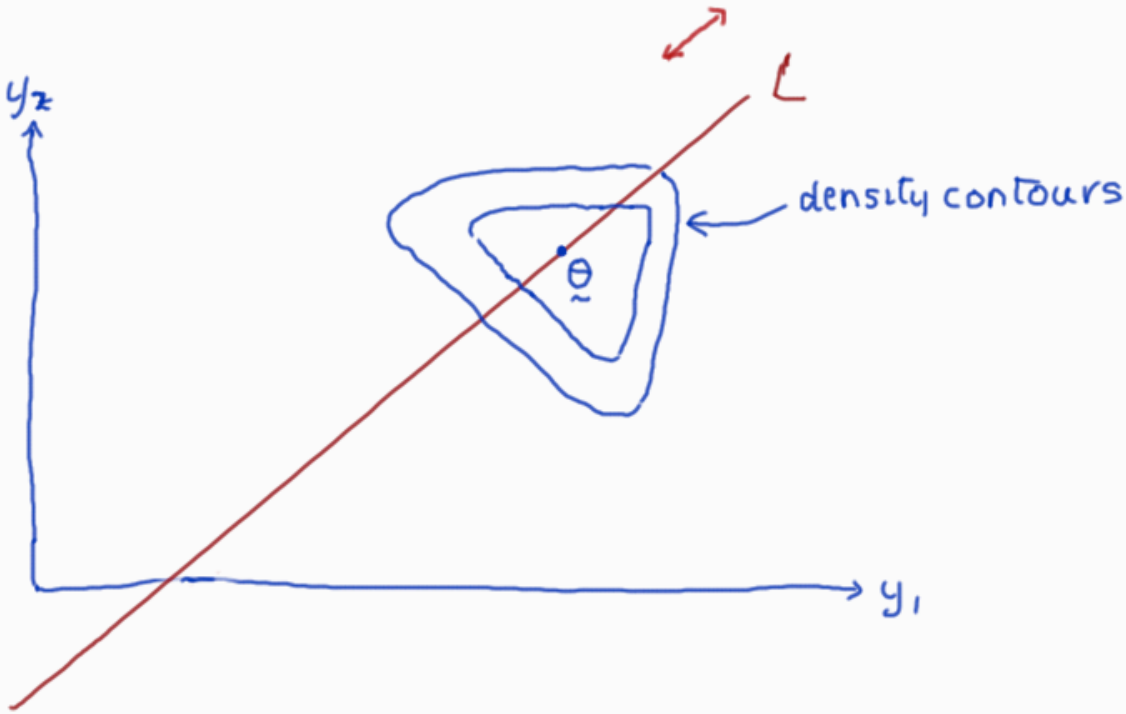
B: Invariance

$$s(\theta) = \int_{\theta}^{\infty} c\,L(t)\,dt$$

Examples: Invariance / neutral re model

<u>Ex 2</u> $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$ $\qquad \underset{\sim}{z} \sim f(\underset{\sim}{z})$

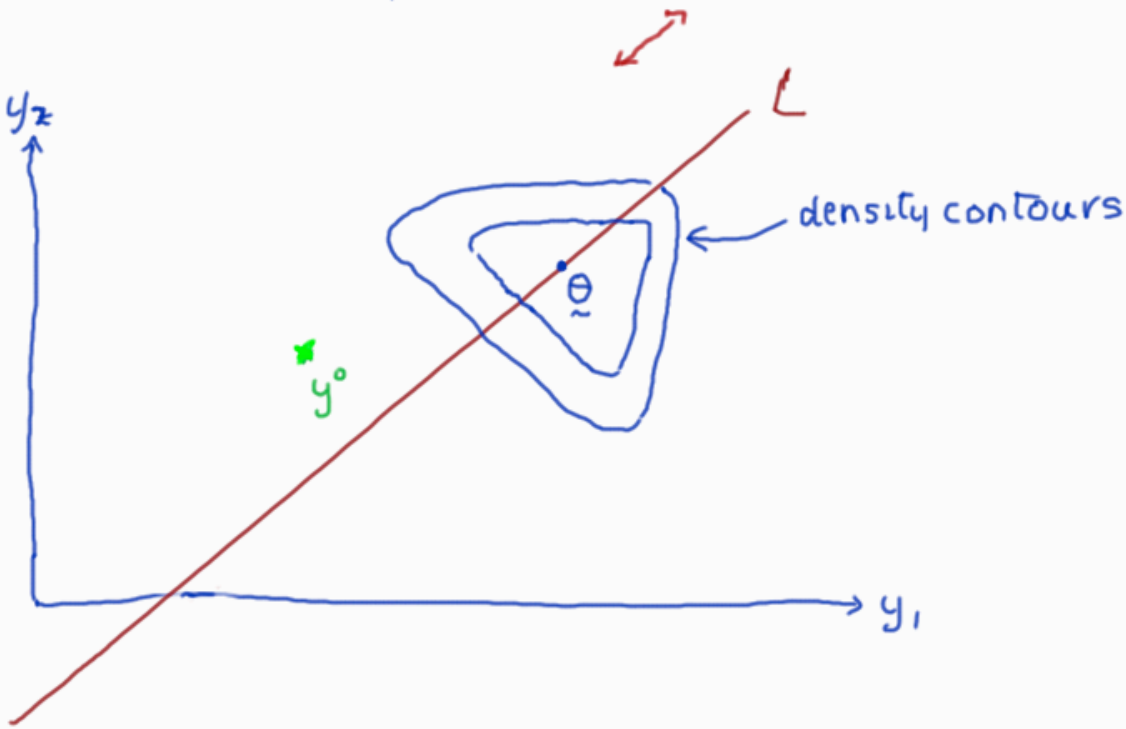
Location on plane; $\underset{\sim}{\theta}$ on line L



Examples: Invariance / neutral re model

<u>Ex 2</u> $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$ $\qquad \underset{\sim}{z} \sim f(\underset{\sim}{z})$

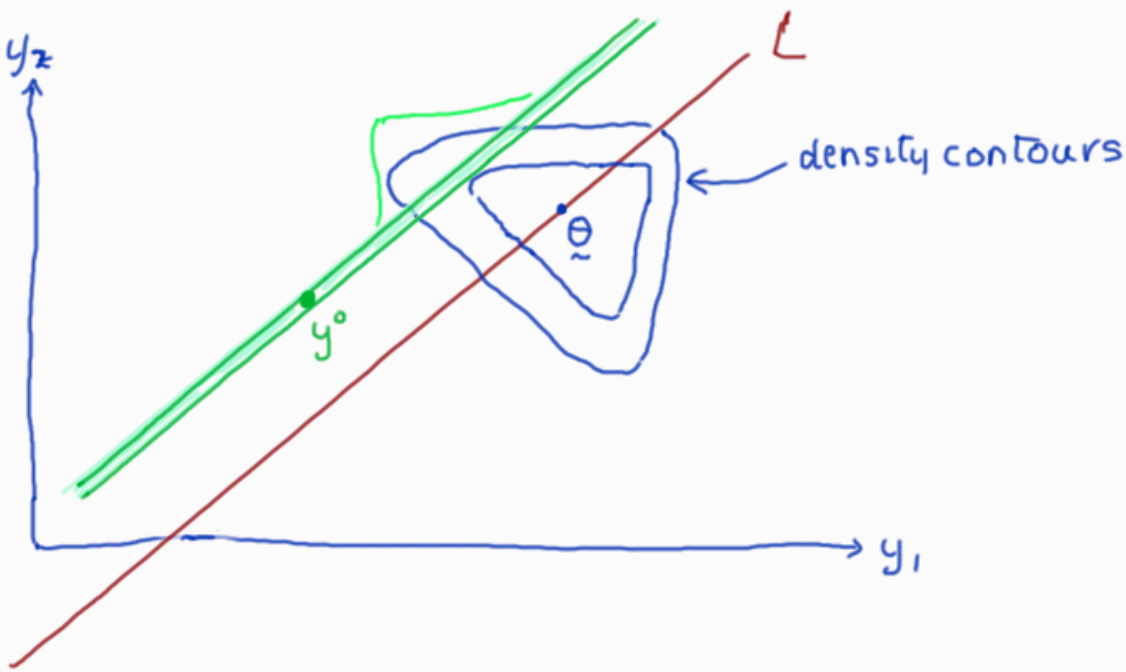
Location on plane; $\underset{\sim}{\theta}$ on line L ; data y°



Examples: Invariance / neutral re model

<u>Ex 2</u> $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$ $\qquad \underset{\sim}{z} \sim f(\underset{\sim}{z})$

Location on plane; $\underset{\sim}{\theta}$ on line L ; data y°



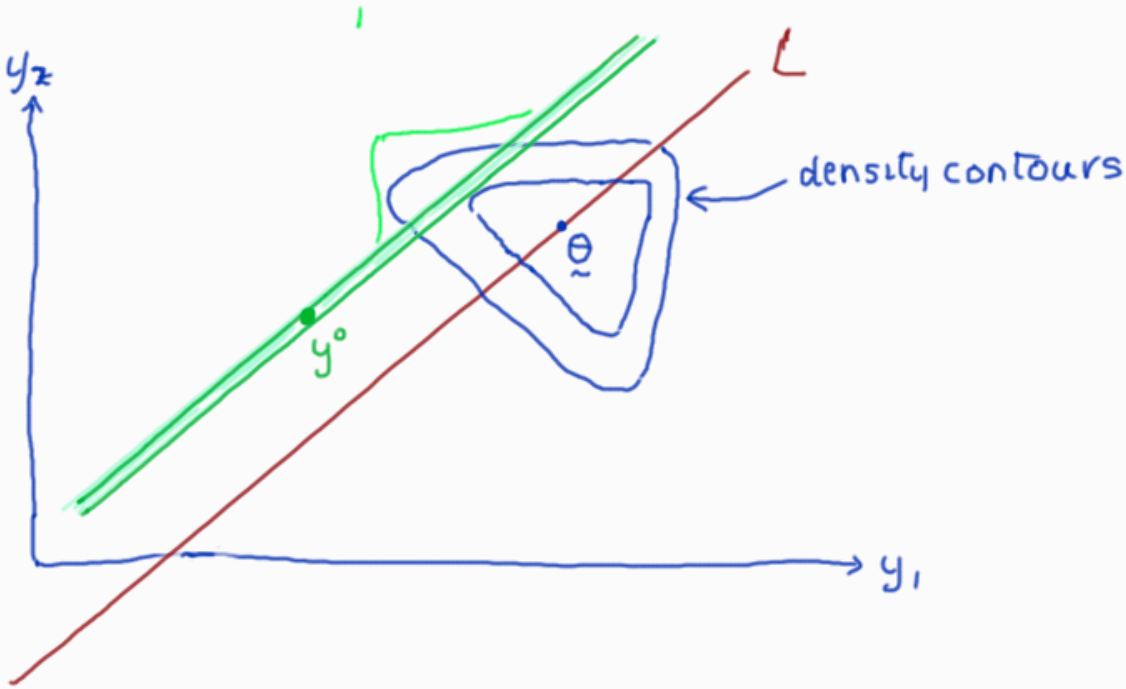
Can <u>identify</u> what part of model is relevant

<u>Condition</u> parallel to line L ; density indicated

Examples: Invariance / neutral re model

<u>Ex 2</u> $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$ $\qquad \underset{\sim}{z} \sim f(\underset{\sim}{z})$

Location on plane; $\underset{\sim}{\theta}$ on line L ; data y°



Can <u>identify</u> what part of model is relevant

<u>Condition</u> parallel to line L ; density indicated

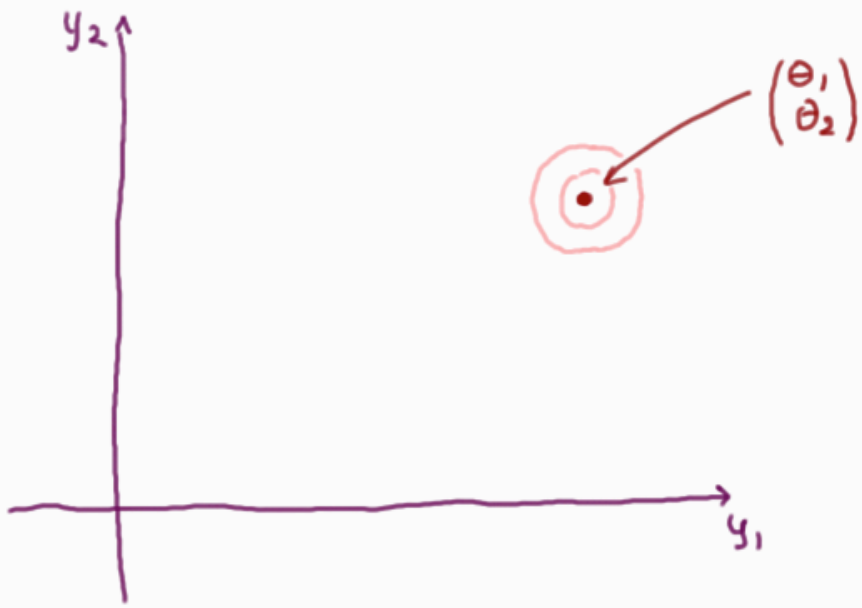
f : Has hang-ups re conditioning ! $\qquad$ He "doesn't get it!"

B : Conditions on data ! $\qquad$ gets it right !

Seems conditioning is "0 - 1" $\qquad$ No - Yes Compromise ?

Simple example: Normal on R^2

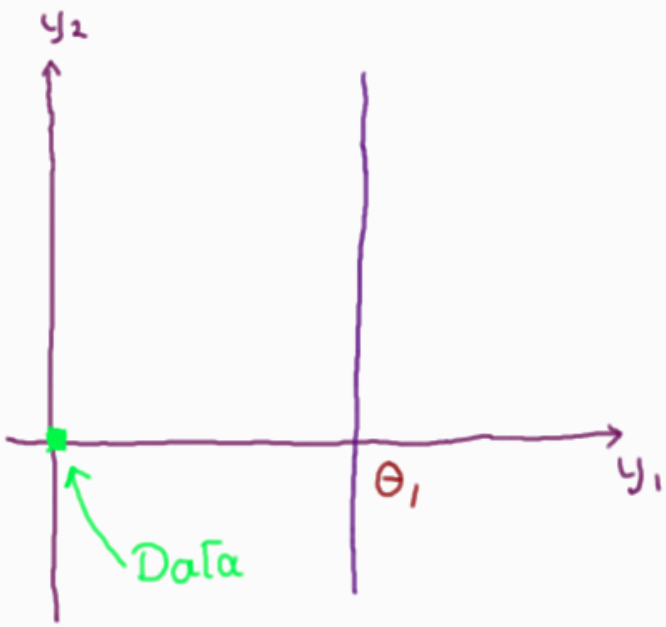
$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \qquad z_i \sim N(0,1)$$



Simple example: Normal on R^2

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \qquad z_i \sim N(0,1) \qquad \begin{pmatrix} y_1^0 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \ldots \text{ say}$$

Interest: $\varphi(\theta) = \theta_1$ ..... linear



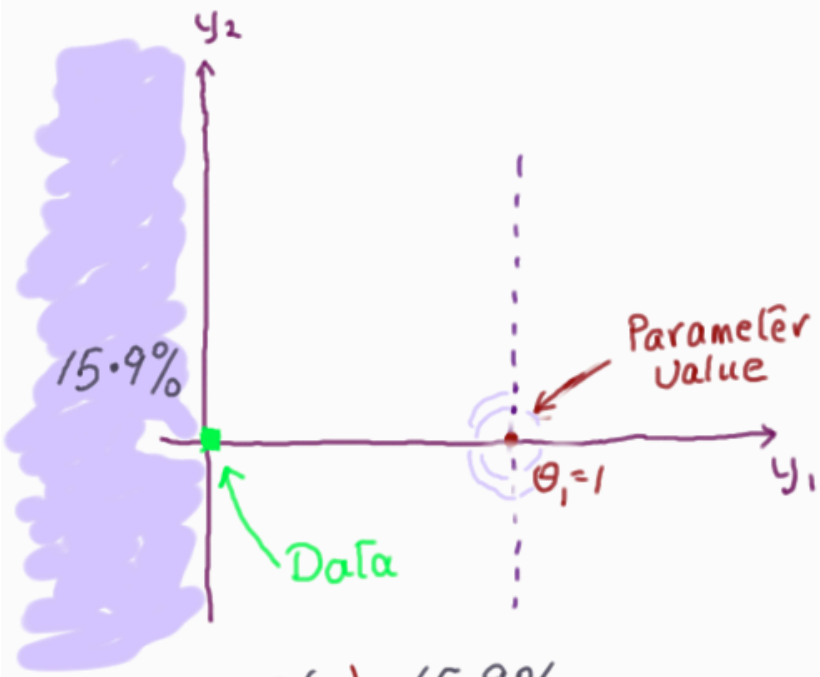
?

Simple example: Normal on R^2

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \qquad z_i \sim N(0,1) \qquad \begin{pmatrix} y_1^0 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \ldots \text{ say}$$

Interest: $\varphi(\theta) = \theta_1$ ..... linear Assess: $\theta_1 = 1$

frequentist: $y_1 = 1 + z$



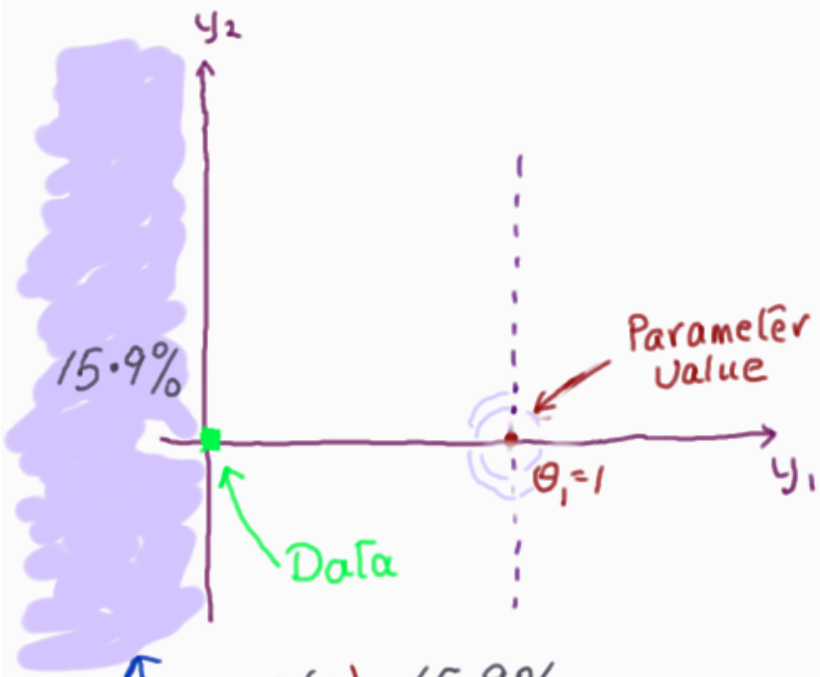
$p(1) = 15.9\%$

Simple example: Normal on R^2

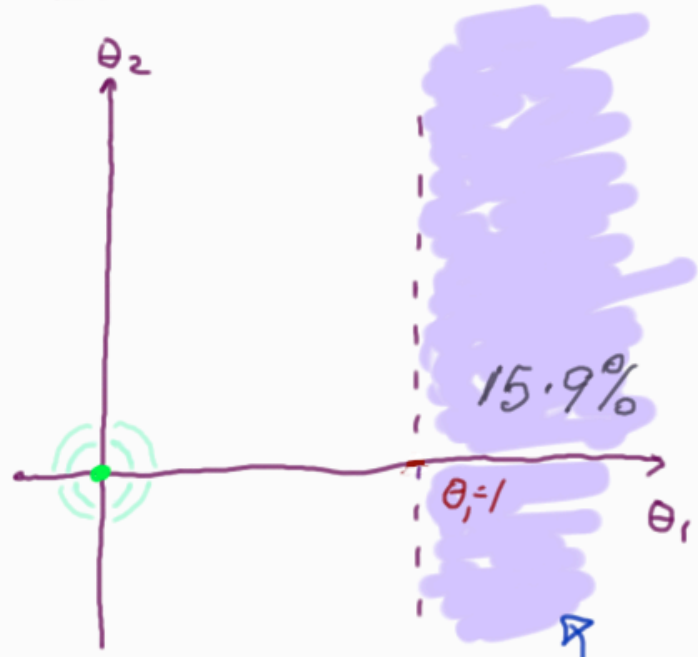
$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \qquad z_i \sim N(0,1) \qquad \begin{pmatrix} y_1^0 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ ... say}$$

Interest: $\varphi(\theta) = \theta_1$ ..... linear Assess: $\theta_1 = 1$

frequentist: $y_1 = 1 + z$ Bayes: $\theta_1 = 0 + z$



$p(1) = 15.9\%$
$= P(y_1 \leq 0)$



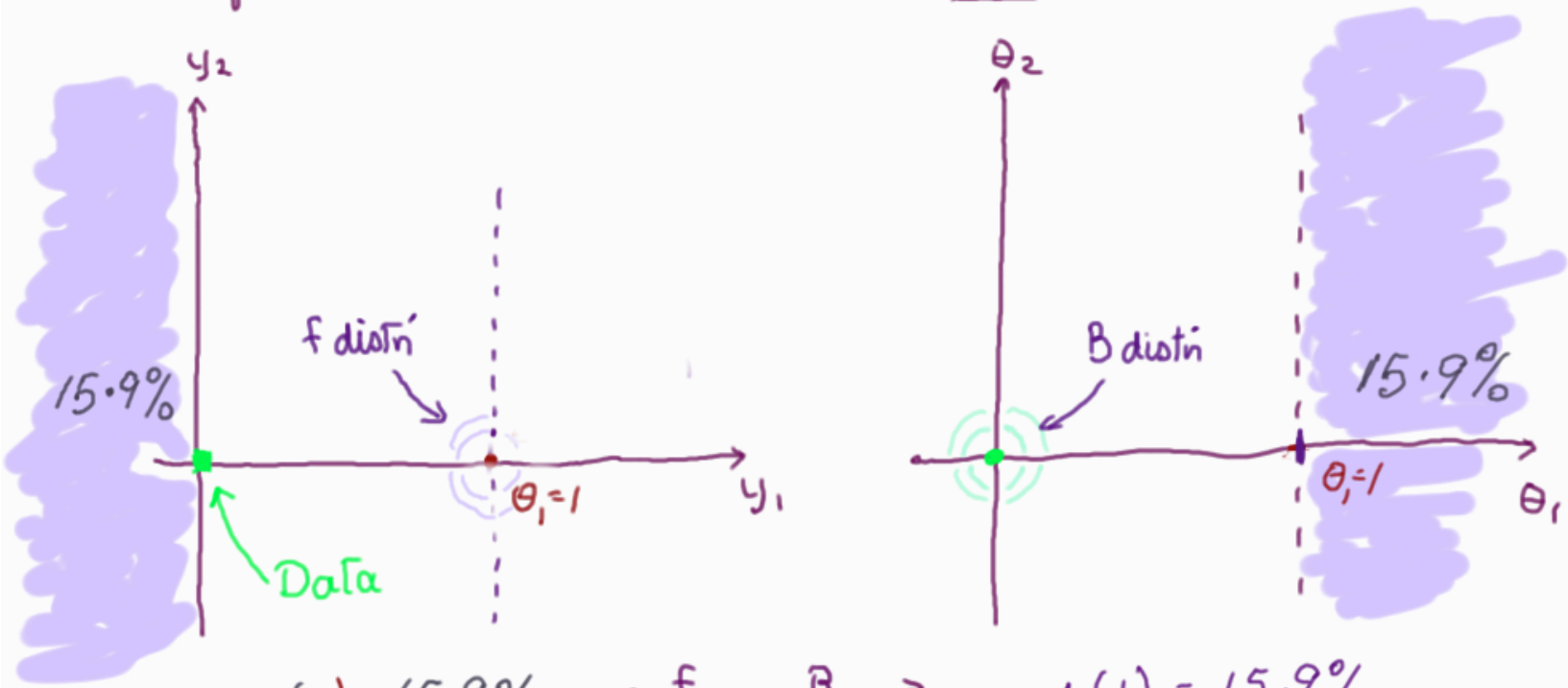
$s(1) = 15.9\%$
$= P(\theta_1 \geq 1)$

Simple example: Normal on R^2

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \qquad z_i \sim N(0,1) \qquad \begin{pmatrix} y_1^0 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \ \ldots \text{ say}$$

Interest: $\varphi(\theta) = \theta_1$ ..... linear Assess: $\theta_1 = 1$

frequentist: $y_1 = 1 + z$ Bayes: $\theta_1 = 0 + z$



$p(1) = 15.9\%$ ← f B → $s(1) = 15.9\%$

agreement ... for linear parameters !

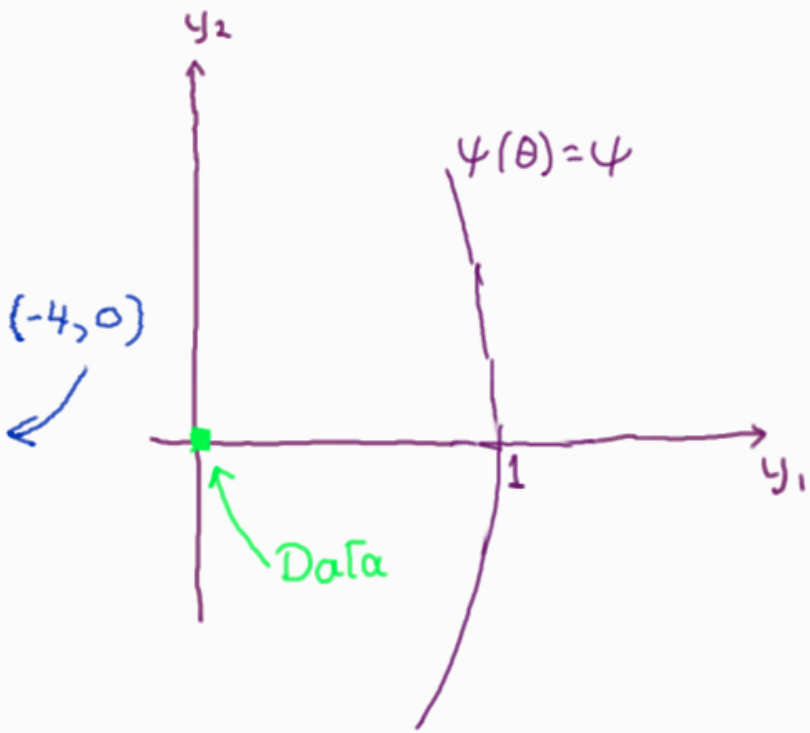
$p(\varphi) = s(\varphi)$ ... All sweetness !

Simple example: Normal on R^2

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \qquad z_i \sim N(0,1) \qquad \begin{pmatrix} y_1^0 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \cdots \text{ say}$$

Curved interest: $\psi(\theta) = \{(4+\theta_1)^2 + \theta_2^2\}^{1/2} - 4$; assess $\psi = 1$

$= \text{distance from } (-4, 0)$

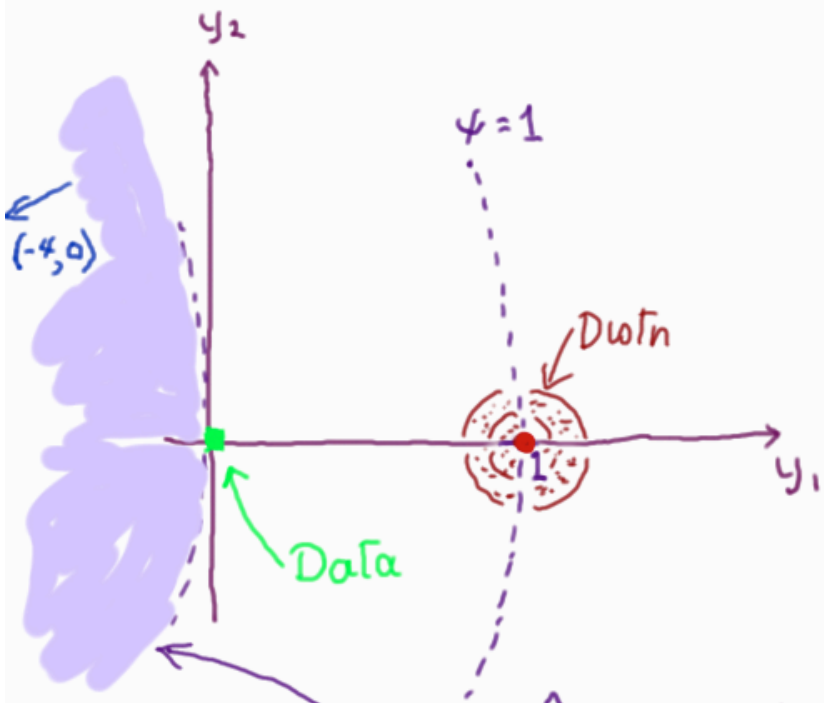


Simple example: Normal on R^2

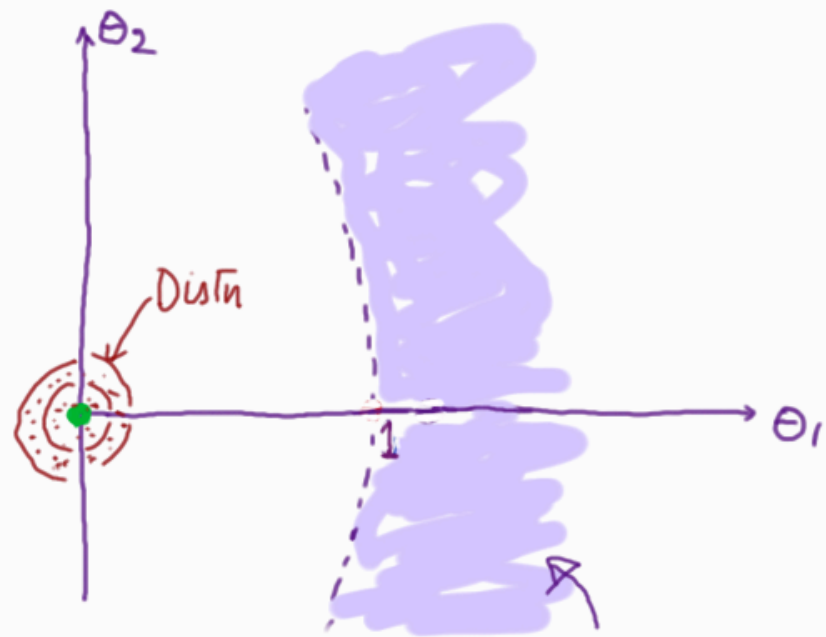
$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \qquad z_i \sim N(0,1) \qquad \begin{pmatrix} y_1^0 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ ... say}$$

Curved interest: $\psi = \{(4+\theta_1)^2 + \theta_2^2\}^{1/2} - 4$; assess $\psi = 1$

f: Distn at $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ B: Post at $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$



$p = P(\hat{\psi} < 0 ; \psi = 1)$
$= 13.5\%$



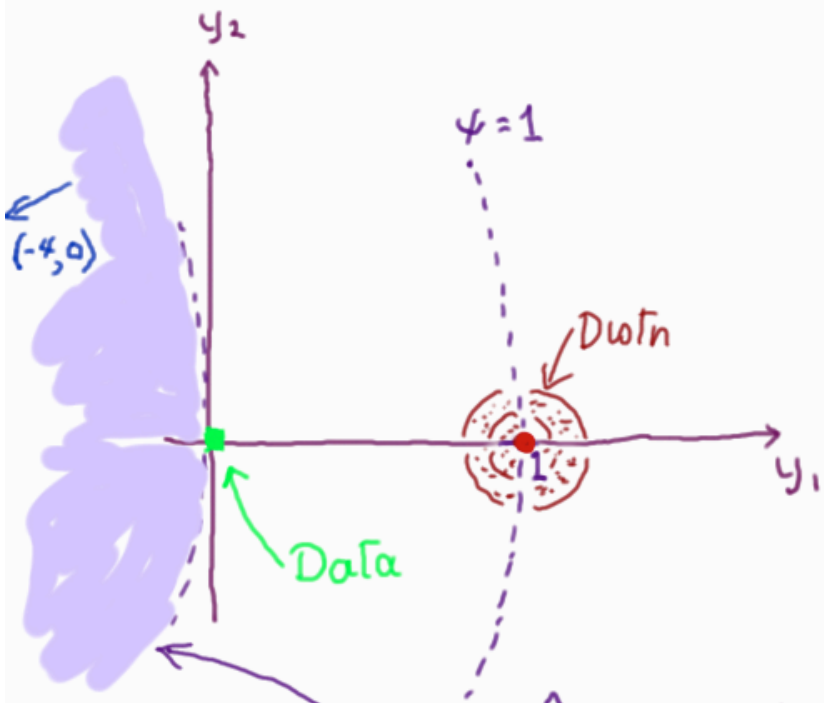
$s = P(\psi > 1 ; y_1 = 0)$
$= 18.7\%$

Simple example: Normal on R^2

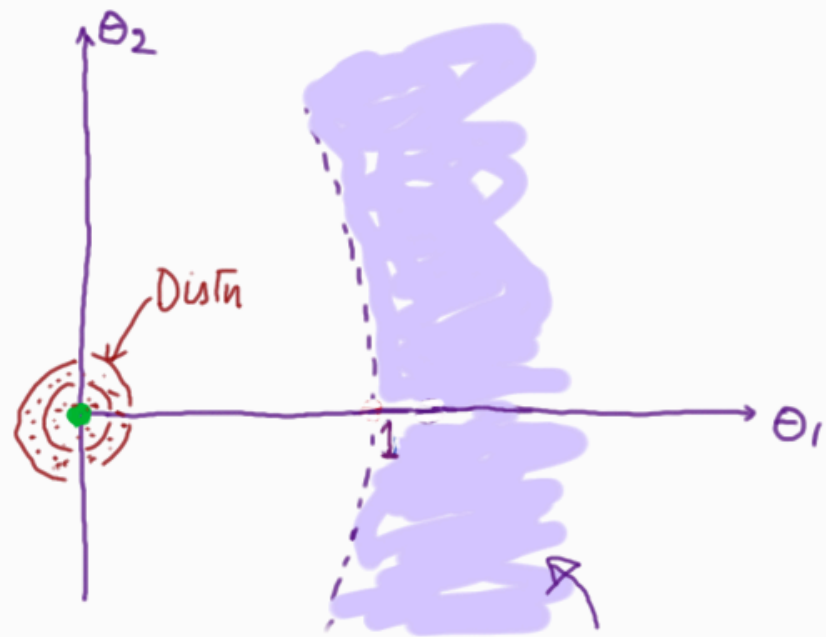
$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \qquad z_i \sim N(0,1) \qquad \begin{pmatrix} y_1^0 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ ... say}$$

Curved interest: $\psi = \{(4+\theta_1)^2 + \theta_2^2\}^{1/2} - 4$; assess $\psi = 1$

f: Distn at $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ B: Post at $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$



$p = P(\hat{\psi} < 0; \psi = 1)$
$= 13.5\%$



$s = P(\psi > 1; y_1 = 0)$
$= 18.7\%$

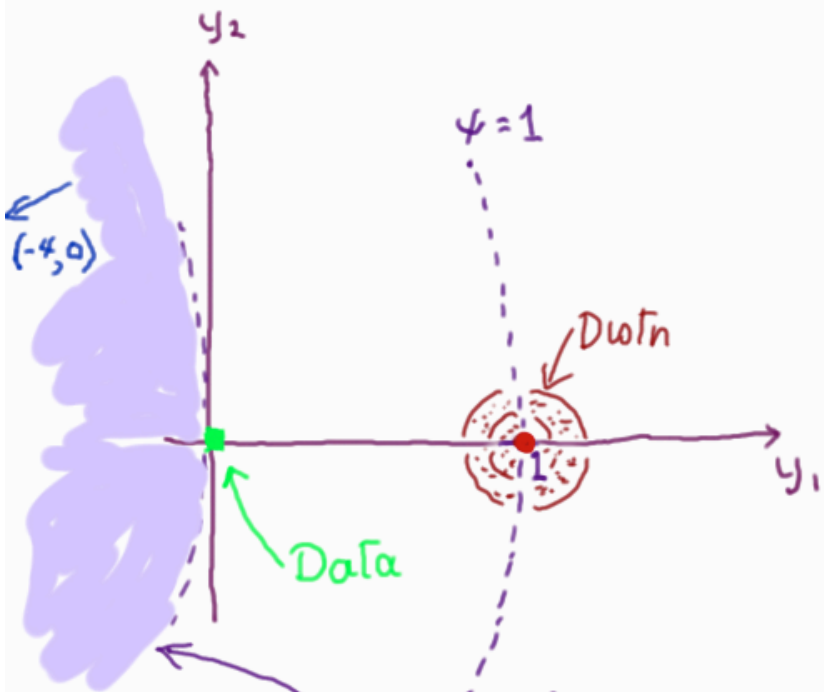
New example / David Stone Zidek 1973

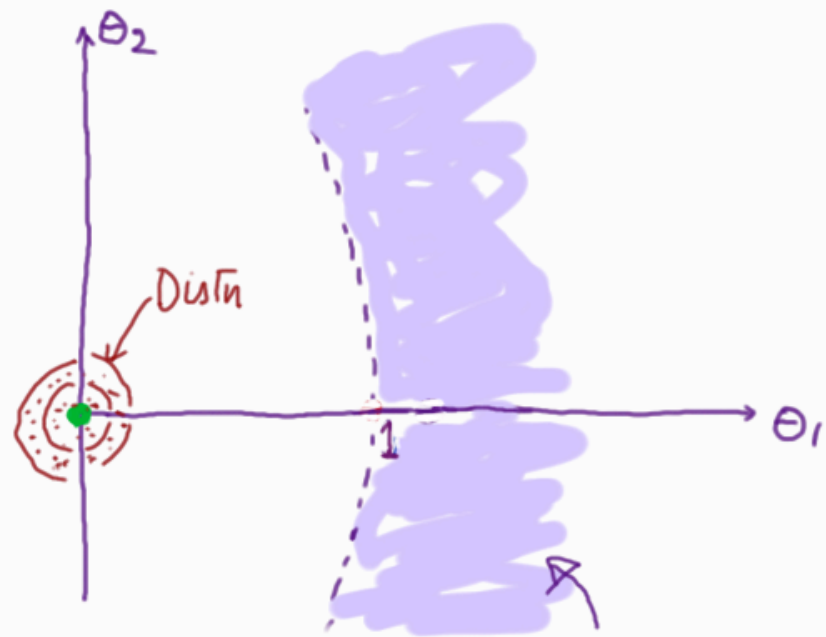
Simple example: Normal on R^2

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \qquad z_i \sim N(0,1) \qquad \begin{pmatrix} y_1^0 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \ldots \text{ say}$$

Curved interest: $\psi = \{(4+\theta_1)^2 + \theta_2^2\}^{1/2} - 4$; assess $\psi = 1$

f: Distn at $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ B: Post at $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$





freq $p = P(\hat{\psi} < 0 ; \psi = 1)$
$= 13.5\%$

B $s = P(\psi > 1 ; y_1 = 0)$
$= 18.7\%$

Different 40%

Linear

curved

Stong matching $\Longrightarrow$ Mis-matching

Examples say:

1) $f(y-\theta)$ $\quad p(\theta) = s(\theta)$ $\quad$ B and f agree :)

Examples say:

1) $f(y-\theta)$ $\quad p(\theta) = s(\theta)$ $\quad$ B and f agree :)

2) $f(\underset{\sim}{y}-\underset{\sim}{\theta})$ $\quad \underset{\sim}{\theta}$ on line L $\quad$ Use identified model :)

Examples say:

1) $f(y-\theta)$ $\quad p(\theta) = s(\theta)$ $\quad$ B and f agree :)

2) $f(\underset{\sim}{y}-\underset{\sim}{\theta})$ $\quad \underset{\sim}{\theta}$ on line L $\quad$ Use identified model :)

3) $f(\underset{\sim}{y}-\underset{\sim}{\theta})$ $\quad$ Interest in $\psi(\theta)$

- If $\psi(\theta)$ linear $\quad p(\theta) = s(\theta)$ $\quad$ B and f are :)

Examples say:

1) $f(y-\theta)$ $\quad p(\theta) = s(\theta)$ $\quad$ B and f agree :)

2) $f(\underset{\sim}{y}-\underset{\sim}{\theta})$ $\quad \underset{\sim}{\theta}$ on line L $\quad$ Use identified model :)

3) $f(\underset{\sim}{y}-\underset{\sim}{\theta})$ $\quad$ Interest in $\psi(\theta)$

- If $\psi(\theta)$ linear $\quad p(\theta) = s(\theta)$ $\quad$ B and f are :)
- If $\psi(\theta)$ curved, then DSZ
  but you can 'measure' $\psi(\theta)$ $\quad$ B & F get together :)

Examples say:

1) $f(y-\theta)$ $\quad p(\theta) = s(\theta)$ $\quad$ B and f agree :)

2) $f(\underset{\sim}{y}-\underset{\sim}{\theta})$ $\quad \underset{\sim}{\theta}$ on line L $\quad$ Use identified model :)

3) $f(\underset{\sim}{y}-\underset{\sim}{\theta})$ $\quad$ Interest in $\psi(\theta)$

- If $\psi(\theta)$ linear $\quad p(\theta) = s(\theta)$ $\quad$ B and f are :)
- If $\psi(\theta)$ curved, then DSZ
  but you can 'measure' $\psi(\theta)$ $\quad$ B & F :)

But: to high accuracy

All/most models are $f(\underset{\sim}{y}-\underset{\sim}{\theta})$

... just requires some work to see it!

... so B and f are :) .... maybe

What does Likelihood theory say?

Model
Data
Continuity

$\Big\} \longrightarrow$?

Moderate regularity, wide generality...

What does Likelihood theory say?

Model
Data
Continuity $\Big\} \longrightarrow$?

Moderate regularity, wide generality, continuity, ..

Model is location re full θ $\qquad f(y-\theta)$

" " " " component $\psi(\theta)$

Just need to see recalibration/reparameterization! $\qquad f(y-\psi)$

What does Likelihood theory say?

Model
Data
Continuity
$\Big\} \longrightarrow$?

Moderate regularity, wide generality, continuity, ..

Model is location re full θ $\qquad f(y-\theta)$
" " " " component $\psi(\theta)$
Just need to see recalibration/reparameterization! $\qquad f(y-\psi)$

$f(y;\theta)$
y^0
"Cont"
$\Big\} \longrightarrow$

$L(\theta) \quad p(\theta) \quad s(\theta) \qquad p(\theta) = s(\theta)$!

$L(\psi) \quad p(\psi) \quad s(\psi) \qquad p(\psi) = s(\psi)$!

Third order, second order

What does Likelihood theory say?

Model
Data
Continuity
$\Big\} \longrightarrow$?

Moderate regularity, wide generality, continuity, ..

Model is location re full θ $\qquad f(y-\theta)$
" " " " component $\psi(\theta)$
Just need to see recalibration/reparameterization! $\qquad f(y-\psi)$

$f(y;\theta)$ Computer $\quad L(\theta) \quad p(\theta) \quad s(\theta) \qquad p(\theta) = s(\theta)$!
y^0 $\longrightarrow$ $L(\psi) \quad p(\psi) \quad s(\psi) \qquad p(\psi) = s(\psi)$!
"Cont"
Computer $\quad$ Third order, second order

Not as simple as: take $L(\theta)$ and add a prior
But it is seeing what "model + data says about theta"

Inference Context

Interested in $\psi(\theta)$	μ	β
What's left over $\lambda(\theta)$	σ^2	α σ^2

Inference Context

Interested in $\psi(\theta)$	μ	β
What's left over $\lambda(\theta)$	σ^2	α σ^2

Two steps:

(i) a measure of departure say $t(y)$ of data from value $\psi(\theta)=\psi$

(ii) Modify $t(y)$ and get p-value $p(\psi)$ free of λ

For example (i) $t(y)=\bar{y}-\mu$ for assessing value μ

(ii) Get $T(y)=\dfrac{\bar{y}-\mu}{s_y/\sqrt{n}}$ free of σ^2 ; $p(\mu)=H_{n-1}(T^{\circ})$

Stud $(n-1)$ df

Inference Context

Interested in $\psi(\theta)$	μ	β
What's left over $\lambda(\theta)$	σ^2	α σ^2

Two steps:

(i) a <u>measure of departure</u> say $t(y)$ of data from value $\psi(\theta)=\psi$

(ii) <u>Modify</u> $t(y)$ and get <u>p-value</u> $p(\psi)$ free of λ

<u>For example</u> (i) $t(y)=\bar{y}-\mu$ for assessing value μ

Stud(n-1) df

(ii) Get $T(y)=\dfrac{\bar{y}-\mu}{s_y/\sqrt{n}}$ free on σ^2 ; $p(\mu)=H_{n-1}(T^{\circ})$

<u>Recent</u>: <u>Step(ii)</u> is "automatic"! From $t(y)$ to $p(\mu)$! f, B :)

Rousseau, F (2004)

① Bayesian: p^*_{cpred}

② Bootstrap: a few times

③ frequentist: from full ancillary dist'n

Get <u>same</u> result

Inference Context

Interested in $\psi(\theta)$	μ	β
What's left over $\lambda(\theta)$	σ^2	α σ^2

Two steps:

(i) a measure of departure say $t(y)$ of data from value $\psi(\theta)=\psi$

(ii) Modify $t(y)$ and get p-value $p(\psi)$ free of λ

For example (i) $t(y)=\bar{y}-\mu$ for assessing value μ

(ii) Get $T(y)=\frac{\bar{y}-\mu}{s_y/\sqrt{n}}$ free on σ^2 ; $p(\mu)=H_{n-1}(T^\circ)$ Stud $(n-1)$ df

Recent: Step(ii) is "automatic"! From $t(y)$ to $p(\mu)$! f, B :)

Rousseau, F (2004)

① Bayesian: p^*_{cpred}

② Bootstrap: a few times

③ frequentist: from full ancillary dist'n

Get same result

Just leaves step (i) at issue!

(i) a measure of departure say $t(y)$ of data from value $\psi(\theta) = \psi$

Likelihood theory again:

Case: $\dim \psi = 1$ $\quad \dim \lambda = p-1$

- Identified model in R^p ... like $(\bar{y}, s_y)$ in R^2
- 3rd order $p(\psi)$ immediate " $H_{n-1}(T^0)$

But No suffcy!

Thus: Model, data, interest $\psi(\theta)$

Unique p-value $p(\psi)$

F. Reid Wu 1999 Biometrika

F Biometrika 2003

(i) Find measure of departure ... re $\psi(\theta) = \psi$

(ii) Modify to get $p(\psi)$ free of λ

f, B, BS agree

(i) <u>Find</u> measure of departure ... re $\psi(\theta) = \psi$

f Need to calibrate parameter

B Seek Targetted prior ... target on $\psi(\theta)$

(ii) <u>Modify</u> to get $p(\psi)$ free of λ

f, B, BS agree

Where to?

1 f: Freely use weighted $L(\theta)$

Welch & Peers 1963 JRSSB
F 1972 Annals Math Stat

Where to?

1 f: Freely use weighted $L(\theta)$

Integrate it, profile it ... you can't go far wrong

It's your likelihood!

Welch & Peers 1963 JRSSB

F 1972 Annals Math Stat

Where to?

f: Freely use weighted $L(\theta)$

B: Use component models that directly measure component parameters

Where to?

f: Freely use weighted $L(\theta)$

B: Use component models that directly measure component parameters

Avoids targetted priors (parameter info available)

Avoids "information processing" (Arnold Z)

- ignores identified model

Big B business

Omnibus technique

Where to?

f : Freely use weighted $L(\theta)$

B : Use component models that directly measure component parameters

f : Feel free to condition; Don't leave it all to the B

Where to?

f: Freely use weighted $L(\theta)$

B: Use component models that directly measure component parameters

f: Feel free to condition; Don't leave it all to the B

- Separates the actual measurement of parameter
  - Can't do by sufficiency

F 2004 Stat Science

Where to?

f : Freely use weighted $L(\theta)$

B : Use component models that directly measure component parameters

f : Feel free to condition; Don't leave it all to the B

B : Condition on data properties... identified model; get right prior

F 2004 Stat Science

Where to?

f : Freely use weighted $L(\theta)$

B : Use component models that directly measure component parameters

f : Feel free to condition; Don't leave it all to the B

B : Condition on data properties... identified model

Less than full conditioning on data; gets prior right!

F 2004 Stat Science

Where to?

f : Freely use weighted $L(\theta)$

B : Use component models that directly measure component parameters

f : Feel free to condition; Don't leave it all to the B

B : Condition on data properties ... identified model

f : Use 'invariant-type' priors $|j(\theta;y)|^{1/2}$, $\ell_\theta / \varphi(\theta)$, $\ell_\psi / \chi(\psi)$

Recalibrates likelihood : location form ; area under L gives p value

F Reid Wong Yun Yi (2003) Valencia 7

F Yun Yi (2003) JISS

Where to?

f : Freely use weighted $L(\theta)$

B : Use component models that directly measure component parameters

f : Feel free to condition; Don't leave it all to the B

B : Condition on data properties... identified model

f : Use 'invariant-type' priors $|j(\theta;y)|^{1/2}$, $\ell_\theta / \varphi(\theta)$, $\ell_\psi / \chi(\psi)$

B: Subjective priors Put alongside final p-value

Where to?

f : Freely use weighted $L(\theta)$

B : Use component models that directly measure component parameters

f : Feel free to condition; Don't leave it all to the B

B : Condition on data properties... identified model

f : Use 'invariant-type' priors $|j(\theta;\psi)|^{1/2}$, $\ell_\theta / \varphi(\theta)$, $\ell_\psi / \chi(\psi)$

B: Subjective priors Put alongside final p-value

Model + Data $\longrightarrow$ $L(\psi)$ $p(\psi)$

Subjective $\pi(\psi)$

Let individual with prior combine as he/she wishes

Let others have their prior...

Where to?

f : Freely use weighted $L(\theta)$

B : Use component models that directly measure component parameters

f : Feel free to condition; Don't leave it all to the B

B : Condition on data properties... identified model

f : Use 'invariant-type' priors $|j(\theta;\psi)|^{1/2}$, $\ell_\theta / \varphi(\theta)$, $\ell_\psi / \chi(\psi)$

B: Subjective priors Put alongside final p-value

Model + Data $\longrightarrow$ $L(\psi)$ $p(\psi)$

Subjective $\pi(\psi)$

Let individual with prior combine as he/she wishes

Let others have their prior...

Some examples:

1 An old enigmatic example Behrens-Fisher (1929) (1934) ; Ghosh Kim CJS (2001)

$y_{11}, \dots, y_{1n}$ N μ_1 σ_1^2

$y_{21}, \dots, y_{2n}$ N μ_2 σ_2^2

Interest $\psi = \mu_1 - \mu_2$

GK $\pi = \sigma_1^{-3} \sigma_2^{-3} \left(\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m} \right)$

Some examples:

1 An old enigmatic example Behrens-Fisher (1929) (1934) ; Ghosh Kim CJS (2001)

$y_{11}, \dots, y_{1n}$ N μ_1 σ_1^2 Interest $\psi = \mu_1 - \mu_2$

$y_{21}, \dots, y_{2n}$ N μ_2 σ_2^2

GK $\pi = \sigma_1^{-3} \sigma_2^{-3} \left(\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m} \right)$

Some simulation numbers: Case

n	m	σ_1^2	σ_2^2	ψ
2	2	2	1	2

Nominal 5% 95% points by various methods

Nominal	5%	95%
Jeffreys	.7%	99.1%
Kim Ghosh	1.7%	97.9%
Lik ratio	13.2%	86.9%
3rd Lik	4.23%	95.8%
Sim 95% limits	(4.86, 5.14)	(94.9, 95.14)

N=100,000

Ex 2 Not-so-old example

Power transformed regression Box Cox (1964, 1982)

$$y_i = (\alpha + \beta x_i + \sigma z_i)^{1/\lambda} \qquad z_i \sim N(0,1) \text{ say}$$

Interest: How y depends on x

Maybe β ?

Chen Lockhart Stephens 2002 CJS

Maybe β/σ ?

Ex 2 Not-so-old example

Power transformed regression Box Cox (1964, 1982)

$$y_i = (\alpha + \beta x_i + \sigma z_i)^{1/\lambda} \qquad z_i \sim N(0,1) \text{ say}$$

Interest: How y depends on x

Maybe β ?

Chen Lockhart Stephens 2002 CJS

Maybe β/σ ?

Current: $\beta\lambda^{-1}(\alpha+\beta x_0)^{\frac{1}{\lambda}-1}$ $\qquad \frac{d}{dx}\tilde{E}(y|x)\Big|_{x_0}$

Simulations in progress...

Complications:

Power transform: $(\text{Big number})^{\text{Power}}$

CPU sweats !

Info matrix goes singular !

Above examples; approx conditioning ; simulations global

Try case where exact conditioning

Ex 3 Current. Non Normal regression

$y_i = \alpha + \beta x_i + \sigma z_i \qquad z_i \sim \text{Student}(7)$

$n = 7$

Data							
x	-3	-2	-1	0	1	2	3
y	-2.68	-4.02	-2.91	$.22$	$.38$	$-.28$	$.03$

from $\boxed{\alpha = 0 \quad \beta = 1 \quad \sigma = 1}$

Bédard, F, Wong

Above examples; approx conditioning ; simulations global

Try case where exact conditioning

Ex 3 Current Non Normal regression

$y_i = \alpha + \beta x_i + \sigma z_i \qquad z_i \sim \text{Student}(7)$

$n = 7$

Data	x	-3	-2	-1	0	1	2	3
	y	-2.68	-4.02	-2.91	$.22$	$.38$	$-.28$	$.03$

from $\boxed{\alpha = 0 \quad \beta = 1 \quad \sigma = 1}$

Test t-statistic $t^0 = -.122178$ $\qquad \pm\left(\frac{SST}{SSE}\right)^{1/2}$

3rd order $p^0 = .105255$

What is <u>true</u> p-value ?

MCMC and interesting things !

Bédard, F, Wong

$y_i = \alpha + \beta x_i + \sigma z_i \qquad z_i \sim \text{Student}(7)$

$n = 7$

Distn of $(\hat{\alpha}, \hat{\beta}, \hat{\sigma}) = (a, b, c)$ LS & residual length

$$f(a,b,s) = c\, s^4 \prod_1^7 \left\{ 1 + [a + b x_i + s d_i]^2 \right\}^{-4}$$

MCMC Proposal $N(0, .35)$

950 dump 50 ; repeat 5000 5M

Record whether " $t \le .05255$ "

$\hat{p} = .10765 \qquad SD = .002$

Recall 3rd order $p^0 = .10525$

$y_i = \alpha + \beta x_i + \sigma z_i \qquad z_i \sim \text{Student}(7)$

$n = 7$

<u>In motion</u> ---

$$\pi(\theta) L(\theta) = \prod_1^7 \left\{ 1 + \frac{1}{7\sigma^2} [y_i - \alpha - \beta x_i]^2 \right\}^{-4} \frac{1}{\sigma^7} \cdot \frac{1}{\sigma}$$

"True $\beta = 1$"

3rd order

Calculate $s(1)$

" $E(\beta)$

" $V(\beta)$

} Likelihood based algorithm

Verify by MCMC (adapted to asymptotic-smooth fn)

How do $f \in B$ fare?

" " 3rd & MCMC do?

Where to ...

More Likelihood & p-values

More conditioning

More priors

Less subjective

Less turf

It's just one discipline .... Tell the Dean!

Thank you