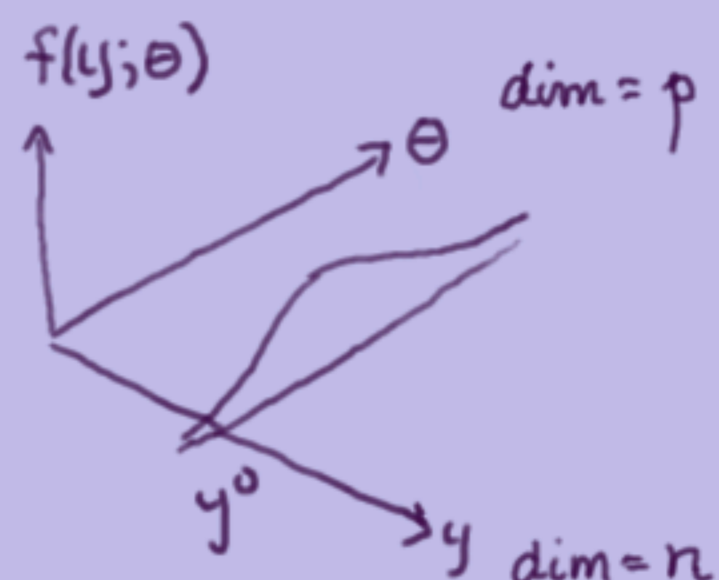
Likelihood, p-values, ancillaries and the vector quantile function

D A S Fraser
Statistics
U of Toronto

2009 May 8
Statistical Laboratory
U of Cambridge
UK

http://fisher.utstat.toronto.edu/dfraser/documents/camb09.pdf
" " /documents/xxx.pdf xxx =

a likelihood & p-values: Simple example - exact
b likelihood & p-values: approximate
c vector Quantile function: Essence of continuity
d How to measure θ? measure?
e How does θ move data? move?
f Two examples: nonlinear regression, Cauchy analysis
g default priors from continuity
h Summary



y^o

y^o-section $\quad f^o(\theta) = f(y^o; \theta)$

$F^o(\theta) = F(y^o; \theta)$

a) likelihood & p-values :

Example: Scalar y, scalar θ, linearity

Model $\quad f(y;\theta) = \sigma_0^{-1}\phi\left(\frac{y-\theta}{\sigma_0}\right) \qquad F(y;\theta) = \Phi\left(\frac{y-\theta}{\sigma_0}\right) \qquad \varphi, \Phi$ N O I or other

a) likelihood & p-value

Example: Scalar y, scalar θ

Model $\quad f(y;\theta) = \sigma_0^{-1}\phi\left(\frac{y-\theta}{\sigma_0}\right) \quad F(y;\theta) = \Phi\left(\frac{y-\theta}{\sigma_0}\right)$ $\quad \Phi$ N O 1 or other

Data: y^0 $\quad f^0(\theta) = c\,\phi\left(\frac{y^0-\theta}{\sigma_0}\right)$

a) likelihood & p-value

Example: Scalar y, scalar θ

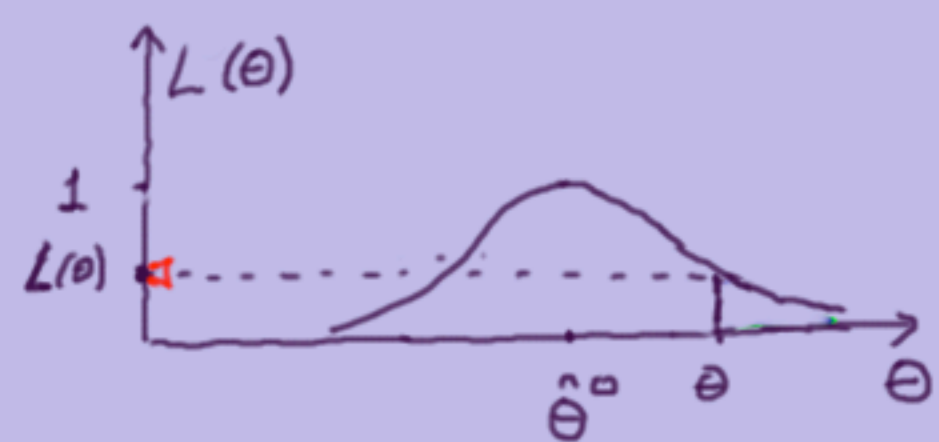
Φ N01 or other

Model $\quad f(y;\theta) = \sigma_0^{-1}\phi\left(\frac{y-\theta}{\sigma_0}\right) \qquad F(y;\theta) = \Phi\left(\frac{y-\theta}{\sigma_0}\right)$

Data y^0 $\quad f^0(\theta) = c\,\phi\left(\frac{y^0-\theta}{\sigma_0}\right)$

The y^0 section of model

Info about θ



Assessment of θ: $L^0(\theta) = f^0(\theta)$

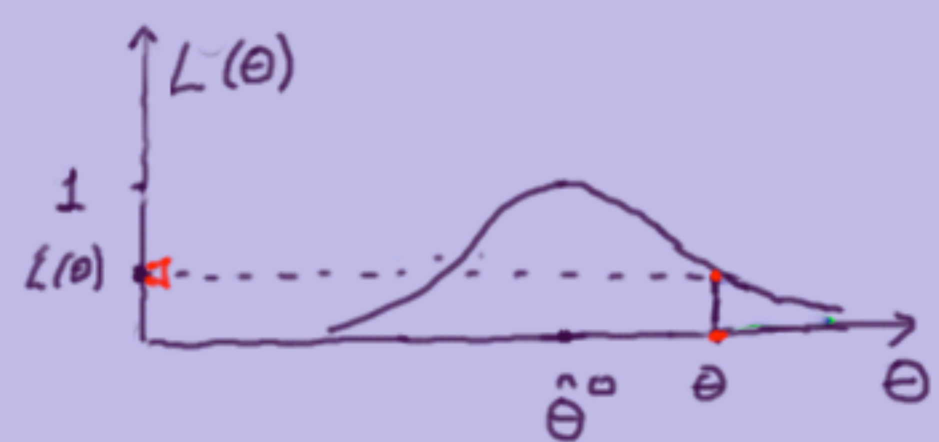
a) likelihood & p-value

Example: Scalar y, scalar θ

Model $\quad f(y;\theta) = \sigma_0^{-1}\phi\left(\frac{y-\theta}{\sigma_0}\right) \quad F(y;\theta) = \Phi\left(\frac{y-\theta}{\sigma_0}\right)$ $\quad \Phi$ N 0 1 or other

Data y^0 $\quad f^0(\theta) = c\,\phi\left(\frac{y^0-\theta}{\sigma_0}\right) \quad F^0(\theta) = \Phi\left(\frac{y^0-\theta}{\sigma_0}\right)$

Info about θ



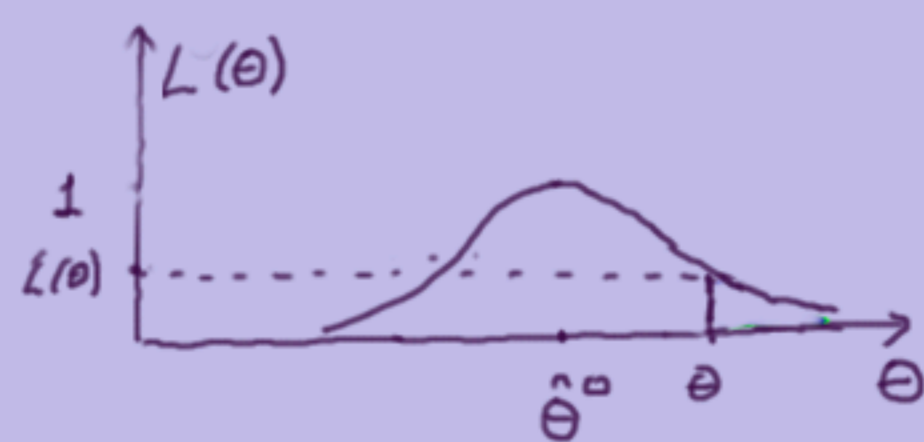
a) likelihood & p-value

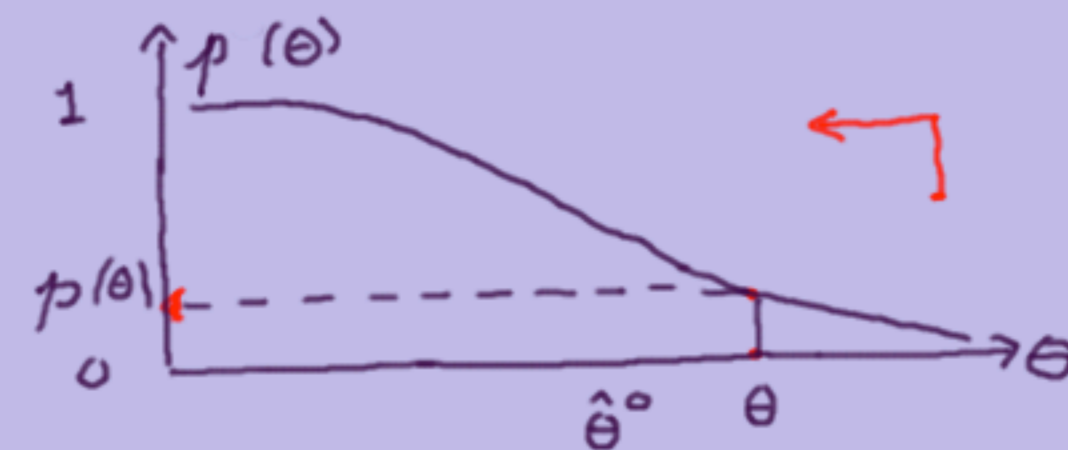
Example: Scalar y, scalar θ

Model $\quad f(y;\theta) = \sigma_0^{-1}\phi\left(\frac{y-\theta}{\sigma_0}\right) \quad F(y;\theta) = \Phi\left(\frac{y-\theta}{\sigma_0}\right)$ $\quad \Phi$ N01 or other

Data y^0 $\quad f^0(\theta) = c\,\phi\left(\frac{y^0-\theta}{\sigma_0}\right) \quad F^0(\theta) = \Phi\left(\frac{y^0-\theta}{\sigma_0}\right) = p(\theta)$

Info about θ





a) Data position (%age) re θ: $\quad p(\theta) = F^0(\theta)$

a) likelihood & p-value

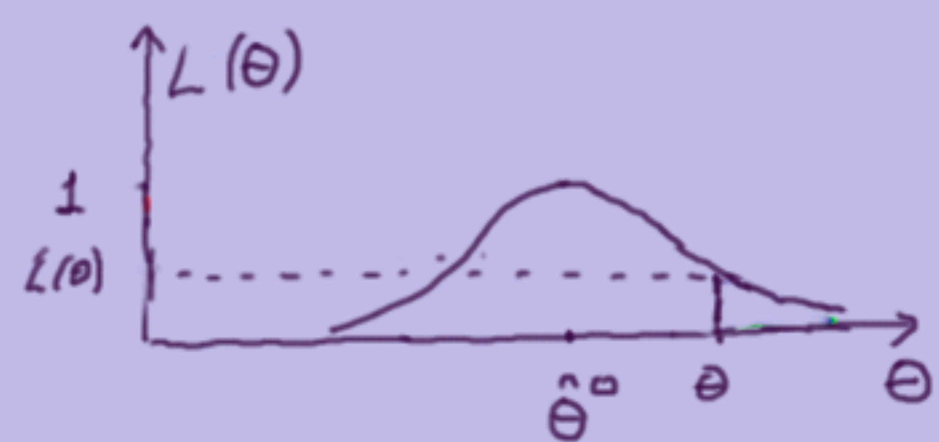
Example: Scalar y, scalar θ

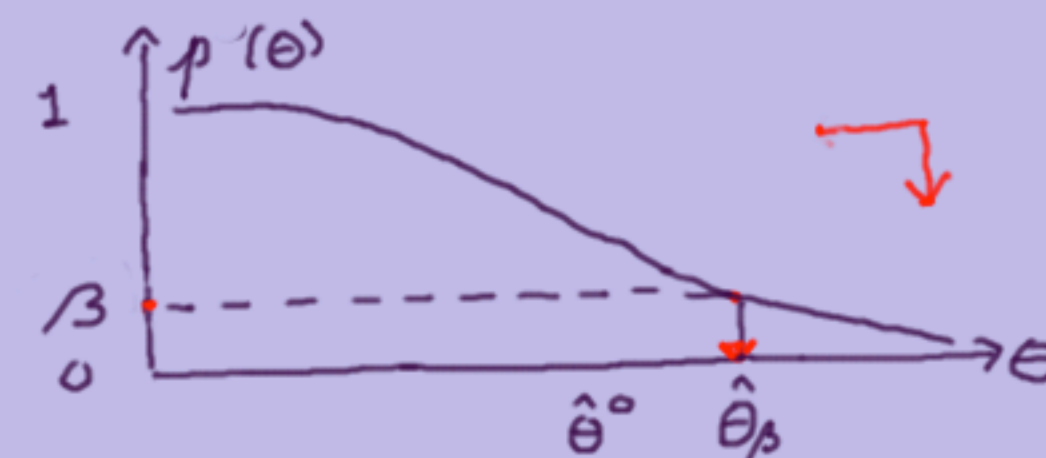
Model $\quad f(y;\theta) = \sigma_0^{-1}\phi\left(\frac{y-\theta}{\sigma_0}\right) \qquad F(y;\theta) = \Phi\left(\frac{y-\theta}{\sigma_0}\right)$

Φ N01 or other

Data $y^0 \quad f^0(\theta) = c\,\phi\left(\frac{y^0-\theta}{\sigma_0}\right) \qquad F^0(\theta) = \Phi\left(\frac{y^0-\theta}{\sigma_0}\right)$

Info about θ





a) Data position (%age) re θ: $\quad p(\theta) = F^0(\theta)$

b) β quantile: $\quad \hat{\theta}_\beta(y^0) = p^{-1}(\beta) = y^0 - \sigma_0 z_\beta \; ||$

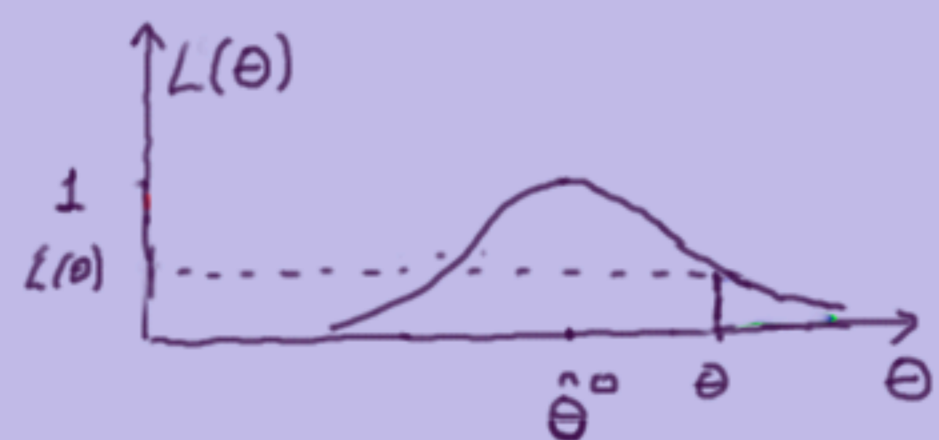
$\beta = \Phi(z_\beta)$

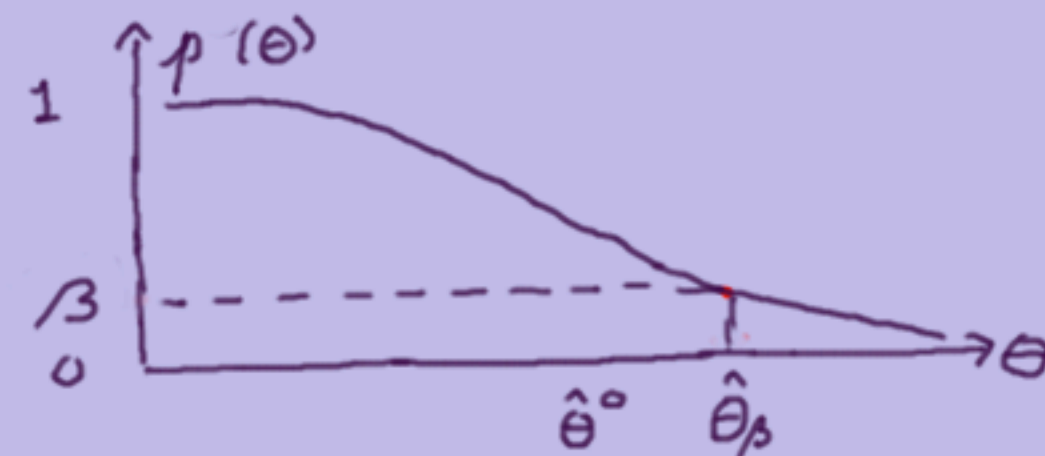
likelihood & p-value

Example: Scalar y, scalar θ

Model	$f(y;\theta) = \sigma_0^{-1}\phi\left(\frac{y-\theta}{\sigma_0}\right)$	$F(y;\theta) = \Phi\left(\frac{y-\theta}{\sigma_0}\right)$	Φ N01 or other
Data y^0	$f^0(\theta) = c\,\phi\left(\frac{y^0-\theta}{\sigma_0}\right)$	$F^0(\theta) = \Phi\left(\frac{y^0-\theta}{\sigma_0}\right)$	

Info about θ





Data assessment of θ

a) Data position (%age) re θ: $p(\theta) = F^0(\theta)$

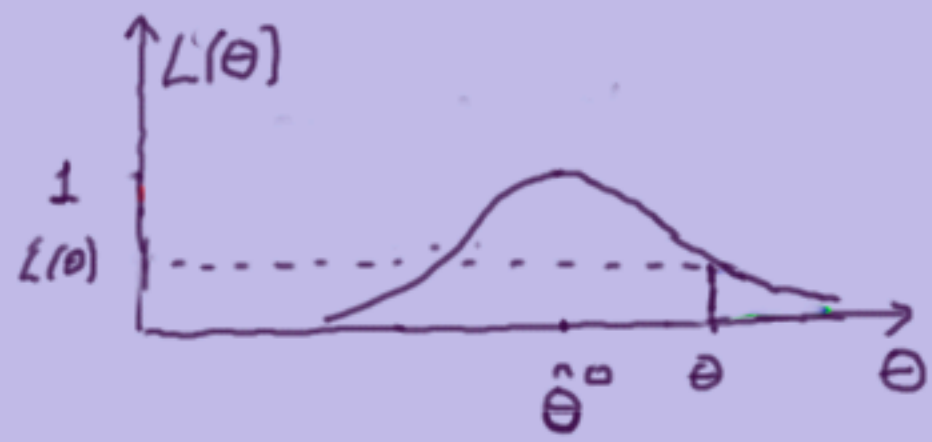
b) β quantile: $\hat\theta_\beta(y^0) = p^{-1}(\beta) = y^0 - \sigma_0 z_\beta$

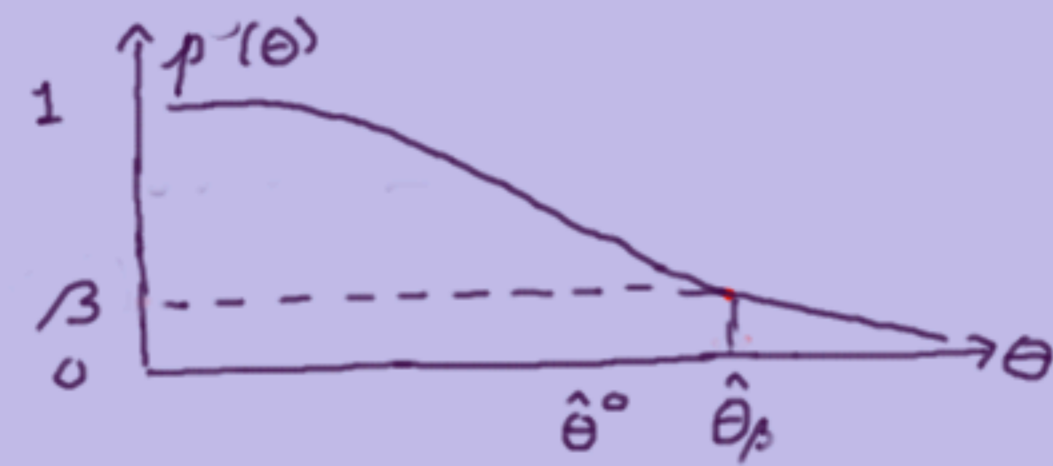
Likelihood
θ-density
...

p-value
posterior df
confidence df
...

$$L(\theta) = c\,\phi\left(\frac{y^0-\theta}{\sigma_0}\right) \qquad p(\theta) = \Phi\left(\frac{y^0-\theta}{\sigma_0}\right)$$

Info about θ





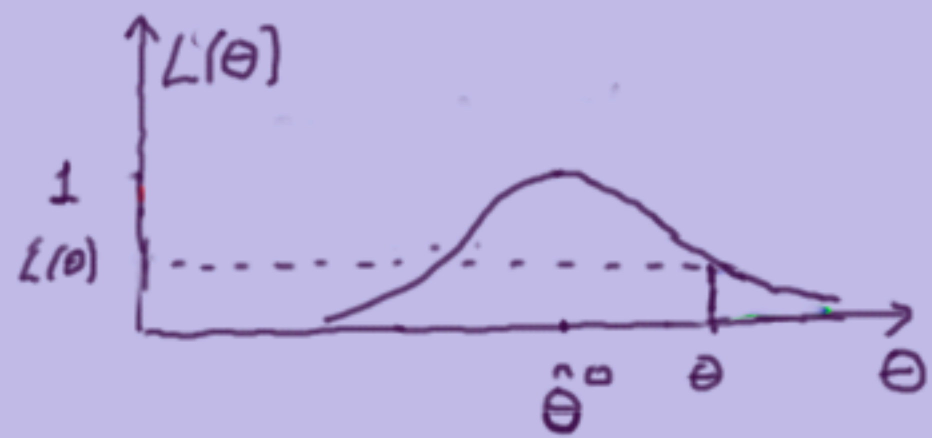
Likelihood → Data assessment $L(\theta)$

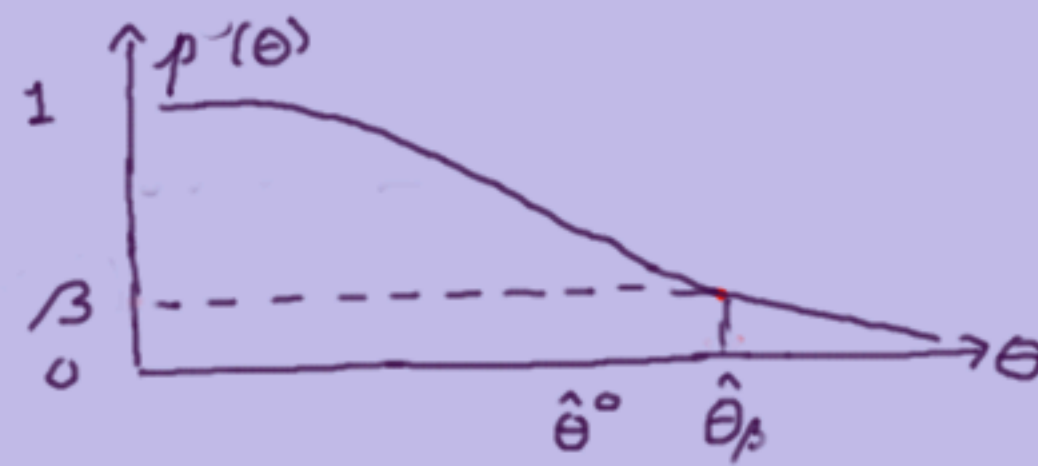
p-value → Data position $\quad p(\theta) = F^0(\theta)$

Bayes (1763) $\quad$ <u>ALL</u> of the above (location model case!)

$$L(\theta) = c\,\phi\left(\frac{y^0-\theta}{\sigma_0}\right) \qquad p(\theta) = \Phi\left(\frac{y^0-\theta}{\sigma_0}\right)$$

Info about θ





Likelihood → Data assessment $L(\theta)$

p-value → Data position $p(\theta) = F^0(\theta)$

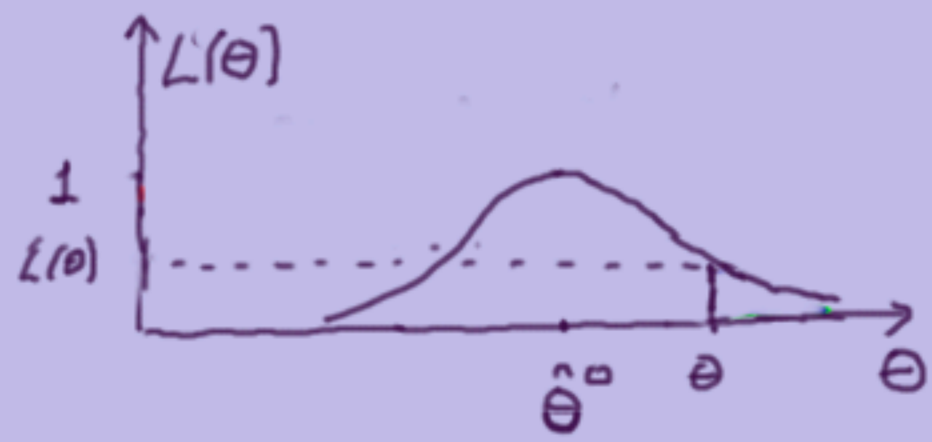
Bayes (1763) <u>ALL</u> of the above (location model case!)

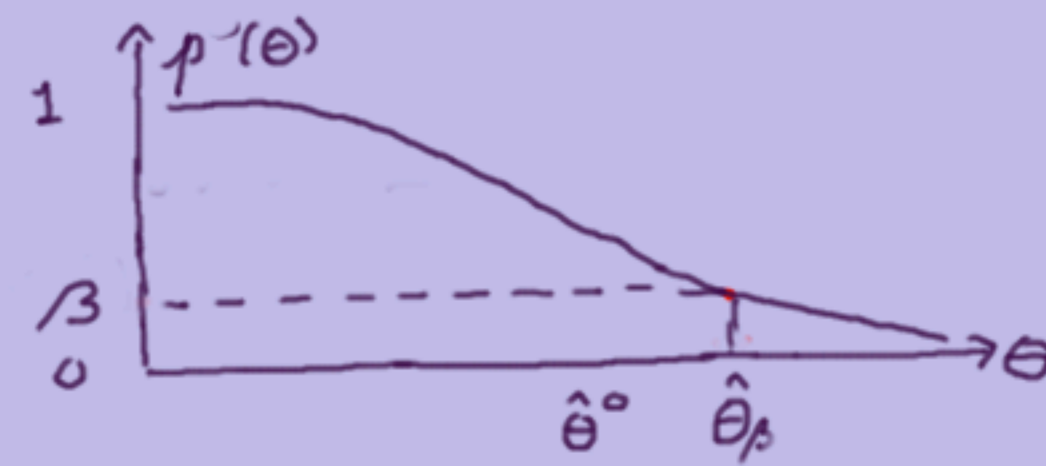
- Quibble about names ?
- Quibble about proofs ?

"All in Bayes (1763)"

$$L(\theta) = c\,\phi\left(\frac{y^{0}-\theta}{\sigma_0}\right) \qquad p(\theta) = \Phi\left(\frac{y^{0}-\theta}{\sigma_0}\right)$$

Info about θ





Likelihood — Data assessment $L(\theta)$

p-value — Data position $p(\theta) = F^{0}(\theta)$

Bayes (1763) <u>ALL</u> of the above (location model case!)

- Quibble about names ?
- Quibble about proofs ?

All in Bayes (1763)

Q: <u>But what if linearity</u> "$y \leftrightarrow \theta$" <u>is missing?</u> 247.pdf

Approximations?

b1

b) Likelihood & p-value: approximate ... available more widely!

(i) Exponential: $f = \exp\{\varphi(\theta) s - K(\theta)\} h(s) \cdot ds$ $s^{\circ} = 0$ +2 us

Asymptotic (p, p) +2+3 mp

b) likelihood & p-value: <u>approximate</u>

(i) <u>Exponential:</u> $f = \exp\{\varphi'(\theta)s - K(\theta)\} h(s) \cdot ds$ $s^{\circ} = 0$

Asymptotic (p, p)

Approx: $\tilde{f} = e^{k/n} \underbrace{\frac{1}{(2\pi)^{p/2}} e^{-r^2(\varphi, s)/2}} |\hat{j}_{\varphi\varphi}|^{-1/2} \cdot ds$ 3rd 1954 Daniels (Fourier Inverse)

Thomas Kuhn

b) likelihood & p-value: approximate

(i) Exponential: $f = \exp\{\varphi(\theta)s - K(\theta)\} h(s) \cdot ds$ $s^0 = 0$

Asymptotic (p, p)

Approx: $\tilde{f} = e^{k/n} \underbrace{\frac{1}{(2\pi)^{p/2}} e^{-r^2(\varphi, s)/2}} |\hat{j}_{\varphi\varphi}|^{-1/2} \cdot ds$ 3rd 1954 Daniels (Fourier Inverse)

Thomas Kuhn

$\tilde{F} = \Phi(r - r^{-1} \log \frac{r}{q})$ scalar θ 3rd 1980 Lugannani & Rice

1986 Barndorff-Nielsen

$r = \operatorname{sgn}(\hat{\theta} - \theta) \cdot [2\{\ell(\hat{\theta}; y) - \ell(\theta; y)\}]^{1/2}$ SLR

$q = \operatorname{sgn}(\hat{\theta} - \theta) \cdot |\hat{j}_{\varphi\varphi}^{1/2} (\hat{\varphi} - \varphi)|$ mle departure φ scale

b) likelihood & p-value: approximate

(i) Exponential: $f = \exp\{\varphi'(\theta) s - K(\theta)\} h(s) \cdot ds \qquad s^\circ = 0$

Asymptotic (p, p)

Approx: $\tilde{f} = e^{k/n} \underbrace{\frac{1}{(2\pi)^{p/2}} e^{-r^2(\varphi, s)/2}} |\hat{j}_{\varphi\varphi}|^{-1/2} \cdot ds$

$\tilde{F} = \Phi(r - r^{-1} \log \frac{r}{q})$ scalar θ

$r = \operatorname{sgn}(\hat{\theta} - \theta) \cdot [2\{\ell(\hat{\theta}; y) - \ell(\theta; y)\}]^{1/2}$ SLR

$q = \operatorname{sgn}(\hat{\theta} - \theta) \cdot |\hat{j}_{\varphi\varphi}^{1/2}(\hat{\varphi} - \varphi)|$ mle departure φ scale

(ii) General: $\varphi(\theta) = \varphi(\theta; y^\circ) = \frac{\partial}{\partial y} \ell(\theta; y)|_{y^\circ}$ 2 + 2 + 2

(asymptotic, (p, p))

$\cancel{\ell(\theta; y)} \Rightarrow \ell(\varphi; s) = \ell^\circ(\theta) + \varphi(\theta) \cdot s \qquad s^\circ = 0$

	y°	ow
$\tilde{f}$ = as above	3rd	2nd
$\tilde{F}$ = as above	3rd	2nd

1980 Barndorff-Nielsen

1986 Barndorff-Nielsen

1990 F 143.pdf

1999 F Reid Wu 196.pdf

b) likelihood & p-value: approximate

(i) Exponential: $f = \exp\{\varphi(\theta)s - K(\theta)\} h(s) \cdot ds$ $\quad s^\circ = 0$

Asymptotic (p, p)

Approx: $\tilde{f} = e^{k/n} \underbrace{\frac{1}{(2\pi)^{p/2}} e^{-r^2(\varphi, s)/2}} |\hat{j}_{\varphi\varphi}|^{-1/2} \cdot ds$

$\tilde{F} = \Phi(r - r^{-1} \log \frac{r}{q})$ $\quad$ scalar θ

$r = \operatorname{sgn}(\hat{\theta} - \theta) \cdot [2\{\ell(\hat{\theta}; y) - \ell(\theta; y)\}]^{1/2}$ $\quad$ SLR

$q = \operatorname{sgn}(\hat{\theta} - \theta) \cdot |\hat{j}_{\varphi\varphi}^{1/2}(\hat{\varphi} - \varphi)|$ $\quad$ mle departure $\quad$ φ scale

(ii) General: $\varphi(\theta) = \varphi(\theta; y^\circ) = \frac{\partial}{\partial y} \ell(\theta; y)\Big|_{y^\circ}$

Asymptotic, (p, p)

$\ell(\cancel{\theta}; y) \Rightarrow \ell(\varphi; s) = \ell^\circ(\theta) + \varphi(\theta) \cdot s$ $\quad s^\circ = 0$

	y°	ow
$\tilde{f}$ = as above	3rd	2nd
$\tilde{F}$ = as above	3rd	2nd

1980 Barndorff-Nielsen

1986 Barndorff-Nielsen

1990 F $\quad$ 143.pdf

1999 F Reid Wu $\quad$ 196.pdf

Moral: log-Likelihood is a cgf, to 3rd order: just Calibrate! (f / B

$\ell^\circ(\theta) \quad \varphi(\theta)$

(iii) General parameters:
Casymptotic (p, p)

Interest $\psi(\theta)$ dim d
Nuisance $\lambda(\theta)$ $p-d$

$\hat{\ell} - \hat{\ell}_\psi$

$$\tilde{f} = e^{k/n} \frac{1}{(2\pi)^{d/2}} \exp\left\{-\frac{r^2(\theta, s)}{2}\right\} \left\{\frac{|j_{(\lambda\lambda)}(\hat{\theta}_\psi)|}{|j_{\varphi\varphi}(\hat{\theta})|}\right\}^{1/2} ds$$

3rd at y^0 2nd ow

F 2003
219.pdf

- $\ell(\varphi; s) = \ell^0(\theta) + \varphi(\theta) \cdot s$ $\surd$
- $\varphi(\theta) = \varphi(\theta; y^0) = \frac{\partial}{\partial y} \ell(\theta; y)\big|_{y^0}$

but have $\hat{j}^0_{\varphi\varphi} = I$

- (..) Derivatives in φ scale

$$r = \pm \left\{2\left[\ell^0(\hat{\psi}, \hat{\lambda}) - \ell^0(\psi, \hat{\lambda}_\psi)\right]\right\}^{1/2}$$

(iii) <u>General parameters:</u> Asymptotic (p, p)

Interest $\psi(\theta)$ dim d
Nuisance $\lambda(\theta)$ $p-d$

$$\tilde{f} = e^{k/n} \frac{1}{(2\pi)^{d/2}} \exp\left\{-\frac{r^2(\theta, s)}{2}\right\} \left\{\frac{|j_{(\lambda\lambda)}(\hat{\theta}_\psi)|}{|j_{\varphi\varphi}(\hat{\theta})|}\right\}^{1/2} ds$$

3rd at y° 2nd ow

- $\ell(\varphi; s) = \ell^\circ(\theta) + \varphi(\theta) \cdot s$

F 2003
219.pdf

- $\varphi(\theta) = \varphi(\theta; y^\circ) = \frac{\partial}{\partial y} \ell(\theta; y)|_{y^\circ}$

but have $\hat{j}^\circ_{\varphi\varphi} = I$

- $(\cdot\cdot)$ Derivatives in φ scale

$$\tilde{F} = \Phi\left(r - r^{-1} \log \frac{r}{q}\right)$$

3rd dim $\psi = 1$

$$r = \pm \left\{2[\ell^\circ(\hat{\psi}, \hat{\lambda}) - \ell^\circ(\psi, \hat{\lambda}_\psi)]\right\}^{1/2}$$

B-N 1986
F Reid 1993
171.pdf

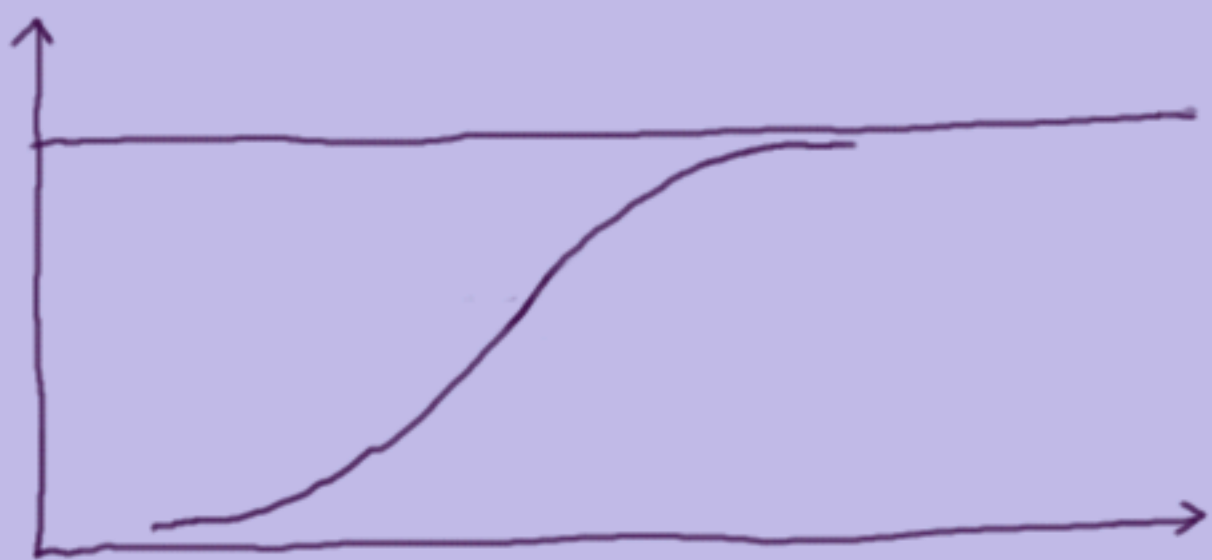
$$q = \pm \left|(\hat{\chi} - \chi)\right| \left\{\frac{|\hat{j}_{\varphi\varphi}|}{|j_{(\lambda\lambda)}(\hat{\theta}_\psi)|}\right\}^{1/2}$$

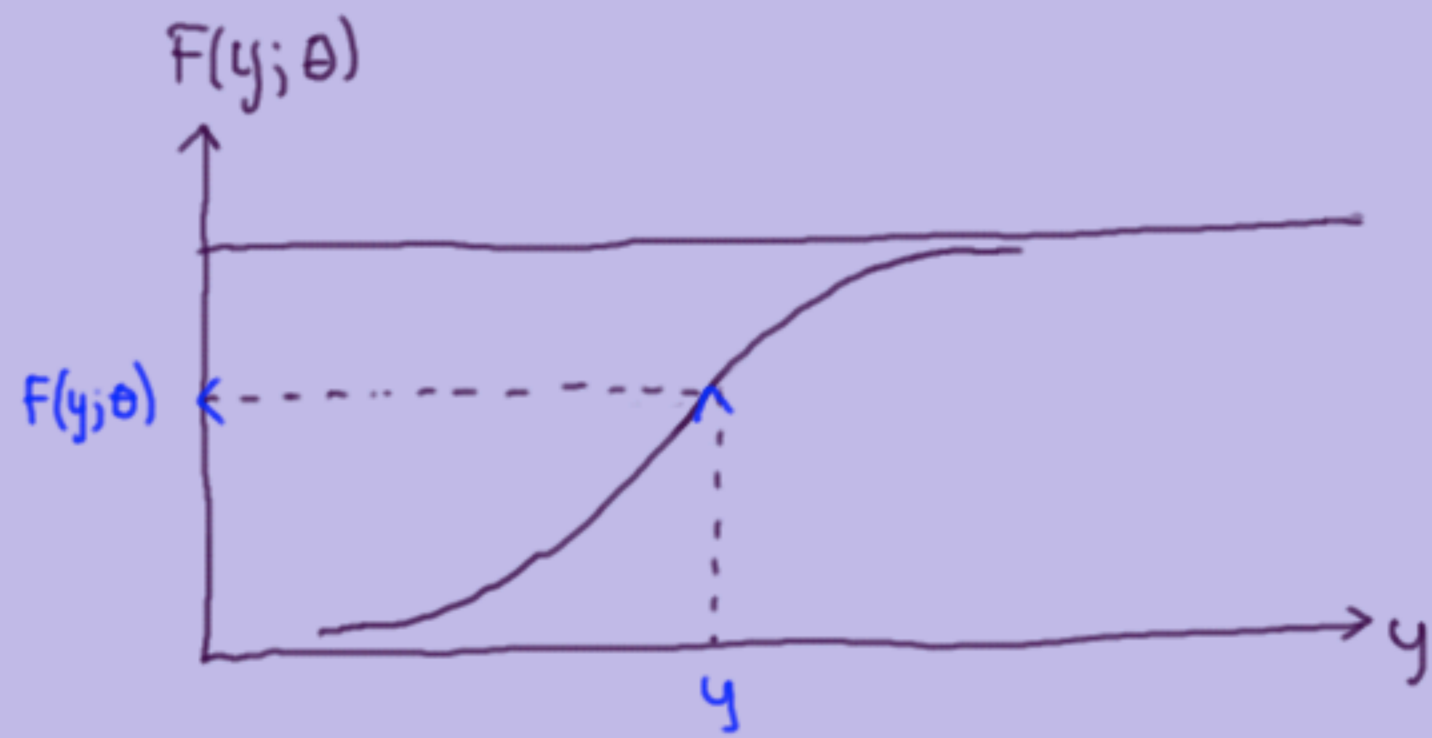
$\pm = \operatorname{sgn}(\hat{\psi} - \psi)$ *

$$\chi = \frac{\psi_\varphi(\hat{\theta}^\circ_\psi)}{|\cdot|} \cdot \varphi$$

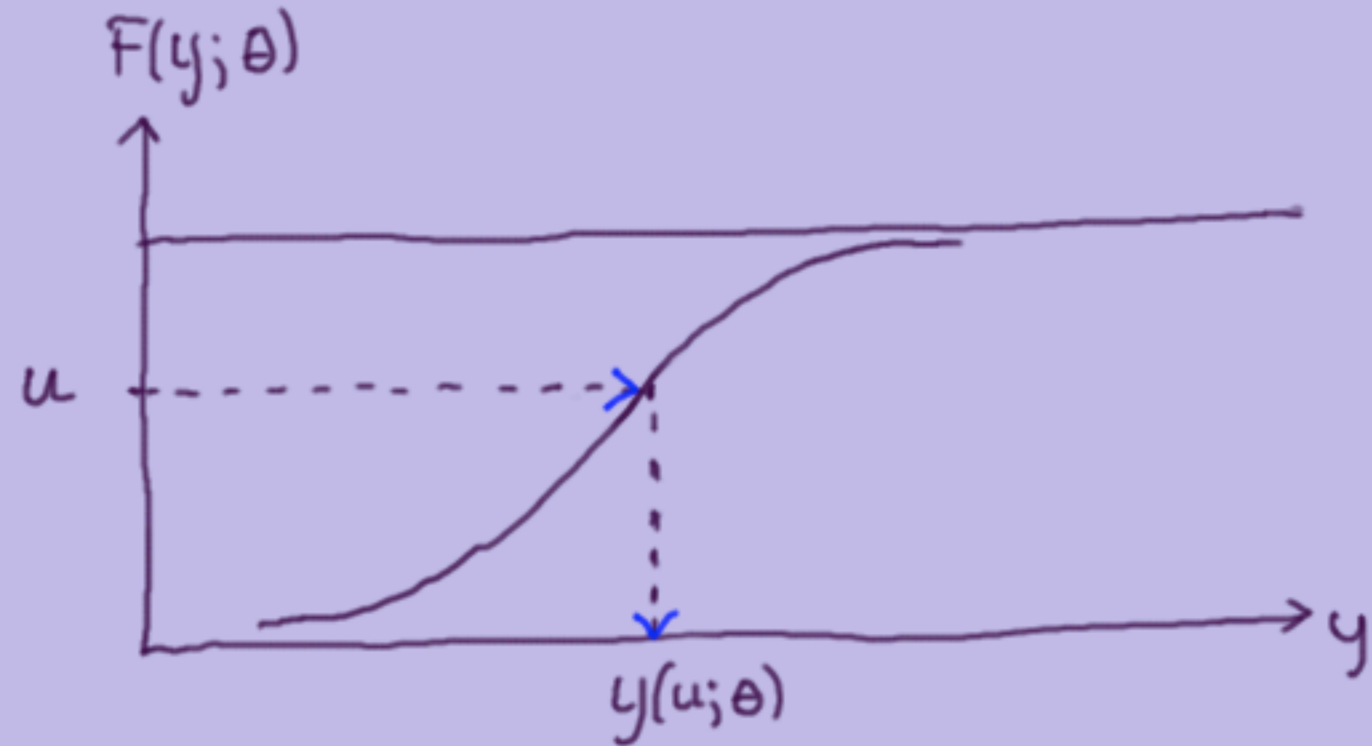
c) Vector quantile function : a different look

What is this?





a) df $\quad u = F(y;\theta) \quad$ p-value



a) df $\quad u = F(y;\theta) \quad$ p-value

b) qf $\quad y = y(u;\theta) \quad$ quantile

More "informative"
" useful
Continuity!

Vector quantile: Case: Indep. coords

$$y_1 = y_1(u_1, \theta)$$
$$\vdots$$
$$y_n = y_n(u_n; \theta)$$

$$\underset{\sim}{y} = \underset{\sim}{y}(\underset{\sim}{u}; \theta)$$ .... How θ <u>affects</u> variable / model <u>Continuity</u>

How θ <u>moves</u> pile of probability !

d) How to measure Θ ? A different look...

R^n

$\cdot\, y^0$

(i) A bucket at data? Go primative

R^n Put a <u>bucket</u> at y^0

y^0

How much prob at data?

$\Pr\{\square\,;\theta\} = f(y^0;\theta)dy = L^0(\theta)\,dy$ ... *Likelihood*

R^n Put a <u>bucket</u> at y^o

y^o

How much prob at data ?

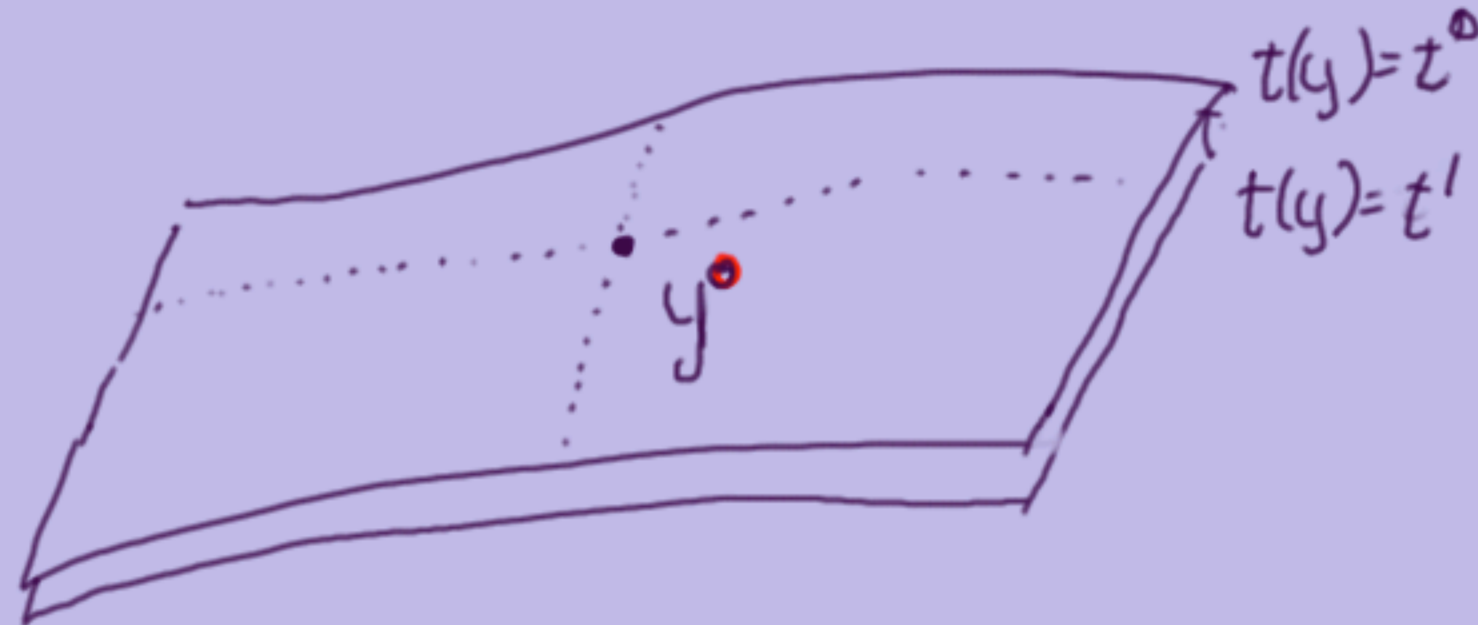
$Pr\{\square\,;\theta\} = f(y^o;\theta)dy = L^o(\theta)\,dy$ ... Likelihood

- Ignore model otherwise? high principle or low concern!

- use	$\omega(\theta)\,L^o(\theta)$	"left wing" ~~reactionary~~ "f"	low c
- Go Bayes	$\pi(\theta)\,L^o(\theta)$	Default "B"	high p
		prior? Sanctions	

(ii) a statistic through data ?

R^n Get a statistic to measure θ !



Traditional: easy ?

Definition elsewhere ?

Distribution ?

(iii) A direction V at data?

R^n

How to get sensible V?

e) How does theta move data?

Vector quantile function $\underset{\sim}{y} = \underset{\sim}{y}(\underset{\sim}{u}; \theta)$ ... continuity, coordinates

Data $\underset{\sim}{y}^{\circ}$ Obs. mle $\hat{\theta}^{\circ}$

Vector quantile function $\underset{\sim}{y} = \underset{\sim}{y}(\underset{\sim}{u}; \theta)$ ... continuity, coordinates

Data y° Obs. mle $\hat{\theta}^\circ$

$\downarrow$

Estimated p-value vector $\hat{u}^\circ$: $y^\circ = y(\hat{u}^\circ; \hat{\theta}^\circ)$

Examine trajectory of $y = y(\hat{u}^\circ; \theta)$ under θ change

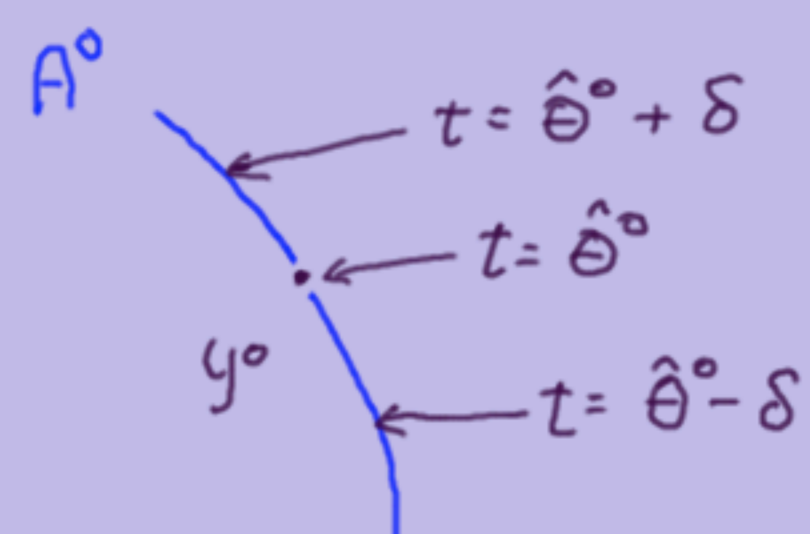
Vector quantile function $\underset{\sim}{y} = \underset{\sim}{y}(\underset{\sim}{u}; \theta)$ ... continuity, coordinates

Data y° Obs. mle $\hat{\theta}^\circ$

Estimated p-value vector $\hat{u}^\circ$: $y^\circ = y(\hat{u}^\circ; \hat{\theta}^\circ)$

Examine trajectory of $y = y(\hat{u}^\circ; \theta)$

Trajectory $= A^\circ = \{ y(\hat{u}^\circ; t) : t \in R^p \}$



What do you get? ... the Intrinsic 2nd order ancillary !

Vector quantile tells the story! ... flow of probability mass: Hydrodynamics!

$$y(\hat{u}^o;\theta) = y^o + \underset{1\times p}{V}\theta + \theta' \underset{p\times p}{W}\theta/2n^{1/2} + \qquad \theta \leftarrow \theta - \hat{\theta}^o \qquad n^{-1/2}$$

Just Taylor!

$V = \dot{y}|_{y^o} = \frac{\partial}{\partial\theta} y(u;\theta)|_{y^o,\hat{\theta}^o}$ = velocity "Elements" are vectors in R^n

$W = \ddot{y}|_{y^o}$ = acceleration Yet use matrix mult. re θ

<u>Vector quantile</u> tells the story! ... <u>flow</u> of probability mass

$$y(\hat{u}^0;\theta) = y^0 + \underset{1\times p}{V}\theta + \theta' \underset{p\times p}{W}\theta/2n^{1/2} + \qquad \theta \leftarrow \theta - \hat{\theta}^0 \qquad n^{-1/2}$$

$V = \dot{y}|_{y^0} = \frac{\partial}{\partial\theta} y(u;\theta)|_{y^0,\hat{\theta}^0}$ = velocity — "Elements" are vectors in R^n

$W = \ddot{y}|_{y^0}$ = acceleration — yet use matrix mult. re θ

<u>Contour</u> for measuring θ: conditionally! — F F Staicu 240. pdf

$A^0 = \{ y(\hat{u}^0; t) : t \in R^p \}$ — Traj from y^0

$= y^0 + \{ Vt + t'Wt/2n^{1/2} : t \in R^p \}$ — Taylor from y^0

<u>Vector quantile</u> tells the story! ... <u>flow</u> of probability mass

$$y(\hat{u}^\circ;\theta) = \underset{1\times p}{y^\circ + V\theta} + \underset{p\times p}{\theta' W\theta}/2n^{1/2} + \qquad \theta \leftarrow \theta - \hat{\theta}^\circ \qquad n^{-1/2}$$

$V = \dot{y}|_{y^\circ} = \frac{\partial}{\partial\theta} y(u;\theta)|_{y^\circ,\hat{\theta}^\circ}$ = velocity — "Elements" are vectors in R^n

$W = \ddot{y}|_{y^\circ}$ = acceleration — yet use matrix mult. re θ

Contour for measuring θ: conditionally!

$A^\circ = \{ y(\hat{u}^\circ; t) : t \in R^p \}$ — Traj from y°

$= y^\circ + \{ Vt + t' W t/2n^{1/2} : t \in R^p \}$ — Taylor from y°

$=$ Solution through y°: "velocity $\frac{dy}{d\theta}|_{y\ \hat{\theta}(y)}$ at y" Frobenius (2nd)

Vector quantile tells the story! ... flow of probability mass

$y(\hat{u}^\circ;\theta) = y^\circ + \underset{1\times p}{V}\theta + \theta' \underset{p\times p}{W}\theta/2n^{1/2} +$ $\qquad \theta \leftarrow \theta - \hat{\theta}^\circ \qquad n^{-1/2}$

$V = \dot{y}|_{y^\circ} = \frac{\partial}{\partial\theta} y(u;\theta)|_{y^\circ,\hat{\theta}^\circ}$ = velocity $\qquad$ "Elements" are vectors in R^n

$W = \ddot{y}|_{y^\circ}$ = acceleration $\qquad$ yet use matrix mult. re θ

Contour for measuring θ: conditionally!

$A^\circ = \{ y(\hat{u}^\circ; t) : t \in R^p \}$ $\qquad$ Traj from y°

$= y^\circ + \{ Vt + t'Wt/2n^{1/2} : t \in R^p \}$ $\qquad$ Taylor from y°

= Solution through y°; velocity $\frac{dy}{d\theta}|_{y\,\hat{\theta}(y)}$ at y $\qquad$ Frobenius (2nd)

Same to 2nd order

Contour: prob. free of θ ... 2nd

Ancillary ... 2nd

"Condition on continuity"

F, F, Staicu

240.pdf

Condition on $y^\circ + Vt + t'Wt/2n^{1/2}$ 2nd order ancillary

Inference $\varphi(\theta) = \ell_{;V}(\theta; y^\circ)$

= likelihood gradient in full/conditioned model

Use: $\{\ell^\circ(\theta), \varphi(\theta)\}$ Act as if Exptl model F. Reid (1993) 178.pdf

Inference for any $\psi(\theta)$ ... 3rd

log-likelihood is a cgf to 3rd order "Extended" Daniels 3rd

Ex1 Non linear regression

Ex. $y \sim N\{\underset{\sim}{x}(\theta) ; I\}$ Solution surface $= S = \{\underset{\sim}{x}(\theta)\}$ $N; I/n$

Ex1 Non linear regression

Ex. $y \sim N\{\underset{\sim}{x}(\theta) ; I\}$ Solution surface = S = $\{\underset{\sim}{x}(\theta)\}$ I/n

Normal on Circle

$\sim N\left\{\begin{matrix}\rho_0 \cos\theta \\ \rho_0 \sin\theta\end{matrix} ; I\right\}$

Fit $\hat{y}^o$

Residual $y^o - \hat{y}^o$

Ex1 Non linear regression

<u>Ex</u>. $y \sim N\{\underset{\sim}{x}(\theta) ; I\}$

Solution surface $= S = \{\underset{\sim}{x}(\theta)\}$

I/n

Quantile conditioning contour:

$A^\circ = S + (y^\circ - \hat{y}^\circ)$

Normal on Circle

$\sim N\left\{\begin{matrix}\rho_0 \cos\theta \\ \rho_0 \sin\theta\end{matrix} ; I\right\}$

Ex1 Non linear regression

<u>Ex</u>. $y \sim N\{\underset{\sim}{x}(\theta);\ I\}$ Solution surface $= S = \{\underset{\sim}{x}(\theta)\}$ I/n

Quantile conditioning contour:

$$A^\circ = S + (y^\circ - \hat{y}^\circ)$$

Normal on Circle

$$\sim N\left\{\begin{matrix}\rho_0 \cos\theta \\ \rho_0 \sin\theta\end{matrix} ;\ I\right\}$$

<u>Severini</u> (2000) p216

- pivot $= y - x(\theta)$
- plug in $\hat{\theta} = a = \tan^{-1}(y_2/y_1)$
- residual $= y - \hat{y} = (r - \rho\ )\begin{pmatrix}\cos a \\ \sin a\end{pmatrix}$ ancillary ?

$y \leftrightarrow (r, a)$

but <u>one-one</u> To $r\begin{bmatrix}\cos a \\ \sin \theta\end{bmatrix} = y$, data itself not ancillary!

Ex1 Non linear regression

<u>Ex</u>. $y \sim N\{\underset{\sim}{x}(\theta); I\}$ Solution surface $= S = \{\underset{\sim}{x}(\theta)\}$ I/n

Quantile conditioning contour:

$$A^{\circ} = S + (y^{\circ} - \hat{y}^{\circ})$$

Normal on Circle

$\sim N\left\{\begin{matrix} \rho_0 \cos\theta \\ \rho_0 \sin\theta \end{matrix} ; I\right\}$

Severini (2000) p216

- pivot $= y - x(\theta)$
- plug in $\hat{\theta}$
- residual $= y - \hat{y} = (r - \rho)\begin{pmatrix} \cos a \\ \sin a \end{pmatrix}$ ancillary ?

$y \leftrightarrow (r, a)$

but <u>one-one</u> to $r\begin{pmatrix} \cos a \\ \sin \theta \end{pmatrix} = y$, data itself not ancillary !

<u>But</u>: 2nd order ancillary gets "Circle through y°"

"Right !"

Ex 2 Cauchy (μ, σ)

$$y_1 = \mu + \sigma z_1$$
$$\vdots$$
$$y_n = \mu + \sigma z_n$$

$$\left.\frac{dy}{d\mu, \sigma}\right|_{y^\circ} = V(\theta) = (1, \hat{z}^\circ)$$

Ex 2 Cauchy (μ, σ)

$$y_1 = \mu + \sigma z_1 \qquad \left.\frac{dy}{d\mu,\sigma}\right|_{y^\circ} = V(\hat{\theta}^\circ) = (1, \hat{z}^\circ)$$

$$\vdots$$

$$y_n = \mu + \sigma z_n$$

Condition $\quad y = \mu \underset{\sim}{1} + \sigma \hat{\underset{\sim}{z}}^\circ$

$$A^\circ = \{ a \underset{\sim}{1} + c \hat{\underset{\sim}{z}}^\circ \} = \mathcal{L}(1, \hat{z}^\circ)$$

usual

ie configuration!

Ex 2 Cauchy (μ, σ)

$$y_1 = \mu + \sigma z_1 \qquad \left.\frac{dy}{d\mu,\sigma}\right|_{y^0} = V(\hat{\theta}^0) = (1, \hat{z}^0)$$

$$\vdots$$

$$y_n = \mu + \sigma z_n$$

Condition $\quad y = \mu \underset{\sim}{1} + \sigma \underset{\sim}{\hat{z}}^0$

$$A^0 = \{ a\underset{\sim}{1} + c\,\underset{\sim}{\hat{z}}^0 \} = \mathcal{L}(1, \hat{z}^0) \qquad \text{ie configuration}$$

McCullagh (1992)

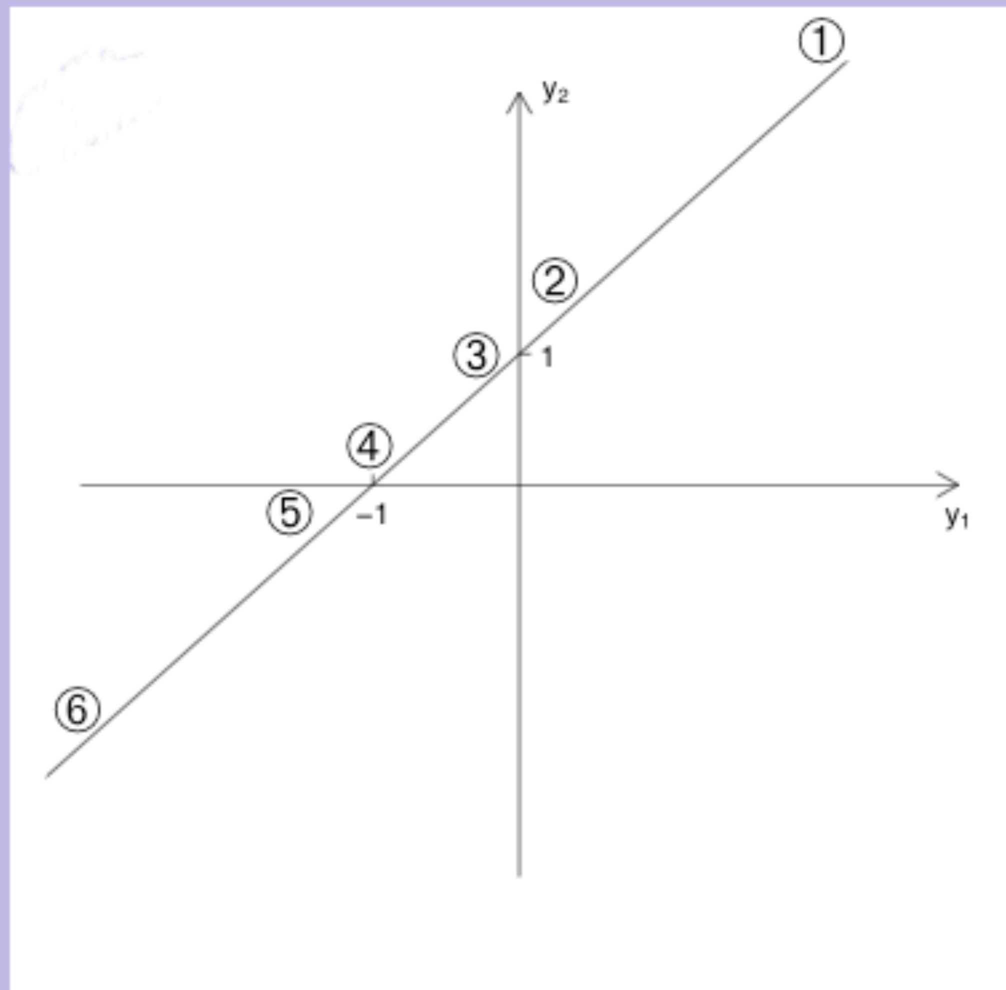
$$y \ \text{Cauchy}(\mu, \sigma) \qquad \tilde{\mu} = \frac{\mu}{\mu^2 + \sigma^2}$$

$$y^{-1} \ \text{Cauchy}(\tilde{\mu}, \tilde{\sigma}) \qquad \tilde{\sigma} = \frac{\sigma}{\mu^2 + \sigma^2}$$

Proposed: Use ancillary for $(y_1^{-1}, \ldots, y_n^{-1})$

Get different conditioning!

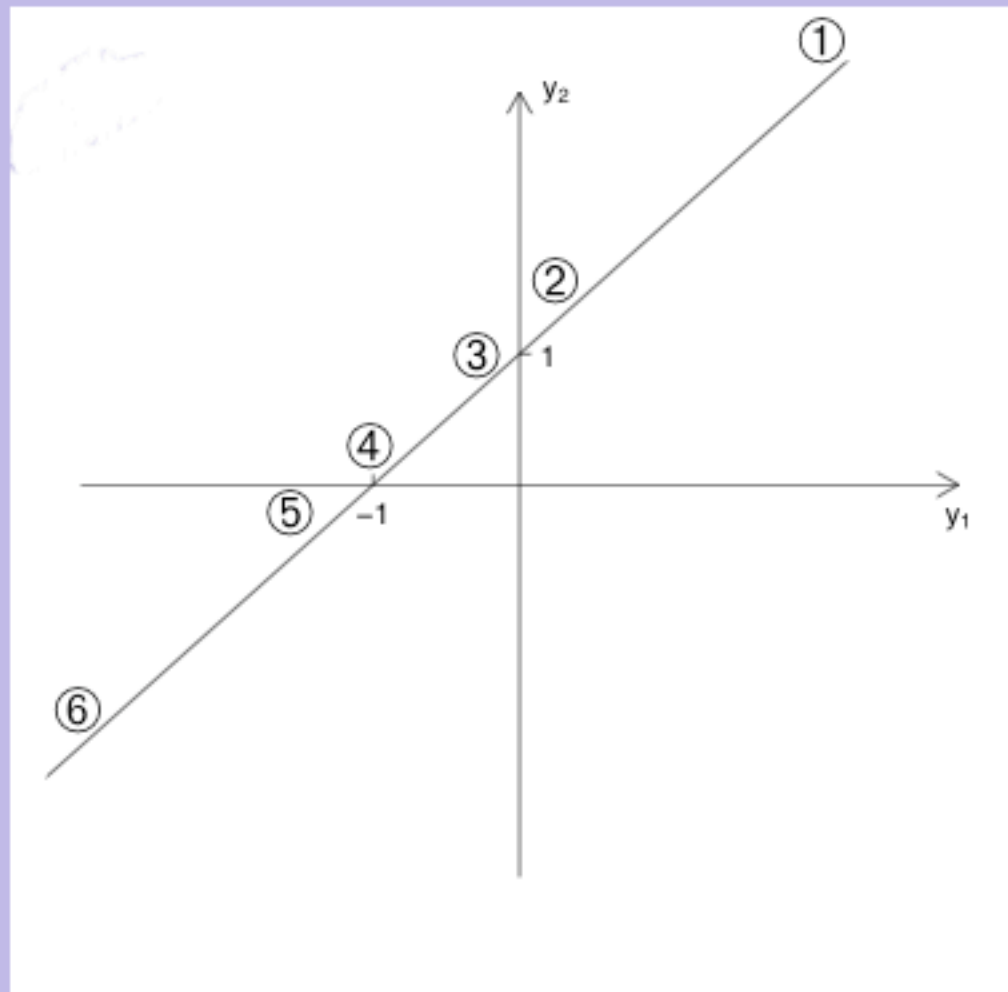
Elegant ... but ..!



a line ①⋯⑥

on usual ancillary surface

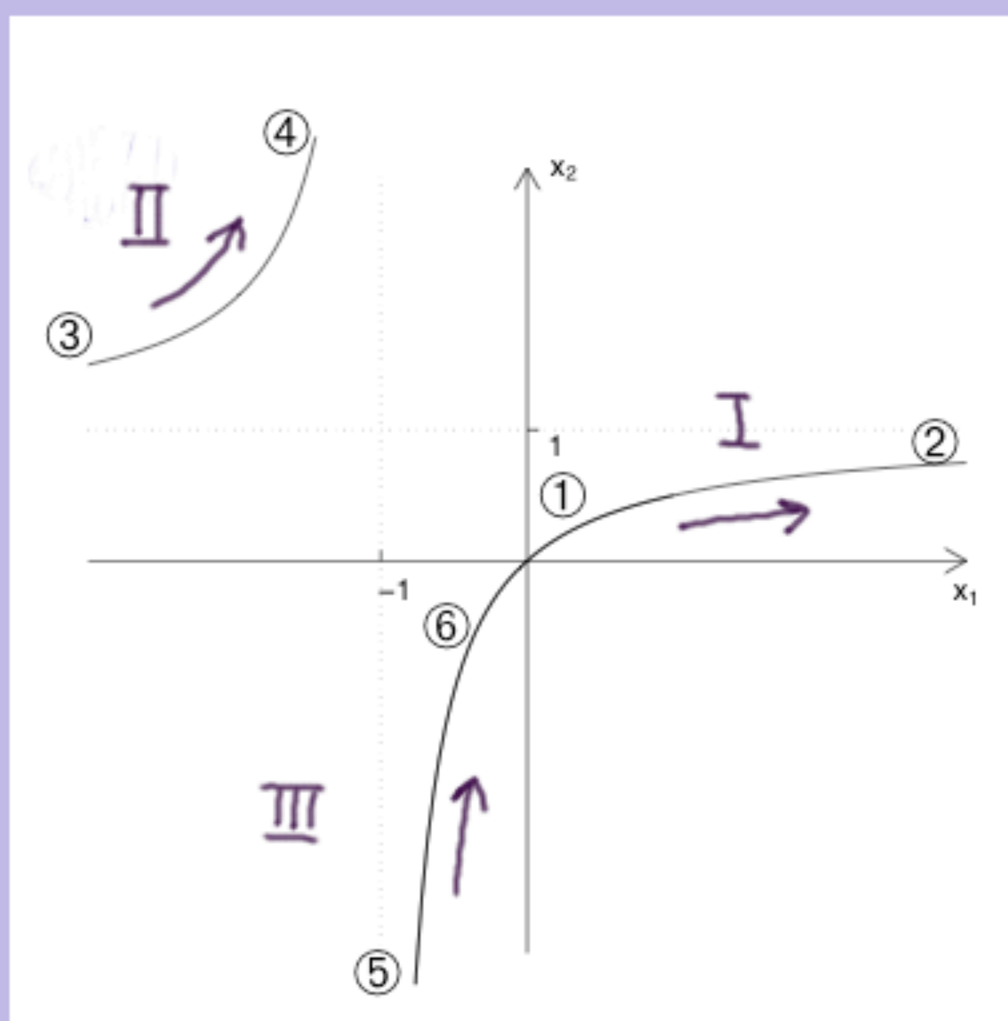
$\mathcal{L}^{+}(1; y^{\circ})$



a line ①···⑥ $n=2$

on usual ancillary surface

$$\mathcal{L}^+(1;y^\circ)$$



After "reciprocal" mapping

get 3 segments n segments!

①→② I

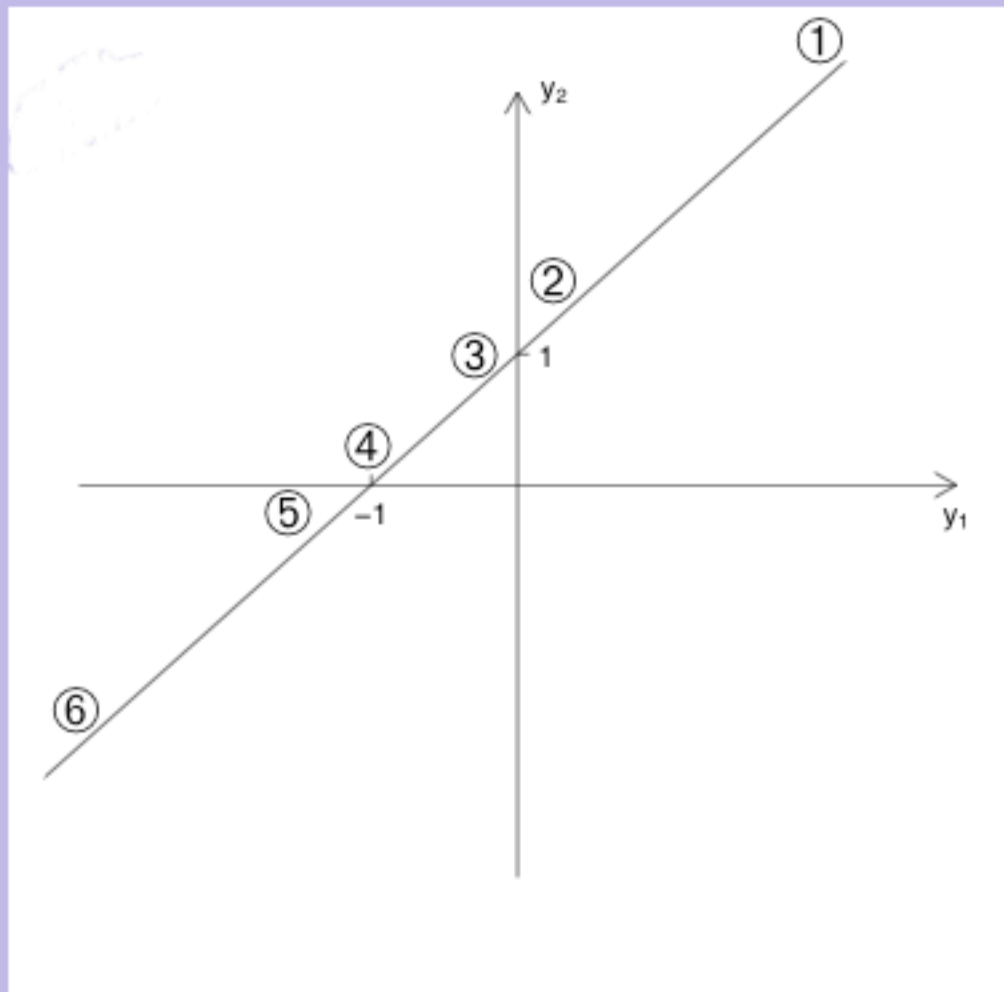
③→④ II

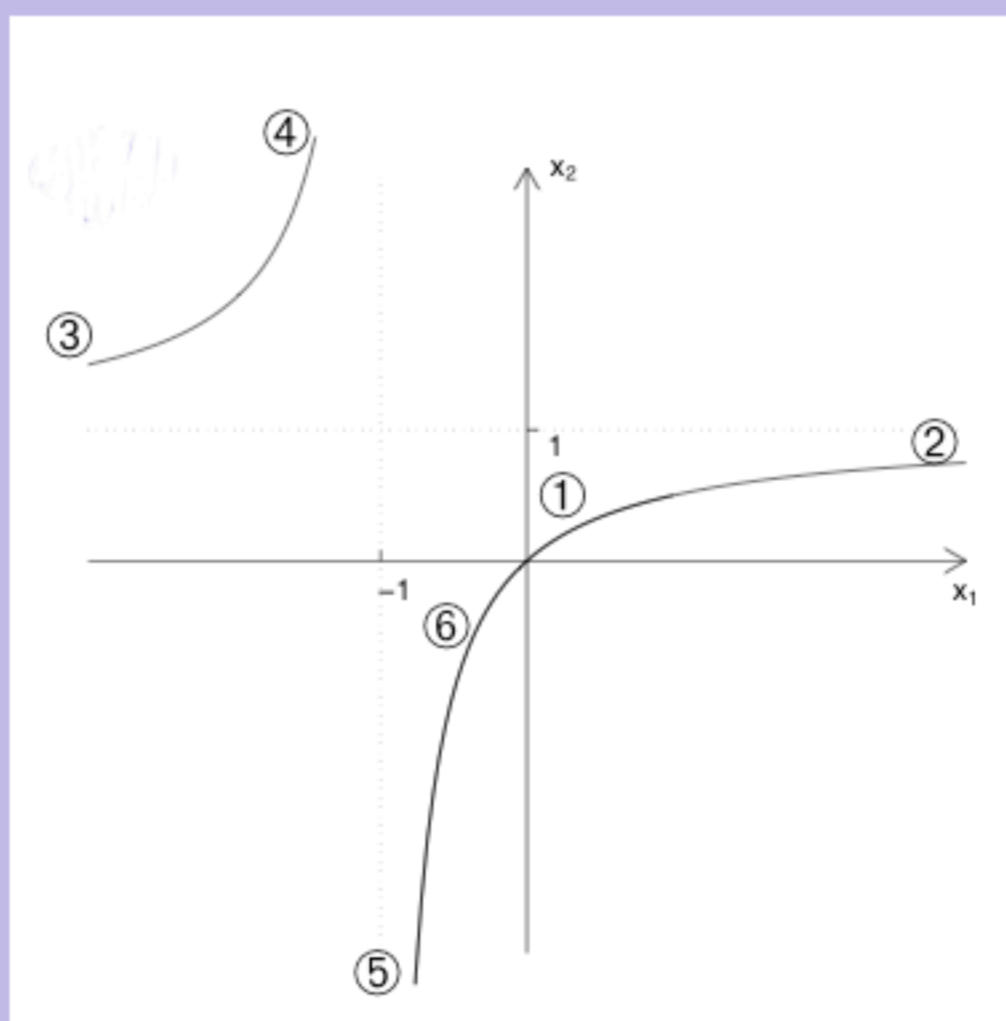
⑤→⑥ III

?

Contuity ?

Data near a boundary?





Did Fisher mention continuity for ancillaries?

I don't think so!

Would he have considered a discontinuous ancillary?

I don't think so!

Vector quantile resolves issue!

f Default priors

$f(y-\theta)$: Bayes (1763) ... all There! Confidence = fiducial = posterior

$$\int^{y^\circ} f(y-\theta)\,dy \equiv \int_\theta f(y^\circ-\alpha)\,d\alpha$$

$$p(\theta) \equiv s(\theta)$$

245 BP

f Default priors

$f(y-\theta)$: Bayes (1763) ... all There!

Confidence = fiducial = posterior

$$\int^{y^\circ} f(y-\theta)\,dy \equiv \int_\theta f(y^\circ-\alpha)\,d\alpha$$

$$p(\theta) \equiv s(\theta)$$

Quantile fn: $y = \theta + z$ $\quad z \sim f(z)$

$dy|_{y^\circ} = d\theta$ $\quad$ fixed p-value.

Integration from sample space to parameter space

f Default priors

$f(y-\theta)$: Bayes (1763) ... all there! Confidence = fiducial = posterior

$$\int^{y^0} f(y-\theta)\,dy \equiv \int_\theta f(y^0-\alpha)\,d\alpha$$

$$p(\theta) \equiv s(\theta)$$

Quantile fn: $y = \theta + z$ $\quad z \sim f(z)$

$dy|_{y^0} = d\theta$ $\quad$ fixed p-value.

Vector Quantile $y = y(u;\theta)$

$dy|_{y_0} = V(\theta)\,d\theta$ $\quad$ fixed "vector p-value"

$|dy| = |V(\theta)|\,d\theta$

[Prior $= |V(\theta)|$ gives corresponding volume for θ $\quad O(n^{-1})$]

f Default priors

$f(y-\theta)$: Bayes (1763) ... all There! Confidence = fiducial = posterior

$$\int^{y^0} f(y-\theta)\,dy \equiv \int_\theta f(y^0-\alpha)\,d\alpha$$

$$p(\theta) \equiv s(\theta)$$

Quantile fn: $y = \theta + z$ $\quad z \sim f(z)$

$dy|_{y^0} = d\theta$ $\quad$ fixed p-value.

Vector Quantile $y = y(u;\theta)$

F Reid
Marras, Y
xxx = 239.pdf

$dy|_{y_0} = V(\theta)\,d\theta$ $\quad$ fixed vector p-value

$dy = |V(\theta)|\,d\theta$

[Prior $|V(\theta)|$ gives corresponding volume for θ $\quad O(n^{-1})$]

Straight "approx. Bayes" & 3rd order accurate

Generalizes: right invariant prior... widely

Finetune: Sample to parameter space $O(n^{-1})$

$$dy = |V(\theta)|\, d\theta$$

$$|d\hat{\theta}| = |W(\theta)|\, d\theta$$

$$W(\theta) = \hat{j}_{\theta\theta}^{-1} \, \mathring{\hat{l}}_{\theta;y} V(\theta)$$

$$\tilde{W}(\theta) = \hat{j}_{\theta\theta}^{-\frac{1}{2}} \, \mathring{\hat{l}}_{\theta;y} V(\theta)$$

(ψ, λ) Targetted on ψ

$$\pi_\psi(\theta)\, d\psi\, d\lambda = |\tilde{W}_{\psi.\lambda}(\hat{\theta}_\psi)|\, d\psi \; |\tilde{W}_\lambda(\theta)|\, d\lambda$$

3rd for ψ

xxx=239.pdf

Data dependent?
The price for
parameter space integration!

h) Summary

a likelihood & p-value

$f(y-\theta)$ y^0 $p(\theta) = s(\theta)$

b likelihood & p-value: approximate / 3rd

Exptl $\{\ell^0(\theta), \varphi(\theta)\}$ $\varphi(\theta) = \frac{\partial}{\partial y} \ell(\theta; y) \big|_{y^0}$

c vector Quantile function

Invert distn fns $\underset{\sim}{y} = \underset{\sim}{y}(\underset{\sim}{u}; \theta)$

d How to measure θ?

Marginal or Conditional

e How does θ move data? Ancillary (2nd)

In direction $V(\hat{\theta}^0)$ $\frac{dy}{d\theta}\big|_{y^0 \hat{\theta}^0}$

f Two examples

Nonlinear regression

Cauchy

g Default priors

$\pi(\theta) = \left|\frac{dy}{d\theta}\right|_{y^0 \hat{\theta}^0}$

Quantile presents continuity

Continuity gives "definition"... widely

Thank you