Continuity
and Statistical Methodology

D A S Fraser
Statistics
U Toronto

2nd Princeton Day of Statistics
Princeton
New Jersey

October 22 2010

http://www.utstat.toronto.edu/dfraser/documents/pu2010.pdf
... / xxx.pdf

1. Politics, Census, Accuracy 19/20

2. fermilab LHC Accuracy

3. Continuity: ... Scalar

4. ... Scalar & vectors

5. What does it do for you?

6. Example

A ⇒ Conditioning on data information

B ⇒ Default priors : Nec. & Suff conditions

C ⇒ Bayes & parameter linearity

10. Curvature & Directions

1. Politics Census Accuracy 19/20

Politics:

Statistics Canada.... "Compulsory Long Form"

downgraded to a survey!

1. Politics Census Accuracy 19/20

Politics:

Statistics Canada.... "Compulsory Long Form"

downgraded to a survey!

Public, business, media, statisticians.....

Concerns: ... Accuracy .....

Chief Statistician resigned ---

JSM: Canada UK Australia

1. Politics Census Accuracy 19/20

Politics:

Statistics Canada.... "Compulsory Long Form"

downgraded to a survey!

Public, business, media, statisticians.....

Concerns: ... Accuracy .....

Chief Statistician resigned ---

JSM: Canada UK Australia

Surprising: Extreme: concern for Accuracy

that 19/20 should mean something!

1. Politics, Census, Accuracy 19/20

Politics:

Statistics Canada.... "Compulsory Long Form"

downgraded to a survey!

Public, business, media, statisticians.....

Concerns: ... Accuracy .....

Chief Statistician resigned - - -

JSM: Canada UK Australia

Surprising: Extreme: concern for Accuracy

that 19/20 should mean something!

that ·000013 should mean something!

1. Politics, Census, Accuracy 19/20

Politics:

Statistics Canada.... "Compulsory Long Form"

downgraded to a survey!

Public, business, media, statisticians.....

Concerns: ... Accuracy .....

Chief Statistician resigned ---

JSM: Canada UK Australia

Surprising: Extreme: concern for Accuracy

that 19/20 should mean something!

that ·000013 should mean something!

That reported results --- not be just casual "exploring"!

2. fermilab LHC Accuracy

1994 Evidence for top quark HEP ... Discussion by Sineruo

Poisson (θ) $\theta \geq \theta_0$ representing background radiation

"New particle" $\longleftrightarrow$ $\theta > \theta_0$

abe et al 1994 Phys Rev Lett 73 .... Sinerua + '200'

2. fermilab LHC Accuracy

1994 Evidence for top quark HEP ... Discussion by Sineruo

Poisson (θ) $\theta \geq \theta_0$ representing background radiation

"New particle" $\longleftrightarrow$ $\theta > \theta_0$

High Energy Physicists ... Kendall & Stuart

2-sided Conf Intervals

... use of Lik ratio criteria ... moved interals away from θ_0 !

abe et al 1994 Phys Rev Lett 73 .... Sinerva + '200'

Mandelkern 2002 Stat Sc

2. fermilab LHC Accuracy

1994 Evidence for top quark HEP ... Discussion by Sinervo

Poisson (θ) $\theta \geq \theta_0$ representing background radiation

"New particle" $\longleftrightarrow$ $\theta > \theta_0$

High Energy Physicists ... Kendall & Stuart

2-sided Conf Intervals

... use of Lik ratio criteria ... moved intervals away from θ_0 !

Resolution ...: Use "confidence lower quantiles"

abe et al 1994 Phys Rev Lett 73 .... Sinerva + '200'
Mandelkern 2002 Stat Sc
F Reid Wong 2004 Phys Rev D

2. fermilab LHC Accuracy

1994 Evidence for top quark HEP ... Discussion by Sinervo

Poisson (θ) $\theta \geq \theta_0$ representing background radiation

"New particle" $\longleftrightarrow$ $\theta > \theta_0$

High Energy Physicists ... Kendall & Stuart

2-sided Conf Intervals

... use of Lik ratio criteria ... moved intervals away from θ_0!

Resolution...: Use "confidence lower quantiles"

BIRS workshops Davison Sartori 2009 +others

Planned: 2011

abe et al 1994 Phys Rev Lett 73 ... Sinervo + 200

Mandelkern 2002 Stat Sc

F Reid Wong 2004 Phys Rev D

2. fermilab LHC Accuracy

1994 Evidence for top quark HEP ... Discussion by Sinervo

Poisson (θ) $\theta \geq \theta_0$ representing background radiation

"New particle" $\longleftrightarrow$ $\theta > \theta_0$

High Energy Physicists ... Kendall & Stuart

2-sided Conf Intervals

... use of Lik ratio criteria ... moved interals away from θ_0 !

Resolution...: Use "confidence lower quantiles"

BIRS workshops Davison Sartori 2009 +others

Planned: 2011

Why? $\dfrac{\text{Accuracy}}{\text{Quantiles}}$ Want Discover / Verify New particle

abe et al 1994 Phys Rev Lett 73 .... Sinerva + '200'

Mandelkern 2002 Stat Sc

F Reid Wong 2004 Phys Rev D

3. Continuity:

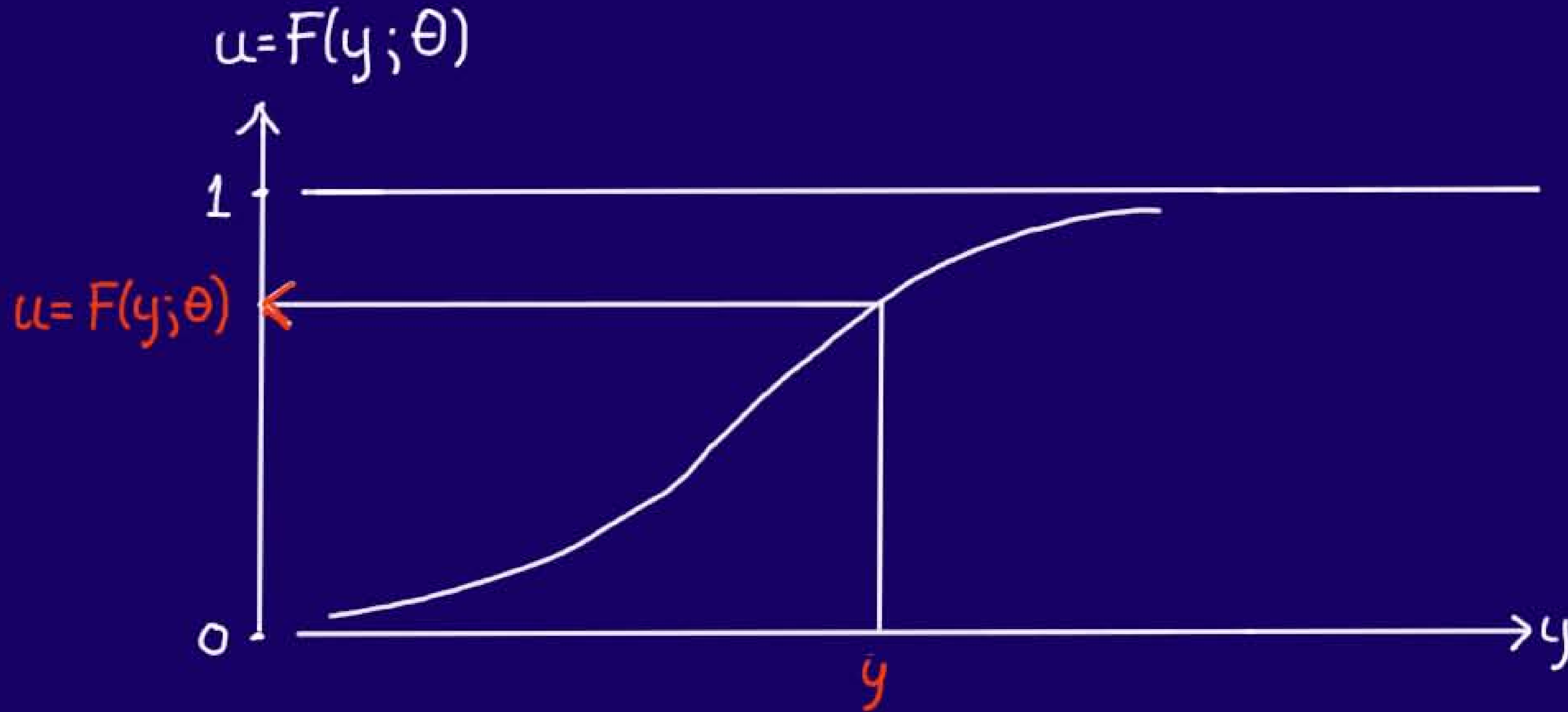
Case: Scalar

Distn fn Continuous

3. Continuity:

Case: Scalar

Distn fn Continuous

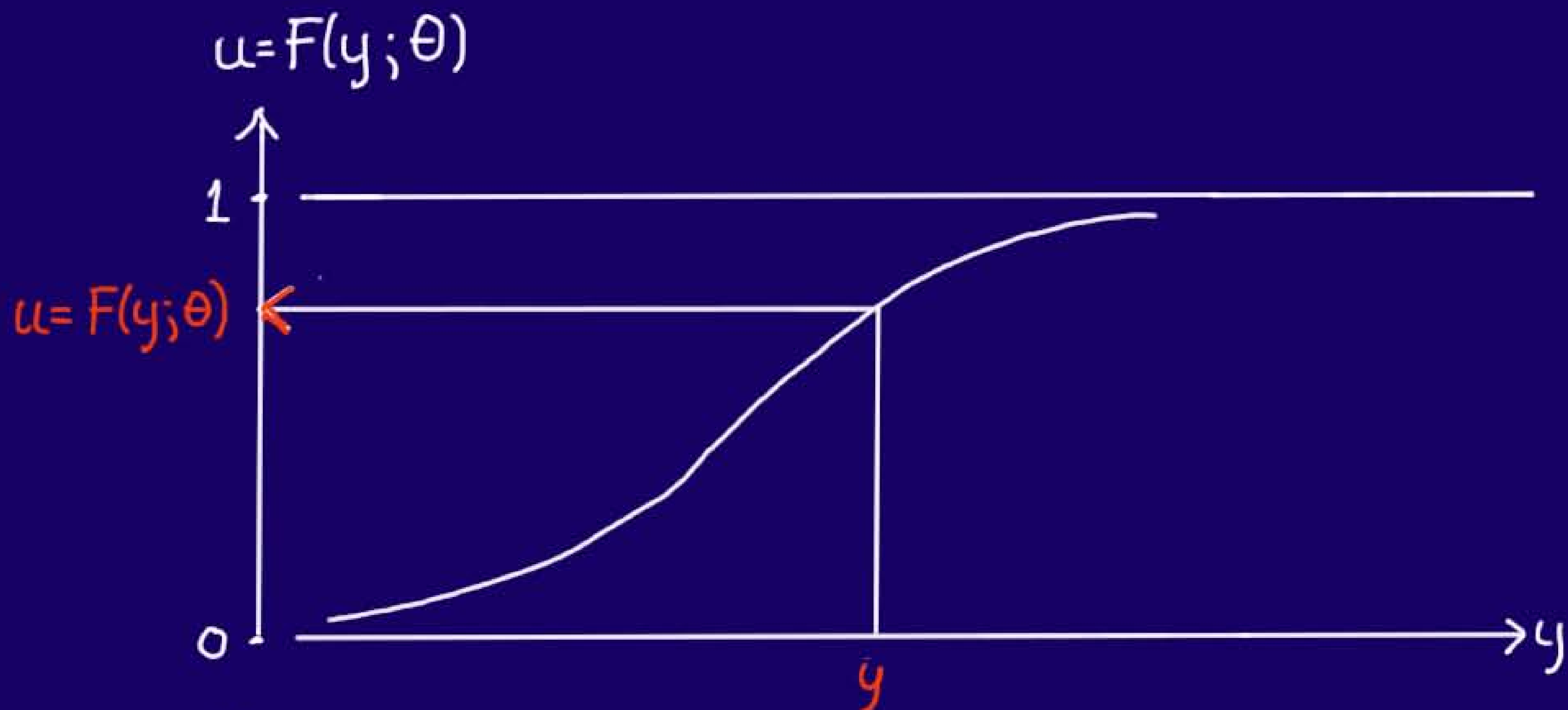


p-value dist'n fn

3. Continuity:

Case: Scalar

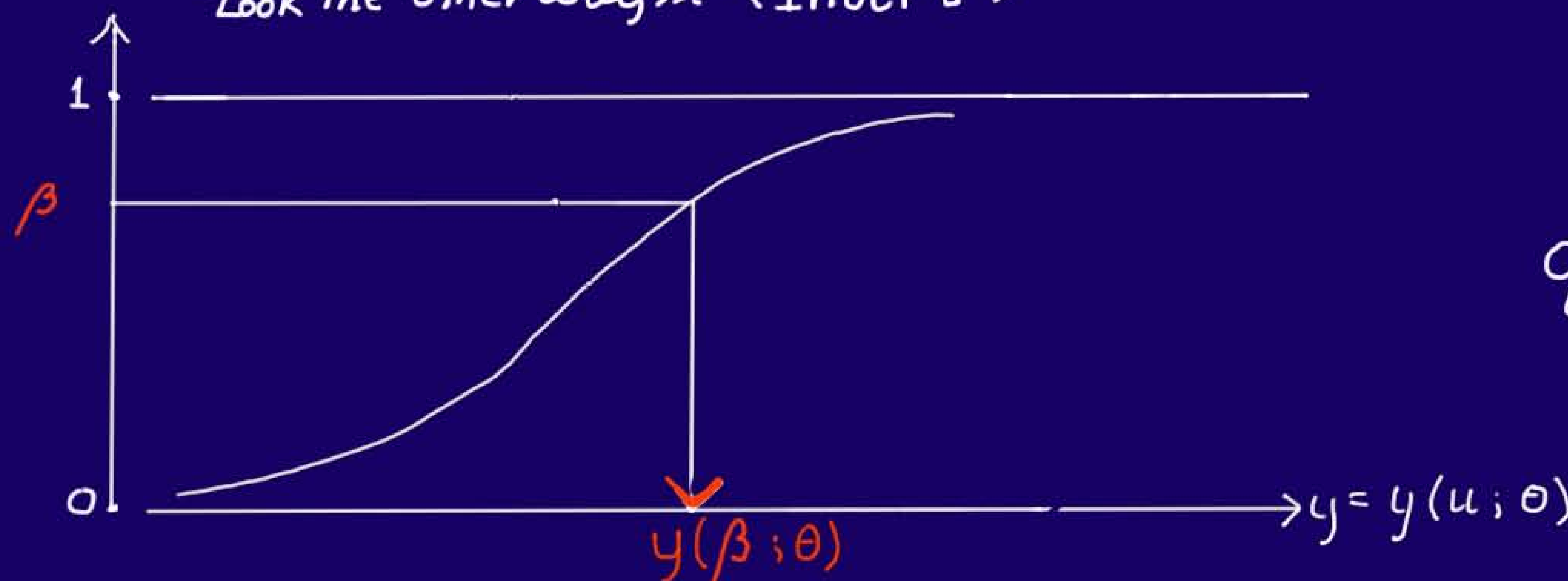
Distn fn Continuous



p-value dist'n fn

u

"Look the other way..." <Invert>

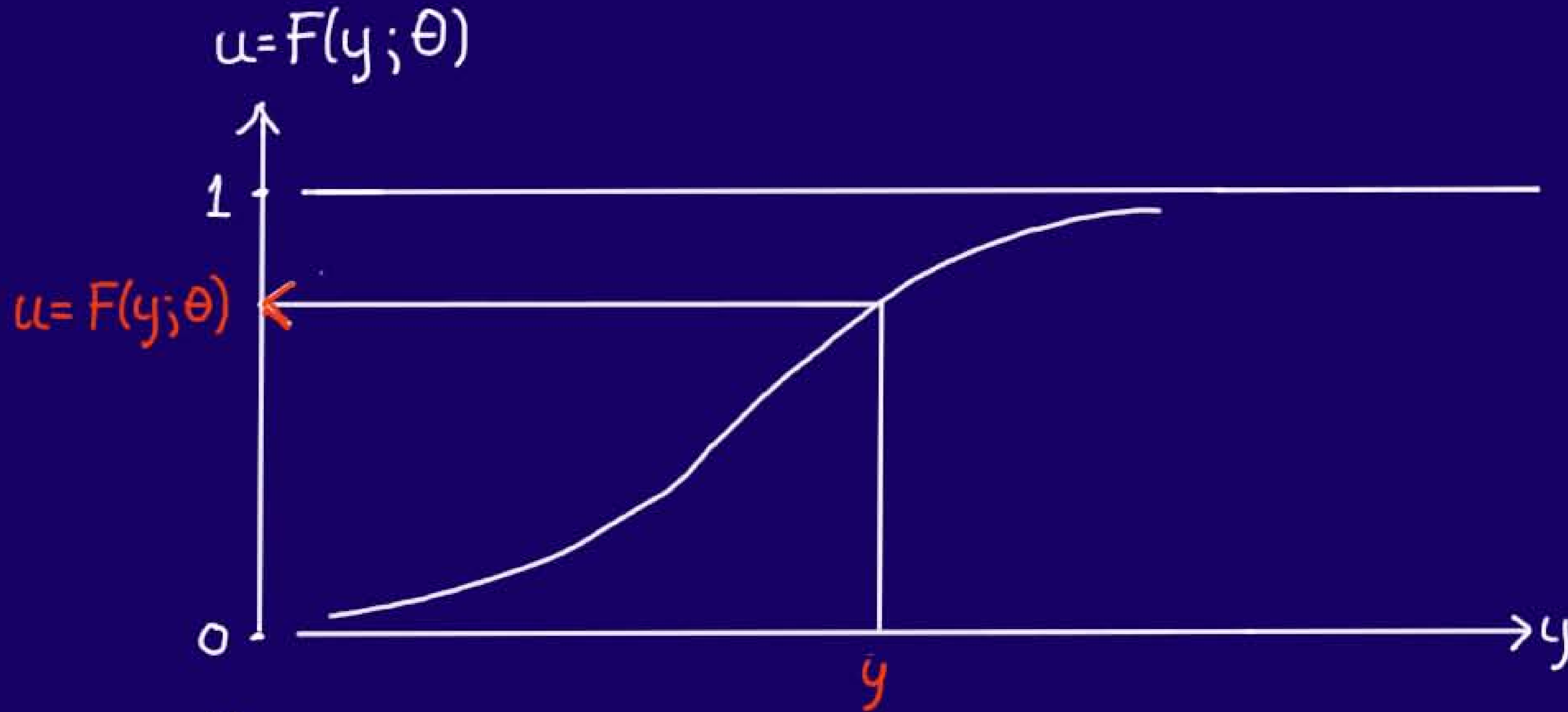


q-value quantile

3. Continuity:

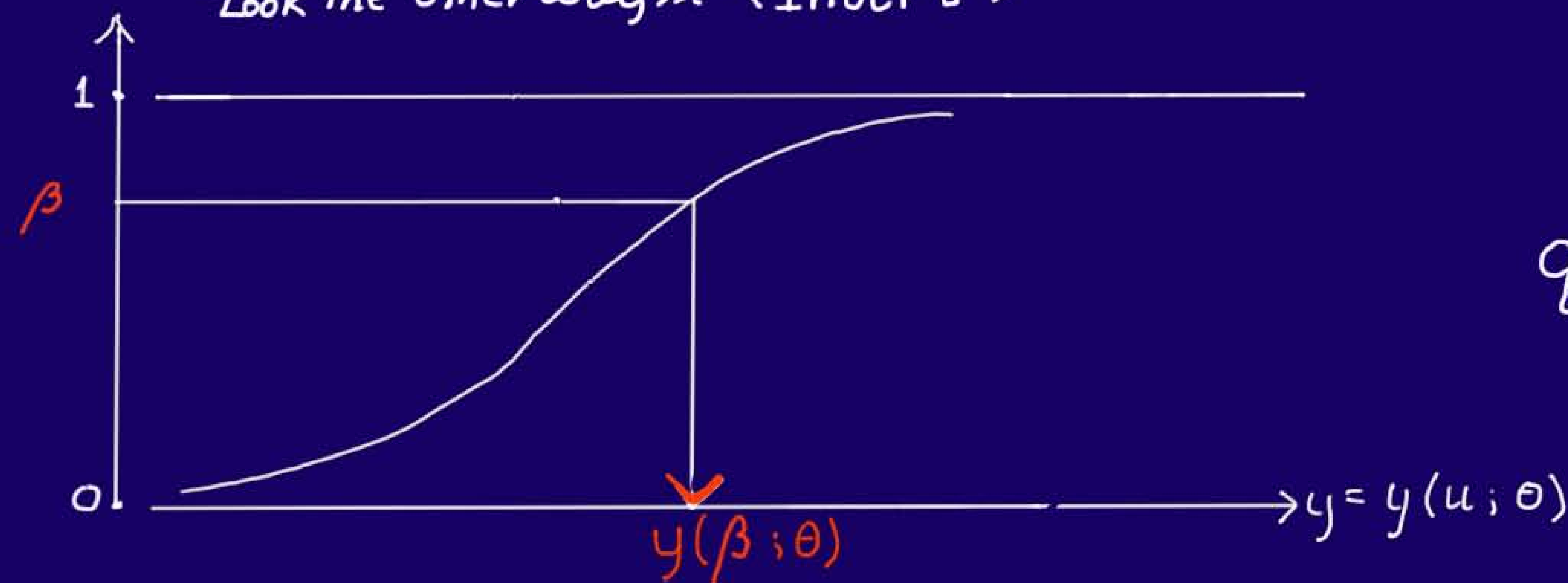
Case: <u>Scalar</u>

Distn fn Continuous



u

"Look the other way..." <Invert>



p-value dist'n fn

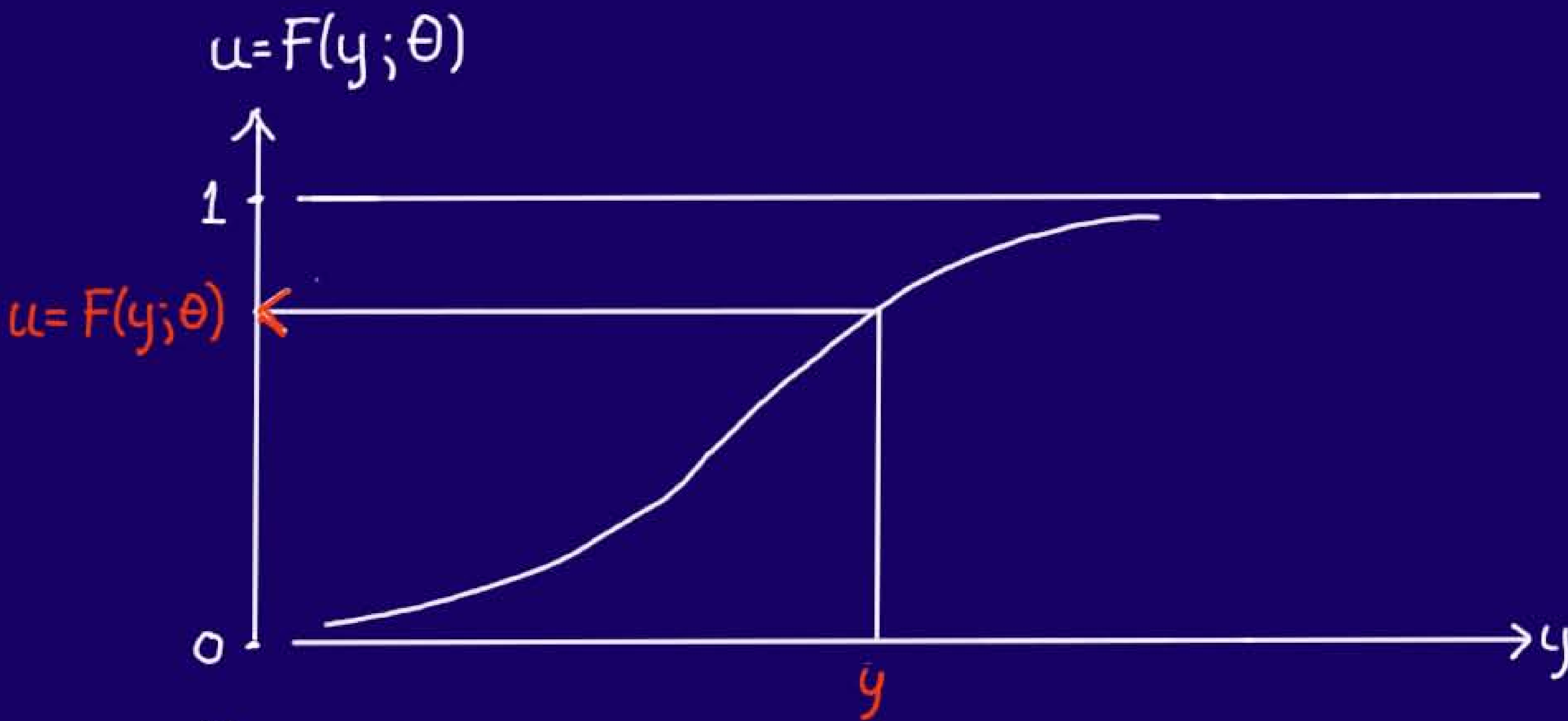
p-q or q-p plots

q-value quantile

3. Continuity:

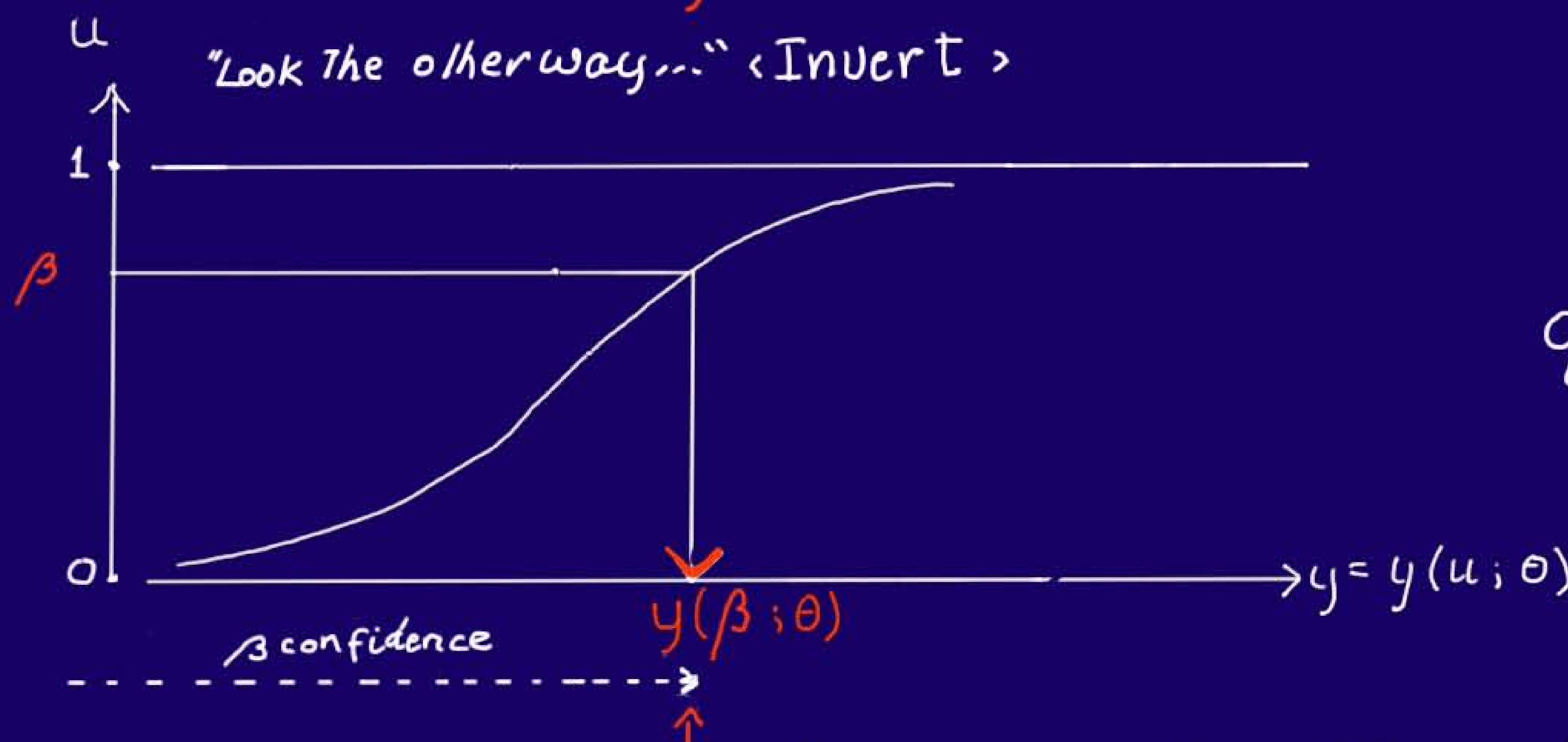
Case: Scalar

Distn fn Continuous



p-value   dist'n fn

"Look the other way..." <Invert>



q-value ← quantile

β-confidence upper quantile

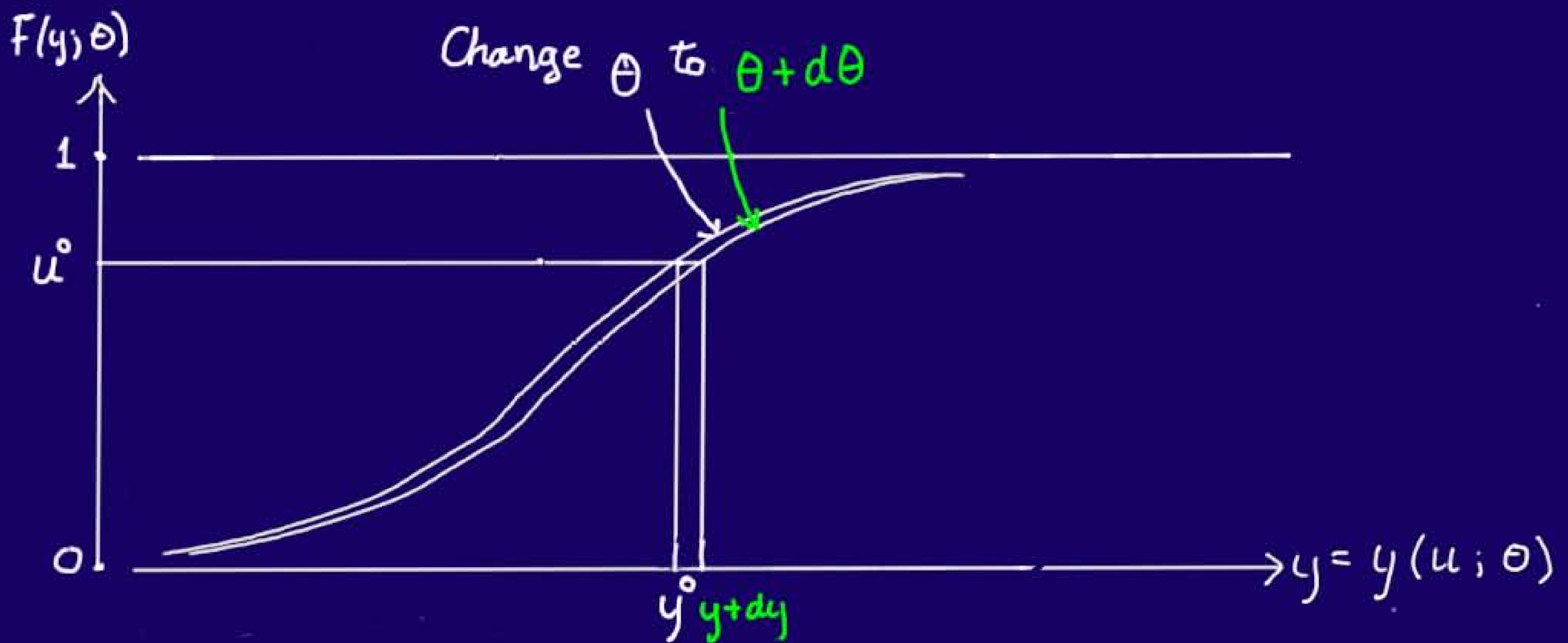
Quantiles ... of the essense

4. Continuity ... more

Case: Scalar and vector

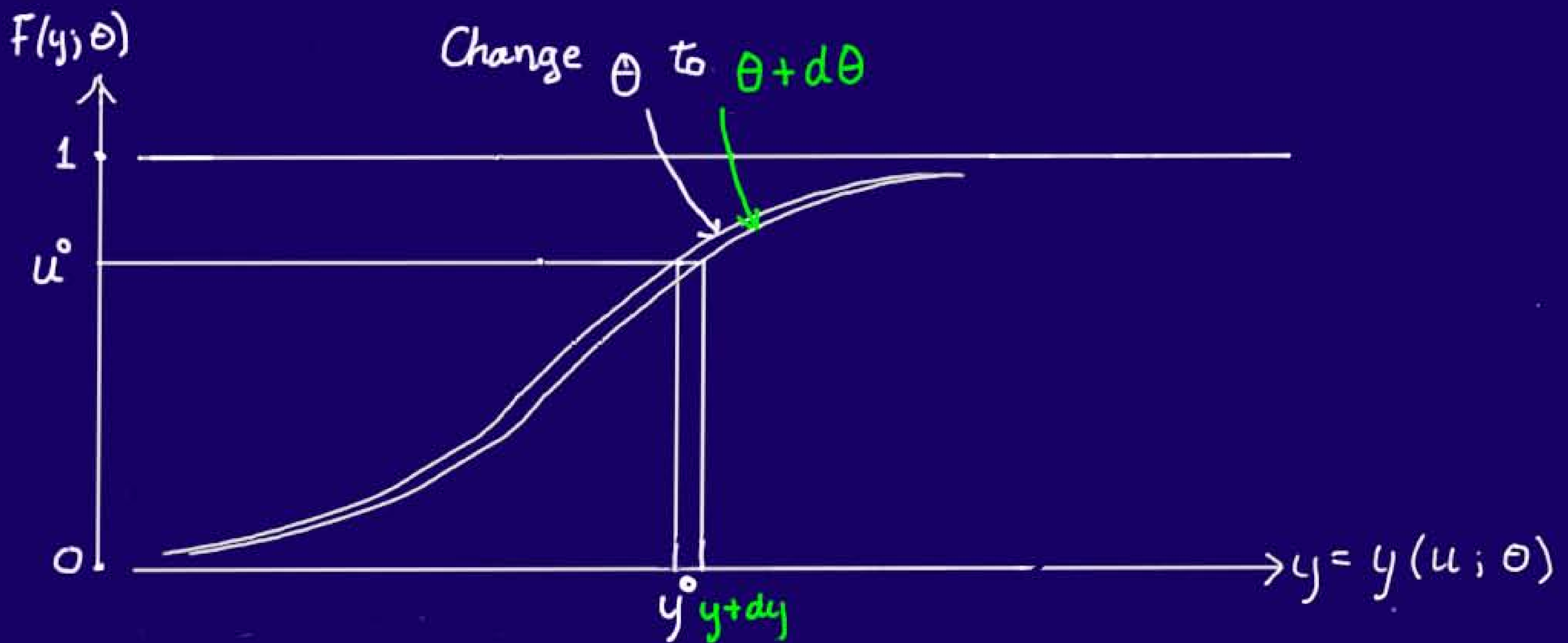
4. Continuity

Case: <u>Scalar</u> and <u>vector</u>



4. Continuity

Case: <u>Scalar</u> and <u>Vector</u>

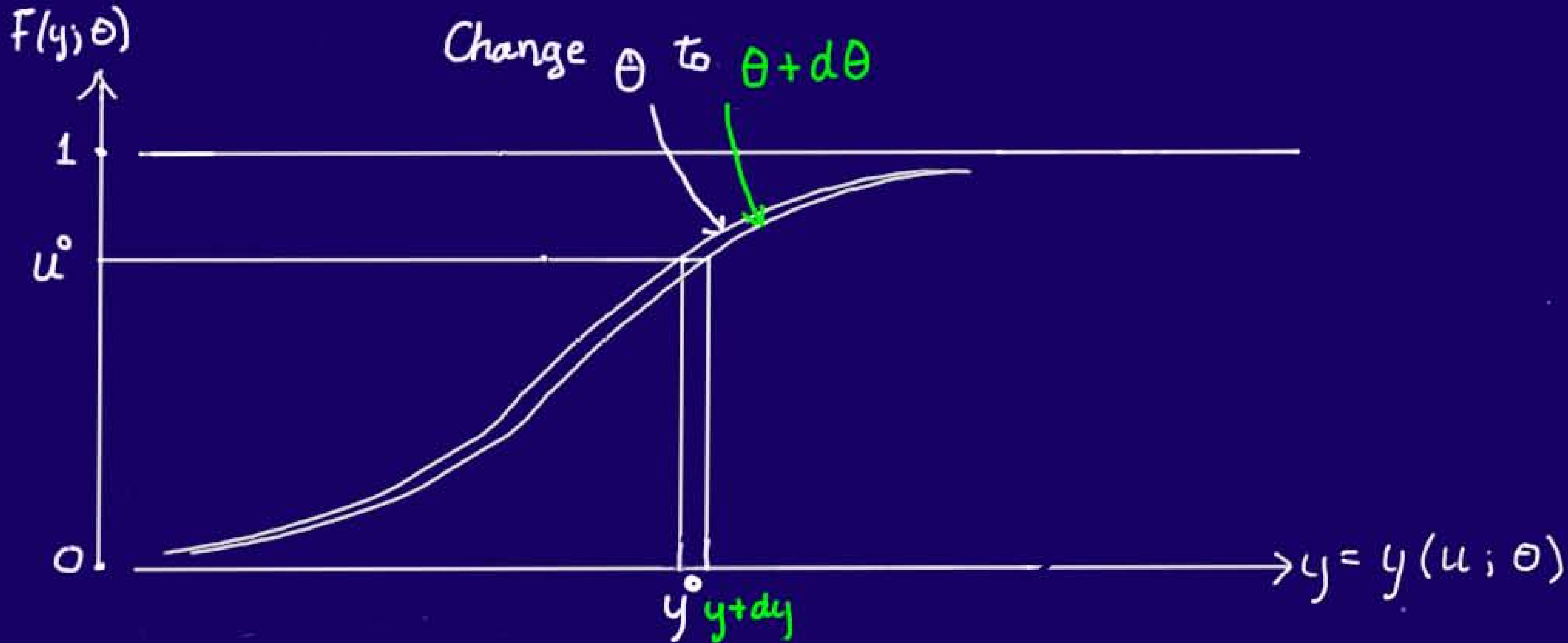


Fixed p-value u°
Change $\theta \to \theta + d\theta$

$$dy = \left.\frac{\partial y}{d\theta}\right|_{y^\circ} d\theta = v(\theta)\, d\theta$$

4. Continuity

Case: Scalar and vector



Fixed p-value u°
Change $\theta \to \theta + d\theta$

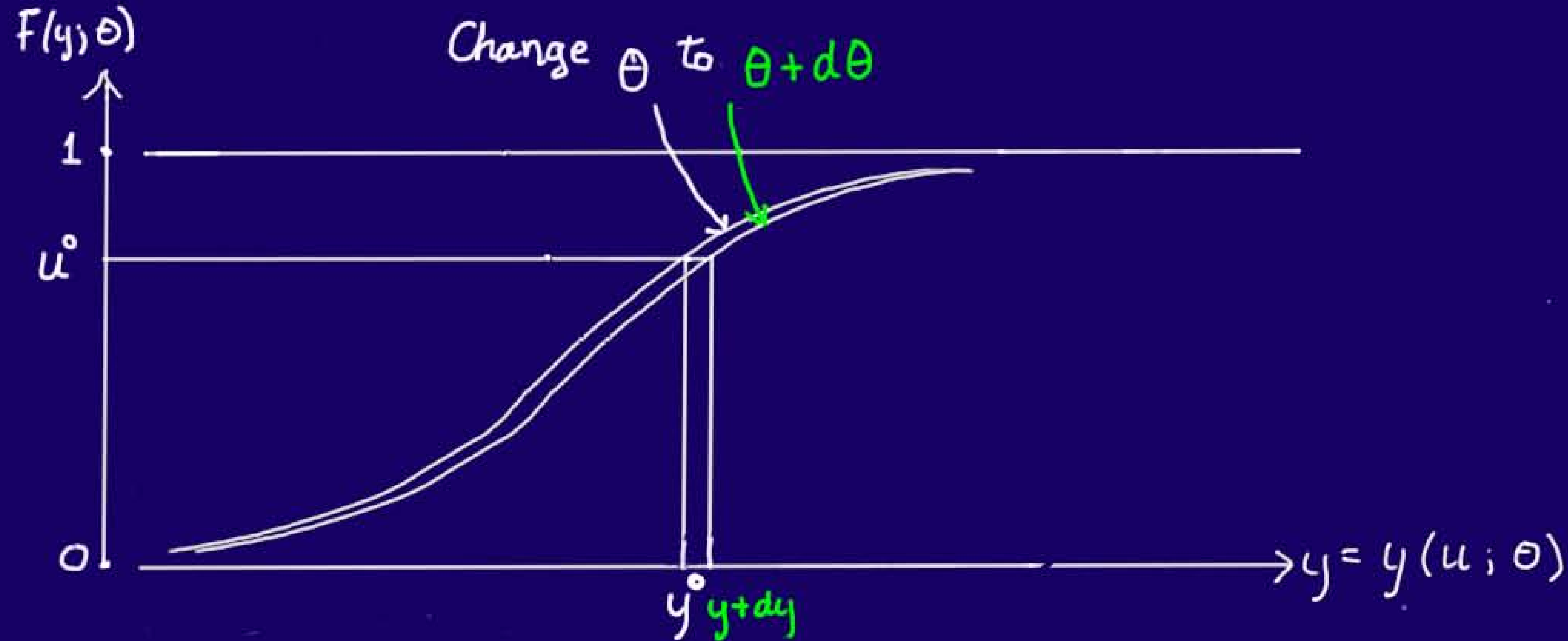
$$dy = \frac{\partial y}{d\theta}\bigg|_{y^\circ} d\theta = \underline{\underline{v(\theta)}}\, d\theta$$

Call it ?

$v(\theta) =$ velocity ?
sensitivity ?
"How θ moves data ?"

4. Continuity

Case: <u>Scalar</u> and <u>vector</u>



Fixed p-value u°

Change $\theta \to \theta + d\theta$

$$dy = \left.\frac{\partial y}{d\theta}\right|_{y^\circ} d\theta = \underline{\underline{v(\theta)}}\, d\theta$$

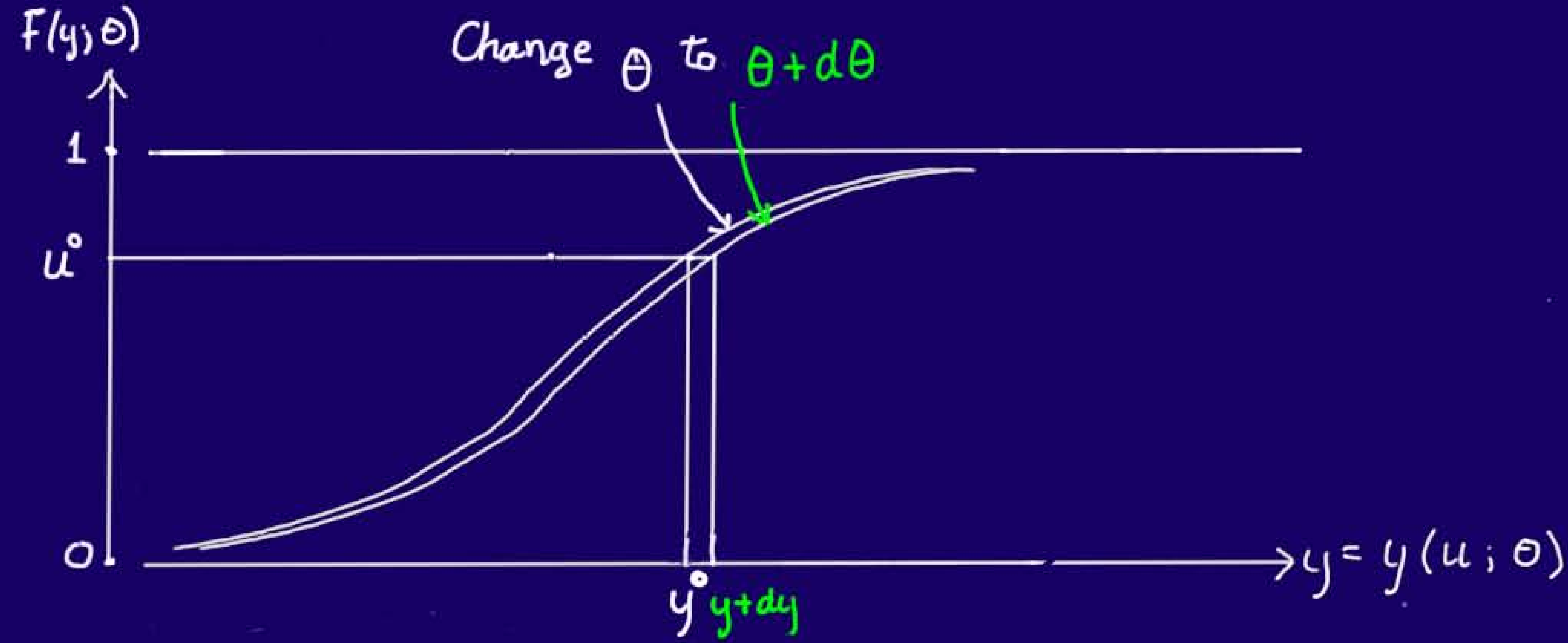
Call it ?

$v(\theta) =$ velocity ? sensitivity ? "How θ moves data ?"

"Unit <u>probability</u> <u>mass</u>" How does it move? A consequence of continuity!

4. Continuity

Case: <u>Scalar</u> and <u>vector</u>



Fixed p-value u°
Change $\theta \to \theta + d\theta$

$$dy = \frac{\partial y}{d\theta}\Big|_{y^\circ} d\theta = \underline{\underline{v(\theta)}}\, d\theta$$

Call it ? $v(\theta) =$ velocity ? sensitivity ? "How θ moves data ?"

"Unit <u>probability</u> <u>mass</u>" How does it move? A consequence of continuity!

<u>Now R^n</u>

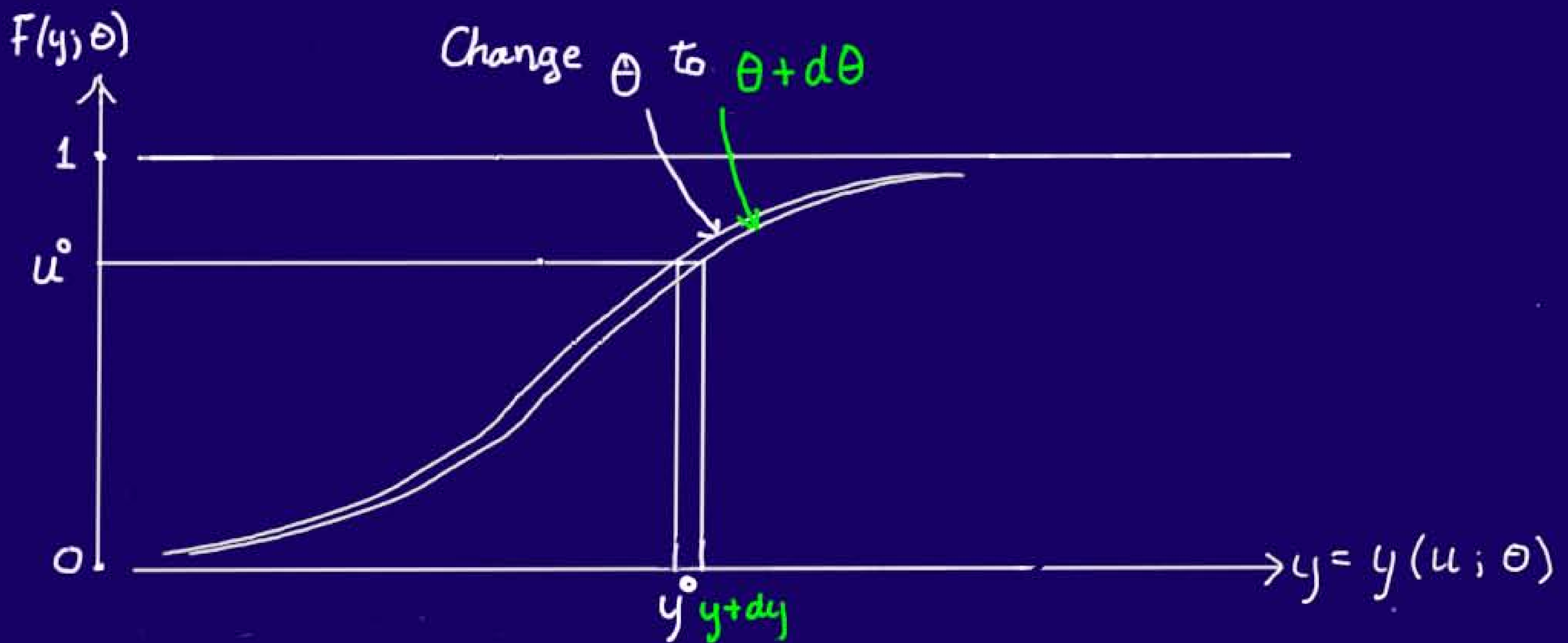
$f(y;\theta)$ dim $y = n$
dim $\theta = p$
Indep coord's say

Data y°

y°

4. Continuity

Case: Scalar and vector



Fixed p-value u°
Change $\theta \to \theta + d\theta$

$$dy = \frac{\partial y}{d\theta}\Big|_{y^\circ} d\theta = \underline{\underline{v(\theta)\, d\theta}}$$

Call it? $v(\theta)$ = velocity? sensitivity? "How θ moves data?"

"Unit probability mass" How does it move? A consequence of continuity!

R^n $f(y;\theta)$ $\dim y = n$
$\dim \theta = p$
Indep coord's say

Change θ: $\theta \to \theta + d\theta$
$y^\circ + dy$

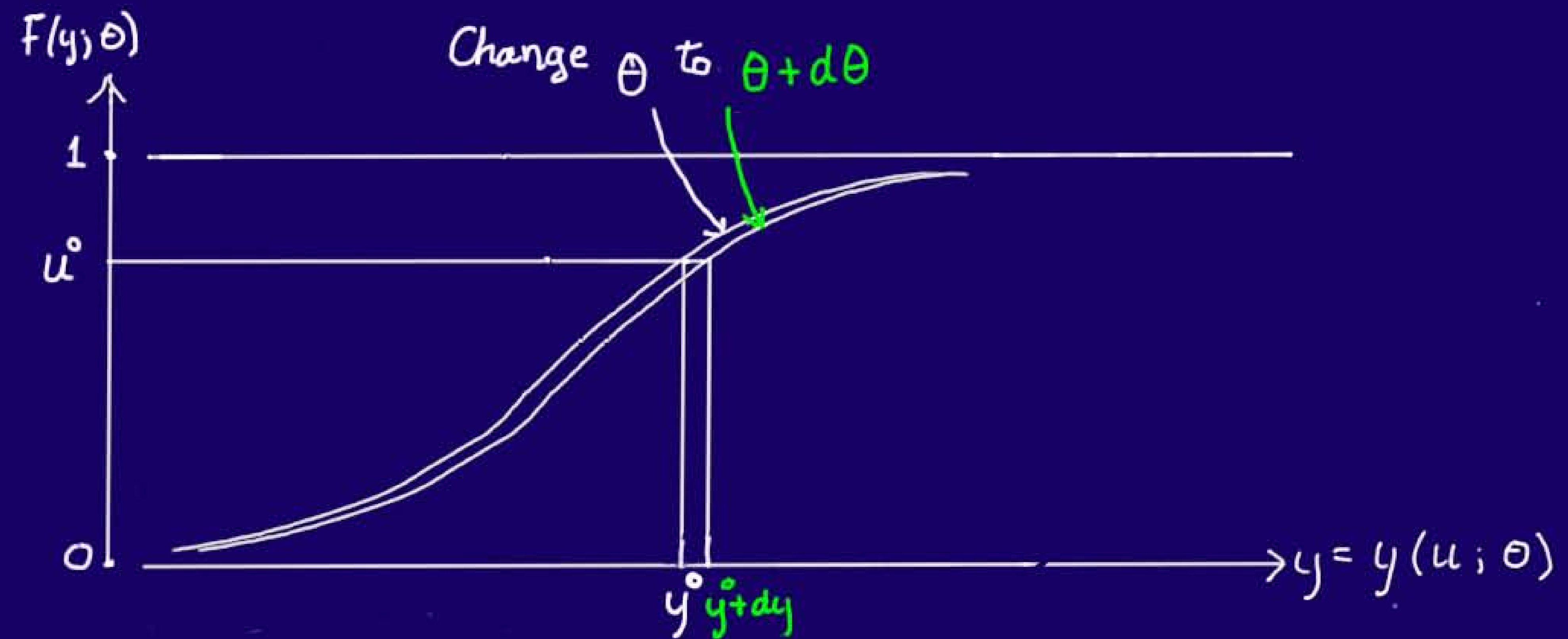
$$\frac{dy}{d\theta} = \underset{n \times p}{V(\theta)} = (\underset{\sim}{v}_1(\theta) \cdots \underset{\sim}{v}_p(\theta))$$

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

4. Continuity

Case: <u>Scalar</u> and <u>vector</u>

!



Fixed p-value u°

Change $\theta \to \theta + d\theta$

$$dy = \left.\frac{\partial y}{d\theta}\right|_{y^\circ} d\theta = v(\theta)\, d\theta$$

R^n $\quad f(y;\theta)$ $\quad \dim y = n$

$\dim \theta = p$

Indep coord's say

Change $\theta: \theta \to \theta + d\theta$

$\nearrow y^\circ + dy$

$$\frac{dy}{d\theta} = \underset{n \times p}{V(\theta)} = (\underset{\sim}{v}_1(\theta) \cdots \underset{\sim}{v}_p(\theta))$$

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

"Ultimate consequence of continuity! $\quad V(\theta) \quad V = V(\hat{\theta}^\circ)$

What's it do for you?

$y^0 \to y^0 + dy$

$$\underset{n\times 1}{dy} = \underset{n\times p}{V(\theta)}\,\underset{p\times 1}{d\theta}$$

What's it do for you ?

$y^0 \to y^0 + dy$

$$\underset{n\times 1}{dy} = \underset{n\times p}{V(\theta)}\ \underset{p\times 1}{d\theta}$$

A Determines Accuracy for inference: $O(n^{-3/2})$

Fraser F Slaicu (2010) Bernoulli 1208-1223

What's it do for you?

$y^0 \to y^0 + dy$

$$\underset{n\times 1}{dy} = \underset{n\times p}{V(\theta)}\,\underset{p\times 1}{d\theta}$$

A Determines Accuracy for inference: $O(n^{-3/2})$

Gives Highly-Accurate SP-type tests of $\psi(\theta)=\psi$ $O(n^{-3/2})$
(cgf-type)

Fraser F Slaicu (2010) Bernoulli 1208-1223
Brazzale Davison Reid (2007) Applied Asymptotics Cambridge

What's it do for you?

$y^0 \to y^0 + dy$

$$\underset{n\times 1}{dy} = \underset{n\times p}{V(\theta)}\ \underset{p\times 1}{d\theta}$$

A Determines Accuracy for inference: $O(n^{-3/2})$

Gives Highly-Accurate SP-type tests of $\psi(\theta) = \psi$ $O(n^{-3/2})$
(cgf-type)

B Nec. & Suff. conditions for Default priors $O(n^{-1})$
(Accurate Bayes)

Fraser F Slaicu (2010) Bernoulli 1208-1223
Brazzale Davison Reid (2007) Applied Asymptotics Cambridge
F Reid Marras Yi (2010) JRRSB 631-654

What's it do for you ?

$y^0 \to y^0 + dy$

$$\underset{n\times 1}{dy} = \underset{n\times p}{V(\theta)}\, \underset{p\times 1}{d\theta}$$

A Determines Accuracy for inference: $O(n^{-3/2})$

Gives Highly-Accurate SP-type tests of $\psi(\theta)=\psi$ $O(n^{-3/2})$
(cgf-type)

B Nec. & Suff. conditions for Default priors $O(n^{-1})$
(Accurate Bayes)

C Linear parameters re Bayes & Curved

Fraser F Slaicu (2010) Bernoulli 1208-1223
Brazzale Davison Reid (2007) Applied Asymptotics Cambridge
F Reid Marras Yi (2010) JRRSB 631-654
Fraser F Fraser (2010) J. Statist Res Efron vol. 44 to appear

An example:

$$y^0 \to y^0 + dy$$

$$\underset{n\times 1}{dy} = \underset{n\times p}{V(\theta)}\ \underset{p\times 1}{d\theta}$$

An example:

$y^0 \rightarrow y^0 + dy$

$$\underset{n\times 1}{dy} = \underset{n\times p}{V(\theta)}\ \underset{p\times 1}{d\theta}$$

Ex: Regression $y = X\beta + \sigma Z$ say Non-N error $f_i(z_i)$

An example:

$y^0 \to y^0 + dy$

$$\underset{n\times 1}{dy} = \underset{n\times p}{V(\theta)}\ \underset{p\times 1}{d\theta}$$

Differentiate quantile fn

Ex: Regression $y = X\beta + \sigma z$ say Non-N error $f_i(z_i)$

at y^0

$$\underset{n\times 1}{dy} = \Big(\underset{n\times k}{X}\quad \underset{n\times 1}{z^0(\beta,\sigma)}\Big)\underset{(k+1)\times 1}{\begin{pmatrix} d\beta \\ d\sigma \end{pmatrix}}$$

F Reid Marras Yi (2010) JRRSB 631-654

An example:

$y^0 \to y^0 + dy$

$$\underset{n\times 1}{dy} = \underset{n\times p}{V(\theta)}\;\underset{p\times 1}{d\theta}$$

Differentiate quantile fn

Ex: Regression $y = X\beta + \sigma z$ say Non-N error $f_i(z_i)$

at y^0 $$\underset{n\times 1}{dy} = \Big(\underset{n\times p}{X} \quad \underset{n\times 1}{z^0(\beta,\sigma)}\Big)\underset{(p+1)\times 1}{\begin{pmatrix} d\beta \\ d\sigma \end{pmatrix}}$$

$z^0(\beta\sigma) = (y^0 - X\beta)/\sigma$ = Data residual re β σ

F Reid Marras Yi (2010) JRRSB 631-654

An example:

$y^0 \to y^0 + dy$

$$\underset{n\times 1}{dy} = \underset{n\times p}{V(\theta)}\ \underset{p\times 1}{d\theta}$$

Differentiate quantile fn

Ex: Regression $y = X\beta + \sigma z$ say Non-N error $f_i(z_i)$

at y^0
$$\underset{n\times 1}{dy} = \left(\underset{n\times p}{X} \quad \underset{n\times 1}{z^0(\beta,\sigma)}\right)\underset{(p+1)\times 1}{\begin{pmatrix} d\beta \\ d\sigma \end{pmatrix}}$$

$z^0(\beta\sigma) = (y^0 - X\beta)/\sigma$ = Data residual re β σ

$\Phi(\downarrow)$ = p-value vector re β, σ

F Reid Marras Yi (2010) JRRSB 631-654

An example:

$y^0 \to y^0 + dy$

$$\underset{n\times 1}{dy} = \underset{n\times p}{V(\theta)}\, \underset{p\times 1}{d\theta}$$

Differentiate quantile fn

Ex: Regression $y = X\beta + \sigma z$ say Non-N error $f_i(z_i)$

at y^0 $\underset{n\times 1}{dy} = \left(\underset{n\times r}{X} \quad \underset{n\times 1}{z^0(\beta,\sigma)} \right) \underset{(r+1)\times 1}{\begin{pmatrix} d\beta \\ d\sigma \end{pmatrix}}$

$z^0(\beta\sigma) = (y^0 - X\beta)/\sigma$ = Data residual re β σ

$\Phi(\downarrow)$ = p-value vector re β, σ

Can put in terms of $d\hat\beta$, $d\hat\sigma$ at Observed ...

at y^0 $d\hat\beta$ =

$d\hat\sigma$ =

F Reid Marras Yi (2010) JRRSB 631-654

An example:

$y^0 \to y^0 + dy$

$$\underset{n\times 1}{dy} = \underset{n\times p}{V(\theta)}\ \underset{p\times 1}{d\theta}$$

Differentiate quantile fn

Ex: Regression $\quad y = X\beta + \sigma z \qquad$ say Non-N error $f_i(z_i)$

at y^0 $\quad \underset{n\times 1}{dy} = \left(\underset{n\times p}{X} \quad \underset{n\times 1}{z^0(\beta,\sigma)}\right)\underset{(p+1)\times 1}{\begin{pmatrix} d\beta \\ d\sigma \end{pmatrix}}$

$z^0(\beta\sigma) = (y^0 - X\beta)/\sigma$ = Data residual re β σ

$\Phi(\ \downarrow\ )$ = p-value vector re β, σ

Can put in terms of $d\hat\beta, d\hat\sigma$ at Observed ...

at y^0
$$d\hat\beta = d\beta + (\hat\beta^0 - \beta)\, d\sigma/\sigma$$
$$d\hat\sigma = \qquad \hat\sigma\, d\sigma/\sigma$$

F Reid Marras Yi (2010) JRRSB 631-654

An example:

$y^0 \to y^0 + dy$

$$\underset{n\times 1}{dy} = \underset{n\times p}{V(\theta)}\, \underset{p\times 1}{d\theta}$$

Differentiate quantile fn

Ex: Regression $y = X\beta + \sigma z$ say Non-N error $f_i(z_i)$

at y^0 $$\underset{n\times 1}{dy} = \left(\underset{n\times r}{X} \quad \underset{n\times 1}{z^0(\beta,\sigma)}\right) \underset{(r+1)\times 1}{\begin{pmatrix} d\beta \\ d\sigma \end{pmatrix}}$$

$z^0(\beta\sigma) = (y^0 - X\beta)/\sigma$ = Data residual re β σ

$\Phi(\downarrow)$ = p-value vector re β, σ

Can put in terms of $d\hat{\beta}$, $d\hat{\sigma}$ at Observed...

at y^0 $$d\hat{\beta} = d\beta + (\hat{\beta}^0 - \beta)\, d\sigma/\sigma$$
$$d\hat{\sigma} = \hat{\sigma}\, d\sigma/\sigma$$

$$\begin{pmatrix} d\hat{\beta} \\ d\hat{\sigma} \end{pmatrix} = \begin{bmatrix} I & (\hat{\beta}^0 - \beta)/\sigma \\ 0 & \hat{\sigma}/\sigma \end{bmatrix} \begin{pmatrix} d\beta \\ d\sigma \end{pmatrix}$$

F Reid Marras Yi (2010) JRRSB 631-654

An example:

$y^0 \to y^0 + dy$

$$\underset{n\times 1}{dy} = \underset{n\times p}{V(\theta)}\,\underset{p\times 1}{d\theta}$$

Differentiate quantile fn

Ex: Regression $y = X\beta + \sigma z$ say Non-N error $f_i(z_i)$

at y^0 $\underset{n\times 1}{dy} = \left(\underset{n\times r}{X}\ \ \underset{n\times 1}{z^0(\beta,\sigma)}\right)\underset{(r+1)\times 1}{\begin{pmatrix} d\beta \\ d\sigma \end{pmatrix}}$

$z^0(\beta\sigma) = (y^0 - X\beta)/\sigma$ = Data residual re $\beta\ \sigma$

$\Phi(\downarrow)$ = p-value vector re β, σ

Can put in terms of $d\hat\beta, d\hat\sigma$ at Observed...

at y^0

$$d\hat\beta = d\beta + (\hat\beta^0 - \beta)\,d\sigma/\sigma$$
$$d\hat\sigma = \hat\sigma\,d\sigma/\sigma$$

$$\begin{pmatrix} d\hat\beta \\ d\hat\sigma \end{pmatrix} = \begin{bmatrix} I & (\hat\beta^0 - \beta)/\sigma \\ 0 & \hat\sigma/\sigma \end{bmatrix}\begin{pmatrix} d\beta \\ d\sigma \end{pmatrix}$$

Absolute value $|d\hat\theta| = \left|\ \downarrow\ \right| d\theta = \underset{\text{prior}}{\pi(\theta)}\,d\theta = \dfrac{\hat\sigma}{\sigma}\,d\beta\,d\sigma$

F Reid Marras Yi (2010) JRRSB 631-654

An example:

$y^0 \to y^0 + dy$

$$\underset{n\times 1}{dy} = \underset{n\times p}{V(\theta)}\ \underset{p\times 1}{d\theta}$$

Differentiate quantile fn

Ex: Regression $y = X\beta + \sigma z$ say Non-N error $f_i(z_i)$

at y^0 $\underset{n\times 1}{dy} = \begin{pmatrix} \underset{n\times r}{X} & \underset{n\times 1}{z^0(\beta,\sigma)} \end{pmatrix} \underset{(r+1)\times 1}{\begin{pmatrix} d\beta \\ d\sigma \end{pmatrix}}$

$z^0(\beta\sigma) = (y^0 - X\beta)/\sigma$ = Data residual re $\beta\ \sigma$

$\Phi(\downarrow)$ = p-value vector re β, σ

Can put in terms of $d\hat\beta, d\hat\sigma$ at Observed...

at y^0

$$d\hat\beta = d\beta + (\hat\beta^0 - \beta)\, d\sigma/\sigma$$

$$d\hat\sigma = \hat\sigma\, d\sigma/\sigma$$

$$\begin{pmatrix} d\hat\beta \\ d\hat\sigma \end{pmatrix} = \begin{bmatrix} I & (\hat\beta^0 - \beta)/\sigma \\ 0 & \hat\sigma/\sigma \end{bmatrix} \begin{pmatrix} d\beta \\ d\sigma \end{pmatrix}$$

Absolute value $|d\hat\theta| = \left| \downarrow \right| d\theta = \underset{\text{prior}}{\pi(\theta)}\, d\theta = \frac{\hat\sigma}{\sigma}\, d\beta\, d\sigma$

"Standard" Bayesian right-invariant prior

Also: in general $\Rightarrow$ 2nd order prior

F Reid Marras Yi (2010) JRRSB 631-654

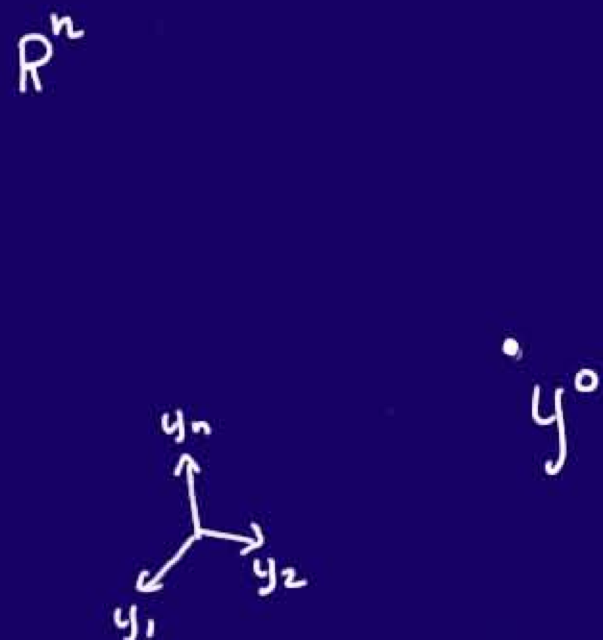
A Conditioning on data information

Model: $f(y;\theta)$ y in R^n $\dim\theta = p$

A Conditioning on data information

Model: $f(y;\theta)$ $\quad$ y in R^n $\quad$ $\dim \theta = p$

Data: y°

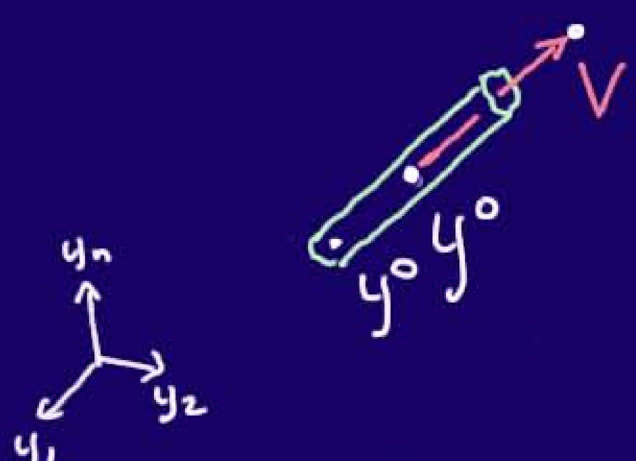


A Conditioning on data information

Model: $f(y;\theta)$ $\quad$ y in R^n $\quad$ $\dim\theta = p$

Data: y^0

R^n



Continuity available:

Take θ from $\hat\theta^0$ to $\hat\theta^0 + d\theta$

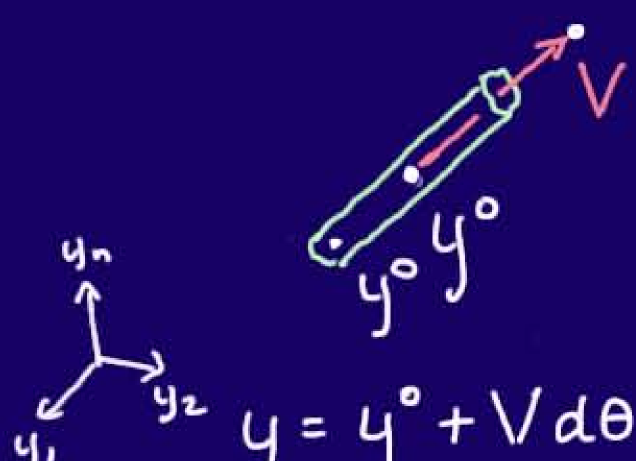
$$dy = V\, d\theta \qquad V = V(\hat\theta^0) = \left.\frac{dy}{d\theta}\right|_{y^0\,\hat\theta^0}$$

A Conditioning on data information

Model: $f(y;\theta)$ $\quad y$ in R^n $\quad \dim\theta = p$

Data: y^0

R^n



$y = y^0 + V\,d\theta$

Continuity available:

Take θ from $\hat\theta^0$ to $\hat\theta^0 + d\theta$

$$dy = V\,d\theta \qquad V = V(\hat\theta^0) = \left.\frac{dy}{d\theta}\right|_{y^0\,\hat\theta^0}$$

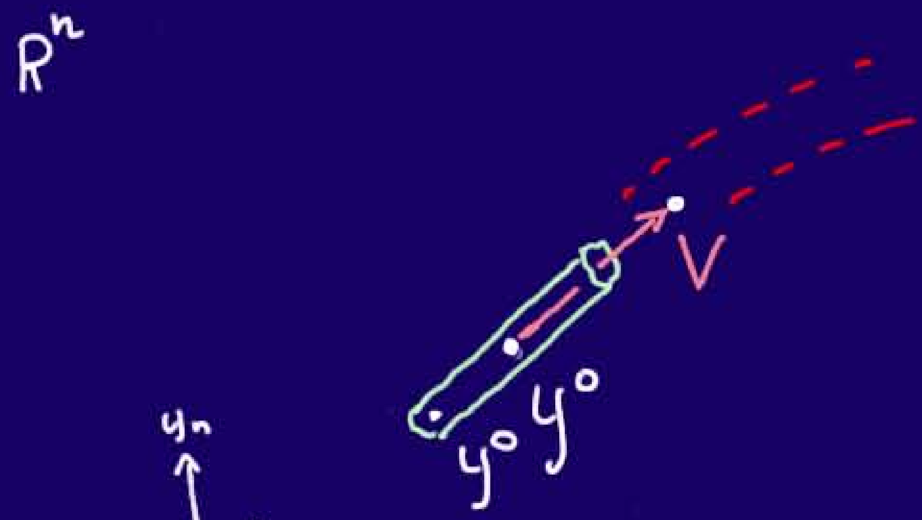
Distribution moves in tube

Fraser F Staicu (2010) Bernoulli 1208-1223

A Conditioning on data information

Model: $f(y;\theta)$ $\quad$ y in R^n $\quad$ $\dim \theta = p$

Data: y°



Continuity available:

Take θ from $\hat\theta^\circ$ to $\hat\theta^\circ + d\theta$

$dy = V\,d\theta \qquad V = V(\hat\theta^\circ) = \frac{dy}{d\theta}\Big|_{y^\circ \hat\theta^\circ}$

Distribution moves <u>in</u> tube

$\left\{ \begin{array}{l} W = V_\theta(\hat\theta^\circ) \;\; p\times p \text{ of } p \\ \text{Frobenius} \end{array} \right.$

$y_1 \quad y_2 \quad y = y^\circ + V d\theta \;+\; d\theta' W d\theta\, n^{-1/2} = y(d\theta, \hat u^\circ)$ $\qquad$ W = curvature $= V_\theta(\hat\theta^\circ)$

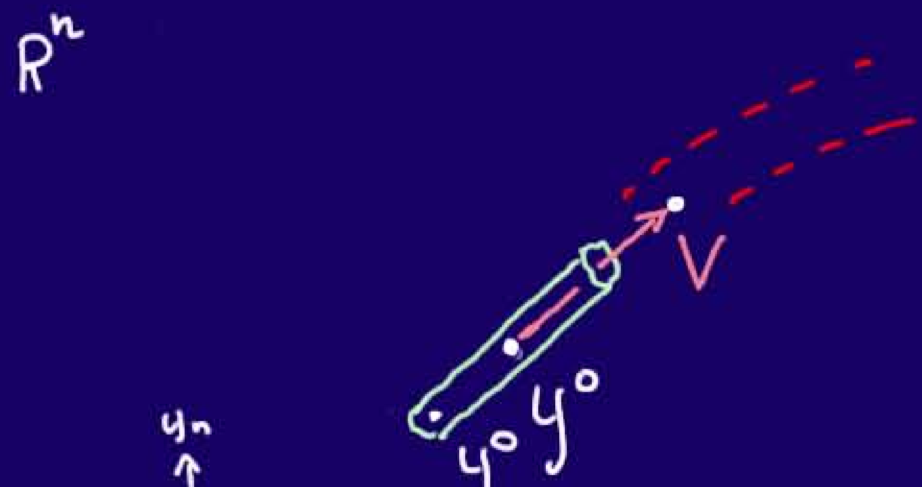
Fraser F Staicu (2010) Bernoulli 1208-1223

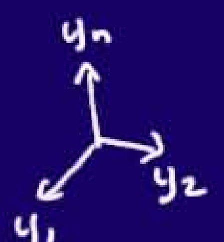
A Conditioning on data information

Model: $f(y;\theta)$ $\quad y$ in R^n $\quad \dim\theta = p$

Data: y°

Continuity available:





Take θ from $\hat\theta^\circ$ to $\hat\theta^\circ + d\theta$

$$dy = V\, d\theta \qquad V = V(\hat\theta^\circ) = \frac{dy}{d\theta}\Big|_{y^\circ \hat\theta^\circ}$$

Distribution moves $\underline{\underline{\text{in}}}$ tube

$\left\{ \begin{array}{l} W = V_\theta(\hat\theta^\circ) \quad p\times p \text{ of } p \\ \text{Frobenius} \end{array} \right.$

$$y = y^\circ + V\,d\theta + d\theta'\,W\,d\theta\, n^{-1/2} = y(d\theta, \hat u^\circ) \qquad W = \text{curvature} = V_\theta(\hat\theta^\circ)$$

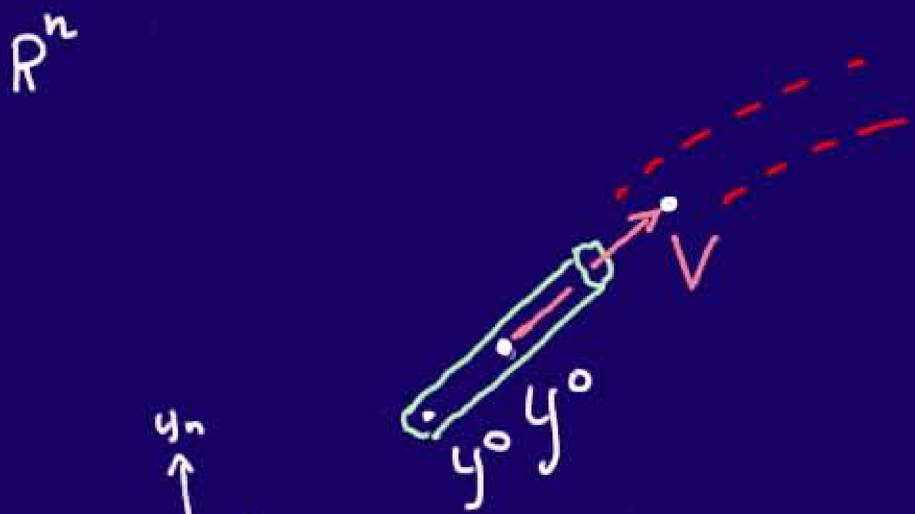
Revised model is Conditional in tube

Fraser F Staicu (2010) Bernoulli 1208-1223

A Conditioning on data information

Model: $f(y;\theta)$ $\quad y$ in R^n $\quad \dim \theta = p$

Data: y°



Continuity available:

Take θ from $\hat{\theta}^\circ$ to $\hat{\theta}^\circ + d\theta$

$$dy = V\,d\theta \qquad V = V(\hat{\theta}^\circ) = \frac{dy}{d\theta}\Big|_{y^\circ \hat{\theta}^\circ}$$

Distribution moves in tube

$\left\{ \begin{array}{l} W = V_\theta(\hat{\theta}^\circ) \;\; p{\times}p \text{ of } p \\ \text{Frobenius} \end{array} \right.$

$$y = y^\circ + V\,d\theta + d\theta' W\,d\theta\, n^{-1/2} = y(d\theta, \hat{u}^\circ) \qquad W = \text{curvature} = V_\theta(\hat{\theta}^\circ)$$

Revised model is Conditional in tube

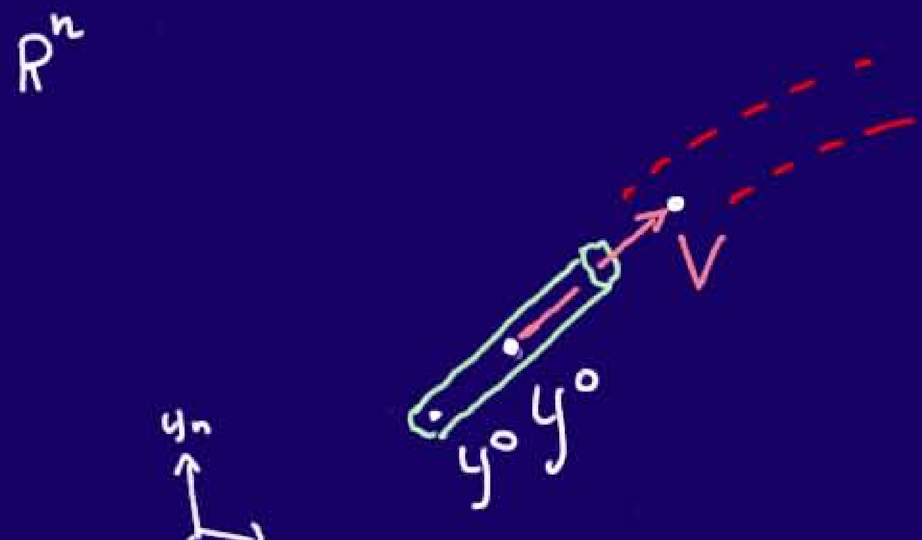
Asymptotics: (i) Just calculate: $\varphi(\theta) = \frac{\partial}{\partial V} \log f(y;\theta)\Big|_{y^\circ}$

Fraser F Staicu (2010) Bernoulli 1208-1223

A Conditioning on data information

Model: $f(y;\theta)$ $\quad y$ in R^n $\quad$ dim $\theta = p$

Data: y^0



Continuity available:

Take θ from $\hat\theta^0$ to $\hat\theta^0 + d\theta$

$$dy = V\,d\theta \qquad V = V(\hat\theta^0) = \left.\frac{dy}{d\theta}\right|_{y^0\,\hat\theta^0}$$

Distribution moves in tube

$\left\{ W = V_\theta(\hat\theta^0) \right.$ $p \times p$ of p
Frobenius

$$y = y^0 + V\,d\theta \; + \; d\theta' \, W \, d\theta \, n^{-1/2} = y(d\theta, \hat u^0) \qquad W = \text{curvature} = V_\theta(\hat\theta^0)$$

Revised model is Conditional in tube

Asymptotics: (i) Just calculate: $\varphi(\theta) = \left.\frac{\partial}{\partial V} \log f(y;\theta)\right|_{y^0}$

(ii) Use $g(s;\varphi) = \exp\{\ell(\theta) + \varphi(\theta)\,s\}\,|\hat\jmath_{\varphi\varphi}(s)|^{-1/2} \cdot ds$ $\qquad$ SP

$s^0 = 0$

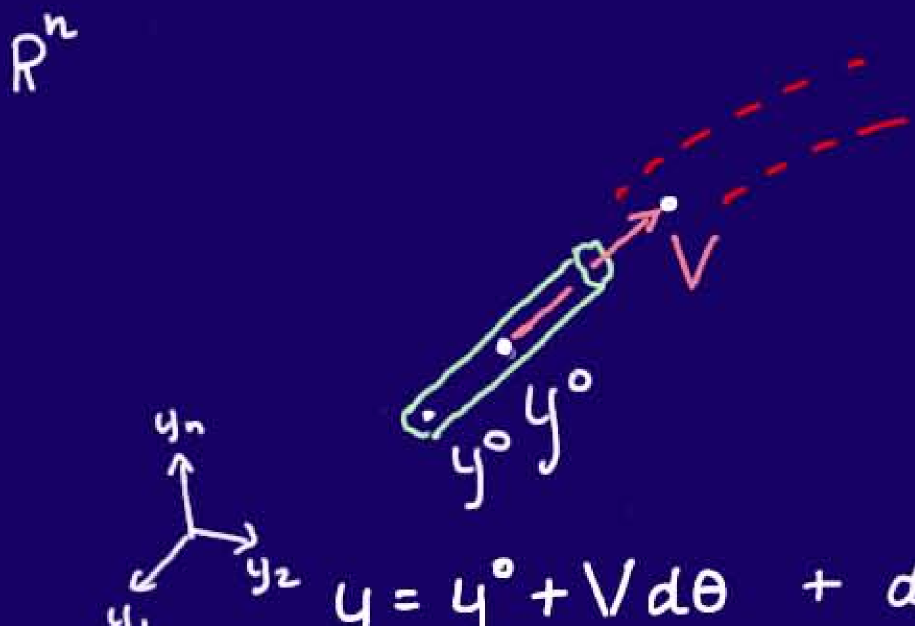
Just an "Exp model"... Easy... $O(n^{-3/2})$

Reid F (2010) Biometrika 159-170
Fraser F Staicu (2010) Bernoulli 1208-1223

A Conditioning on data information

Model: $f(y;\theta)$ $\quad$ y in R^n $\quad$ $\dim\theta = p$

Data: y^0



Continuity available:

Take θ from $\hat\theta^0$ to $\hat\theta^0 + d\theta$

$$dy = V\,d\theta \qquad V = V(\hat\theta^0) = \frac{dy}{d\theta}\Big|_{y^0\hat\theta^0}$$

Distribution moves in tube

$\{ W = V_\theta(\hat\theta^0)$ $p \times p$ of p
Frobenius

$$y = y^0 + V\,d\theta \;+\; d\theta' W\,d\theta\, n^{-1/2} = y(d\theta, \hat u^0) \qquad W = \text{curvature} = V_\theta(\hat\theta^0)$$

Revised model is Conditional in tube

Asymptotics: (i) Just calculate: $\varphi(\theta) = \dfrac{\partial}{\partial V}\log f(y;\theta)\Big|_{y^0}$

(ii) Use $g(s;\varphi) = \exp\{\ell(\theta) + \varphi(\theta)\,s\}\,|\hat\jmath_{\varphi\varphi}(s)|^{-1/2}\cdot ds$ $\quad$ SP

$s^0 = 0$

Just an "Exp model"... Easy... $O(n^{-3/2})$

All p-values $p(\psi)$ for scalar $\psi(\theta)$ available

Reid F (2010) Biometrika 159-170
Fraser F Staicu (2010) Bernoulli 1208-1223

B Default priors : "Nec. & Suff" conditions

Case: Location $f(\underset{\sim}{y}-\underset{\sim}{\beta}(\theta))$

B Default priors : "Nec. & Suff" conditions

Case: Location $f(\underset{\sim}{y} - \underset{\sim}{\beta}(\theta))$ Laplace / Obvious / routine ... "flat in β"

$$\underline{\pi(\theta)}\, d\theta = d\beta = \underline{|\beta_\theta(\theta)|} \cdot d\theta$$

B Default priors : "Nec. & Suff" conditions

Case: Location $f(\underset{\sim}{y} - \underset{\sim}{\beta}(\theta))$

Laplace / Obvious / routine ... "flat in β"

$$\pi(\theta)\,d\theta = d\beta = |\beta_\theta(\theta)| \cdot d\theta$$

frequentist p-value $p(\psi)$ } EQUAL iff

Bayes posterior survivor $s(\psi)$ } ψ linear in β

B Default priors : "Nec. & Suff" conditions

Case: Location $f(\underset{\sim}{y} - \underset{\sim}{\beta}(\theta))$

Laplace / Obvious / routine ... "flat in β"

$$\underline{\pi(\theta)}\, d\theta = d\beta = \underline{|\beta_\theta(\theta)|} \cdot d\theta$$

frequentist p-value $p(\psi)$ } EQUAL iff
Bayes posterior survivor $s(\psi)$ } ψ linear in β

General: $f(y;\theta)$ y°

Continuity $\Rightarrow V(\theta)$

B Default priors : "Nec. & Suff" conditions

Case: Location $f(\underset{\sim}{y} - \underset{\sim}{\beta}(\theta))$ Laplace / Obvious / routine ... "flat in β"

$$\underline{\pi(\theta)}\, d\theta = d\beta = \underline{|\beta_\theta(\theta)|} \cdot d\theta$$

frequentist p-value $p(\psi)$ } EQUAL iff
Bayes posterior survivor $s(\psi)$ } ψ linear in β

General: $f(y;\theta)$ y°

Continuity $\Rightarrow V(\theta)$

Default prior: $\pi(\theta) d\theta = |V(\theta)|\, d\theta$ Or switch $y \longrightarrow \hat{\theta}$

Examples: Non Normal $y = X\beta + \sigma z$ $\pi = |X\, \hat{z}^\circ(\beta,\sigma)| = \sigma^{-1}$ Widely prefered

Non Linear $y = x(\beta) + \sigma z$ $\pi = |V(\theta)|$

F Reid Marras Yi (2010) JRRSB 631-654

B Default priors : "Nec. & Suff" conditions

Case: Location $f(\underset{\sim}{y} - \underset{\sim}{\beta}(\theta))$ Laplace / Obvious / routine ... "flat in β"

$$\pi(\theta)\, d\theta = d\beta = |\beta_\theta(\theta)| \cdot d\theta$$

frequentist p-value $p(\psi)$ / Bayes posterior survivor $s(\psi)$ } EQUAL iff ψ linear in β

same

General: $f(y;\theta)$ y°

Continuity $\Rightarrow V(\theta)$

Default prior: $\pi(\theta)d\theta = |V(\theta)|\, d\theta$

Or switch $y \longrightarrow \hat{\theta}$

$|V(\theta)|\, d\theta$ gives 2nd order Accurate for linear parameters ψ

F Reid Marras Yi (2010) JRRSB 631-654

B Default priors : "Nec. & Suff" conditions

Case: Location $f(\underset{\sim}{y} - \underset{\sim}{\beta}(\theta))$

Laplace / Obvious / routine ... "flat in β"

$$\pi(\theta)\, d\theta = d\beta = |\beta_\theta(\theta)| \cdot d\theta$$

frequentist p-value $p(\psi)$
Bayes posterior survivor $s(\psi)$
} EQUAL iff ψ linear in β

same

General: $f(y;\theta)$ y°

Continuity $\Rightarrow V(\theta)$

Default prior: $\pi(\theta)d\theta = |V(\theta)|\, d\theta$

Or switch $y \longrightarrow \hat{\theta}$

$|V(\theta)|\, d\theta$ gives 2nd order Accurate for linear parameters ψ

Examples: Non Normal $y = X\beta + \sigma z$ $\pi = |X\, \hat{z}^\circ(\beta,\sigma)| = \sigma^{-1}$ widely prefered

Non Linear $y = x(\beta) + \sigma z$ $\pi = |V(\theta)|$

F Reid Marras Yi (2010) JRRSB 631-654

C Parameters: Linear and <u>Curved</u>

C Parameters: Linear and Curved

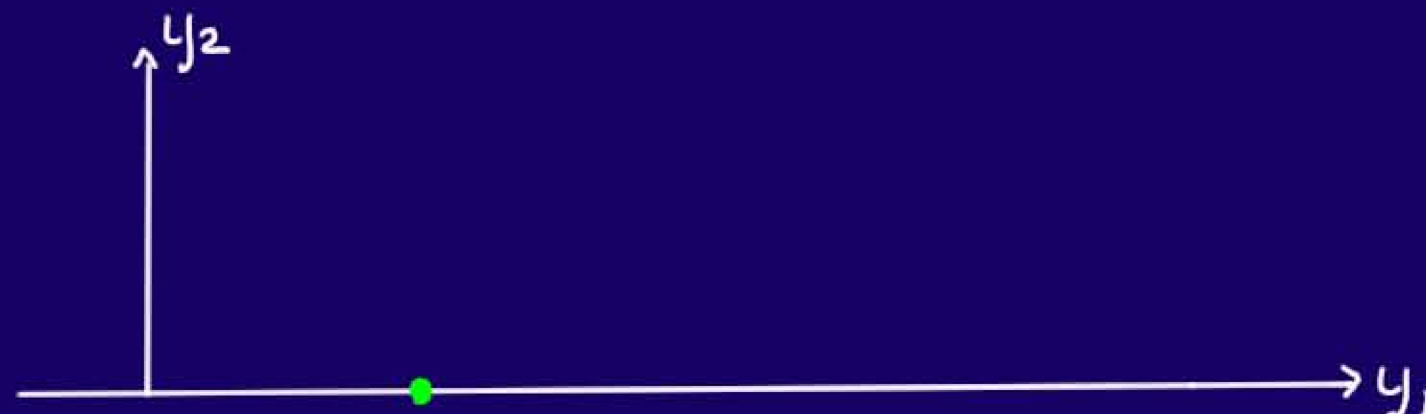
Case: Location model, say N $\quad \phi(y_1 - \theta_1, y_2 - \theta_2)$

Data: y^0 $\quad$ Assess: $\psi(\theta_1, \theta_2)$

C Parameters: Linear and Curved

Case: Location model, say N $\phi(y_1 - \theta_1, y_2 - \theta_2)$

Data: y^0 Assess: $\psi(\theta_1, \theta_2)$

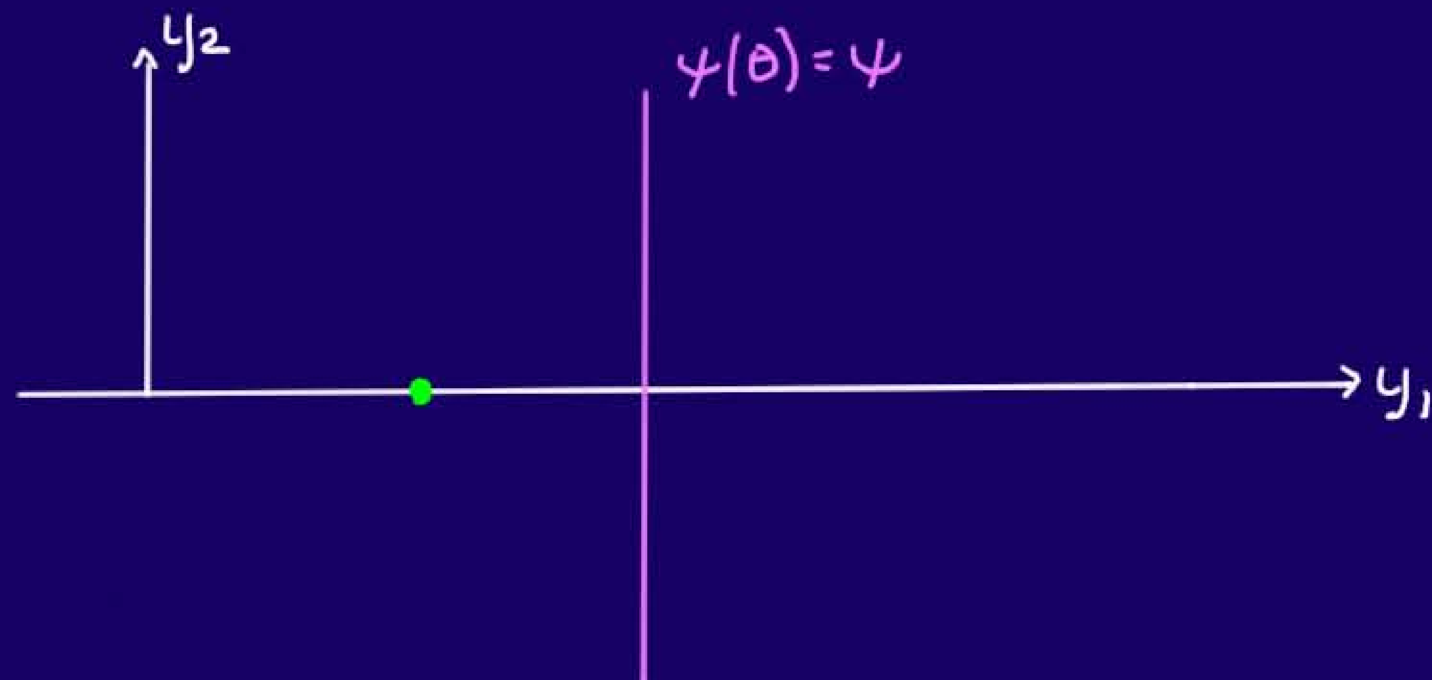


C Parameters: Linear and Curved

Case: Location model, say N $\phi(y_1-\theta_1, y_2-\theta_2)$

Data: y^0 Assess: $\psi(\theta_1, \theta_2)$

① Linear: Assess $\psi(\theta_1, \theta_2) = \theta_1$ say

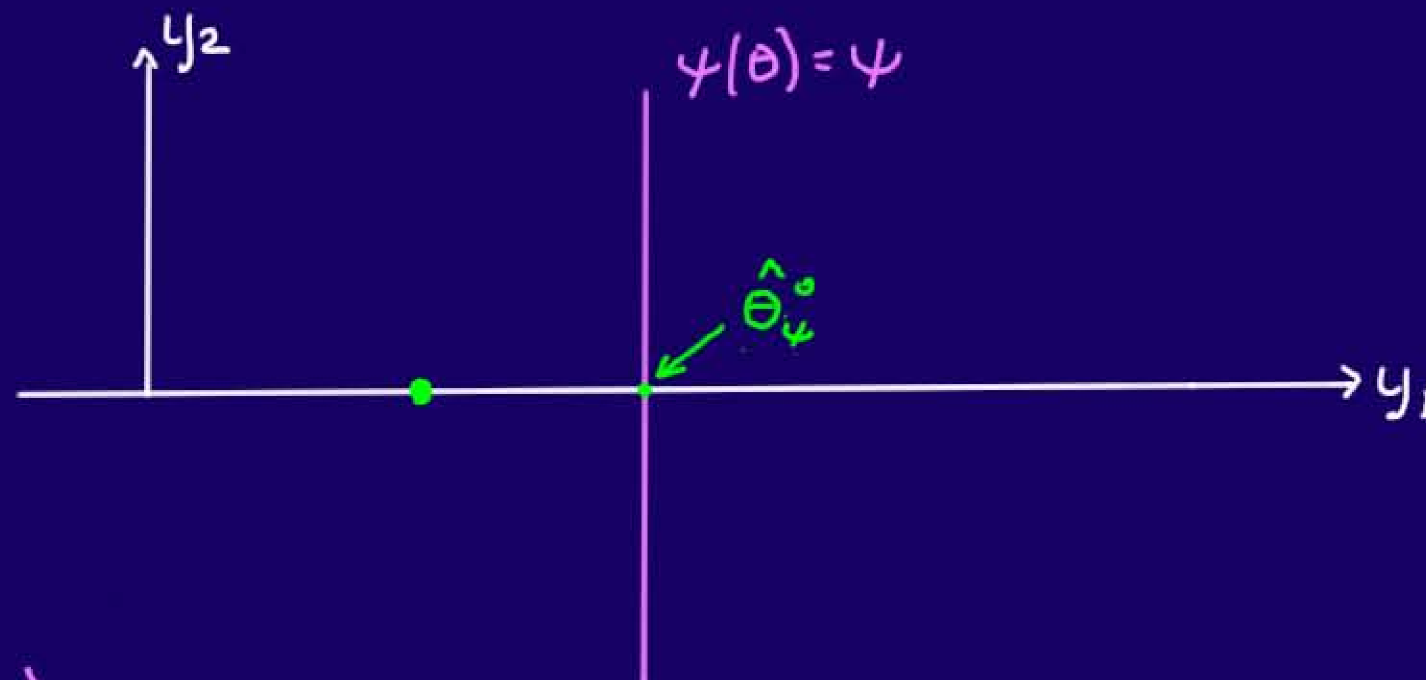


C Parameters: Linear and Curved

Case: Location model, say N $\quad \phi(y_1-\theta_1, y_2-\theta_2)$

Data: y^0 $\quad$ Assess: $\psi(\theta_1, \theta_2)$

① Linear: Assess $\psi(\theta_1, \theta_2) = \theta_1$ say



$\Phi(-1)$

p-value $= p(\psi) = .16$

s-value $= s(\psi) = .16$

freq p-value and Bayes posterior survivor equal

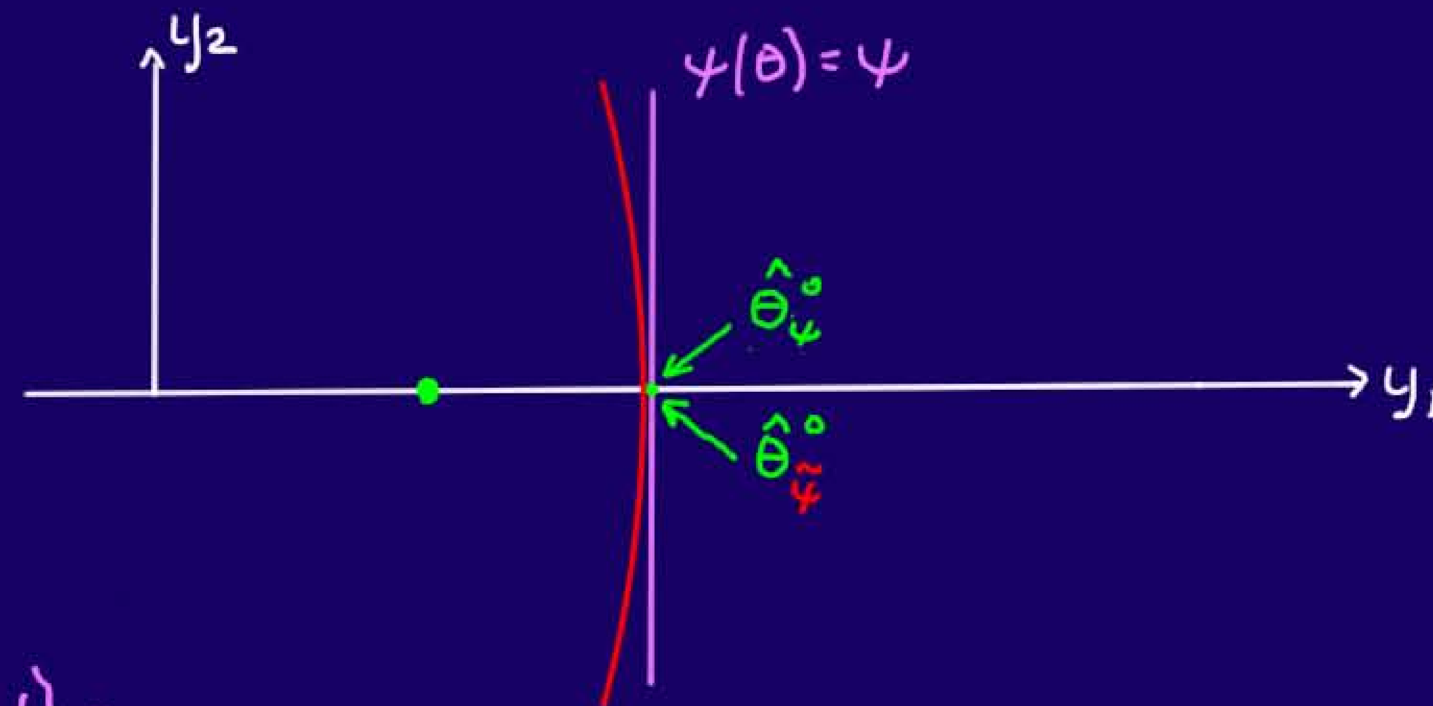
C Parameters: Linear and Curved

Case: Location model, say N $\quad \phi(y_1-\theta_1, y_2-\theta_2)$

Data: y^0 $\quad$ Assess: $\psi(\theta_1, \theta_2)$

② Curved

Assess $\tilde{\psi}(\theta_1, \theta_2) = \{(\theta_1+5)^2 + \theta_2^2\}^{1/2} = \psi$



$\Phi(-1)$

p-value $= p(\psi) = .16$

s-value $= s(\psi) = .16$

freq p-value and Bayes posterior survivor equal

p-value $= .13$ ↓

s-value $= .21$ ↑

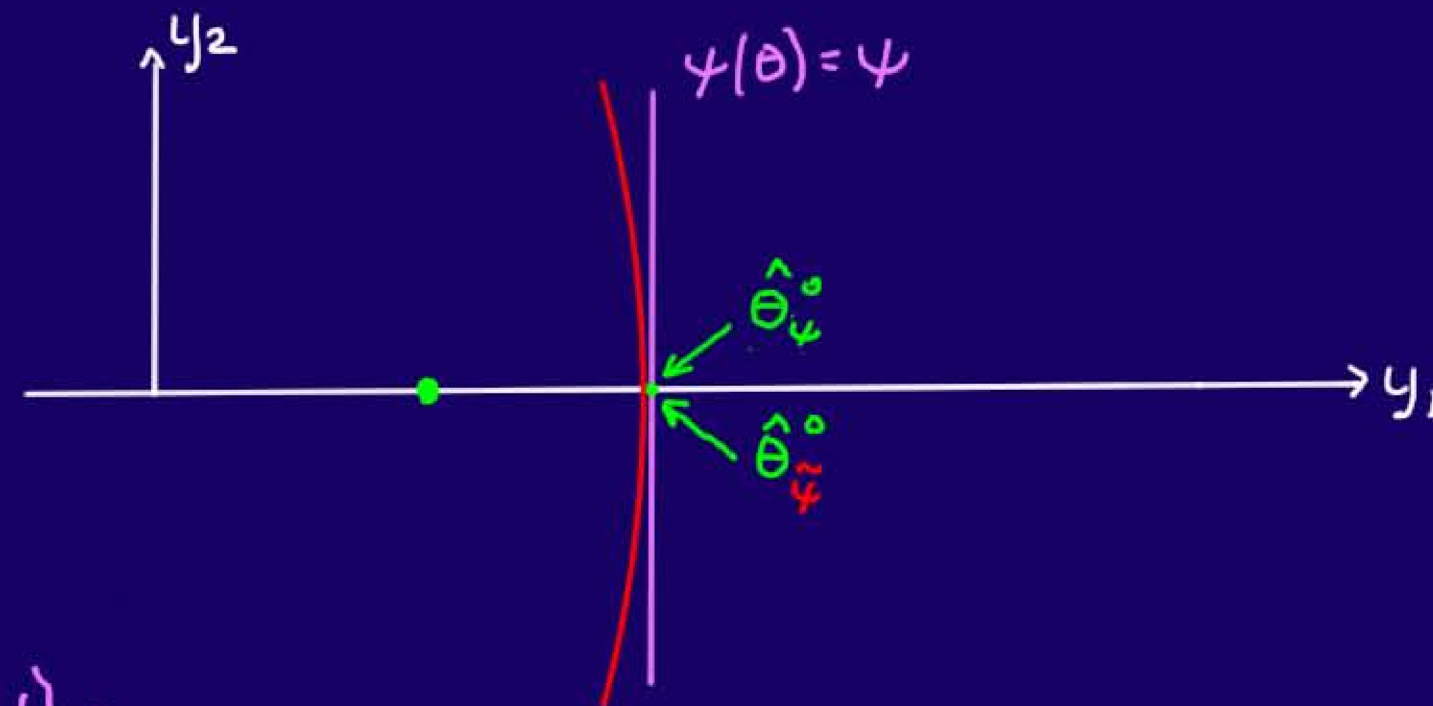
pvalue & s-value are not equal

C Parameters: Linear and Curved

Case: Location model, say N $\phi(y_1-\theta_1, y_2-\theta_2)$

Data: y^0 Assess: $\psi(\theta_1, \theta_2)$

(2) Curved Assess $\tilde{\psi}(\theta_1,\theta_2) = \{(\theta_1+5)^2 + \theta_2^2\}^{1/2} = \psi$



$\Phi(-1)$

p-value $= p(\psi) = .16$

s-value $= s(\psi) = .16$

freq p-value and Bayes posterior survivor equal

p-value $= .13$ ↓

s-value $= .21$ ↑

pvalue & s-value are not equal

p-value & s-value move in opposite directions

Accuracy?

10. Curvature of parameters $O(n^{-1})$ and its Bias on Bayes

Asymptotics says:

1) Model $O(n^{-1})$ is $\exp\{\ell(\theta) + \varphi'(\theta)s\}|\hat{j}_{\varphi\varphi}(s)|^{-\frac{1}{2}} \cdot ds$

Reid F (2010) Biometrika

10. Curvature of parameters $O(n^{-1})$ and its Bias on Bayes

Asymptotics says:

1) Model $O(n^{-1})$ is

$$\exp\{\underbrace{\ell(\theta)}_{\text{Obs } \ell} + \underbrace{\varphi'(\theta)}_{\varphi \text{ "repar"}} s\} \, |\underbrace{\hat{j}_{\varphi\varphi}(s)}_{\text{Obs info}}|^{-\frac{1}{2}} \cdot ds \quad \leftarrow \text{SP}$$

Reid F (2010) Biometrika

10. Curvature of parameters $O(n^{-1})$ and its Bias on Bayes

Asymptotics says:

Obs ℓ — φ "repar" — Obs info

1) Model $O(n^{-1})$ is $\quad \exp\{\ell(\theta) + \varphi'(\theta)s\}\,|\hat{j}_{\varphi\varphi}(s)|^{-\frac{1}{2}} \cdot ds \quad$ SP

2) Can find linear par. β near $\hat{\varphi}^0$

Reid F (2010) Biometrika

Fraser F Fraser (2010) J. Stat. Res. Efron Vol.

10. Curvature of parameters $O(n^{-1})$ and its Bias on Bayes

Asymptotics says:

Obs ℓ $\quad$ φ "repar" $\quad$ Obs info

1) Model $O(n^{-1})$ is $\quad \exp\{\ell(\theta) + \varphi'(\theta)s\}\,|\hat{j}_{\varphi\varphi}(s)|^{-\frac{1}{2}} \cdot ds \quad$ SP

2) Can find linear par. β near $\hat{\varphi}^o$

Continuity gives $\quad d\hat{\varphi} = M(\varphi)\, d\varphi$

Reid F (2010) Biometrika
Fraser F Fraser (2010) J. Stat. Res. Efron Vol.

10. Curvature of parameters $O(n^{-1})$ and its Bias on Bayes

Asymptotics says:

Obs ℓ φ "repar"   Obs info

1) Model $O(n^{-1})$ is $$\exp\{\ell(\theta) + \varphi'(\theta)s\}\,|\hat{j}_{\varphi\varphi}(s)|^{-\frac{1}{2}}\cdot ds$$ SP

2) Can find linear par. β near $\hat{\varphi}^0$

Continuity gives $d\hat{\varphi} = M(\varphi)\,d\varphi$ $\mathbb{M}^0 = M_\varphi(\hat{\varphi}^0)$

Integrate $O(n^{-1})$ $\beta(\varphi) = \underset{p\times 1}{M^0} d\varphi + \frac{1}{2n^{1/2}}\, d\varphi' \underset{p\times p}{\mathbb{M}^0} d\varphi$ arrays of vectors

Reid F (2010) Biometrika

Fraser F Fraser (2010) J. Stat. Res. Efron Vol.

10. Curvature of parameters $O(n^{-1})$ and its Bias on Bayes

Asymptotics says:

1) Model $O(n^{-1})$ is

$$\exp\{\ell(\theta) + \varphi'(\theta) s\} \, |\hat{j}_{\varphi\varphi}(s)|^{-\frac{1}{2}} \cdot ds$$

Obs ℓ ; φ "repar" ; Obs info ; SP

2) Can find linear par. β near $\hat{\varphi}^0$

Continuity gives $d\hat{\varphi} = M(\varphi)\, d\varphi$ $\qquad \mathbb{M}^0 = M_{\varphi}(\hat{\varphi}^0)$

Integrate $O(n^{-1})$ $\quad \beta(\varphi) = \underset{p\times 1}{M^0} d\varphi + \frac{1}{2n^{1/2}} d\varphi' \underset{p\times p}{\mathbb{M}^0} d\varphi$ $\qquad$ arrays of vectors

Get linear contours ... So that Vector Bayes works!

Reid F (2010) Biometrika

Fraser F Fraser (2010) J. Stat. Res. Efron Vol.

// Summary & Directions

Have: Sample space conditioning

Have 3rd order p-value inference

Have default priors

11 Summary & Directions

Have: Sample space conditioning

Have 3rd order p-value inference

Have default priors

But They don't handle parameter curvature DSZ

Dawid Stone Zidek (1973)

// Summary & Directions

Have: Sample space conditioning

Have 3rd order p-value inference

Have default priors

But They don't handle parameter curvature DSZ

Now have: Curvature measure $\gamma(\hat{\theta}^\circ)$

Dawid Stone Zidek (1973)

// Summary & Directions

Have: Sample space conditioning

Have 3rd order p-value inference

Have default priors

But They don't handle parameter curvature DSZ

Now have: Curvature measure $\gamma(\hat{\theta}^0)$

<u>Develop</u>: Bayes adjustment for curvature

$s(\psi) +$ "<u>Curv. corr.</u>" .... $O(n^{-1})$

Removes $O(n^{-1/2})$ DSZ bias

F. Sun (2010)

Dawid Stone Zidek (1973)

Thank you

Thank you

and Thank you 

Thank you

and Thank you

and Thank you