Priors and Inference: A differential view

D A S Fraser
Statistics U Toronto

November 8 2012

Department of Statistics
University of Oxford

www.utstat.toronto.edu/dfraser/documents/oxford2012.pdf

/xxx Index for some references

Outline:

1 Priors & Inference: Is there a problem?

2 Data dependent prior? Why?

3 Welch Peers: Exponential as location model

4 Exponential model: as an approximation

5 Saddlepoint: for exponential & for general

6 Conditional / Marginal for linear parameter component

7 Directional Jeffreys

8 Curved parameter: How Bayes misses!

Some Q & A at talk

Some Q & A at discussion later

1 Priors & Inference: Is there a problem?

- Bayes & freq. give different results!
- a discipline with 2 contradictory methodologies!

1 Priors & Inference: Is there a problem?

- Bayes & freq. give different results!
- a discipline with 2 contradictory methodologies!
- Does accuracy matter?

1 <u>Priors & Inference</u>: Is there a <u>problem</u>?

- Bayes & freq. give different results!
- a discipline with 2 contradictory methodologies!
- Does <u>accuracy</u> matter?
- Does the discipline <u>care</u>?

1 Priors & Inference: Is there a problem? 1

- Bayes & freq. give different results!
- A discipline with 2 contradictory methodologies!
- Does accuracy matter?
- Does the discipline care?

Merck, Vioxx, \$10b profit, Recalled in 2006

FDA penalty \$5b Est. # deaths 40K /uwo12.pdf

Q Need for accuracy?

A Available statistics had shown high risks for Vioxx. Some information had been suppressed, other information had been ignored, as in the "Shuttle's" O-rings; so Accuracy is needed broadly

Both: In choice and use of statistics & In the external messages from statistics

as with Vioxx / O-rings mentioned above.

1 Priors & Inference: Is there a problem?

- Bayes & freq. give different results!
- A discipline with 2 contradictory methodologies!
- Does accuracy matter?
- Does the discipline care?

Merck, Vioxx, \$10b profit, Recalled in 2006

FDA penalty \$5b Est. # deaths 40K /uwo12.pdf

· Take a simple case: scalar

* If $f(y-\theta)$

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

B $s(\theta) = \int_{\theta}^{\infty} \pi(\theta)\, f(y^0-\theta)\,d\theta$

1 Priors & Inference: Is there a problem?

- Bayes & freq. give different results!
- A discipline with 2 contradictory methodologies!
- Does accuracy matter?
- Does the discipline care?

Merck, Vioxx, \$10b profit, Recalled in 2006
FDA penalty \$5b Est. # deaths 40K /uwo12.paf

· Take a simple case: scalar

* If $f(y-\theta)$ f $\quad p(\theta) = \int_{-\infty}^{y^0} f(y-\theta)\,dy$ Equal if Nice pure case, but

B $\quad s(\theta) = \int_{\theta}^{\infty} \pi(\theta) f(y^0-\theta)\,d\theta$ $\quad \pi(\theta) = c$ Vector case ⇒ Doesn't work!

A: Paradigm for when Bayes is objectively repetition-valid

1 <u>Priors & Inference</u>: Is there a <u>problem</u>? 1

- Bayes & freq. give different results!
- A discipline with 2 contradictory methodologies!
- Does <u>accuracy</u> matter?
- Does the discipline <u>care</u>?

Merck, Vioxx, \$10b profit, Recalled in 2006
FDA penalty \$5b Est. # deaths 40K /uwo12.pdf

· Take a <u>simple case</u>: <u>scalar</u>

* If $\underline{f(y-\theta)}$

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

B $\quad s(\theta) = \int_{\theta}^{\infty} \pi(\theta)\, f(y^0-\theta)\,d\theta$

<u>Equal</u> if $\pi(\theta) = c$

Nice pure case, but <u>Vector case</u> ⇒ <u>Doesn't work</u>!

← calculus ↘ (bdry conditions)

If $\underline{f(y;\theta)}$ stoch. inc.

$$p(\theta) = F(y^0;\theta) \equiv \int_{\theta}^{\infty} -F_{;\theta}(y^0;\theta)\,d\theta$$

$$s(\theta) = \int_{\theta}^{\infty} \pi(\theta)\, F_y(y^0:\theta)\,d\theta \qquad \text{"agree?"}$$

1 Priors & Inference: Is there a problem?

- Bayes & freq. give different results!
- A discipline with 2 contradictory methodologies!
- Does accuracy matter?
- Does the discipline care?

Merck, Vioxx, \$10b profit, Recalled in 2006
FDA penalty \$5b Est. # deaths 40K /uwo12.pdf

· Take a simple case: scalar

* If $f(y-\theta)$ f $p(\theta)=\int_{-\infty}^{y^0} f(y-\theta)\,dy$ Equal if $\pi(\theta)=c$ Nice pure case, but Vector case ⇒ Doesn't work!

B $s(\theta)=\int_{\theta}^{\infty}\pi(\theta)\,f(y^0-\theta)\,d\theta$

← calculus ↘ (bdry conditions)

If $f(y;\theta)$ stoch. inc.

$$p(\theta)= F(y^0;\theta) \equiv \int_{\theta}^{\infty} -F_{;\theta}(y^0;\theta)\,d\theta$$

$$s(\theta)=\int_{\theta}^{\infty}\pi(\theta)\,F_y(y^0;\theta)\,d\theta \quad \text{"agree?"}$$

Q: Clarify:
A: To have $p(\theta)\equiv s(\theta)$
and thus have Bayes repetition-valid
one needs to have:

Equal if $\pi(\theta)=-\dfrac{F_{;\theta}(y^0;\theta)}{F_y(y^0;\theta)}$ ⟷ $\pi(\theta)=\left.\dfrac{dy}{d\theta}\right|_{y^0\hat\theta^0}$

prior

1 Priors & Inference: Is there a problem?

- Bayes & freq. give different results!
- A discipline with 2 contradictory methodologies!
- Does accuracy matter?
- Does the discipline care?

Merck, Vioxx, \$10b profit, Recalled in 2006
FDA penalty \$5b Est. # deaths 40K /uwo12.pdf

· Take a simple case: scalar

* If $f(y-\theta)$

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

B $s(\theta) = \int_{\theta}^{\infty} \pi(\theta) f(y^0-\theta)\,d\theta$

Equal if $\pi(\theta) = c$

Nice pure case, but Vector case ⇒ Doesn't work!

← calculus ↘ (bdry conditions)

If $f(y;\theta)$ stoch. inc.

$$p(\theta) = F(y^0;\theta) \equiv \int_{\theta}^{\infty} -F_{;\theta}(y^0;\theta)\,d\theta$$

$$s(\theta) = \int_{\theta}^{\infty} \pi(\theta) F_y(y^0;\theta)\,d\theta$$

"agree?"

df | $F(y;\theta) = u$ distn
qf | $y = y(u;\theta)$ quantile

Equal if prior → $\pi(\theta) = -\dfrac{F_{;\theta}(y^0;\theta)}{F_y(y^0;\theta)}$

1 Priors & Inference: Is there a problem?

- Bayes & freq. give different results!
- A discipline with 2 contradictory methodologies!
- Does accuracy matter?
- Does the discipline care?

Merck, Vioxx, \$10b profit, Recalled in 2006
FDA penalty \$5b Est. # deaths 40K /uwo12.pdf

· Take a simple case: scalar

* If $f(y-\theta)$ f $p(\theta) = \int_{-\infty}^{y^0} f(y-\theta)\,dy$ Equal if $\pi(\theta) = c$

B $s(\theta) = \int_{\theta}^{\infty} \pi(\theta) f(y^0-\theta)\,d\theta$

Nice pure case, but Vector case ⇒ Doesn't work!

← calculus ↘ (b'dry conditions)

If $f(y;\theta)$ stoch. inc.

$p(\theta) = F(y^0;\theta) \equiv \int_{\theta}^{\infty} -F_{;\theta}(y^0;\theta)\,d\theta$

$s(\theta) = \int_{\theta}^{\infty} \pi(\theta) F_y(y^0:\theta)\,d\theta$ "agree?"

df | $F(y;\theta) = u$ distn

qf | $y = y(u;\theta)$ quantile

Equal if $\pi(\theta) = -\dfrac{F_{;\theta}(y^0;\theta)}{F_y(y^0;\theta)} = \pi(\theta) = \dfrac{\partial}{\partial\theta} y(u;\theta)\Big|_{y^0\hat{\theta}^0}$ = "How θ moves y near y^0"

prior

⟹ Data-dependent prior!

2 Data-dependent prior? Why?

- Contrary to "classical" Bayes view!

2 Data-dependent prior? Why?

- Contrary to "classical" Bayes view!

- But well recognized ... in places

    Box & Cox 1964

    Wasserman 2000 JRSSB

    F Reid MY 2010 JRSSB

2 Data-dependent prior? Why?

- Contrary to "classical" Bayes view!
- But well recognized ... in places

    Box & Cox 1964

    Wasserman 2000 JRSSB

    FReid MY 2010 JRSSB

- B. & freq. differ? Data-dependent can make it right? Sometimes!

2 Data-dependent prior? Why?

- Contrary to "classical" Bayes view!
- But well recognized ... in places

    Box & Cox 1964

    Wasserman 2000 JRSSB

    F Reid MY 2010 JRSSB

- B. & freq. differ? Data-dependent can make it right? Sometimes!
- How to evaluate B vs freq? Use bounds, confidence vs. posterior

    Confidence: level always right! ← under model!

    Posterior: level right if $\pi(\theta)$ describes source!

    otherwise wrong!

2 Data-dependent prior? Why?

- Contrary to "classical" Bayes view!
- But well recognized... in places
  - Box & Cox 1964
  - Wasserman 2000 JRSSB
  - Fraser Reid MY 2010 JRSSB
- B. & freq. differ? Data-dependent can make it right? Sometimes!
- How to evaluate B vs freq? Use bounds, confidence vs. posterior

  Confidence: level always right!

  Posterior: level right if $\pi(\theta)$ describes source!

  otherwise wrong!

- Opinion prior?

  e.g. Multiple Conf bounds (or Confidence dist'ns maybe?)

  $\Rightarrow$ Get confidence result first; put opinion prior in parallel

  See later

  - Don't use opinion prior to analyze model!

    $\Rightarrow$ an Admission of bankruptcy!

3 Welch-Peers 1963 Scalar Exponential: As a location model

Q Just Exponential?

A also "General asymptotic" in W-P using Expected info $i(\theta)$ but interest here in Exponential model (also see next Section)

3 Welch-Peers 1963 Scalar Exponential: As a location model

Root info prior ⇒ Posterior bound is confidence bound $O(n^{-1})$

dim = n

$$f(y;\theta) = \exp\{s(y)\varphi(\theta) + K(\theta)\}H(y)$$

3 Welch-Peers 1963 Scalar Exponential: As a location model

Root info prior $\Rightarrow$ Posterior bound is confidence bound $O(n^{-1})$

dim = n

$f(y;\theta) = \exp\{s(y)\varphi(\theta) + K(\theta)\} H(y)$

$g(s;\varphi) = \exp\{s\varphi - K(\varphi)\} g(s)$ Sufficiency

3 Welch-Peers 1963 Scalar Exponential: As a location model

Root info prior $\Rightarrow$ Posterior bound is confidence bound $O(n^{-1})$

dim = n

$f(y;\theta) = \exp\{s(y)\varphi(\theta) + K(\theta)\}H(y)$

$g(s;\varphi) = \exp\{s\varphi - K(\varphi)\}\, g(s)$ Sufficiency

$g(s;\varphi) = \exp\{s\varphi + \ell^{0}(\varphi)\}\, g(s)$ $\leftarrow$ Center & scale $j_{\varphi\varphi}^{1/2}(\hat{\varphi}^{0})(\varphi - \hat{\varphi}^{0})$ write as φ; $j_{\varphi\varphi}^{-1/2}(\hat{\varphi}^{0})(s - s^{0})$ write as s; $j^{0}_{\varphi\varphi}$ = obs info

3 Welch-Peers 1963 Scalar Exponential: As a location model

Root info prior $\Rightarrow$ Posterior bound is confidence bound $O(n^{-1})$

dim = n

$f(y;\theta) = \exp\{s(y)\varphi(\theta) + K(\theta)\}H(y)$

Obs. likelihood but in new coordinates

$g(s;\varphi) = \exp\{s\varphi - K(\varphi)\}g(s)$ Sufficiency

$g(s;\varphi) = \exp\{s\varphi + \ell^0(\varphi)\}g(s)$ $\leftarrow$ Center & scale $j_{\varphi\varphi}^{1/2}(\hat{\varphi}^0)(\varphi - \hat{\varphi}^0)$ write as φ; $j_{\varphi\varphi}^{-1/2}(\hat{\varphi}^0)(s - s^0)$ write as s; $j^0_{\varphi\varphi}$ = obs info

$= \exp\{-\frac{\varphi^2}{2} - \gamma\varphi^3/6n^{1/2} + s\varphi\}g(s)$ $\leftarrow$ Taylor expand logmodel

missing terms then come from "pdf" property

3 Welch-Peers 1963 Scalar Exponential: As a location model

Root info prior $\Rightarrow$ Posterior bound is confidence bound $O(n^{-1})$

dim = n

$$f(y;\theta) = \exp\{s(y)\varphi(\theta) + K(\theta)\}H(y)$$

$$g(s;\varphi) = \exp\{s\varphi - K(\varphi)\}g(s)$$ Sufficiency

$$g(s;\varphi) = \exp\{s\varphi + \ell^{0}(\varphi)\}g(s)$$ $\leftarrow$ Center & Scale

$j_{\varphi\varphi}^{1/2}(\hat{\varphi}^{0})(\varphi - \hat{\varphi}^{0})$ write as φ

$j_{\varphi\varphi}^{-1/2}(\hat{\varphi}^{0})(s - s^{0})$ write as s

$j_{\varphi\varphi}^{\theta}$ = obs info

$$= \exp\{-\tfrac{\varphi^2}{2} - \gamma\,\varphi^3/6n^{1/2} + s\varphi\}\,g(s)$$ $\leftarrow$ Taylor expand log model

$$= \phi(s-\varphi)\exp\{-\gamma\frac{\varphi^3}{6n^{1/2}} + \gamma\frac{s^3}{6n^{1/2}}\}\left(1 - \gamma\frac{s}{2n^{1/2}}\right)$$ $\leftarrow$ Get other terms by "pdf property"

Use: If s is $N\,\varphi\,1$ then $Es^3 = \varphi^3 + 3\varphi$

3 Welch-Peers 1963 Scalar Exponential: As a location model

Root info prior $\Rightarrow$ Posterior bound is confidence bound $O(n^{-1})$

dim = n

$f(y;\theta) = \exp\{s(y)\varphi(\theta) + K(\theta)\}H(y)$

$g(s;\varphi) = \exp\{s\varphi - K(\varphi)\}g(s)$ Sufficiency

$g(s;\varphi) = \exp\{s\varphi + \ell^0(\varphi)\}g(s)$ $\leftarrow$ Center & scale $j_{\varphi\varphi}^{1/2}(\hat\varphi^0)(\varphi - \hat\varphi^0)$ write as φ; $j_{\varphi\varphi}^{-1/2}(\hat\varphi^0)(s - s^0)$ write as s; $j^0_{\varphi\varphi}$ = obs info

$= \exp\{-\frac{\varphi^2}{2} - \gamma\varphi^3/6n^{1/2} + s\varphi\}g(s)$ $\leftarrow$ Taylor expand logmodel

$= \phi(s-\varphi)\exp\{-\gamma\frac{\varphi^3}{6n^{1/2}} + \gamma\frac{s^3}{6n^{1/2}}\}(1 - \gamma\frac{s}{2n^{1/2}})$ $\leftarrow$ Get other terms by "pdf property" Use If s is $N\,\varphi\,1$ then $Es^3 = \varphi^3 + 3\varphi$

Get

Root info parameter $\beta = \int_0^\varphi j^{1/2}(\varphi)\,d\varphi = \int_0^\varphi (1 + \gamma\varphi/2n^{1/2})\,d\varphi = \varphi + \gamma\varphi^2/4n^{1/2}$

mle Use: $\hat\varphi = s - \gamma s^2/2n^{1/2}$ $\Rightarrow$ $\hat\beta = s - \gamma s^2/4n$

Location parameter $\beta = \varphi + \gamma\varphi^2/4n^{1/2}$

(a Quadratic adjustment)

3 Welch-Peers 1963 Scalar Exponential: As a location model

Root info prior $\Rightarrow$ Posterior bound is confidence bound $O(n^{-1})$

dim = n

$f(y;\theta) = \exp\{s(y)\varphi(\theta) + K(\theta)\}H(y)$

$g(s;\varphi) = \exp\{s\varphi - \kappa(\varphi)\}\, g(s)$ Sufficiency

$g(s;\varphi) = \exp\{s\varphi + \ell^0(\varphi)\}\, g(s)$ $\leftarrow$ Center & scale $j_{\varphi\varphi}^{1/2}(\hat\varphi^0)(\varphi - \hat\varphi^0)$ write as φ; $j_{\varphi\varphi}^{-1/2}(\hat\varphi^0)(s - s^0)$ write as s; $j^0_{\varphi\varphi}$ = obs info

$= \exp\{-\frac{\varphi^2}{2} - \gamma\varphi^3/6n^{1/2} + s\varphi\}\, g(s)$ $\leftarrow$ Taylor expand log model

$= \phi(s-\varphi)\exp\{-\gamma\frac{\varphi^3}{6n^{1/2}} + \gamma\frac{s^3}{6n^{1/2}}\}(1 - \gamma\frac{s}{2n^{1/2}})$ $\leftarrow$ Get other terms by "pdf property" Use If s is $N\varphi 1$ then $Es^3 = \varphi^3 + 3\varphi$

Get

Root info parameter $\beta = \int_0^\varphi j^{1/2}(\varphi)\,d\varphi = \int_0^\varphi (1 + \gamma\varphi/2n^{1/2})\,d\varphi = \varphi + \gamma\varphi^2/4n^{1/2} = \beta$

mle Use: $\hat\varphi = s - \gamma s^2/2n^{1/2}$ $\Rightarrow$ $\hat\beta = s - \gamma s^2/4n$

$\Rightarrow f(\hat\beta;\beta)\,d\hat\beta = (2\pi)^{-1/2}\exp\{-(\hat\beta-\beta)^2/2 - \gamma(\hat\beta-\beta)^3/12n^{1/2}\}$ = location model $O(n^{-1})$

Q: Should there be cubic power?

A: Yes Thank you! Added above

3 Welch-Peers 1963 Scalar Exponential: As a location model

Root info prior $\Rightarrow$ Posterior bound is confidence bound $O(n^{-1})$

dim = n

$f(y;\theta) = \exp\{s(y)\varphi(\theta) + K(\theta)\}H(y)$

$g(s;\varphi) = \exp\{s\varphi - K(\varphi)\}g(s)$ Sufficiency

$g(s;\varphi) = \exp\{s\varphi + \ell^0(\varphi)\}g(s)$ $\leftarrow$ Center & scale $j^{1/2}_{\varphi\varphi}(\hat\varphi^0)(\varphi - \hat\varphi^0)$ write as φ; $j^{-1/2}_{\varphi\varphi}(\hat\varphi^0)(s - s^0)$ write as s $\quad j^0_{\varphi\varphi}$ = obs info

$= \exp\{-\frac{\varphi^2}{2} - \gamma\varphi^3/6n^{1/2} + s\varphi\}g(s)$ $\leftarrow$ Taylor expand log model

$= \phi(s-\varphi)\exp\{-\gamma\frac{\varphi^3}{6n^{1/2}} + \gamma\frac{s^3}{6n^{1/2}}\}(1 - \gamma\frac{s}{2n^{1/2}})$ $\leftarrow$ Get other terms by "pdf property" $\quad$ Use: If s is $N\ \varphi\ 1$ then $Es^3 = \varphi^3 + 3\varphi$

Get

Root info parameter $\beta = \int_0^\varphi j^{1/2}(\varphi)d\varphi = \int_0^\varphi(1 + \gamma\varphi/2n^{1/2})d\varphi = \varphi + \gamma\varphi^2/4n^{1/2}$

mle $\quad$ Use: $\hat\varphi = s - \gamma s^2/2n^{1/2}$ $\Rightarrow$ $\hat\beta = s - \gamma s^2/4n$

$\Rightarrow f(\hat\beta;\beta)d\hat\beta = (2\pi)^{-1/2}\exp\{-(\hat\beta-\beta)^2/2 - \gamma(\hat\beta-\beta)/12n^{1/2}\}$ = location model $O(n^{-1})$

Bayes = freq (Location) with root-info prior, scalar case (flat in β)

Vector case? Does not work!

4 Exponential Model:

as a building block? or As an approximation to a model $O(n^{-1})$

Consider $f(y;\theta)$ regular:

4 Exponential Model:

as a building block? or As an approximation to a model $O(n^{-1})$

Consider $f(y;\theta)$ regular:

$f(\underset{n}{y};\underset{p}{\theta})$ density, or distn fns, or quantile fns (Often easily available; often used simulations)

4 Exponential Model:

as a building block? or As an approximation to a model $O(n^{-1})$

Consider $f(y;\theta)$ regular:

$f(\underset{n}{y};\underset{p}{\theta})$ density, or distn fns, or quantile fns (Often easily available; often used simulations)

quantile version

$\tilde{y} = \tilde{y}(\tilde{u},\tilde{\theta})$ & data y^0, $\hat{\theta}^0$

$$\underset{n\times p}{V} = (v_1,\ldots,v_p) = \frac{\partial}{\partial\theta} y(u;\theta)\Big|_{y^0\,\hat{\theta}^0}$$

4 Exponential Model:

As a building block? or As an approximation to a model $O(n^{-1})$

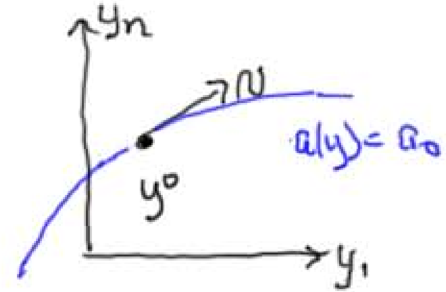
Consider $f(y;\theta)$ regular:

$f(\underset{n}{y};\underset{p}{\theta})$ density, or distn fns, or quantile fns (Often easily available; often used simulations)

quantile version

$\underset{\sim}{y} = \underset{\sim}{y}(\underset{\sim}{u},\underset{\sim}{\theta})$ & data y^0, $\hat{\theta}^0$ but V tangent to ancillary

$\underset{n\times p}{V} = (v_1,\dots,v_p) = \frac{\partial}{\partial\theta} y(u;\theta)\Big|_{y^0\,\hat{\theta}^0}$ (continuity: FF Staicu, Bern. 2010)



4 <u>Exponential Model</u>:

as a <u>building</u> block? or As an approximation to a model $O(n^{-1})$

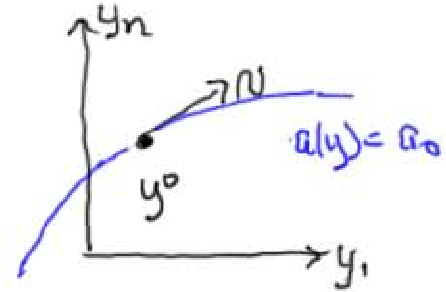
Consider $f(y;\theta)$ regular:

$f(y;\theta)$ density, or distn fns, or <u>quantile fns</u> (Often easily available; often used simulations)

$n \quad p$

quantile version

$\underset{\sim}{y} = \underset{\sim}{y}(\underset{\sim}{u}, \underset{\sim}{\theta})$ & data y^0, $\hat{\theta}^0$ but V tangent to ancillary



$\underset{n \times p}{V} = (v_1, \ldots, v_p) = \frac{\partial}{\partial \theta} y(u;\theta) \Big|_{y^0 \hat{\theta}^0}$ (continuity: FFStaicu, Bern. 2010)

and get $\varphi(\theta) = \frac{\partial}{\partial V} \ell(\theta; y)\Big|_{y^0} \to$ Canonical para. of ExpMod, $O(n^{-1})$ approx.

$$V = \frac{\partial y}{\partial \theta}\Big|_{y^0, \hat{\theta}^0} \text{ fixed "p-value" vector}$$

$$\varphi = \frac{\partial}{\partial V} \ell(\theta; y)\Big|_{y^0}$$

Then use φ-based Exp model for inference!

4 Exponential Model:

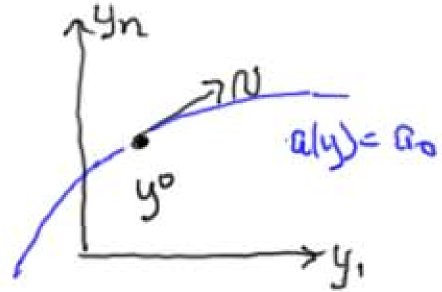
as a building block? or As an approximation to a model $O(n^{-1})$

Consider $f(y;\theta)$ regular:

$f(\underset{n}{y};\underset{p}{\theta})$ density, or distn fns, or quantile fns (Often easily available; often used simulations)

quantile version

$\underset{\sim}{y} = \underset{\sim}{y}(\underset{\sim}{u}, \underset{\sim}{\theta})$ & data y°, $\hat{\theta}^{\circ}$ but V tangent to ancillary



$\underset{n \times p}{V} = (v_1, \ldots, v_p) = \frac{\partial}{\partial \theta} y(u;\theta)\Big|_{y^{\circ}\, \hat{\theta}^{\circ}}$ (continuity: FFStaicu, Bern 2010)

and get $\varphi(\theta) = \frac{\partial}{\partial V} \ell(\theta; y)\big|_{y^{\circ}} \rightarrow$ Canonical para. of ExpMod, $O(n^{-1})$ approx.

$f_T(s;\varphi) = \exp\{\ell(\varphi) + s\varphi\}\, h(s)$ $\quad s^{\circ} = 0$ Conditional on ancillary

4 Exponential Model:

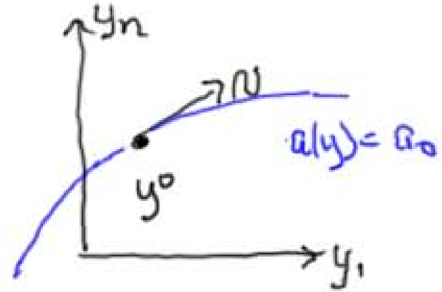
as a building block? or As an approximation to a model $O(n^{-1})$

Consider $f(y;\theta)$ regular:

$f(\underset{n}{y};\underset{p}{\theta})$ density, or distn fns, or quantile fns (Often easily available; often used simulations)

quantile version

$\tilde{y} = \tilde{y}(\tilde{u},\tilde{\theta})$ & data y^0, $\hat{\theta}^0$ but V tangent to ancillary



$\underset{n\times p}{V} = (v_1,\dots,v_p) = \frac{\partial}{\partial\theta} y(u;\theta)\Big|_{y^0\,\hat{\theta}^0}$ (continuity: FF Staicu, Bern. 2010)

and get $\varphi(\theta) = \frac{\partial}{\partial V}\ell(\theta;y)\Big|_{y^0}$ → Canonical para. of ExpMod, $O(n^{-1})$ approx.

$f_T(s;\varphi) = \exp\{\ell(\varphi) + s\varphi\}\,h(s)$ $\quad s^0 = 0$ $\quad$ Conditional on ancillary

$\ell(\varphi;s) = \varphi s + \ell^0(\varphi)$ $\quad$ 2nd order mod devns

$s^0 = 0$

$\boxed{\text{Act as if Expt'l model } O(n^{-1}) \text{ and use SP}}$

4 **Exponential Model:**

as a building block? or As an approximation to a model $O(n^{-1})$

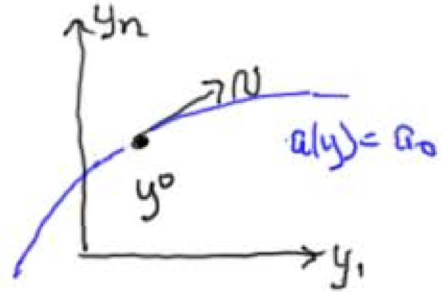
Consider $f(y;\theta)$ regular:

$f(y;\theta)$ density, or distn fns, or quantile fns (Often easily available; often used simulations)

n p

quantile version

$\underset{\sim}{y} = \underset{\sim}{y}(\underset{\sim}{u}, \underset{\sim}{\theta})$ & data y°, $\hat{\theta}^{\circ}$ but V tangent to ancillary



$\underset{n\times p}{V} = (v_1, \ldots, v_p) = \frac{\partial}{\partial\theta} y(u;\theta) \Big|_{y^{\circ}\,\hat{\theta}^{\circ}}$ (continuity: FF Staicu, Bern. 2010)

and get $\varphi(\theta) = \frac{\partial}{\partial V} \ell(\theta; y)\big|_{y^{\circ}}$ → Canonical para. of ExpMod, $O(n^{-1})$ approx.

$f_T(s;\varphi) = \exp\{\ell(\varphi) + s\varphi\}\, h(s)$ $\quad s^{\circ} = 0$ $\quad$ Conditional on ancillary

$\ell(\varphi; s) = \varphi s + \ell^{\circ}(\varphi)$ 2nd order mod devns
$s^{\circ} = 0$

Act as if Exptl model $O(n^{-1})$

$|i_{\theta\theta}(\theta)|^{1/2}$

Does Welch-Peers extend? Yes! but not as Jeffreys hoped!

5 <u>Saddepoint</u> For exponential models (general, by approximation)

Exp. mod: $f(s;\varphi) = \exp\{ s^T\varphi + \ell(\varphi)\}\, h(s)$ Centered coordinates $s^0 = 0$ Daniels 1954 B-N Cox 1979

5 Saddepoint For exponential models (general, by approximation)

Exp. mod: $f(s;\varphi) = \exp\{s^T\varphi + \ell(\varphi)\}\, h(s)$

Centered coordinates $s^0 = 0$

Daniels 1954
B-N Cox 1979

with or without scaling here

SP: $f(s;\varphi)ds = \dfrac{e^{k/n}}{(2\pi)^{p/2}} \exp\left\{-\dfrac{r^2(\varphi;s)}{2}\right\} \cdot \left|\hat{j}_{\varphi\varphi}(\hat{\varphi})\right|^{-1/2} ds$ $\quad O(n^{-3/2})$

$\cdot \left|\hat{j}_{\varphi\varphi}(\hat{\varphi})\right|^{1/2} d\hat{\varphi}$

5 Saddlepoint For exponential models (general, by approximation)

Exp. mod: $f(s;\varphi) = \exp\{s^T\varphi + \ell(\varphi)\}\, h(s)$ Centered coordinates $s^0 = 0$ Daniels 1954 B-N Cox 1979

SP: $f(s;\varphi)ds = \frac{e^{k/n}}{(2\pi)^{p/2}} \exp\left\{-\frac{r^2(\varphi;s)}{2}\right\} \cdot \left|\hat{j}_{\varphi\varphi}(\hat{\varphi})\right|^{-1/2} ds$ $O(n^{-3/2})$ $r^2/2 = \ell(\hat{\varphi};s) - \ell(\varphi;s)$

$\cdot \left|\hat{j}_{\varphi\varphi}(\hat{\varphi})\right|^{1/2} d\hat{\varphi}$ $= \log$ lik ratio

$\hat{\varphi} = \arg\sup \ell(\varphi;s)$

Simple pieces $\Rightarrow$ lik ratio Exp'd info / obs info... re φ

5 Saddepoint For exponential models (general, by approximation)

Exp. mod: $f(s;\varphi) = \exp\{s^T\varphi + \ell(\varphi)\}\, h(s)$ Centered coordinates $s^0 = 0$ Daniels 1954 B-N Cox 1979

SP: $f(s;\varphi)\,ds = \frac{e^{k/n}}{(2\pi)^{p/2}} \exp\left\{-\frac{r^2(\varphi;s)}{2}\right\} \cdot \left|\hat{j}_{\varphi\varphi}(\hat{\varphi})\right|^{-1/2} ds$ $O(n^{-3/2})$

$r^2/2 = \ell(\hat{\varphi};s) - \ell(\varphi;s)$ = log lik ratio

$\cdot \left|\hat{j}_{\varphi\varphi}(\hat{\varphi})\right|^{1/2} d\hat{\varphi}$

$\hat{\varphi} = \text{argsup}\ \ell(\varphi;s)$

Simple pieces ⇒ lik ratio Exp'd info / obs info... re φ

Accuracy & Separation component parameters: Tilt $= s\psi + t\lambda$

SP $f(s,t;\psi,\lambda) = \frac{e^{k/n}}{(2\pi)^{p/2}} \exp\left\{-\frac{r^2(\psi,\lambda;s,t)}{2}\right\} \left|\hat{j}_{\varphi\varphi}(\hat{\psi},\hat{\lambda})\right|^{-\frac{1}{2}}$

dim s = dim ψ = d
dim t = dim λ = $p-d$

5 Saddepoint For exponential models (general, by approximation)

Exp. mod: $f(s;\varphi) = \exp\{ s^T\varphi + \ell(\varphi)\}\, h(s)$ — Centered coordinates $s^0 = 0$ — Daniels 1954, B-N Cox 1979

SP: $f(s;\varphi)\,ds = \dfrac{e^{k/n}}{(2\pi)^{p/2}} \exp\left\{-\dfrac{r^2(\varphi;s)}{2}\right\} \cdot \left|\hat{j}_{\varphi\varphi}(\hat\varphi)\right|^{-1/2} ds$ $\quad O(n^{-3/2})$

$\cdot \left|\hat{j}_{\varphi\varphi}(\hat\varphi)\right|^{1/2} d\hat\varphi$

$r^2/2 = \ell(\hat\varphi;s) - \ell(\varphi;s)$

$= \log$ lik ratio

$\hat\varphi = \arg\sup\, \ell(\varphi;s)$

Simple pieces ⇒ lik ratio; Exp'd info / obs info ... re φ

Accuracy & Separation component parameters: Tilt $s\psi + t\lambda$

SP $\quad f(s,t;\psi,\lambda) = \dfrac{e^{k/n}}{(2\pi)^{p/2}} \exp\left\{-\dfrac{r^2(\psi,\lambda;s,t)}{2}\right\} \left|\hat{j}_{\varphi\varphi}(\hat\psi,\hat\lambda)\right|^{-\frac{1}{2}}$

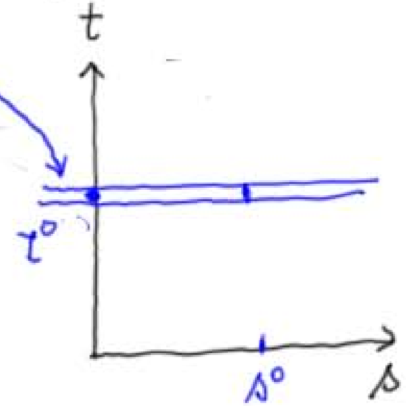
$\dim s = \dim\psi = d$

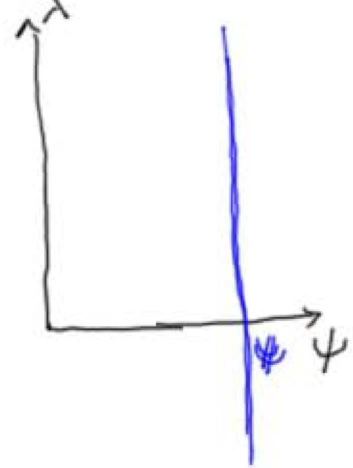
$\dim t = \dim\lambda = p - d$

Interest in ψ $\quad s\,|\,t^0\quad$ Exp'l

ψ linear in φ

e.g. $\varphi = (\psi,\lambda)$





5 Saddlepoint For exponential models (general, by approximation)

Exp. mod: $f(s;\varphi) = \exp\{s^T\varphi + \ell(\varphi)\}\, h(s)$ Centered coordinates $s^0 = 0$ Daniels 1954, B-N Cox 1979

SP: $f(s;\varphi)ds = \dfrac{e^{k/n}}{(2\pi)^{p/2}} \exp\left\{-\dfrac{r^2(\varphi;s)}{2}\right\} \cdot |\hat{j}_{\varphi\varphi}(\hat\varphi)|^{-1/2} ds$ $O(n^{-3/2})$

$\cdot |\hat{j}_{\varphi\varphi}(\hat\varphi)|^{1/2} d\hat\varphi$

$r^2/2 = \ell(\hat\varphi;s) - \ell(\varphi;s)$ = log lik ratio

$\hat\varphi = \arg\sup \ell(\varphi;s)$

Simple pieces ⇒ lik ratio ↥ ↳ Exp'd info / obs info ... re φ

Accuracy & Separation component parameters: ← Tilt $s\psi + t\lambda$

SP $f(s,t;\psi,\lambda) = \dfrac{e^{k/n}}{(2\pi)^{p/2}} \exp\left\{-\dfrac{r^2(\psi,\lambda;s,t)}{2}\right\} |\hat{j}_{\varphi\varphi}(\hat\psi,\hat\lambda)|^{-1/2}$

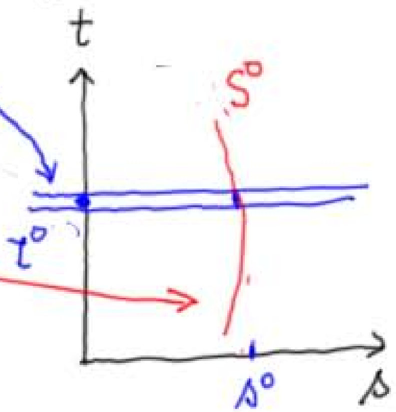
$\dim s = \dim\psi = d$
$\dim t = \dim\lambda = p-d$

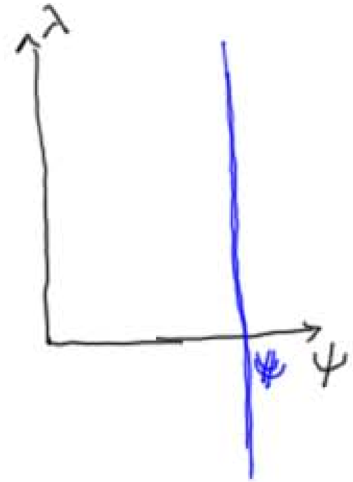
Interest in ψ $s|t^0$ Exp'l

Interest in ψ S 3rd

↳ Here: ψ linear

But S also available for curved ψ !





5 Saddepoint For exponential models (general, by approximation)

Exp. mod: $f(s;\varphi) = \exp\{s^T\varphi + \ell(\varphi)\}\, h(s)$ Centered coordinates $s^0 = 0$ Daniels 1954, B-N Cox 1979

SP: $f(s;\varphi)\,ds = \dfrac{e^{k/n}}{(2\pi)^{p/2}} \exp\left\{-\dfrac{r^2(\varphi;s)}{2}\right\} \cdot \left|\hat{j}_{\varphi\varphi}(\hat\varphi)\right|^{-1/2} ds$ $O(n^{-3/2})$

$\cdot \left|\hat{j}_{\varphi\varphi}(\hat\varphi)\right|^{1/2} d\hat\varphi$

$r^2/2 = \ell(\hat\varphi;s) - \ell(\varphi;s)$

$=$ log lik ratio

$\hat\varphi = \text{argsup}\ \ell(\varphi;s)$

Simple pieces $\Rightarrow$ lik ratio   Exp'd info / obs info ... re φ

Accuracy & Separation component parameters: Tilt $s\psi + t\lambda$

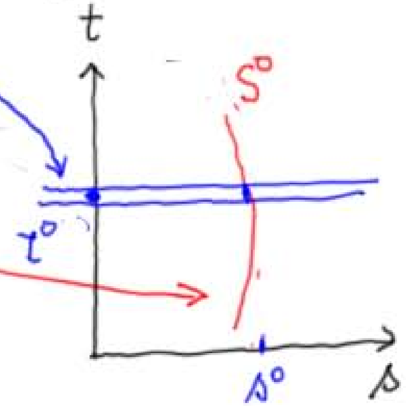
SP $f(s,t;\psi,\lambda) = \dfrac{e^{k/n}}{(2\pi)^{p/2}} \exp\left\{-\dfrac{r^2(\psi,\lambda;s,t)}{2}\right\} \left|\hat{j}_{\varphi\varphi}(\hat\psi,\hat\lambda)\right|^{-\frac{1}{2}}$

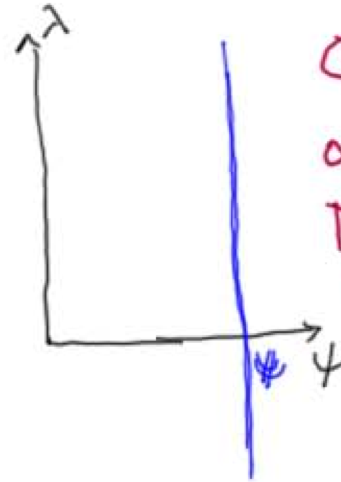
$\dim s = \dim \psi = d$

$\dim t = \dim \lambda = p - d$

Interest in ψ $s|t^0$ Exp'l

Interest in ψ S 3rd





Q: Does S depend on $\psi = \psi_0$ being tested?

3rd order: Yes

2nd order ?

But $s|t^0$, S are equiv $O(n^{-3/2})$

SP makes dist'ns available

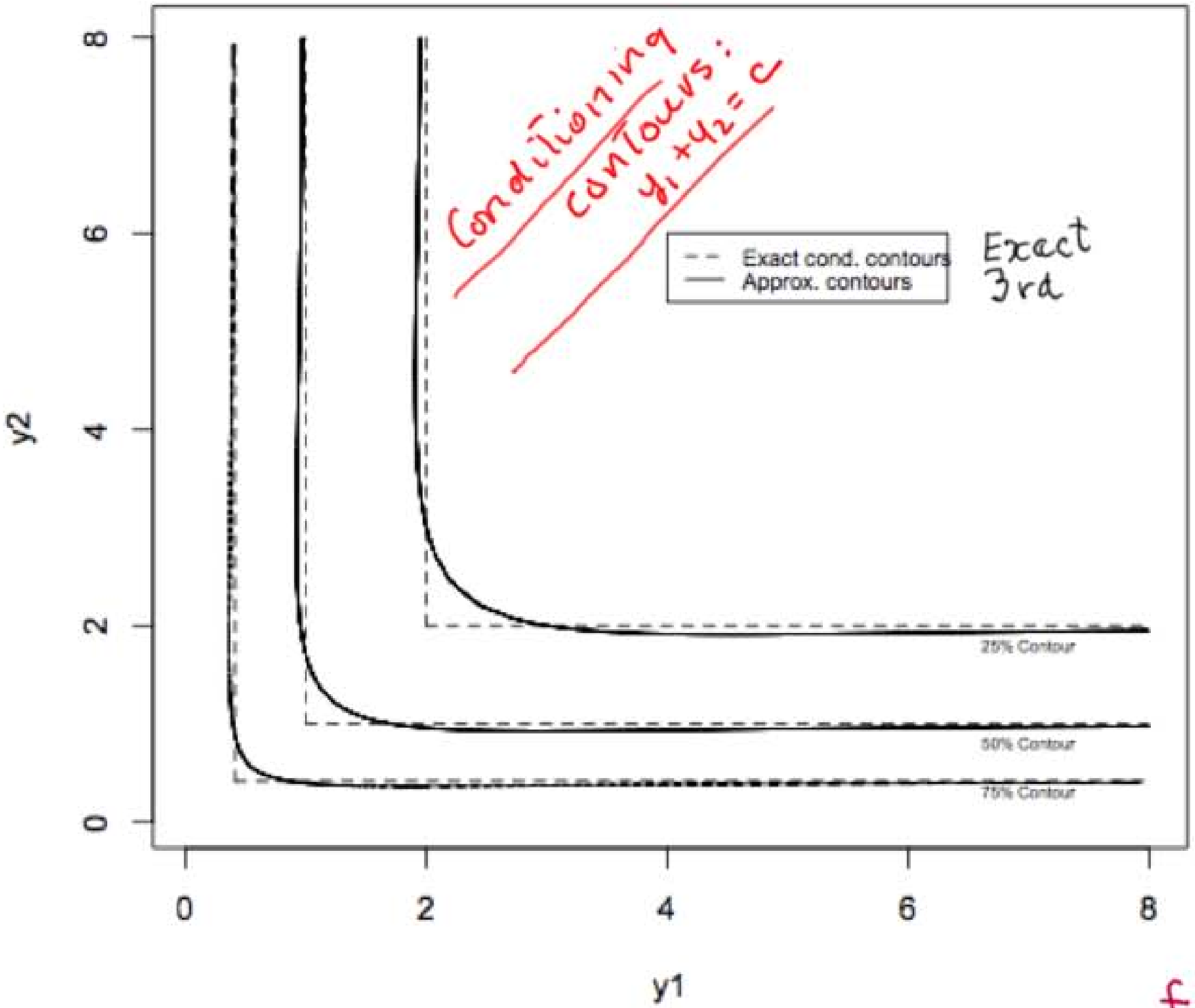


Fig. 1. The exact conditional contours and the third order approximate contours at quantile levels 25%, 50% and 75% for testing $\psi_0 = 0.6931$ in the simple exponential life model

Asymptotics but $n=1$

$$y_1 \sim \varphi_1 e^{-\varphi_1 y_1}$$

$$y_2 \sim \varphi_2 e^{-\varphi_2 y_2}$$

Interest $\psi = \varphi_1 + \varphi_2$

Test $\psi = .4901$

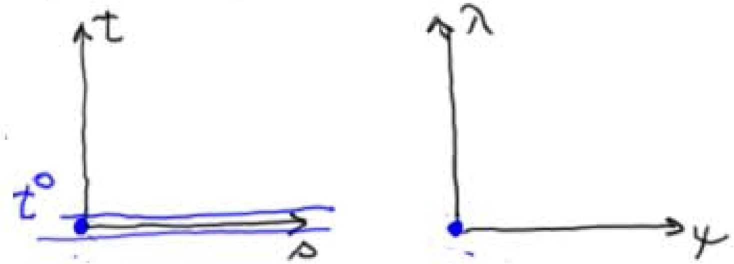
3rd for $s|t$ | Equal
3rd for S |

$$f = e^{\log \varphi_1 + \log \varphi_2 - \varphi_1 y_1 - \varphi_2 y_2}$$

S available as power series

6 Conditional & Marginal for linear component parameter

$$f(s,t;\psi,\lambda)=\frac{e^{k/n}}{(2\pi)^{p/2}}\exp\{s\psi+t\lambda+\ell(\psi,\lambda)\}\,|\hat{j}(s,t)|^{-\frac{1}{2}}$$



get $s \mid t^0$

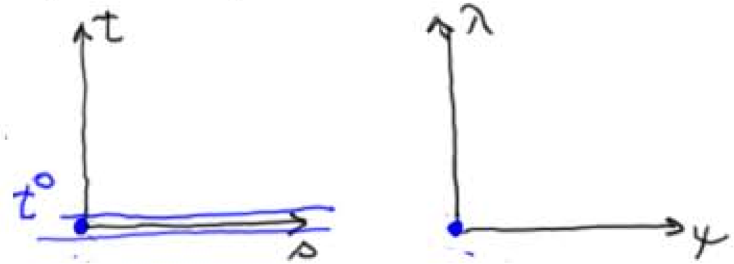
$$\left|\hat{j}_{(\chi,\lambda)(\chi,\lambda)}\{\hat{\chi}(s,t),\hat{\lambda}(s,t)\}\right|^{1/2}$$

2nd derivative info matrix in canonical parameterization.

6 Conditional & Marginal for linear component parameter

$$f(s,t;\psi,\lambda) = \frac{e^{k/n}}{(2\pi)^{p/2}} \exp\{s\psi + t\lambda + \ell(\psi,\lambda)\} \left|\hat{j}(s,t)\right|^{-\frac{1}{2}}$$

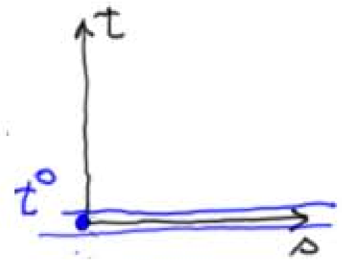
Get: $s \mid t^0$ Exponential in ψ Use SP $\nearrow$

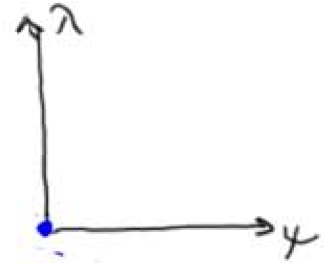


6 Conditional & Marginal for linear component parameter

$$f(s,t;\psi,\lambda) = \frac{e^{k/n}}{(2\pi)^{p/2}} \exp\{s\psi + t\lambda + \ell(\psi,\lambda)\} |\hat{j}(s,t)|^{-\frac{1}{2}}$$

$s \mid t^0$ Exponential in ψ Use SP





Case: $\dim\psi = \dim\lambda = 1$

$\hat{j}$ factors Cond'l info $= \hat{j}_{\lambda\lambda}$

Profile info $= \hat{j}^{p}_{\psi\psi}$

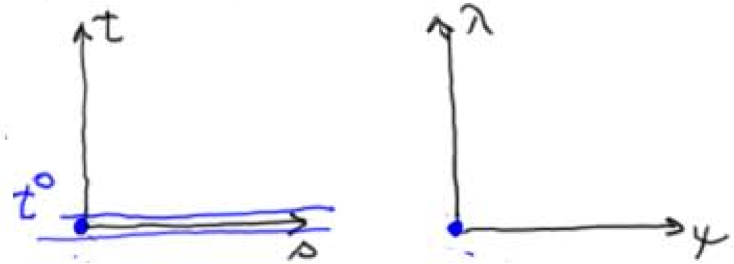
$\hat{\lambda}_\psi$ free of s

$\hat{j}_{\lambda\lambda}$ free of s

6 Conditional & Marginal for linear component parameter

$$f(s,t;\psi,\lambda) = \frac{e^{k/n}}{(2\pi)^{p/2}} \exp\{s\psi + t\lambda + \ell(\psi,\lambda)\} |\hat{j}(s,t)|^{-\frac{1}{2}}$$

$s \mid t^0$ Exponential in ψ Use SP



$\hat{\lambda}_\psi$ free of s
$\hat{j}_{\lambda\lambda}$ free of s

Case: $\dim\psi = \dim\lambda = 1$

$\hat{j}$ factors Cond'l info $= \hat{j}_{\lambda\lambda}$
Profile info $= \hat{j}^p_{\psi\psi}$

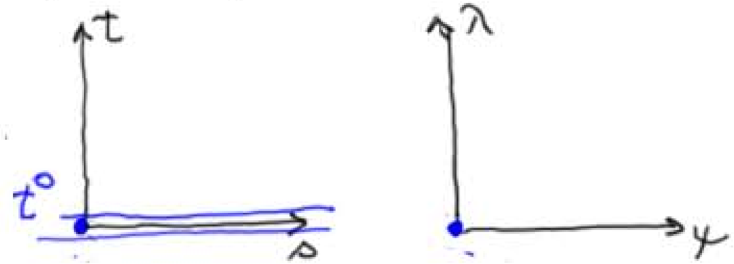
SP gives

$$f(s \mid t^0) = \frac{e^{k/n}}{(2\pi)^{1/2}} e^{-\frac{r_p^2(\psi, s)}{2}} \left| \hat{j}^p_{\psi\psi} \right|^{-1/2}$$

1st: t^0 density
2nd: Likelihood adj.

6 Conditional & Marginal for linear component parameter

$$f(s,t;\psi,\lambda) = \frac{e^{k/n}}{(2\pi)^{p/2}} \exp\{s\psi + t\lambda + \ell(\psi,\lambda)\} \, |\hat{\jmath}(s,t)|^{-\frac{1}{2}}$$



$s \mid t^0$ Exponential in ψ Use SP ↑

Case: $\dim\psi = \dim\lambda = 1$ $\hat{\jmath}$ factors Condn'l info $= \hat{\jmath}_{\lambda\lambda}$

Profile info $= \hat{\jmath}^{p}_{\psi\psi}$

$\hat{\lambda}_\psi$ free of s

$\hat{\jmath}_{\lambda\lambda}$ free of s

SP gives

$$f(s \mid t^0) = \frac{e^{k/n}}{(2\pi)^{1/2}} e^{-\frac{r_p^2(\psi,s)}{2}} \left| \hat{\jmath}^{p}_{\psi\psi} \right|^{-1/2}$$

1st: t^0 density
2nd: Likelihood adj.

Q: and $= f(S)$ for $S = ?$ A: for linear ψ, available as

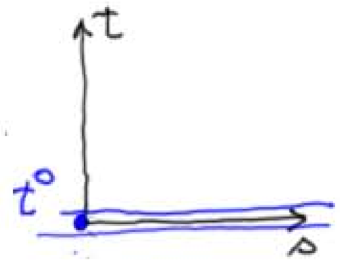
1) Inference for ψ Condn'l and marginal same (3rd) for linear parameters

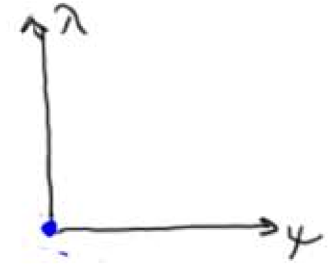
composite of conditional $s \mid t$

6 Conditional & Marginal for linear component parameter

$$f(s,t;\psi,\lambda) = \frac{e^{k/n}}{(2\pi)^{p/2}} \exp\{s\psi + t\lambda + \ell(\psi,\lambda)\} |\hat{j}(s,t)|^{-\frac{1}{2}}$$

$s \mid t^0$ Exponential in ψ · Use SP





Case: $\dim\psi = \dim\lambda = 1$

$\hat{j}$ factors Cond'l info $= \hat{j}_{\lambda\lambda}$

Profile info $= \hat{j}^{p}_{\psi\psi}$

$\hat{\lambda}_\psi$ free of s

$\hat{j}_{\lambda\lambda}$ free of s

SP gives

$$f(s \mid t^0) = \frac{e^{k/n}}{(2\pi)^{1/2}} e^{-\frac{r_p^2(\psi,s)}{2}} \left| \hat{j}^{p}_{\psi\psi} \right|^{-1/2}$$

1st: t^0 density

2nd: Likelihood adj.

1) Inference for ψ Cond'l and marginal same (3rd)

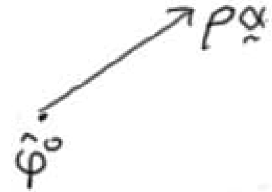
2) Bayes for ψ $|\hat{j}_{\psi\psi}(\psi,\hat{\lambda}_\psi)|^{1/2} d\psi$ Conditional same (2nd) for lin. par.

Same by Welch-Peers

But also in mod. devns $O(n^{-1})$ available from $\ell(\varphi)$; see §7 ahead

7 Directional Jeffreys: $O(n^{-1})$ for linear parameters

$$\varphi = (\varphi_1, \ldots, \varphi_p)$$
$$= \underset{\text{data}}{\hat{\varphi}^0} + \rho \underset{\sim}{\alpha}$$



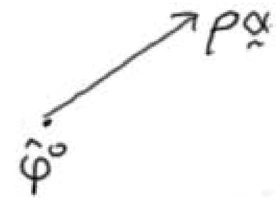
7 Directional Jeffreys: $O(n^{-1})$ for linear parameters

$\varphi = (\varphi_1, \dots, \varphi_p)$

$= \underset{\text{data}}{\hat{\varphi}^0} + \rho \underset{\sim}{\alpha}$

Polar coord's; direction $\underset{\sim}{\alpha}$

"Rotated φ" as ρ; data $\hat{\rho}^0 = 0$



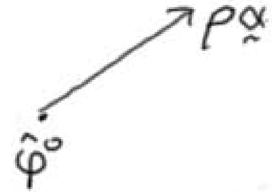
7 Directional Jeffreys: $O(n^{-1})$ for linear parameters

$\varphi = (\varphi_1, \ldots, \varphi_p)$

$= \underset{\text{data}}{\hat{\varphi}^0} + \rho \underset{\sim}{\alpha}$

Polar coord's; direction $\underset{\sim}{\alpha}$

"Rotated φ" as ρ; data $\hat{\rho}^0 = 0$



From preceding: Marginal distn re ρ = Condnl re ρ

& both equal the Bayes by Welch-Peers

$$-\frac{\partial^2 \ell(\varphi)}{\partial \rho^2}$$

- Canonical coordinates
- Diff twice re ρ
  with ρ defined re $\hat{\varphi}^0$ (original φ coords)

7 Directional Jeffreys: $O(n^{-1})$ for linear parameters

$\varphi = (\varphi_1, \ldots, \varphi_p)$ Polar coord's; direction α $\quad \hat\varphi^0 \to \rho\alpha$

$= \underset{\text{data}}{\hat\varphi^0} + \rho\alpha$ "Rotated ψ" as ρ; data $\hat\rho^0 = 0$

From preceding: Marginal distn re ρ = Condnl re ρ

& both equal the Bayes by Welch-Peers

Bayes is $\pi(\theta) = j^{1/2}_{\rho\rho}(\rho\alpha)\cdot d\rho$ on line; repetition validity $O(n^{-1})$

Polar coords re obs mle in φ

Interpretation for "derivation": mod devns about $(s^0, \hat\varphi^0)$

thus include just $n^{-1/2}$ terms in expansions

Derivatives effectively of the cum. gen. fn.

Q: Would this be used only on line for particular ψ of interest?

A: That would be easiest integration!

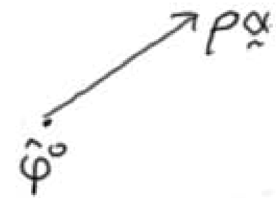
Q: Or used on full φ-space?

A: Indications are yes! Details elsewhere.

7 Directional Jeffreys: $O(n^{-1})$ for linear parameters

$\varphi = (\varphi_1, \ldots, \varphi_p)$ Polar coord's; direction $\underset{\sim}{\alpha}$

$= \underset{\text{data}}{\hat{\varphi}^0} + \rho \underset{\sim}{\alpha}$ "Rotated φ" as ρ; data $\hat{\rho}^0 = 0$



From preceding: Marginal distn re ρ = Condnl re ρ

& both equal the Bayes by Welch-Peers

Bayes is $\pi(\theta) = j_{\rho\rho}^{1/2}(\rho\underset{\sim}{\alpha}) \cdot d\rho$ on line repetition validity $O(n^{-1})$

Does right things: $N\ \mu\ \sigma^2$ $d\mu\, d\sigma/\sigma$

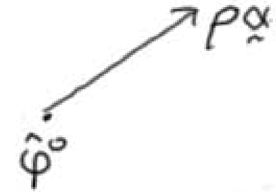
$y = X\beta + \sigma z$ $d\beta\, d\sigma/\sigma$

7 Directional Jeffreys: $O(n^{-1})$ for linear parameters

$\varphi = (\varphi_1, \dots, \varphi_p)$ Polar coord's; direction $\underset{\sim}{\alpha}$

$= \hat{\varphi}^0 + \rho\underset{\sim}{\alpha}$ "Rotated φ" as ρ; data $\hat{\rho}^0 = 0$

data



From preceding: Marginal distn re ρ = Condnl re ρ

& both equal the Bayes by Welch-Peers

Bayes is $\pi(\theta) = \int_{pee}^{1/2}(\rho\underset{\sim}{\alpha}) \cdot d\rho$ on line repetition validity $O(n^{-1})$

Does right things: $N\ \mu\ \sigma^2$ $d\mu\, d\sigma/\sigma$

$y = X\beta + \sigma z$ $d\beta\, d\sigma/\sigma$

In general: Data dependent ... necessarily so !

Priors need to be data - dependent to extract all available info from likelihood function.

Q Fieller-Creasy problem: any light on this?

A Creasy believed she was following Fisher but the angle (here $\underset{\sim}{\alpha}$) has no $-\infty, +\infty$

Don't think Creasy solution would fit present analysis

Why should a Bayesian "sell himself short" by using only a prior with definition free of the data point ?

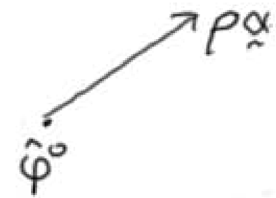
7 Directional Jeffreys: $O(n^{-1})$ for linear parameters

$\varphi = (\varphi_1, \ldots, \varphi_p)$

$= \underset{\text{data}}{\hat{\varphi}^0} + \rho \underset{\sim}{\alpha}$

Polar coord's; direction $\underset{\sim}{\alpha}$

"Rotated ψ" as ρ; data $\hat{\rho}^0 = 0$



From preceding: Marginal distn re ρ = Condnl re ρ

& both equal the Bayes by Welch-Peers

Bayes is $\pi(\theta) = j_{\rho\rho}^{1/2}(\rho\underset{\sim}{\alpha}) \cdot d\rho$ on line   repetition validity $O(n^{-1})$

Does right things: $N\ \mu\ \sigma^2 \qquad d\mu\, d\sigma/\sigma$

$y = X\beta + \sigma z \qquad d\beta\, d\sigma/\sigma$

In general: Data dependent ... necessarily so!

But just for linear parameters

Curved parameters ⇒ Bayes fails

AQ: Even with data-dependent priors? Yes! To work with curved parameters, the prior needs to be specially designed for the curved parameter: draft paper available.

See Section 8

8 Parameter curvature How Bayes misses!

$y_1 \; N \; \psi \; 1$ Interest in ψ Data y^0 Suppose $y_1^0 - \psi_0 = 1$

$y_2 \; N \; \lambda \; 1$ Nuisance λ Test ψ_0

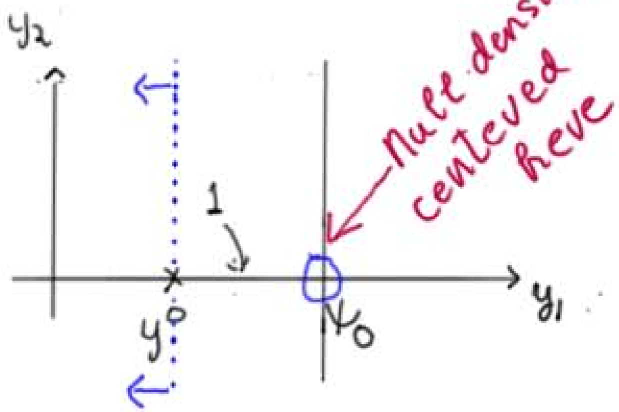
8 Parameter curvature How Bayes misses!

$y_1 \; N \; \psi \; 1$ Interest in ψ Data y^0

$y_2 \; N \; \lambda \; 1$ Nuisance λ Test ψ_0

Suppose $y_1^0 - \psi_0 = 1$

f



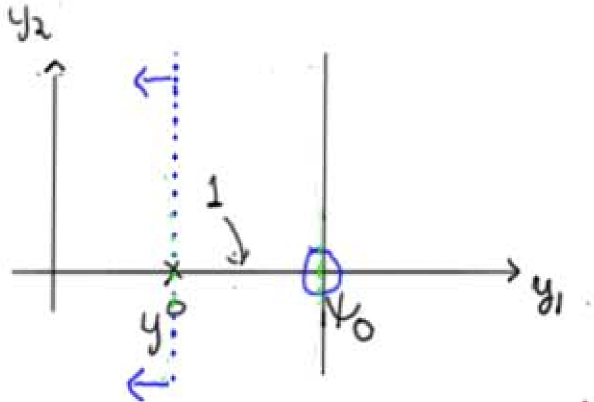
$p(\psi_0) = 17\%$

8 Parameter curvature How Bayes misses!

$y_1 \; N \; \psi \; 1$ Interest in ψ Data y° Suppose $y_1^\circ - \psi_0 = 1$

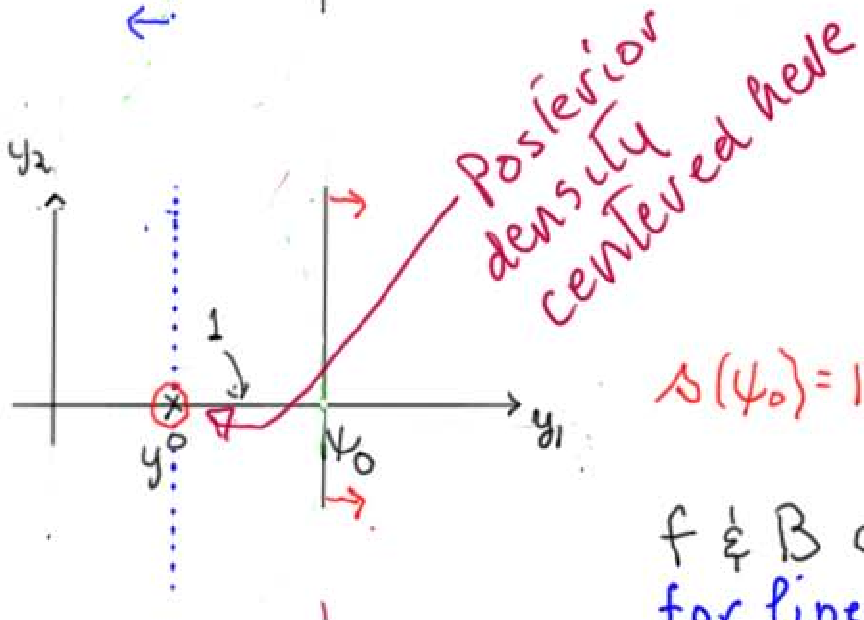
$y_2 \; N \; \lambda \; 1$ Nuisance λ Test ψ_0

f



$p(\psi_0) = 17\%$

B



$s(\psi_0) = 17\%$

f & B agree
for linear par.

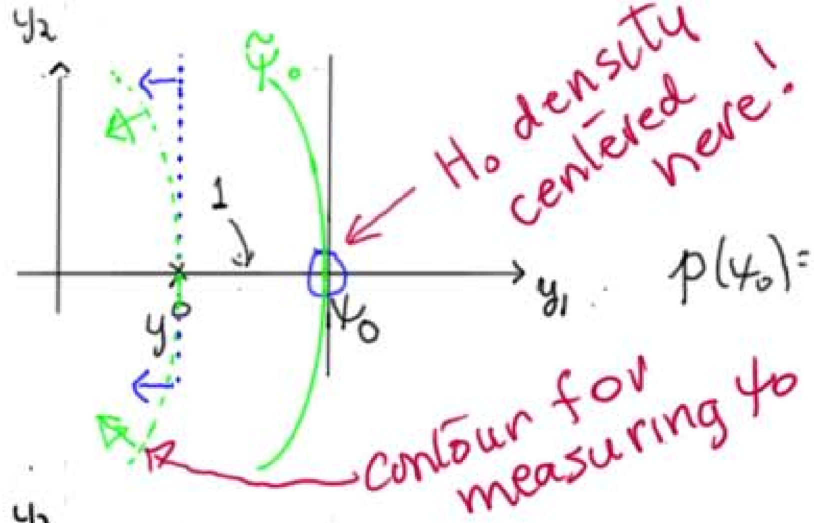
8 Parameter curvature How Bayes misses!

$y_1 \; N \; \psi \; 1$ Interest in ψ Data y^0 Suppose $y_1^0 - \psi_0 = 1$

$y_2 \; N \; \lambda \; 1$ Nuisance λ Test ψ_0

Now consider curved $\tilde{\psi} = \tilde{\psi}_0$

f

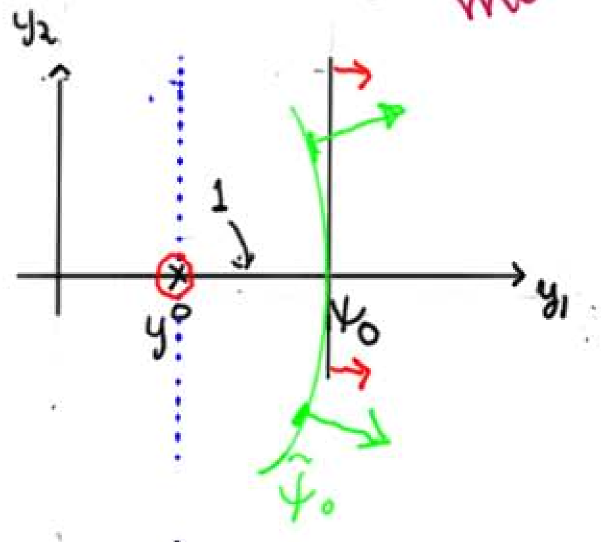


$p(\psi_0) = 17\%$

p-value decreases

$p(\tilde{\psi}_0) < 17\%$

B



$s(\psi_0) = 17^\circ$

f & B agree

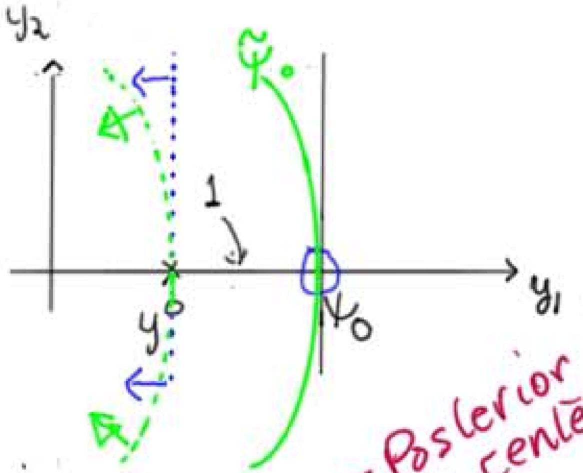
8 Parameter curvature How Bayes misses!

$y_1 \; N \; \psi \; 1$ Interest in ψ Data y^0 Suppose $y_1^0 - \psi_0 = 1$

$y_2 \; N \; \lambda \; 1$ Nuisance λ Test ψ_0

Now consider curved $\tilde{\psi} = \tilde{\psi}_0$

f

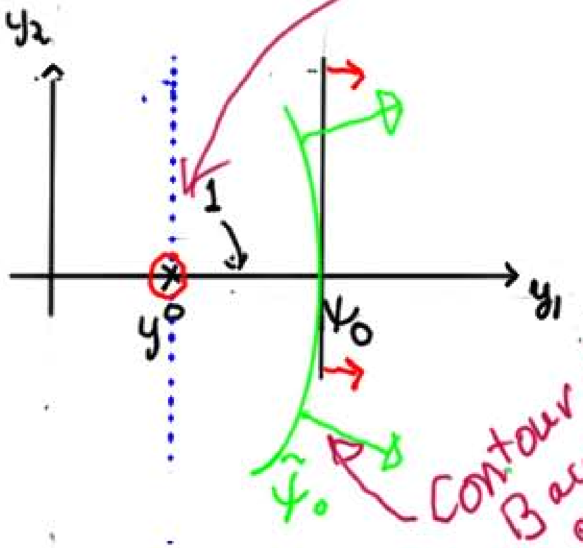


$p(\psi_0) = 17\%$ p-value decreases

$p(\tilde{\psi}_0) < 17\%$

Posterior centered here

B



$s(\psi_0) = 17°$ s-value increases

$s(\tilde{\psi}_0) > 17\%$

Contour for Bayes evaluation

f & B agree

f & B disagree

Go in different directions
Bayes is wrong
for curved parameters

1. Priors & Inference: Is there a problem?
2. Data-dependent prior? Why?
3. Welch Peers: Exponential as location model
4. Exponential model: as an approximation
5. Saddlepoint: for exponential & for general
6. Conditional / Marginal for linear parameter component
7. Directional Jeffreys
8. Curved parameter: How Bayes misses!

Thank you