How Saddlepoint and Continuity determine Statistical Inference

D A S Fraser
Statistical Sciences
Univ Toronto

Conference Honoring Malay Ghosh
College Park
2014 May 29-31

www.utstat.toronto.edu/dfraser/documents/Malay2014.pdf

/ xxx Some references

1 Familiar inference tools

2 Laplace

3 p^*

4 SP

5 Tangent Exponential Model

6 Test $\psi = \psi_0$

7 Test scalar ψ

What can SP do for Bayes ?

8 a) Welch - Peers

b) Example 1

c) Bayes : Vector parameter

d) Ex 2 Ex 3 Ex 4 (linear, rotating, curved)

Discussion

1 Familiar Inference tools:

1 Standardized departure

- Departure: Estimate - parameter $= t - \theta$

1 Familiar Inference tools:

1 Standardized departure

- Departure: Estimate - parameter $= t - \theta$

- Standardize: Stz Dep $= (t - \theta)/SD$

(1) Familiar Inference tools:

1 Standardized departure

- Departure: Estimate - parameter $= t - \theta$
- Standardize: StzDep $= (t-\theta)/SD$
- Stat. position: $p(\theta) = \Phi(\text{StzDep})$ ... Asy Normality

  $= P_{BS}(\text{StzDep})$ ... Bootstrap

1 Familiar Inference tools:

1 Standardized departure

Regression Model

- Departure: Estimate - parameter $= t - \theta$ $\qquad b_r - \beta_r$
- Standardize: St Dep = $(t - \theta)/SD$ $\qquad (b_r - \beta_r)/s_{b_r}$
- Stat. position: $p(\theta) = \Phi(\text{StDep})$ $\qquad \Phi\,(b_r - \beta_r)/s_{b_r}$

  $= P_{BS}(\text{StDep})$

(1) Familiar Inference tools:

1 Standardized departure

Regression Model

- Departure: Estimate - parameter $= t - \theta$ $\qquad b_r - \beta_r$

- Standardize: StzDep $=$ $(t-\theta)/SD$ $\qquad (b_r - \beta_r)/s_{b_r}$

- Stat. position: $p(\theta) = \Phi(\text{StzDep})$ $\qquad \Phi\,(b_r - \beta_r)/s_{b_r}$

  $= P_{BS}(\text{StzDep})$

2 Weighted Likelihood

- Likelihood $L(\theta) = c f(y^\circ; \theta)$

(1) Familiar Inference tools:

1 Standardized departure

Regression Model

- Departure: Estimate - parameter $= t - \theta$ $\qquad b_r - \beta_r$

- Standardize: Stz Dep $= (t-\theta)/SD$ $\qquad (b_r - \beta_r)/s_{b_r}$

- Stat. position: $p(\theta) = \Phi(\text{StzDep})$ $\qquad \Phi\,(b_r - \beta_r)/s_{b_r}$

  $= P_{BS}(\text{StzDep})$

2 Weighted Likelihood

- Likelihood $\quad L(\theta) = c f(y^\circ; \theta)$

- Weighted Likelihood $\quad \pi(\theta; y^\circ) L(\theta)$

  (To get "better info from likelihood")

1 Familiar Inference tools:

1 Standardized departure

Regression Model

- Departure: Estimate - parameter $= t - \theta$ — $b_r - \beta_r$
- Standardize: Stz Dep $= (t-\theta)/SD$ — $(b_r - \beta_r)/s_{b_r}$
- Stat. position: $p(\theta) = \Phi(\text{StzDep})$ — $\Phi\,(b_r - \beta_r)/s_{b_r}$

  $= P_{BS}(\text{StzDep})$

2 Weighted Likelihood

- Likelihood $L(\theta) = c f(y^\circ; \theta)$
- Weighted Likelihood $\pi(\theta; y^\circ) L(\theta)$

  (To get "better info from likelihood")

3 Other Sufficiency ?

- Glamorous for the Normal

(1) Familiar Inference tools:

1 Standardized departure

Regression Model

- Departure: Estimate - parameter $= t - \theta$ $\qquad b_r - \beta_r$

- Standardize: Stz Dep $\qquad (t-\theta)/SD$ $\qquad (b_r - \beta_r)/s_{b_r}$

- Stat. position: $p(\theta) = \Phi(\text{StzDep})$ $\qquad \Phi\, (b_r - \beta_r)/s_{b_r}$
$= P_{BS}(\text{StzDep})$

2 Weighted Likelihood

- Likelihood $\quad L(\theta) = c f(y^\circ; \theta)$

- Weighted Likelihood $\quad \pi(\theta; y^\circ) L(\theta)$
(To get "better info from likelihood")

3 Other Sufficiency ?

- Glamorous for the Normal
- Useless otherwise ?

1 Familiar Inference tools:

1 Standardized departure

			Regression Model
- Departure:	Estimate - parameter	$= t - \theta$	$b_r - \beta_r$
- Standardize:	Stz Dep =	$(t-\theta)/SD$	$(b_r - \beta_r)/s_{b_r}$
- Stat. position:	$p(\theta) = \Phi(\text{StzDep})$ $= P_{BS}(\text{StzDep})$		$\Phi\,(b_r - \beta_r)/s_{b_r}$

2 Weighted Likelihood

- Likelihood $\quad L(\theta) = c f(y^\circ; \theta)$

- Weighted Likelihood $\quad \pi(\theta; y^\circ) L(\theta)$

(To get "better info from likelihood")

3 Other Sufficiency?

- glamorous for the Normal
- useless otherwise?

4 Accuracy

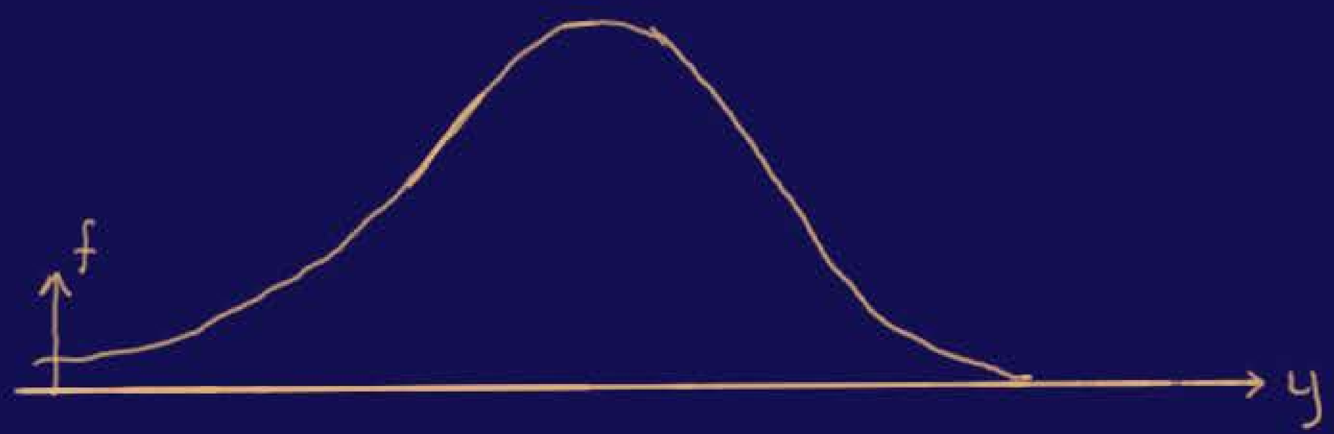
Mostly $O(n^{-1/2})$

Omnibus, but...

(2) Laplace: "substance"

Have a function $f(y)$

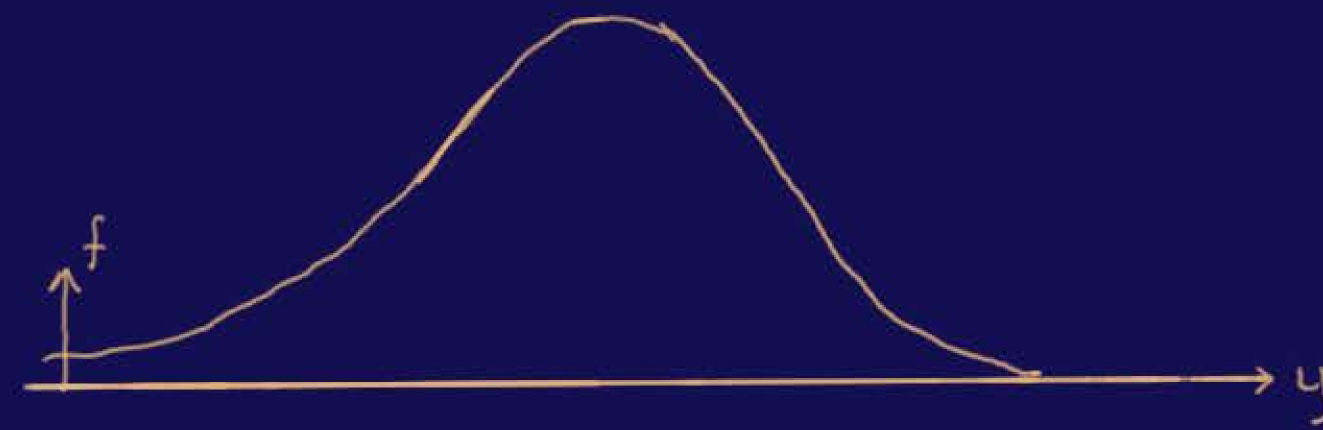
Want to integrate it:



(2) Laplace: "substance"

Have a function $f(y)$

Want to integrate it:



Assume: - $\log f(y)$ is $O(n)$

- smooth
- unique max

② Laplace: "substance"

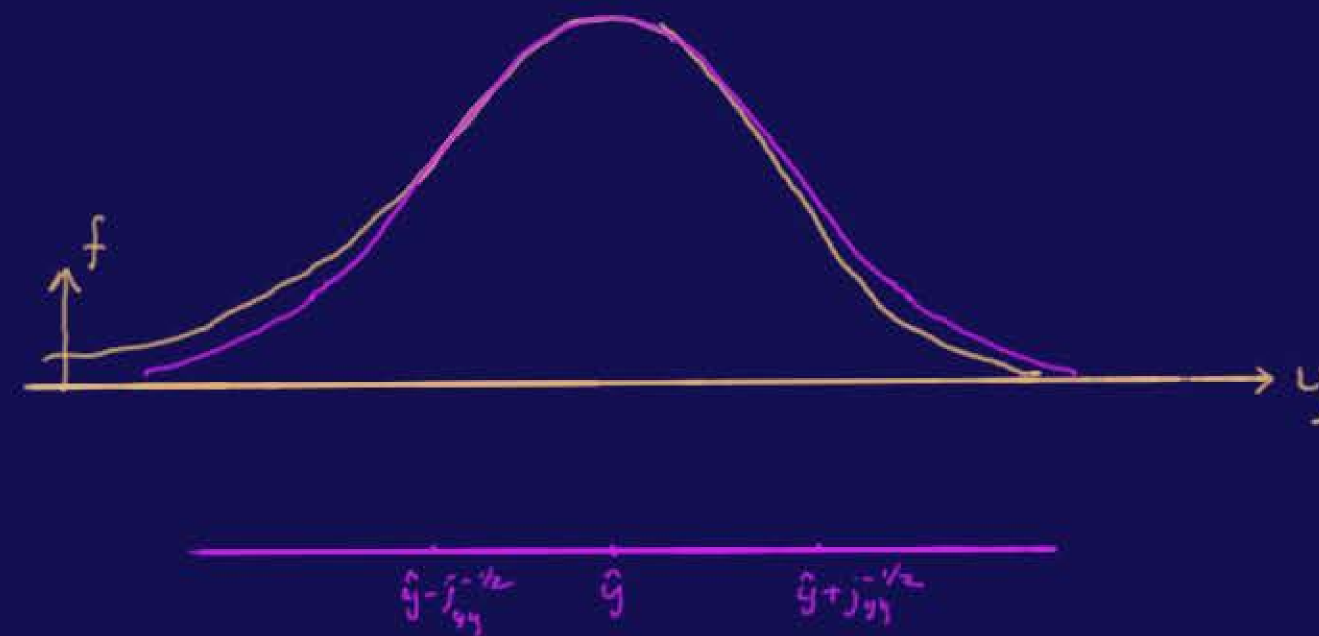
Have a function $f(y)$

Want to integrate it:

Assume: - $\log f(y)$ is $O(n)$

- smooth
- unique max

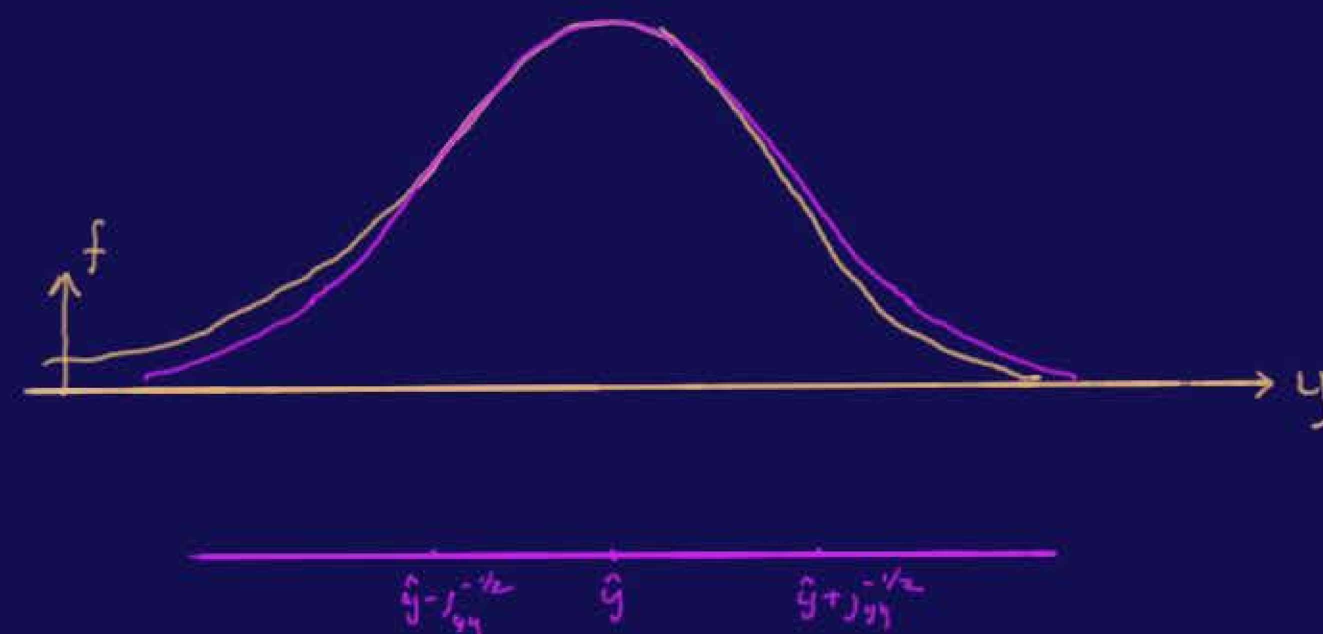
Fit Normal at max



② Laplace: "substance"

Have a function $f(y)$

Want to integrate it:



Assume: - $\log f(y)$ is $O(n)$

- smooth
- unique max

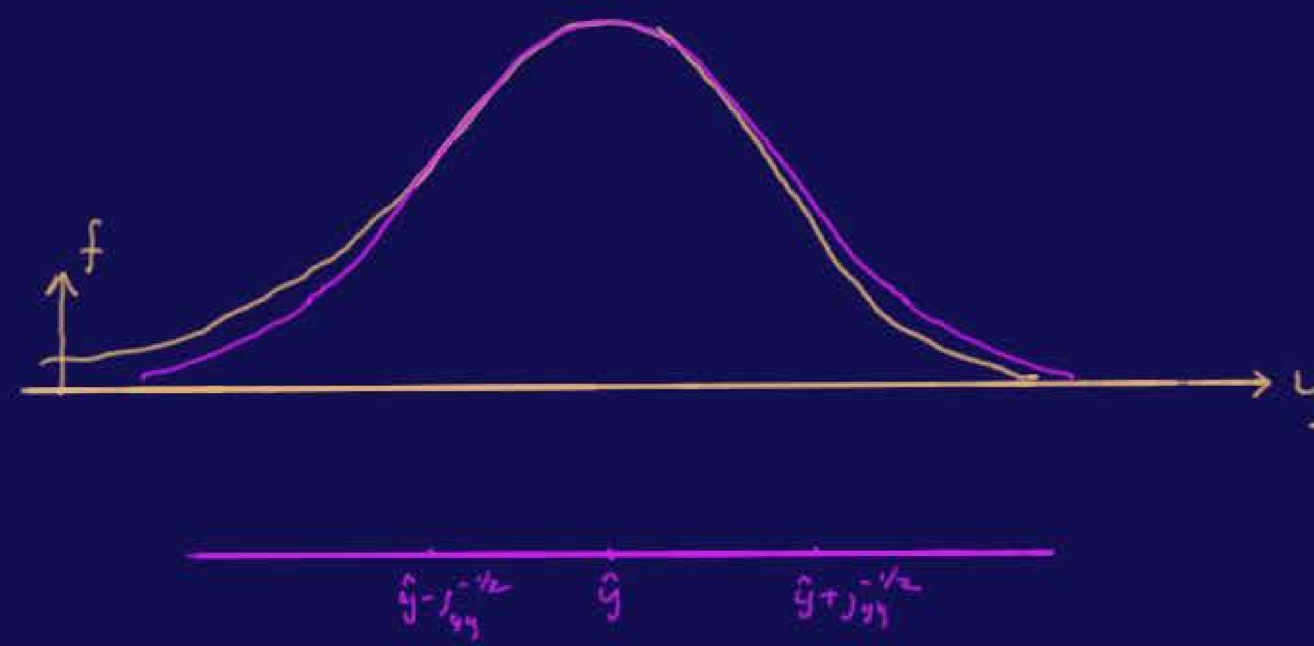
Fit Normal at max

Normal: Height $= \dfrac{|\hat{J}_{yy}|^{1/2}}{(2\pi)^{p/2}}$ Integral $= 1$

(2) Laplace: "substance"

Have a function $f(y)$

Want to integrate it:



Assume: - $\log f(y)$ is $O(n)$
- smooth
- unique max

Fit Normal at max

Normal: Height $= \dfrac{|\hat{J}_{yy}|^{1/2}}{(2\pi)^{p/2}}$ Integral $= 1$

fn $f(y)$: Height $= f(\hat{y})$ Integral $= \int f(y)\, dy$

(2) Laplace: "substance"

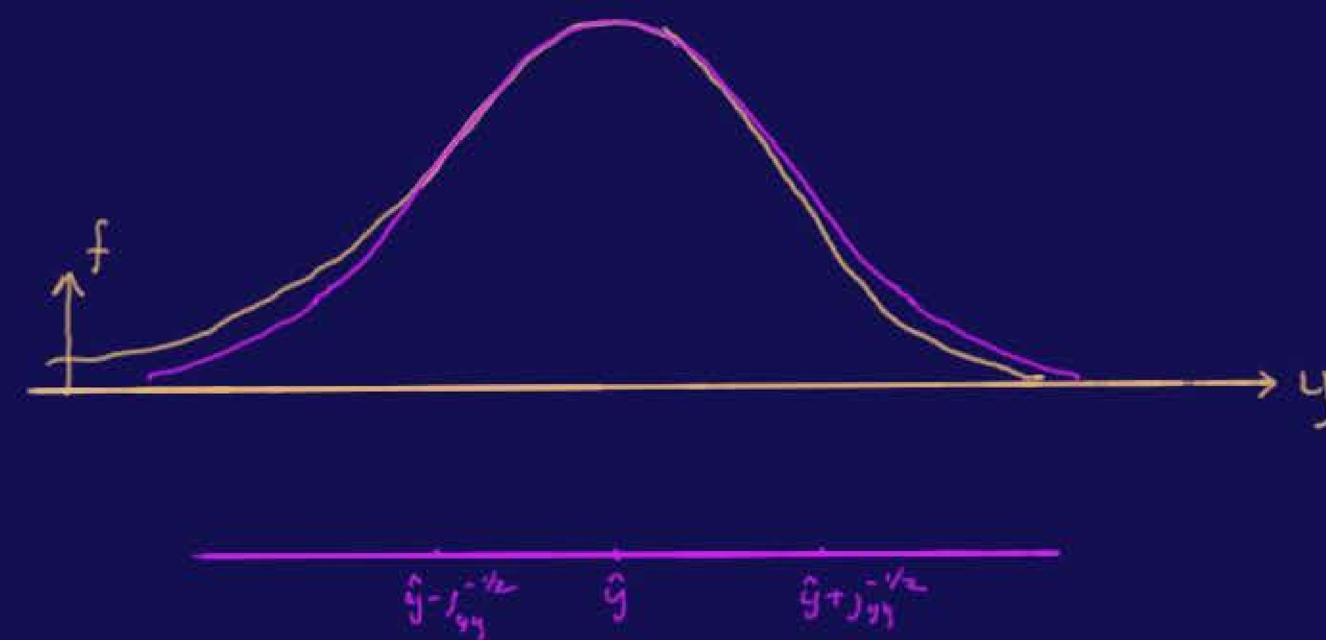
Have a function $f(y)$

Want to integrate it:

Assume: - $\log f(y)$ is $O(n)$

- smooth
- unique max

Fit Normal at max



Normal: Height $= \dfrac{|\hat{j}_{yy}|^{1/2}}{(2\pi)^{p/2}}$ Integral $= 1$

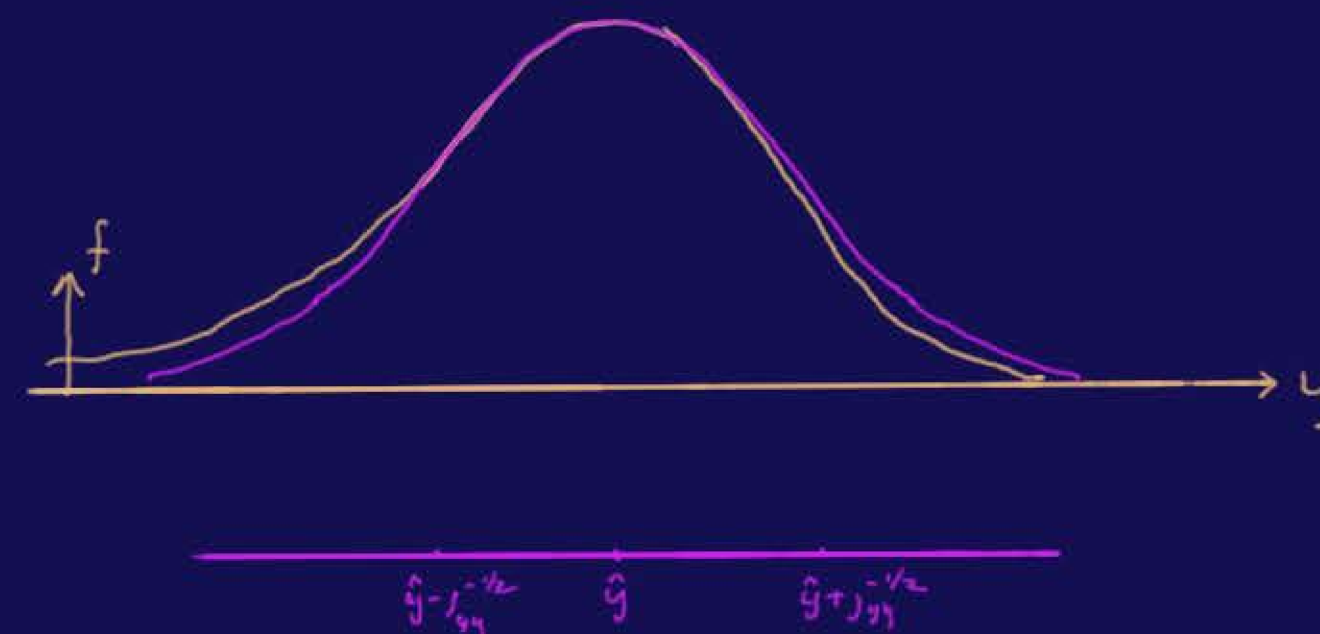
Fn $f(y)$: Height $= f(\hat{y})$ Integral $= \int f(y)\,dy$

Normal approx: $\boxed{\int f(y)\,dy = f(\hat{y}) \dfrac{(2\pi)^{p/2}}{|\hat{j}_{yy}|^{1/2}}}$ $\Leftarrow$ Laplace Integral (1774)

② Laplace: "substance"

Have a function $f(y)$

Want to integrate it:



Assume:
- $\log f(y)$ is $O(n)$
- smooth
- unique max

Fit Normal at max

Normal: Height $= \dfrac{|\hat{J}_{yy}|^{1/2}}{(2\pi)^{p/2}}$ Integral $= 1$

Fn $f(y)$: Height $= f(\hat{y})$ Integral $= \int f(y)\,dy$

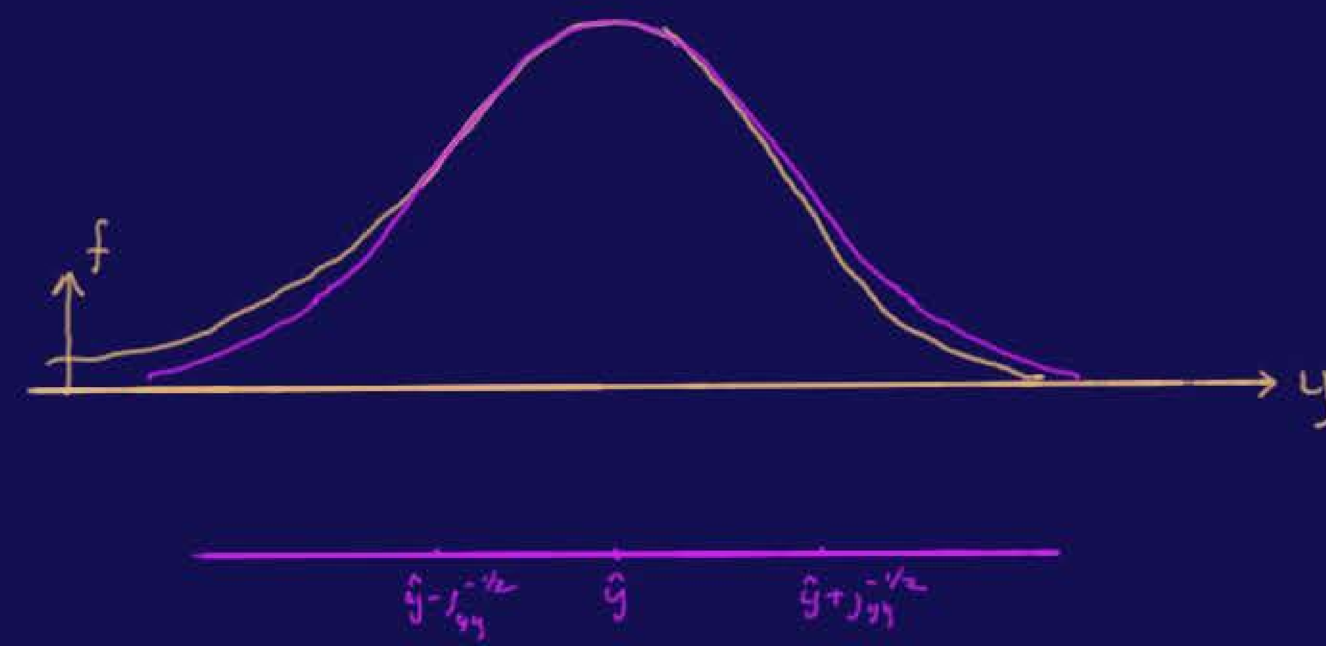
Normal approx: $\boxed{\int f(y)\,dy = f(\hat{y}) \dfrac{(2\pi)^{p/2}}{|\hat{J}_{yy}|^{1/2}}}$ $\Leftarrow$ Laplace Integral (1774)

Why does it work? $c\, e^{-z^2/2 + a_3 z^3/6n^{1/2} + a_4 z^4/24n}$

(2) Laplace: "substance"

Have a function $f(y)$

Want to integrate it:



Assume: - $\log f(y)$ is $O(n)$

- smooth
- unique max

Fit Normal at max

Normal: Height $= \dfrac{|\hat{J}_{yy}|^{1/2}}{(2\pi)^{p/2}}$ Integral $= 1$

Fn $f(y)$: Height $= f(\hat{y})$ Integral $= \int f(y)\,dy$

Normal approx: $\boxed{\int f(y)\,dy = f(\hat{y})\dfrac{(2\pi)^{p/2}}{|\hat{J}_{yy}|^{1/2}}}$ $\Leftarrow$ Laplace Integral (1774)

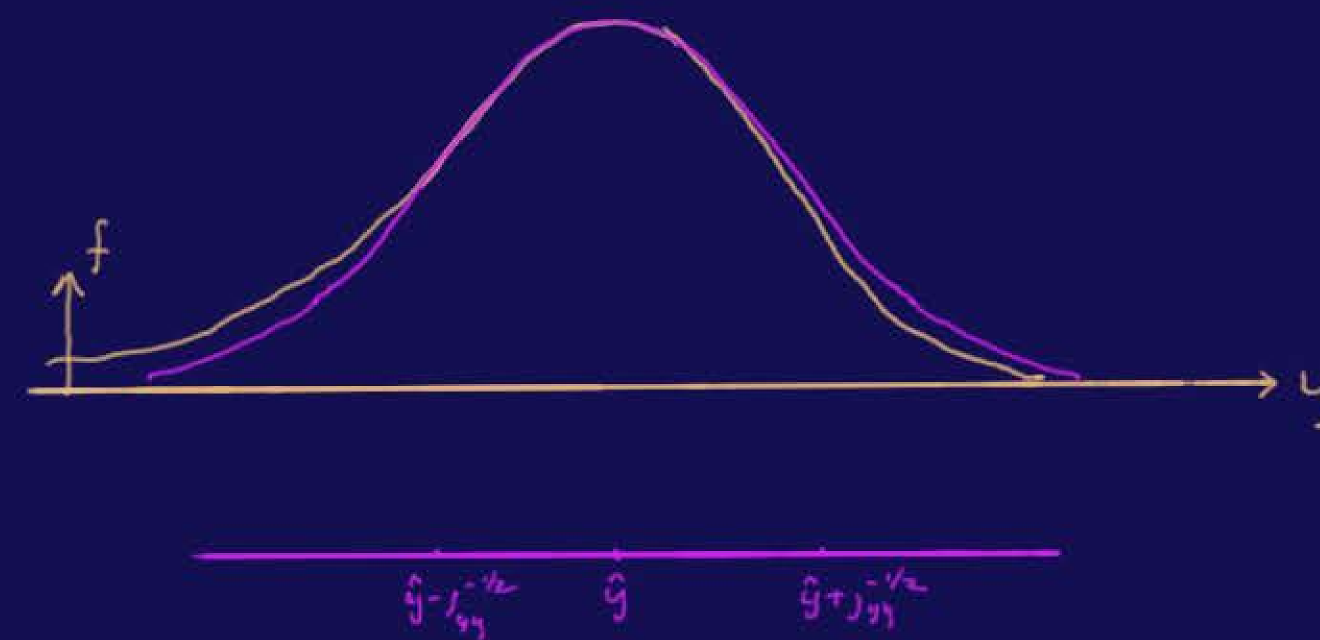
Why does it work? $c\, e^{-z^2/2 + a_3 z^3/6n^{1/2} + a_4 z^4/24n}$

$= c\, e^{-z^2/2}\{1 + a_3 z^3/6n^{1/2} + a_4 z^4/24n + a_3^2 z^6/72n \ldots\}$

② Laplace: "substance"

Have a function $f(y)$

Want to integrate it:



Assume: - $\log f(y)$ is $O(n)$
- smooth
- unique max

Fit Normal at max

Normal: Height $= \dfrac{|\hat{j}_{yy}|^{1/2}}{(2\pi)^{p/2}}$ Integral $= 1$

Fn $f(y)$: Height $= f(\hat{y})$ Integral $= \int f(y)\,dy$

Normal approx: $\boxed{\int f(y)\,dy = f(\hat{y})\dfrac{(2\pi)^{p/2}}{|\hat{j}_{yy}|^{1/2}}}$ $\Leftarrow$ Laplace Integral (1774)

Why does it work? $c\, e^{-z^2/2 + a_3 z^3/6n^{1/2} + a_4 z^4/24n}$

$= c\, e^{-z^2/2}\{1 + a_3 z^3/6n^{1/2} + a_4 z^4/24n + a_3^2 z^6/72n \ldots\}$

$= e^{-\frac{3a_4 + 5a_3^2}{24n}}\,\phi(z)\, e^{a_3 z^3/6n^{1/2} + a_4 z^4/24n}$ $O(n^{-2})$ $\leftarrow$

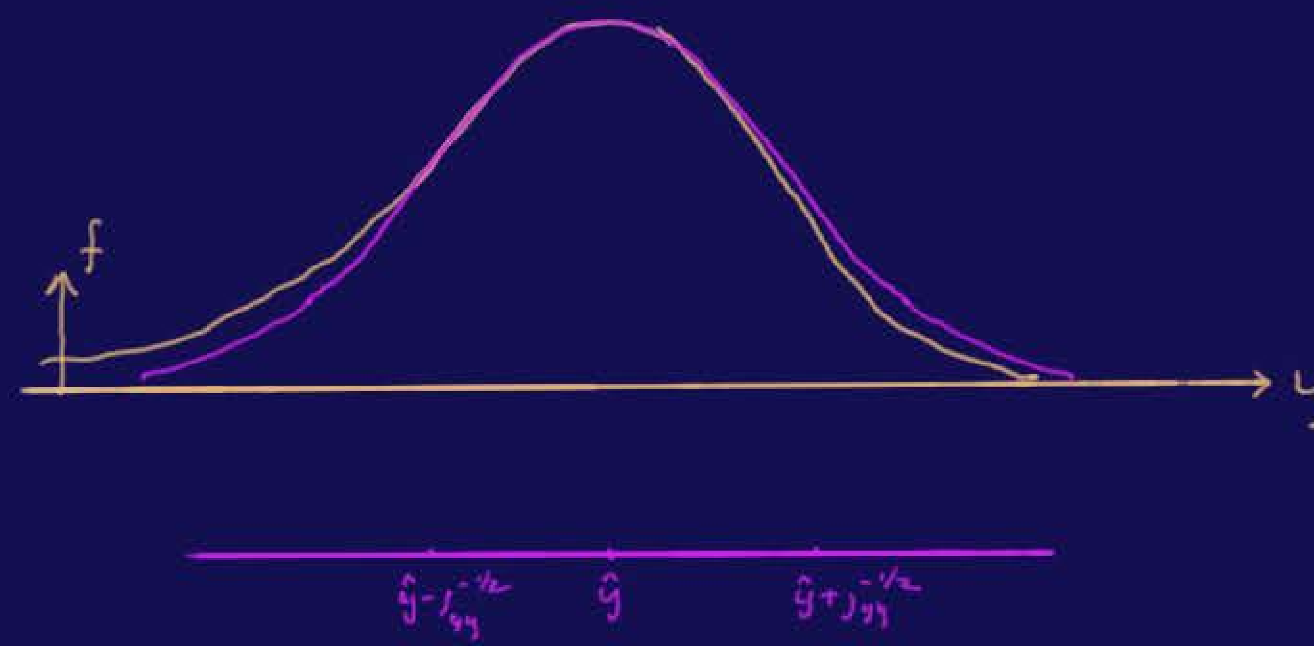
$N(0,1)$
$E(z^3) = 0$
$E(z^4) = 3$
$E(z^6) = 5 \cdot 3$

② Laplace: "substance"

Have a function $f(y)$

Want to integrate it:

Assume: - $\log f(y)$ is $O(n)$
- smooth
- unique max



Fit Normal at max

Normal: Height $= \dfrac{|\hat{j}_{yy}|^{1/2}}{(2\pi)^{p/2}}$ Integral $= 1$

Fn $f(y)$: Height $= f(\hat{y})$ Integral $= \int f(y)\,dy$

Normal approx: $\boxed{\int f(y)\,dy = f(\hat{y}) \dfrac{(2\pi)^{p/2}}{|\hat{j}_{yy}|^{1/2}}}$ $\Leftarrow$ Laplace Integral (1774)

Why does it work? $c\, e^{-z^2/2 + a_3 z^3/6n^{1/2} + a_4 z^4/24n}$

$= c\, e^{-z^2/2} \{ 1 + a_3 z^3/6n^{1/2} + a_4 z^4/24n + a_3^2 z^6/72n \ldots \}$

$= e^{\frac{-3a_4 + 5a_3^2}{24n}} \phi(z)\, e^{a_3 z^3/6n^{1/2} + a_4 z^4/24n}$ $O(n^{-2})$

$\leftarrow$ $N(0,1)$: $E(z^3) = 0$, $E(z^4) = 3$, $E(z^6) = 5 \cdot 3$

Simple Accurate Powerful

3 p^* Have a model $f(y;\theta)$: Want to integrate it!

③ p^* Have a model $f(y;\theta)$: Want to integrate it!

assume - $\log f(y;\theta)$ is $O(n)$
- smooth
- unique max

③ p^* Have a model $f(y;\theta)$: Want to integrate it!

assume - $\log f(y;\theta)$ is $O(n)$

- smooth
- unique max

dim $y = p$

- Marginally ...
- Condn'lly given $a(y)$ $\begin{cases} \text{distn} \\ \text{free of } \theta \end{cases}$

③ p^* Have a model $f(y;\theta)$: Want to integrate it!

assume - $\log f(y;\theta)$ is $O(n)$ $\quad$ $\dim y = p$ $\quad$ · Marginally .... or

- smooth $\quad$ · Condn'lly given $a(y)$ $\left\{ \begin{array}{l} \text{distn free} \\ \text{of } \theta \end{array} \right.$

- unique max

Write in terms of $\hat{\theta}$:

$$g(\hat{\theta};\theta)$$

③ p^* Have a model $f(y;\theta)$: Want to integrate it!

assume - $\log f(y;\theta)$ is $O(n)$
- smooth
- unique max

dim $y = p$ $\left|\begin{array}{l} \cdot \text{ \underline{Marginally} ... or} \\ \cdot \text{ \underline{Condn'lly} given } a(y) \begin{cases} \text{distn free} \\ \text{of } \theta \end{cases} \end{array}\right.$

Write in terms of $\hat{\theta}$:

$g(\hat{\theta};\theta)$ Try approx: $= c\exp\{\ell(\theta;\hat{\theta}) - \ell(\hat{\theta};\hat{\theta})\}\dots$

③ p^* Have a model $f(y;\theta)$: Want to integrate it!

assume - $\log f(y;\theta)$ is $O(n)$

- smooth
- unique max

dim $y = p$ | · Marginally .... or
| · Cond'n'lly given $a(y)$ $\left\{ \begin{array}{l} \text{dist'n free} \\ \text{of } \theta \end{array} \right.$

Get likelihood right for each $\hat{\theta}$, re max value

| Just: std log lik ratio

Write in terms of $\hat{\theta}$:

$g(\hat{\theta};\theta)$ Try approx: $c\exp\{\ell(\theta;\hat{\theta}) - \ell(\hat{\theta};\hat{\theta})\}$ ...

③ p^* Have a model $f(y;\theta)$: Want to integrate it!

assume - $\log f(y;\theta)$ is $O(n)$ $\dim y = p$

- smooth
- unique max

$\cdot$ Marginally ... or

$\cdot$ Condn'lly given $a(y)$ $\left\{ \begin{array}{l} \text{distn free} \\ \text{of } \theta \end{array} \right.$

Write in terms of $\hat{\theta}$:

Get likelihood right for each $\hat{\theta}$, re max value

Just std log lik ratio

$g(\hat{\theta};\theta)$ Try approx: $c\exp\{\ell(\theta;\hat{\theta}) - \ell(\hat{\theta};\hat{\theta})\}\ldots$

Use Laplace $\theta = \hat{\theta}$:
so $\int g(\hat{\theta};\theta)\,d\hat{\theta} = 1$

get: $\dfrac{e^{k/n}}{(2\pi)^{p/2}} e^{\ell(\theta)-\ell(\hat{\theta})} \left| J_{\hat{\theta}\hat{\theta}}(\hat{\theta}) \right|^{1/2} \cdot d\hat{\theta}$ $O(n^{-3/2})$

(3) p^* Have a model $f(y;\theta)$: Want to integrate it!

assume - $\log f(y;\theta)$ is $O(n)$

- smooth
- unique max

dim $y = p$

- Marginally ... or
- Cond'n'lly given $a(y)$ $\left\{ \begin{array}{l} \text{dist'n free} \\ \text{of } \theta \end{array} \right.$

Write in terms of $\hat\theta$:

Get likelihood right for each $\hat\theta$, re max value

Just std log lik ratio

$g(\hat\theta;\theta)$ Try approx: $c\exp\{\ell(\theta;\hat\theta) - \ell(\hat\theta;\theta)\}\ldots$

Use Laplace $\theta = \hat\theta$, so $\int g(\hat\theta;\theta)\,d\hat\theta = 1$

get: $$\frac{e^{k/n}}{(2\pi)^{p/2}} e^{\ell(\theta)-\ell(\hat\theta)} \left| J_{\hat\theta\hat\theta}(\hat\theta) \right|^{1/2} \cdot d\hat\theta \qquad O(n^{-3/2})$$

Minor calculation

$$\boxed{\frac{e^{k/n}}{(2\pi)^{p/2}} e^{\ell(\theta)-\ell(\hat\theta)} \left| J_{\theta\theta}(\hat\theta) \right|^{1/2} \cdot d\hat\theta} \Longleftarrow p^* \text{ formula}$$

Barndorff-Nielsen 1981

(3) p^* Have a model $f(y;\theta)$: Want to integrate it!

assume - $\log f(y;\theta)$ is $O(n)$ dim $y = p$ | · Marginally ... or
- smooth | · Cond'n'lly given $a(y)$ { distn free of θ
- unique max

Get likelihood right for each $\hat\theta$, re max value | Just std log lik ratio

Write in terms of $\hat\theta$:

$g(\hat\theta;\theta)$ Try approx: $c\exp\{\ell(\theta;\hat\theta)-\ell(\hat\theta;\theta)\}\ldots$

Use Laplace $\theta=\hat\theta$, so $\int g(\hat\theta;\theta)\,d\hat\theta=1$

get: $$\frac{e^{k/n}}{(2\pi)^{p/2}}\, e^{\ell(\theta)-\ell(\hat\theta)}\left|J_{\hat\theta\hat\theta}(\hat\theta)\right|^{1/2}\cdot d\hat\theta \qquad O(n^{-3/2})$$

Minor calculations

$$\boxed{\frac{e^{k/n}}{(2\pi)^{p/2}}\, e^{\ell(\theta)-\ell(\hat\theta)}\left|J_{\theta\theta}(\hat\theta)\right|^{1/2}\cdot d\hat\theta} \quad \Longleftarrow \quad p^* \text{ formula}$$

Barndorff-Nielsen 1981

Simple Accurate Powerful?

③ p^* Have a model $f(y;\theta)$: Want to integrate it!

assume - $\log f(y;\theta)$ is $O(n)$
- smooth
- unique max

dim $y = p$ · Marginally .... or

· Condn'lly given $a(y)$ $\left\{ \begin{array}{l} \text{distn free} \\ \text{of } \theta \end{array} \right.$

Get likelihood right for each $\hat{\theta}$, re max value

Just std log lik ratio

Write in terms of $\hat{\theta}$:

$g(\hat{\theta};\theta)$ Try approx: $c\exp\{\ell(\theta;\hat{\theta}) - \ell(\hat{\theta};\hat{\theta})\} \ldots$

Use Laplace $\theta = \hat{\theta}$, so $\int g(\hat{\theta};\theta)\,d\hat{\theta} = 1$

get: $\dfrac{e^{k/n}}{(2\pi)^{p/2}} e^{\ell(\theta) - \ell(\hat{\theta})} \left| J_{\hat{\theta}\hat{\theta}}(\hat{\theta}) \right|^{1/2} \cdot d\hat{\theta}$ $O(n^{-3/2})$

Minor calculations

$$\boxed{\frac{e^{k/n}}{(2\pi)^{p/2}} e^{\ell(\theta) - \ell(\hat{\theta})} \left| J_{\theta\theta}(\hat{\theta}) \right|^{1/2} \cdot d\hat{\theta}}$$

$\Longleftarrow$ p^* formula Barndorff-Nielsen 1981

Simple Accurate Powerful?

But: How to get $J_{\theta\theta}(\hat{\theta})$ for points other than data y°?

Widely: Not available!

Limited usefulness! ... but

4 Saddlepoint: for full exponential models

$$f(y;\theta) = \exp\{\varphi'(\theta)\,u(y) - k(\theta)\}\,h(y)$$

④ Saddlepoint : for full exponential models

$$f(y;\theta) = \exp\left\{\varphi'(\theta)u(y) - k(\theta)\right\} h(y)$$

$$q(u;\varphi) = \exp\left\{\varphi' u - K(\varphi)\right\} g(u)$$

- Can rewrite with $g(u)$ as a pdf:

④ Saddlepoint: for full exponential models

$$f(y;\theta) = \exp\{\varphi'(\theta)u(y) - k(\theta)\}\, h(y)$$

$$q(u;\varphi) = \exp\{\varphi' u - K(\varphi)\}\, g(u)$$

- Can rewrite with $g(u)$ as a pdf:

$\exp\{K(t)\}$ = mgf of $g(u)$ $\qquad$ $K(t)$ = cgf of $g(u)$

(4) Saddlepoint : for full exponential models

$$f(y;\theta) = \exp\{\varphi'(\theta)u(y) - k(\theta)\}\, h(y)$$

$$g(u;\varphi) = \exp\{\varphi' u - K(\varphi)\}\, g(u)$$

- Can rewrite with $g(u)$ as a pdf:

$\exp\{K(t)\}$ = mgf of $g(u)$ $\qquad$ $K(t)$ = cgf of $g(u)$

"Saddlepoint approx": <u>Apply p^*</u>

$$SP \Rightarrow \frac{e^{k/n}}{(2\pi)^{p/2}}\, e^{\ell(\varphi;u) - \ell(\hat{\varphi};u)}\, |J_{\varphi\varphi}(\hat{\varphi})|^{1/2}\, d\hat{\varphi}$$

④ Saddlepoint : for full exponential models

$$f(y;\theta) = \exp\left\{\varphi'(\theta)u(y) - k(\theta)\right\} h(y)$$

$$g(u;\varphi) = \exp\left\{\varphi' u - K(\varphi)\right\} g(u) \qquad \ell(\varphi;u) = \varphi' u - K(\varphi)$$

· Can rewrite with $g(u)$ as a pdf:

$\exp\{K(t)\}$ = mgf of $g(u)$ $\qquad K(t)$ = cgf of $g(u)$

Saddlepoint approx: Apply p^*

log lik ratio

Info at $\hat{\varphi}$ now available from $K(\varphi)$

$$SP \Rightarrow \frac{e^{k/n}}{(2\pi)^{p/2}} e^{\ell(\varphi;u) - \ell(\hat{\varphi};u)} \left| j_{\varphi\varphi}(\hat{\varphi}) \right|^{1/2} d\hat{\varphi}$$

④ Saddlepoint : for full exponential models

$$f(y;\theta) = \exp\{\varphi'(\theta)u(y) - k(\theta)\}\, h(y)$$

$$g(u;\varphi) = \exp\{\varphi' u - K(\varphi)\}\, g(u) \qquad \ell(\varphi;u) = \varphi' u - K(\varphi)$$

- Can rewrite with $g(u)$ as a pdf:

$\exp\{K(t)\}$ = mgf of $g(u)$ $\qquad$ $K(t)$ = cgf of $g(u)$

Saddlepoint approx: Apply p^*

log lik ratio

Info at $\hat{\varphi}$ now available from $K(\varphi)$

$$SP \Rightarrow \frac{e^{k/n}}{(2\pi)^{p/2}}\, e^{\ell(\varphi;u) - \ell(\hat{\varphi};u)}\, |J_{\varphi\varphi}(\hat{\varphi})|^{1/2}\, d\hat{\varphi}$$

$$SP \Rightarrow \underbrace{\frac{e^{k/n}}{(2\pi)^{p/2}}\, e^{\ell(\varphi;u) - \ell(\hat{\varphi};u)}\, |J_{\varphi\varphi}(\hat{\varphi})|^{-1/2}\, du}$$

$$\ell_{\varphi}(\hat{\varphi};u) = 0$$
$$\ell_{\varphi\varphi}(\hat{\varphi};u)\,d\hat{\varphi} + du = 0$$

④ Saddlepoint : for full exponential models

$$f(y;\theta) = \exp\left\{\varphi'(\theta)u(y) - k(\theta)\right\} h(y)$$

$$g(u;\varphi) = \exp\{\varphi' u - K(\varphi)\} g(u) \qquad \ell(\varphi;u) = \varphi' u - K(\varphi)$$

- Can rewrite with $g(u)$ as a pdf:

$\exp\{K(t)\}$ = mgf of $g(u)$ $\qquad$ $K(t)$ = cgf of $g(u)$

Saddlepoint approx: Apply p^*

log lik ratio

Info at $\hat\varphi$ now available from $K(\varphi)$

$$SP \Rightarrow \frac{e^{k/n}}{(2\pi)^{p/2}} e^{\ell(\varphi;u) - \ell(\hat\varphi;u)} \left| J_{\varphi\varphi}(\hat\varphi)\right|^{1/2} d\hat\varphi$$

$$SP \Rightarrow \underbrace{\frac{e^{k/n}}{(2\pi)^{p/2}} e^{\ell(\varphi;u) - \ell(\hat\varphi;u)} \left| J_{\varphi\varphi}(\hat\varphi)\right|^{-1/2} du}$$

$$\ell_\varphi(\hat\varphi;u) = 0$$
$$\ell_{\varphi\varphi}(\hat\varphi;u)\,d\hat\varphi + du = 0$$

Can. var

Can. par

(4) Saddlepoint: for full exponential models

$$f(y;\theta) = \exp\{\varphi'(\theta)\,u(y) - k(\theta)\}\,h(y)$$

$$g(u;\varphi) = \exp\{\varphi' u - K(\varphi)\}\,g(u) \qquad \ell(\varphi;u) = \varphi' u - K(\varphi)$$

- Can rewrite with $g(u)$ as a pdf:

$\exp\{K(t)\}$ = mgf of $g(u)$ $\qquad$ $K(t)$ = cgf of $g(u)$

Saddlepoint approx: Apply p^*

log lik ratio

Info at $\hat\varphi$ now available from $K(\varphi)$

$$SP \Rightarrow \frac{e^{k/n}}{(2\pi)^{p/2}}\, e^{\ell(\varphi;u) - \ell(\hat\varphi;u)}\, |j_{\varphi\varphi}(\hat\varphi)|^{1/2}\, d\hat\varphi$$

$$SP \Rightarrow \underline{\frac{e^{k/n}}{(2\pi)^{p/2}}\, e^{\ell(\varphi;u) - \ell(\hat\varphi;u)}\, |j_{\varphi\varphi}(\hat\varphi)|^{-1/2}\, du}$$

$$\ell_\varphi(\hat\varphi;u) = 0$$
$$\ell_{\varphi\varphi}(\hat\varphi;u)\,d\hat\varphi + du = 0$$

Can. var

Can. par

Daniels 1954

B-N, Cox 1979 $\quad$ 25 $\quad$ 60

Now: $\quad$ Not in most texts

(5) Tangent Exponential Model : Model $f(y;\theta)$ Data y°

assume $\log f(y;\theta)$ is $O(n)$, on R^n

smooth

unique max

⑥ Tangent Exponential Model :

Model $f(y;\theta)$ Data y°

assume $\log f(y;\theta)$ is $O(n)$, on R^n

smooth

unique max

What do you need for Accurate 3rd order inference?

(5) Tangent Exponential Model : Model $f(y;\theta)$ Data y°

assume $\log f(y;\theta)$ is $O(n)$, on R^n

smooth

unique max

What do you need for Accurate 3rd order inference?

1) $\ell^\circ(\theta) = \log f(y^\circ;\theta)$ Obs log-likelihood .... of course!

5) Tangent Exponential Model :   Model $f(y;\theta)$ Data y°

assume $\log f(y;\theta)$ is $O(n)$, on R^n
smooth
unique max

What do you need for Accurate 3rd order inference?

1) $\ell^\circ(\theta) = \log f(y^\circ;\theta)$ Obs log-likelihood .... of course!

2) $\varphi(\theta) = \frac{d}{dV} \ell(\theta;y)\Big|_{y^\circ}$ "How $\ell(\theta;y)$ changes" at data ... in ...

(5) Tangent Exponential Model : Model $f(y;\theta)$ Data y°

assume $\log f(y;\theta)$ is $O(n)$, on R^n
smooth
unique max

What do you need for Accurate 3rd order inference?

1) $\ell^\circ(\theta) = \log f(y^\circ;\theta)$ Obs log-likelihood .... of course!

2) $\varphi(\theta) = \frac{d}{dV}\, \ell(\theta;y)\big|_{y^\circ}$ "How $\ell(\theta;y)$ changes "at data"

Could be just an .... Expt'l Model!

$$\bar{f}(y;\theta) \doteq e^{\ell^\circ(\theta) - \ell^\circ(\hat\theta^\circ) + \varphi'(\theta)\, u} \,.\, ds \qquad \text{Data } u^\circ = 0$$

Category theory!
Why not?

5 Tangent Exponential Model :   Model $f(y;\theta)$ Data y°

assume $\log f(y;\theta)$ is $O(n)$, on R^n
smooth
unique max

What do you need for Accurate 3rd order inference?

1) $\ell^\circ(\theta) = \log f(y^\circ;\theta)$ Obs log-likelihood   .... of course!

2) $\varphi(\theta) = \frac{d}{dV}\,\ell(\theta;y)\big|_{y^\circ}$ "How $\ell(\theta;y)$ changes "at data""

Could be just an .... Expt'l Model!

$$\bar{f}(y;\theta) \doteq e^{\ell^\circ(\theta)-\ell^\circ(\hat\theta^\circ) + \varphi'(\theta)u}\,.\,ds \qquad \text{Data } u^\circ = 0$$

Would make no difference for 3rd order inference ... !

So;
-"act as if" model is Expontial $L^\circ(\theta)$, can par $\varphi(\theta)$

(5) Tangent Exponential Model : Model $f(y;\theta)$ Data y°

assume $\log f(y;\theta)$ is $O(n)$, on R^n
smooth
unique max

What do you need for Accurate 3rd order inference?

1) $\ell^{\circ}(\theta) = \log f(y^{\circ};\theta)$ Obs log-likelihood .... of course!

2) $\varphi(\theta) = \frac{d}{dV}\ell(\theta;y)\big|_{y^{\circ}}$ "How $\ell(\theta;y)$ changes "at data"

Could be just an .... Expt'l Model!

$$\bar{f}(y;\theta) \doteq e^{\ell^{\circ}(\theta)-\ell^{\circ}(\hat{\theta}^{\circ}) + \varphi'(\theta)u}\,.\,ds \qquad \text{Data } u^{\circ}=0$$

Would make no difference for 3rd order inference ... !

So;

"act as if" model Expontial $\underline{L^{\circ}(\theta), \text{can par } \varphi(\theta)}$

F Reid (1995)

F Reid Wu (1999)

Call it : "Tangent exponential Model" TEM

Replaces sufficiency (when available)

but Works in wide generality

Transparent!

6 Test $\psi(\theta) = \psi_0$ Scalar, vector?

No problem!

6 Test $\psi(\theta) = \psi_0$ Scalar, vector?

No problem! $\dim y = n$

Just get $\ell(\theta)$, $\varphi(\theta)$ $\dim \varphi = p$ very general

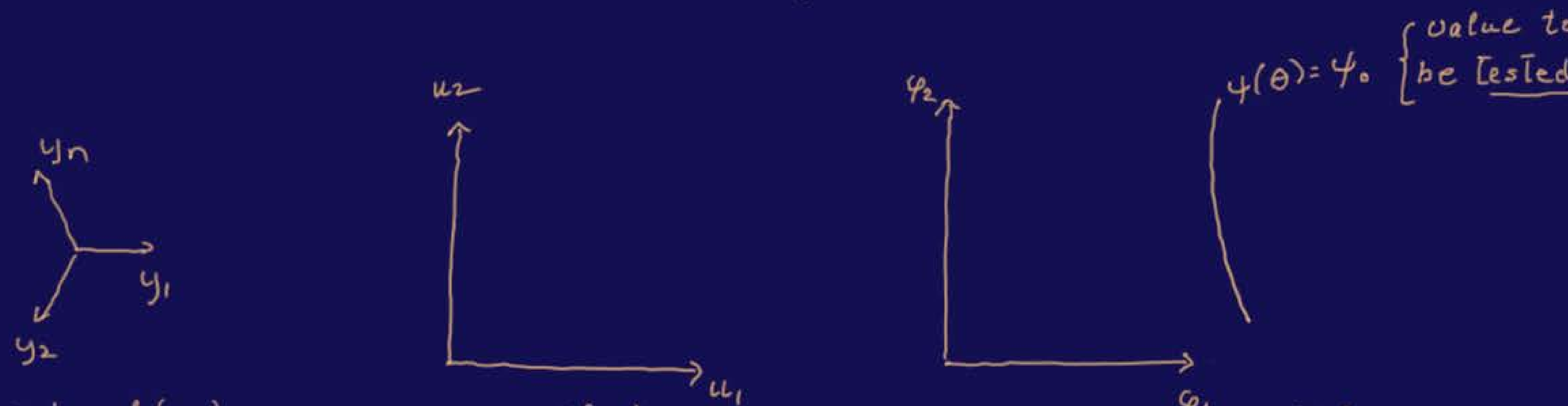
$\dim \psi = d$

⑥ Test $\psi(\theta) = \psi_0$ Scalar, vector?

No problem! $\dim y = n$

Just get $\ell(\theta)$, $\varphi(\theta)$ $\dim \varphi = p$ very general

$\dim \psi = d$



original (n) can var (p) can par (p) (d)

(6) <u>Test</u> $\psi(\theta) = \psi_0$

Scalar, vector?

<u>No</u> problem!

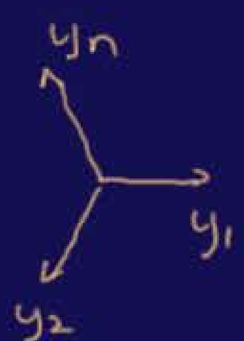
<u>Just</u> get $\ell(\theta)$, $\varphi(\theta)$

dim $y = n$

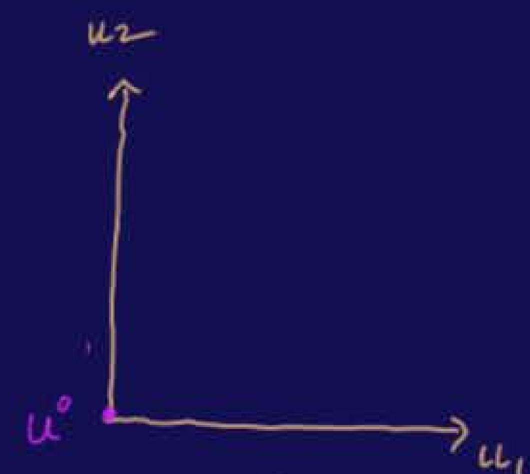
dim $\varphi = p$

dim $\psi = d$

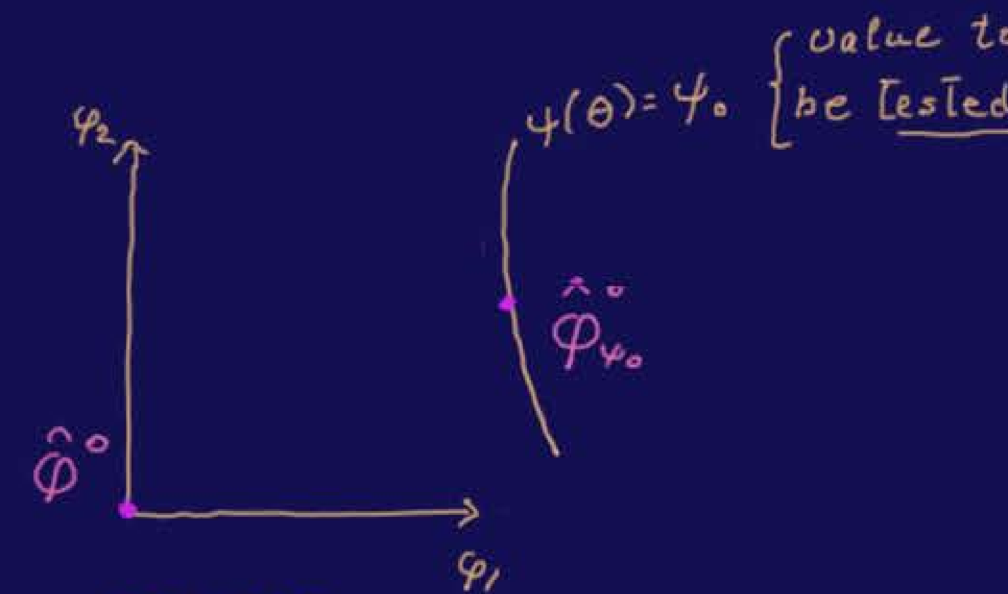
very general



<u>original</u>



<u>can var</u>



<u>can par</u>

Observed y^0: $u^0 = 0$

mle Overall

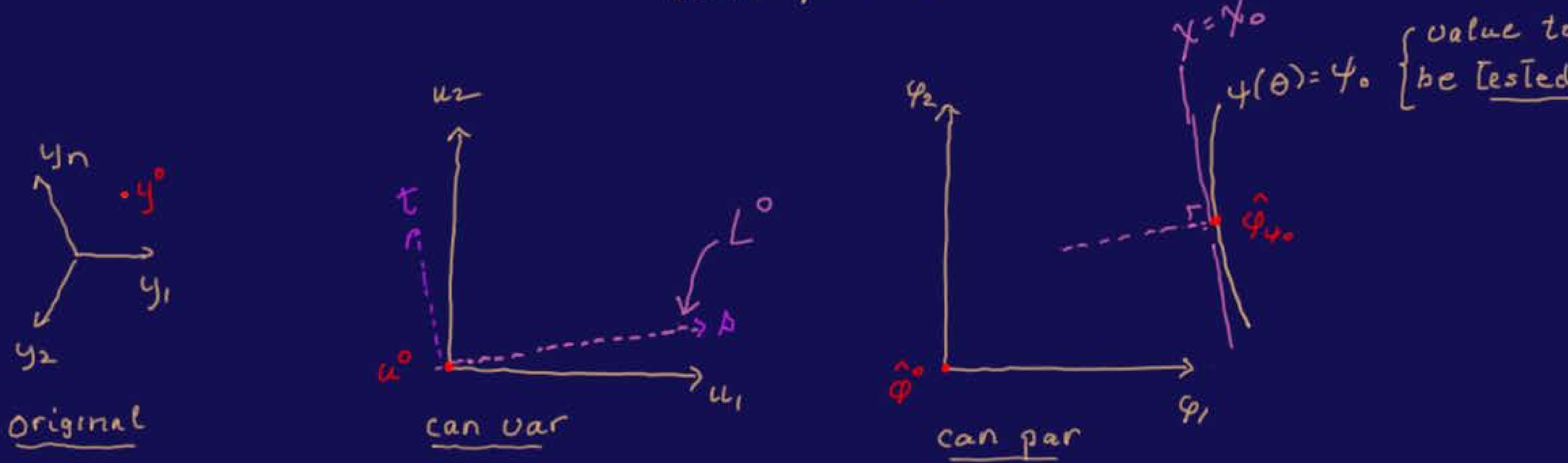
Constrained (under $\psi = \psi_0$) $\hat\varphi^0_{\psi_0} = \tilde\varphi^0$

⑥ Test $\psi(\theta) = \psi_0$ — Scalar, vector?

No problem! $\dim y = n$

Just get $\ell(\theta)$, $\varphi(\theta)$ — $\dim \varphi = p$ — very general

$\dim \psi = d$



Observed y^0: $u^0 = 0$

mle — Overall $\hat{\varphi}^0$

Constrained (under $\psi = \psi_0$) $\hat{\varphi}^0_{\psi_0} = \tilde{\varphi}^0$

... equiv. linear par. $\chi(\theta)$

... perp line L^0

⑥ Test $\psi(\theta) = \psi_0$ Scalar, vector?

No problem!

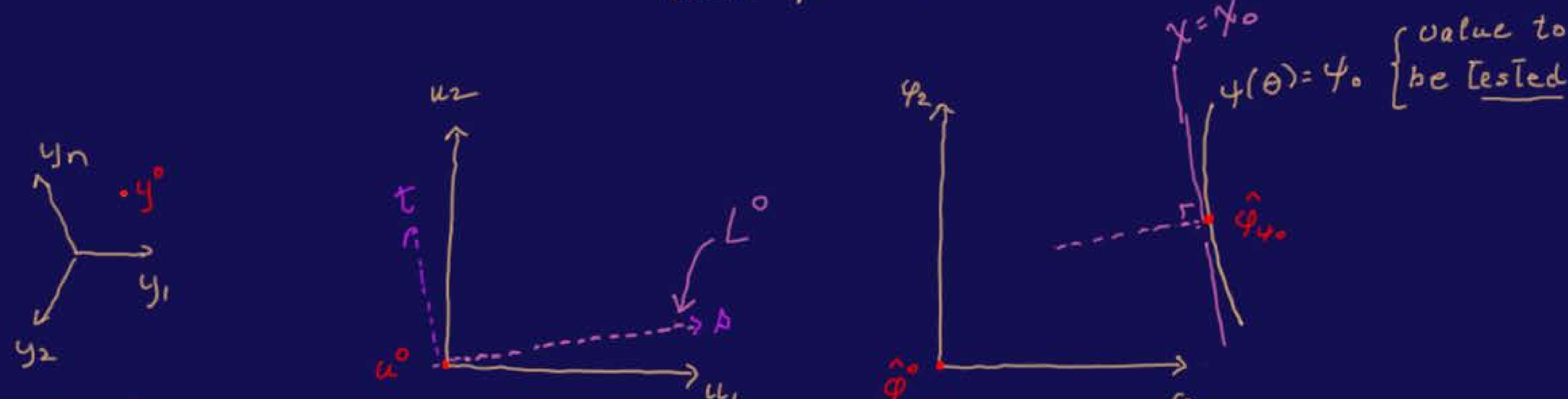
Just get $\ell(\theta)$, $\varphi(\theta)$

$\dim y = n$

$\dim \varphi = p$

$\dim \psi = d$

very general



original can var can par

Observed y^0: $u^0 = 0$

mle Overall $\hat{\varphi}^0$

Constrained (under $\psi = \psi_0$) $\hat{\varphi}^0_{\psi_0} = \tilde{\varphi}^0$

... equiv. linear par. $\chi(\theta)$

... perp line L^0

Everything happens on line/plane L^0

Explicit

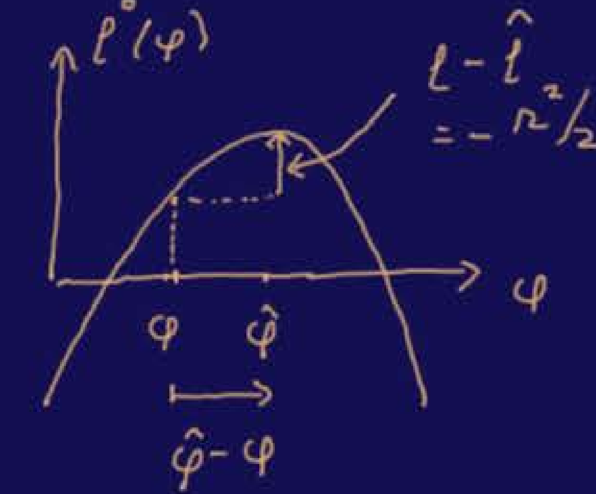
full 3rd order

$$\frac{e^{k/n}}{(2\pi)^{d/2}} e^{-r^2/2} \left| j_{\varphi\varphi}(\hat{\varphi}) \right|^{-1/2} \frac{\left| j_{(\lambda\lambda)}(\hat{\varphi}_\psi) \right|^{1/2}}{\left| j_{[\lambda\lambda]}(\hat{\varphi}_\psi) \right|^{1/2}} \cdot ds$$

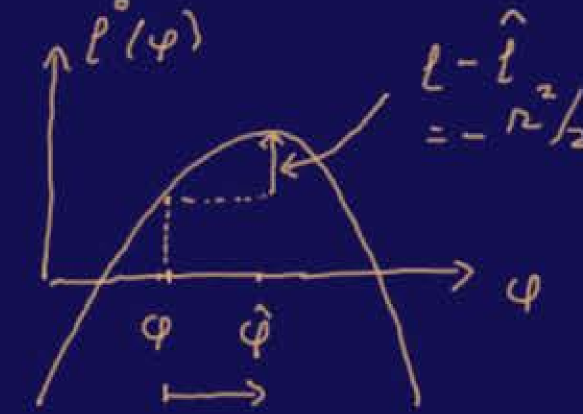
xxx = 260-Jan12-1.pdf

7 Scalar interest case: $\varphi = \varphi(\theta)$ scalar

$$\text{SP} \quad \frac{e^{k/n}}{(2\pi)^{1/2}} \, e^{\ell - \hat{\ell}} \, j_{\varphi\varphi}^{-1/2}(\hat{\varphi}) \cdot du$$



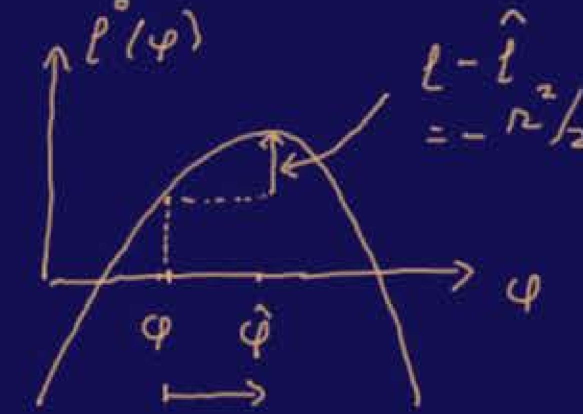
7 Scalar interest case : $\varphi = \varphi(\theta)$ scalar



SP $\dfrac{e^{k/n}}{(2\pi)^{1/2}} e^{\ell - \hat{\ell}} \, j_{\varphi\varphi}^{-1/2}(\hat{\varphi}) \cdot du$

Horizontal departure $q = (\hat{\varphi} - \varphi) \hat{j}_{\varphi\varphi}^{1/2}$ = mle departure

(7) Scalar interest case : $\varphi = \varphi(\theta)$ scalar



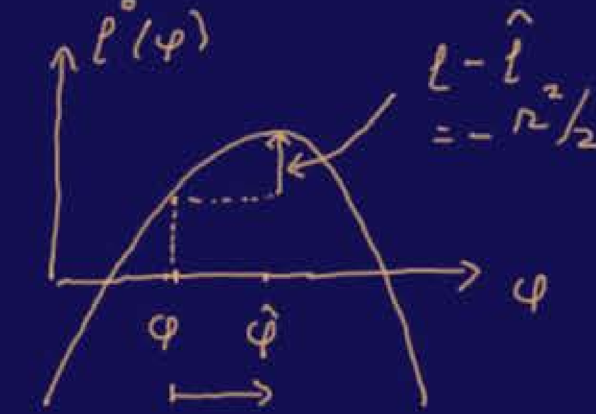
SP $\quad \dfrac{e^{k/n}}{(2\pi)^{1/2}} e^{\ell - \hat{\ell}} \, j_{\varphi\varphi}^{-1/2}(\hat{\varphi}) \cdot du$

Horizontal departure $\quad q = (\hat{\varphi} - \varphi)\, \hat{j}_{\varphi\varphi}^{1/2}$ = mle departure

Vertical " $\quad \hat{\ell} - \ell = r^2/2$ = loglik ratio

$r = \pm\sqrt{2(\hat{\ell} - \ell)}$ = SLR = signed lik. root

(7) Scalar interest case : $\varphi = \varphi(\theta)$ scalar



SP $\quad \dfrac{e^{k/n}}{(2\pi)^{1/2}}\, e^{\ell - \hat{\ell}}\, \hat{j}_{\varphi\varphi}^{-1/2}(\hat{\varphi}) \cdot du$

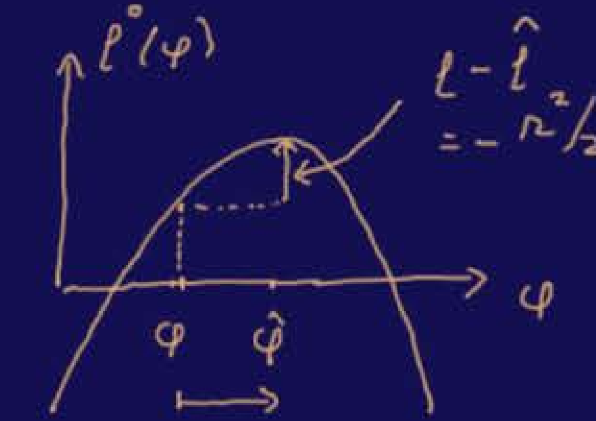
Horizontal departure $\quad q = (\hat{\varphi} - \varphi)\, \hat{j}_{\varphi\varphi}^{1/2}$ = mle departure

Vertical " $\quad \hat{\ell} - \ell = r^2/2$ = loglik ratio

$r = \pm\sqrt{2(\hat{\ell} - \ell)}$ = SLR = signed lik. root

Density re r $\quad f(r) \cdot dr = \underbrace{e^{k/n}\, \phi(r)\, \frac{r}{q}\, dr}$ $\quad \phi(r) = \frac{1}{\sqrt{2\pi}} e^{-r^2/2}$ Calculus

7 Scalar interest case : $\varphi = \varphi(\theta)$ scalar



SP $\quad \dfrac{e^{k/n}}{(2\pi)^{1/2}} \, e^{\ell - \hat{\ell}} \, j_{\varphi\varphi}^{-1/2}(\hat{\varphi}) \cdot du$

Horizontal departure $\quad q = (\hat{\varphi} - \varphi)\, \hat{j}_{\varphi\varphi}^{1/2}$ = mle departure

Vertical " $\quad \hat{\ell} - \ell = r^2/2$ = loglik ratio

$r = \pm\sqrt{2(\hat{\ell} - \ell)}$ = SLR = signed lik. root

Density re r $\quad f(r) \cdot dr = \underbrace{e^{k/n} \phi(r) \frac{r}{q} \, dr}$

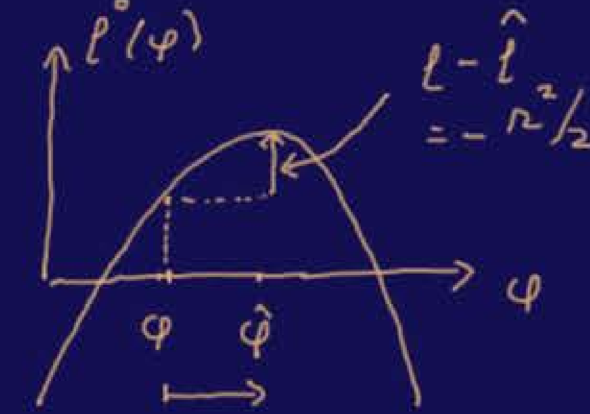
$\phi(r) = \dfrac{1}{\sqrt{2\pi}} e^{-r^2/2}$ Calculus

df re r $\quad F(r) = \Phi\left(r - r^{-1} \log \frac{r}{q}\right) = \Phi(r^*)$

$\Phi(r) = \int_{-\infty}^{r} \phi(r)\, dr$

B-N (1990)

(7) Scalar interest case : $\varphi = \varphi(\theta)$ scalar



SP $\quad \dfrac{e^{k/n}}{(2\pi)^{1/2}} e^{\ell - \hat{\ell}} \, j_{\varphi\varphi}^{-1/2}(\hat{\varphi}) \cdot du$

Horizontal departure $\quad q = (\hat{\varphi} - \varphi)\, \hat{j}_{\varphi\varphi}^{1/2}$ = mle departure

Vertical " $\quad \hat{\ell} - \ell = r^2/2$ = loglik ratio

$r = \sqrt{\pm 2(\hat{\ell} - \ell)}$ = SLR = signed lik. root

Density re r $\quad f(r) \cdot dr = \underbrace{e^{k/n} \phi(r) \dfrac{r}{q} dr}$ $\qquad \phi(r) = \dfrac{1}{\sqrt{2\pi}} e^{-r^2/2}$ Calculus

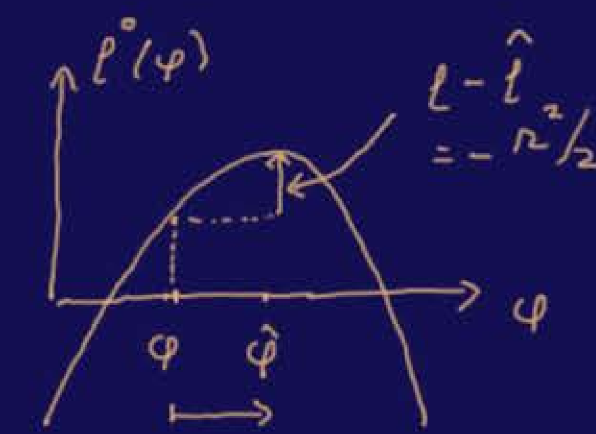
df re r $\quad F(r) = \Phi\left(r - r^{-1} \log \dfrac{r}{q}\right) = \Phi(r^*)$ $\qquad \Phi(r) = \int_{-\infty}^{r} \phi(r)\, dr$

B-N (1990)

High accuracy $\quad O(n^{-3/2})$

Widely useful

(7) Scalar interest case : $\varphi = \varphi(\theta)$ scalar



SP $\quad \dfrac{e^{k/n}}{(2\pi)^{1/2}} e^{\ell - \hat{\ell}} \, j_{\varphi\varphi}^{-1/2}(\hat{\varphi}) \cdot du$

Horizontal departure $\quad q = (\hat{\varphi} - \varphi)\, \hat{j}_{\varphi\varphi}^{1/2}$ = mle departure

Vertical " $\quad \hat{\ell} - \ell = r^2/2$ = loglik ratio

$r = \pm\sqrt{2(\hat{\ell} - \ell)}$ = SLR = signed lik. root

Density re r $\quad f(r) \cdot dr = \underbrace{e^{k/n} \phi(r) \dfrac{r}{q} dr}$ $\qquad \phi(r) = \dfrac{1}{\sqrt{2\pi}} e^{-r^2/2}$ Calculus

df re r $\quad F(r) = \Phi\left(r - r^{-1} \log \dfrac{r}{q}\right) = \Phi(r^*)$ $\qquad \Phi(r) = \int_{-\infty}^{r} \phi(r)\, dr$

B-N (1990)

High accuracy $\quad O(n^{-3/2})$

Widely useful

Above by: Simple change of variable
" integration

Just calculus !

8) What can SP do for Bayes? a) Scalar parameter: Welch Peers

8) What can SP do for Bayes? a) Scalar parameter: Welch Peers

Have likelihood $L^{\circ}(\theta)$

add weight fn $\pi(\theta)$

"Treat" $\pi(\theta)$ as probability Get: $\pi(\theta \mid y^{\circ}) = c\, L^{\circ}(\theta)\, \pi(\theta)$?

(8) What can SP do for Bayes? a) Scalar parameter: Welch Peers

Have likelihood $L^{\circ}(\theta)$

add weight fn $\pi(\theta)$

"Treat" $\pi(\theta)$ as probability   Get: $\pi(\theta \mid y^{\circ}) = c\, L^{\circ}(\theta)\, \pi(\theta)$?

Expand exponent $$f(s;\varphi) = \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{(s-\varphi)^2}{2} - \gamma \varphi^3/6n^{1/2} + \gamma s^3/6n^{1/2} \right\} (1 - \gamma s/2n^{1/2})$$

8) What can SP do for Bayes? a) Scalar parameter: <u>Welch Peers</u>

Have likelihood $L^{\circ}(\theta)$

add weight fn $\pi(\theta)$

"Treat" $\pi(\theta)$ as <u>probability</u> Get: $\pi(\theta \mid y^{\circ}) = c\, L^{\circ}(\theta)\, \pi(\theta)$?

Expand exponent $f(s;\varphi) = \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{(s-\varphi)^2}{2} - \gamma \varphi^3/6n^{1/2} + \gamma s^3/6n^{1/2} \right\} (1 - \gamma s/2n^{1/2})$

Try: $\beta = \varphi + \varphi^2/2n^{1/2}$ $f(\hat\beta;\beta) = \frac{1}{\sqrt{2\pi}} \exp\left\{ \frac{(\hat\beta-\beta)^2}{2} - \gamma (\hat\beta-\beta)^3/6n^{1/2} \right\} \cdot d\hat\beta$ Location model

(8) What can SP do for Bayes? a) Scalar parameter: Welch Peers

Have likelihood $L^{\circ}(\theta)$

add weight fn $\pi(\theta)$

"Treat" $\pi(\theta)$ as probability Get: $\pi(\theta \mid y^{\circ}) = c\, L^{\circ}(\theta)\, \pi(\theta)$?

Expand exponent $\quad f(s;\varphi) = \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{(s-\varphi)^2}{2} - \gamma \varphi^3/6n^{1/2} + \gamma s^3/6n^{1/2} \right\} (1 - \gamma s/2n^{1/2})$

Use $\beta = \varphi + \varphi^2/2n^{1/2}$ $\quad f(\hat{\beta};\beta) = \frac{1}{\sqrt{2\pi}} \exp\left\{ \frac{(\hat{\beta}-\beta)^2}{2} - \gamma (\hat{\beta}-\beta)^3/6n^{1/2} \right\} \cdot d\hat{\beta}$ $\quad$ Location model

Posterior (loc'n flip) $\quad f(\beta \mid \hat{\beta}) = \frac{1}{\sqrt{2\pi}} \exp\left\{ \frac{(\hat{\beta}-\beta)^2}{2} - \gamma (\hat{\beta}-\beta)^3/6n^{1/2} \right\} \cdot d\beta$ $\quad d\beta = \underbrace{(1 + \gamma\varphi/n^{1/2})}_{\text{root information prior}} d\varphi$

(8) What can SP do for Bayes? a) Scalar parameter: Welch Peers

Have likelihood $L^{\circ}(\theta)$

add weight fn $\pi(\theta)$

"Treat" $\pi(\theta)$ as probability

Get: $\pi(\theta \mid y^{\circ}) = c\, L^{\circ}(\theta)\, \pi(\theta)$?

Expand exponent $\quad f(s;\varphi) = \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{(s-\varphi)^2}{2} - \gamma \varphi^3/6n^{1/2} + \gamma s^3/6n^{1/2} \right\} (1 - \gamma s/2n^{1/2})$

Use $\beta = \varphi + \varphi^2/2n^{1/2}$ $\quad f(\hat\beta;\beta) = \frac{1}{\sqrt{2\pi}} \exp\left\{ \frac{(\hat\beta-\beta)^2}{2} - \gamma (\hat\beta-\beta)^3/6n^{1/2} \right\} \cdot d\hat\beta$ $\quad$ Location model

Posterior (loc'n flip) $\quad f(\beta \mid \hat\beta) = \frac{1}{\sqrt{2\pi}} \exp\left\{ \frac{(\hat\beta-\beta)^2}{2} - \gamma (\hat\beta-\beta)^3/6n^{1/2} \right\} \cdot d\beta$ $\quad d\beta = \underline{(1 + \gamma\varphi/n^{1/2})}\, d\varphi$

root information prior

(Jeffreys prior)

Welch Peers 1963

Brown Cai DasGupta 2001
(binomial)

(8) What can SP do for Bayes? a) Scalar parameter: Welch Peers

Have likelihood $L^{\circ}(\theta)$

add weight fn $\pi(\theta)$

"Treat" $\pi(\theta)$ as probability

Get: $\pi(\theta \mid y^{\circ}) = c\, L^{\circ}(\theta)\, \pi(\theta)$?

Expand exponent $f(s;\varphi) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(s-\varphi)^2}{2} - \gamma \varphi^3/6n^{1/2} + \gamma s^3/6n^{1/2}\right\}(1 - \gamma s/2n^{1/2})$

Use $\beta = \varphi + \varphi^2/2n^{1/2}$ $f(\hat\beta;\beta) = \frac{1}{\sqrt{2\pi}} \exp\left\{\frac{(\hat\beta-\beta)^2}{2} - \gamma(\hat\beta-\beta)^3/6n^{1/2}\right\} \cdot d\hat\beta$ Location model

Posterior (loc'n flip) $f(\beta \mid \hat\beta) = \frac{1}{\sqrt{2\pi}} \exp\left\{\frac{(\hat\beta-\beta)^2}{2} - \gamma(\hat\beta-\beta)^3/6n^{1/2}\right\} \cdot d\beta$ $d\beta = \underline{(1 + \gamma\varphi/n^{1/2})}\, d\varphi$

root information prior

(Jeffreys prior)

Welch Peers 1963

Brown Cai DasGupta 2001
(binomial)

If Bayes was always so easy?

(8) What can SP do for Bayes? a) Scalar parameter: Welch Peers

Have likelihood $L^{\circ}(\theta)$

add weight f'n $\pi(\theta)$

"Treat" $\pi(\theta)$ as probability   Get: $\pi(\theta|y^{\circ}) = c\, L^{\circ}(\theta)\, \pi(\theta)$?

Expand exponent $\quad f(s;\varphi) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(s-\varphi)^2}{2} - \gamma\varphi^3/6n^{1/2} + \gamma s^3/6n^{1/2}\right\}(1-\gamma s/2n^{1/2})$

Use $\beta = \varphi + \varphi^2/2n^{1/2}$ $\quad f(\hat\beta;\beta) = \frac{1}{\sqrt{2\pi}} \exp\left\{\frac{(\hat\beta-\beta)^2}{2} - \gamma(\hat\beta-\beta)^3/6n^{1/2}\right\} \cdot d\hat\beta$ $\quad$ Location model

Posterior (loc'n flip) $\quad f(\beta|\hat\beta) = \frac{1}{\sqrt{2\pi}} \exp\left\{\frac{(\hat\beta-\beta)^2}{2} - \gamma(\hat\beta-\beta)^3/6n^{1/2}\right\} \cdot d\beta$ $\quad d\beta = \underline{(1+\gamma\varphi/n^{1/2})}\, d\varphi$

root information prior

(Jeffreys prior)

Welch Peers 1963

Brown Cai DasGupta 2001
(binomial)

If Bayes was always so easy?

but maybe it is!

8b

Scalar parameter

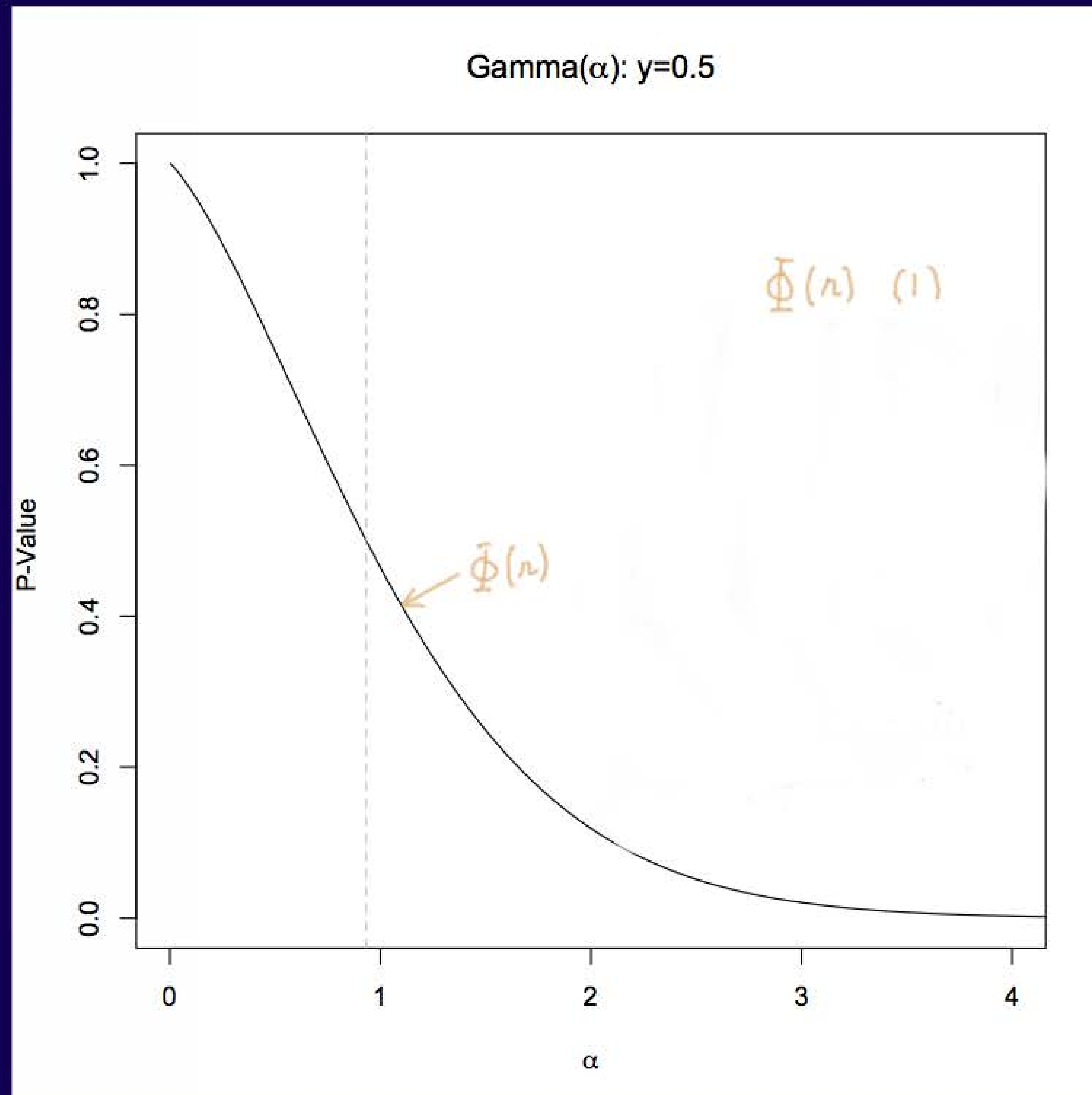
Example 1

Setup

- Model: $Y \sim Gamma(\alpha, \beta)$ where α is shape and β is rate
  pdf is $\Gamma^{-1}(\alpha)\, y^{\alpha-1} e^{-y} \cdot dy$ ... with $\beta = 1$

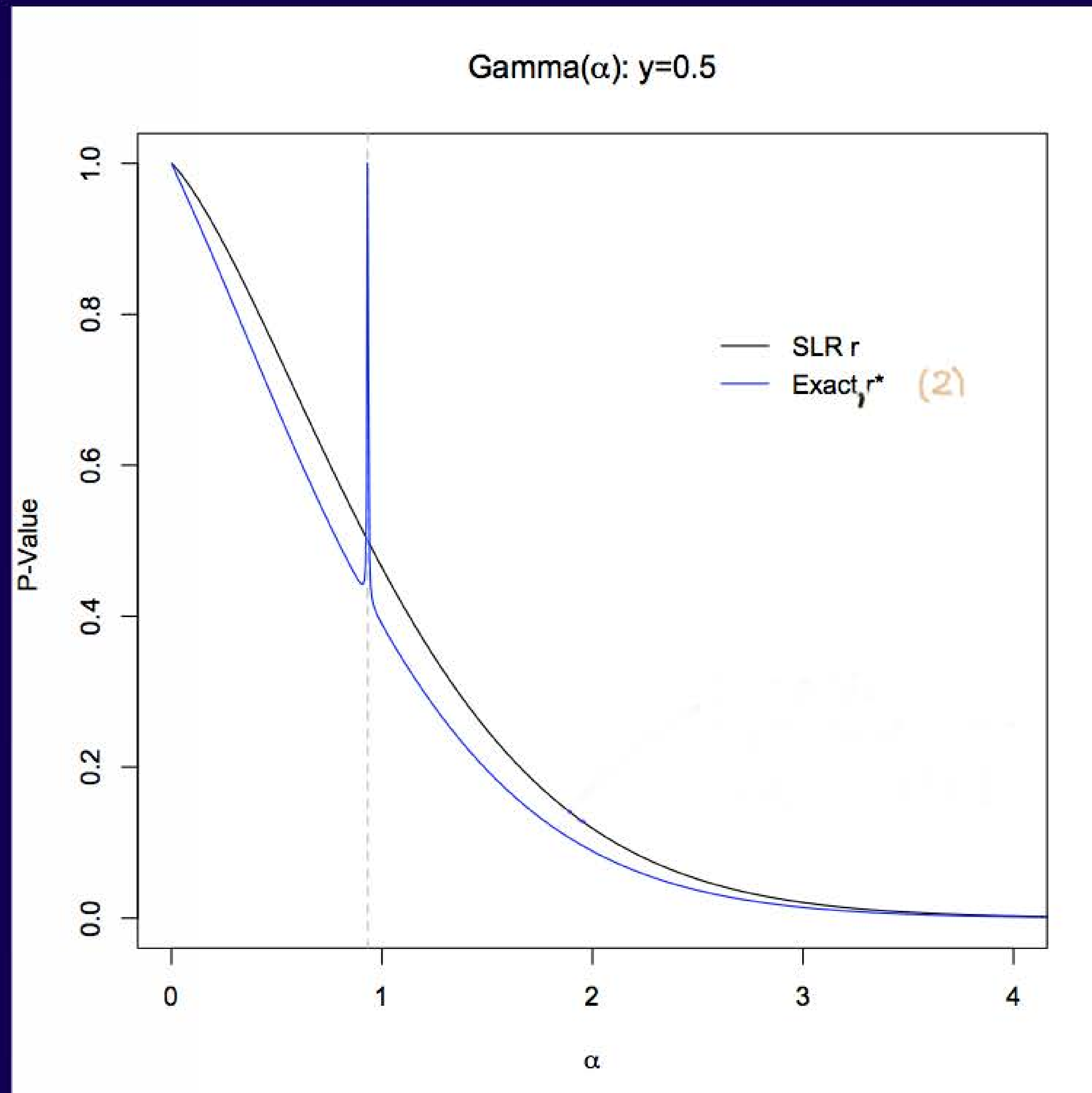
  Let $n = 1$ $y^{\circ} = .5$

  Parameter of interest is α



(1)
What likelihood says: $\Phi(r)$

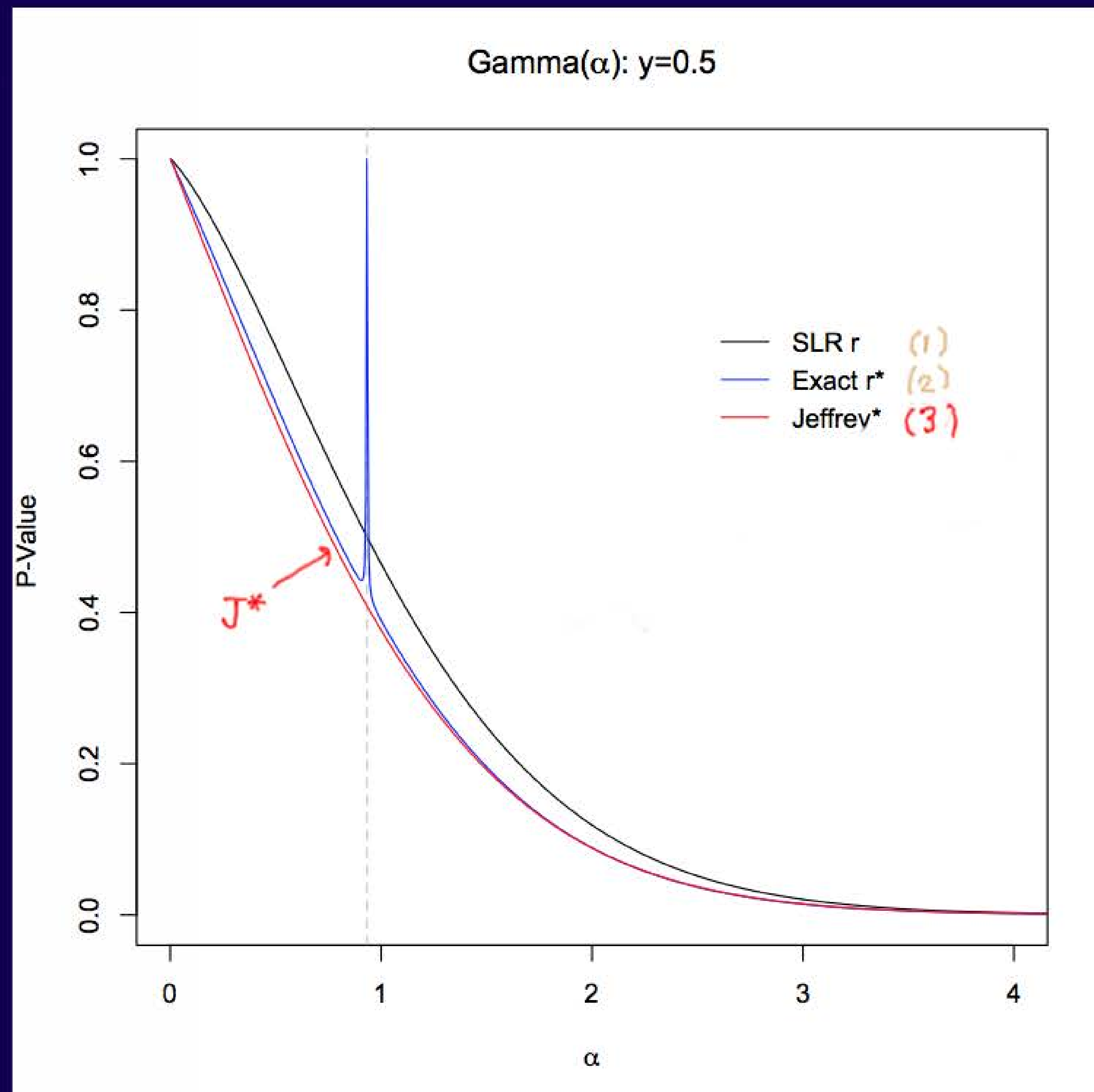
r = SLR $\quad \Phi$ df N01



(1)
What likelihood says: $\Phi(r)$

r = SLR $\quad \Phi$ df N01

(2)
Third order r^*
Exact



(1)

What likelihood says: $\Phi(r)$

$r = \text{SLR}$ $\quad \Phi$ df N01

(2)

Third order r^*

Exact

(3)

Jeffreys "$p = 1$"

Closely hugs exact

J^* is reproducible

9a Bayes: Vector parameter

Interest $\psi = \psi(\varphi)$ Scalar

Nuisance $\lambda = \psi(\varphi)$

9a Bayes: Vector parameter

Interest $\psi = \psi(\varphi)$ Scalar

Nuisance $\lambda = \psi(\varphi)$

SP separates: $$f(s;\psi)\, ds \cdot \frac{|j_{(\lambda\lambda)}(\hat{\varphi}_\psi)|^{-1/2}}{(2\pi)^{\frac{p-1}{2}}} \exp\{\ell - \tilde{\ell}\}\, dt$$

9a Bayes: Vector parameter

Interest $\psi = \psi(\varphi)$ Scalar

Nuisance $\lambda = \psi(\varphi)$

SP separates: $\quad f(s;\psi)\, ds \cdot \dfrac{|j_{(\lambda\lambda)}(\hat{\varphi}_\psi)|^{-1/2}}{(2\pi)^{\frac{p-1}{2}}} \exp\{\ell - \hat{\ell}\}\, dt$

Get rid of nuisance

(on profile of constrained model) $\quad |j_{(\lambda\lambda)}(\hat{\varphi}_\psi)|^{1/2} \cdot 1$

9a Bayes: Vector parameter

Interest $\psi = \psi(\varphi)$ Scalar

Nuisance $\lambda = \psi(\varphi)$

SP separates: $f(s;\psi)\, ds \cdot \dfrac{|j_{(\lambda\lambda)}(\hat{\varphi}_\psi)|^{-1/2}}{(2\pi)^{\frac{p-1}{2}}} \exp\{\ell - \tilde{\ell}\}\, dt$

Get rid of nuisance
(on profile of constrained model) $|j_{(\lambda\lambda)}(\hat{\varphi}_\psi)|^{1/2} \cdot 1$

Welch Peers for ψ $|j^{p}_{(\psi\psi)}(\psi)|^{1/2}$

9a Bayes: Vector parameter

Interest $\psi = \psi(\varphi)$ Scalar

Nuisance $\lambda = \psi(\varphi)$

SP separates: $f(s;\psi)\, ds \cdot \dfrac{|j_{(\lambda\lambda)}(\hat{\varphi}_\psi)|^{-1/2}}{(2\pi)^{\frac{p-1}{2}}} \exp\{\ell - \tilde{\ell}\}\, dt$

Get rid of nuisance
(on profile of constrained model) $|j_{(\lambda\lambda)}(\hat{\varphi}_\psi)|^{1/2} \cdot 1$

Welch Peers for ψ $|j^{p}_{(\psi\psi)}(\psi)|^{1/2}$

Combine (full prior)
- On profile
- One dimensional

$$|j_{\varphi\varphi}(\hat{\varphi}_\psi)|^{1/2} \frac{|j_{(\lambda\lambda)}(\hat{\varphi}_\psi)|^{1/2}}{|j_{[\lambda\lambda]}(\hat{\varphi}_\psi)|^{1/2}}$$

9a Bayes: Vector parameter

Interest $\psi = \psi(\varphi)$ Scalar

Nuisance $\lambda = \psi(\varphi)$

SP separates: $f(s;\psi)\, ds \cdot \dfrac{|j_{(\lambda\lambda)}(\hat{\varphi}_\psi)|^{-1/2}}{(2\pi)^{\frac{p-1}{2}}} \exp\{\ell - \hat{\ell}\}\, dt$

Get rid of nuisance
(on profile of constrained model) $|j_{(\lambda\lambda)}(\hat{\varphi}_\psi)|^{1/2} \cdot 1$

Welch Peers for ψ $|j^{p}_{(\psi\psi)}(\psi)|^{1/2}$

Combine (full prior)
- On profile
- One dimensional

$$\underbrace{|j_{\varphi\varphi}(\hat{\varphi}_\psi)|^{1/2}}_{\text{Regular Jeffreys}} \cdot \underbrace{\frac{|j_{(\lambda\lambda)}(\hat{\varphi}_\psi)|^{1/2}}{|j_{[\lambda\lambda]}(\hat{\varphi}_\psi)|^{1/2}}}_{\text{Nonlinearity adjustment}}$$

Root nuisance info re ψ

Root nuisance info re linear χ

9a Bayes: Vector parameter

Interest $\psi = \psi(\varphi)$ Scalar

Nuisance $\lambda = \psi(\varphi)$

SP separates: $f(s;\psi)\, ds \cdot \dfrac{|j_{(\lambda\lambda)}(\hat{\varphi}_\psi)|^{-1/2}}{(2\pi)^{\frac{p-1}{2}}} \exp\{\ell - \tilde{\ell}\}\, dt$

Get rid of nuisance
(<u>on profile</u> of constrained model) $|j_{(\lambda\lambda)}(\hat{\varphi}_\psi)|^{1/2} \cdot 1$

Welch Peers for ψ $|j^{p}_{(\psi\psi)}(\psi)|^{1/2}$

Combine (full prior)
- On profile
- <u>One</u> dimensional

$$|j_{\varphi\varphi}(\hat{\varphi}_\psi)|^{1/2} \cdot \frac{|j_{(\lambda\lambda)}(\hat{\varphi}_\psi)|^{1/2}}{|j_{[\lambda\lambda]}(\hat{\varphi}_\psi)|^{1/2}}$$

Root nuisance info re ψ

Root nuisance info re linear χ

Regular Jeffreys

Nonlinearity adjustment

Second order

Jeffreys works <u>but use on one dim. contour</u>

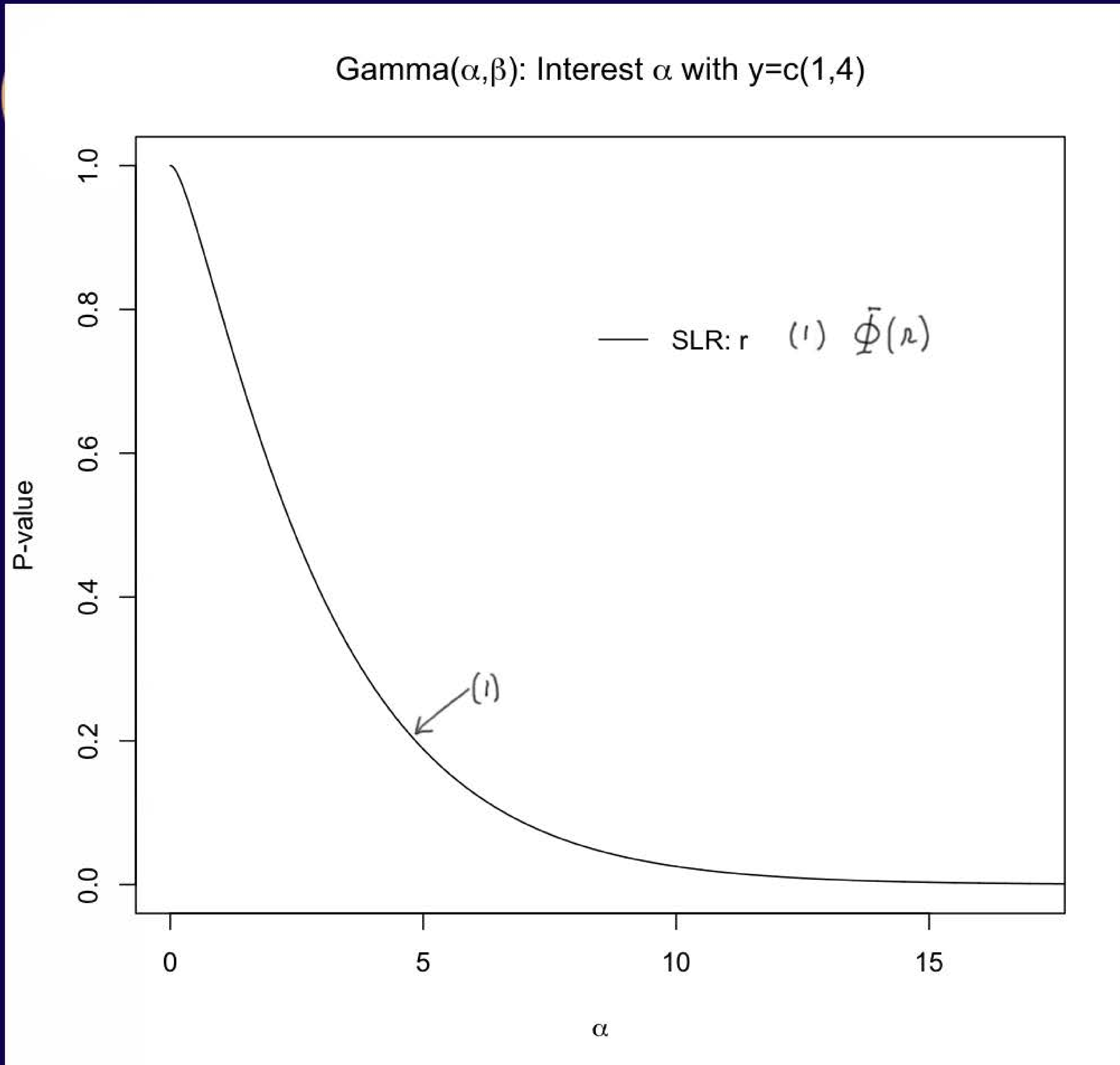
9b Vector parameter (α, β); Scalar linear interest α

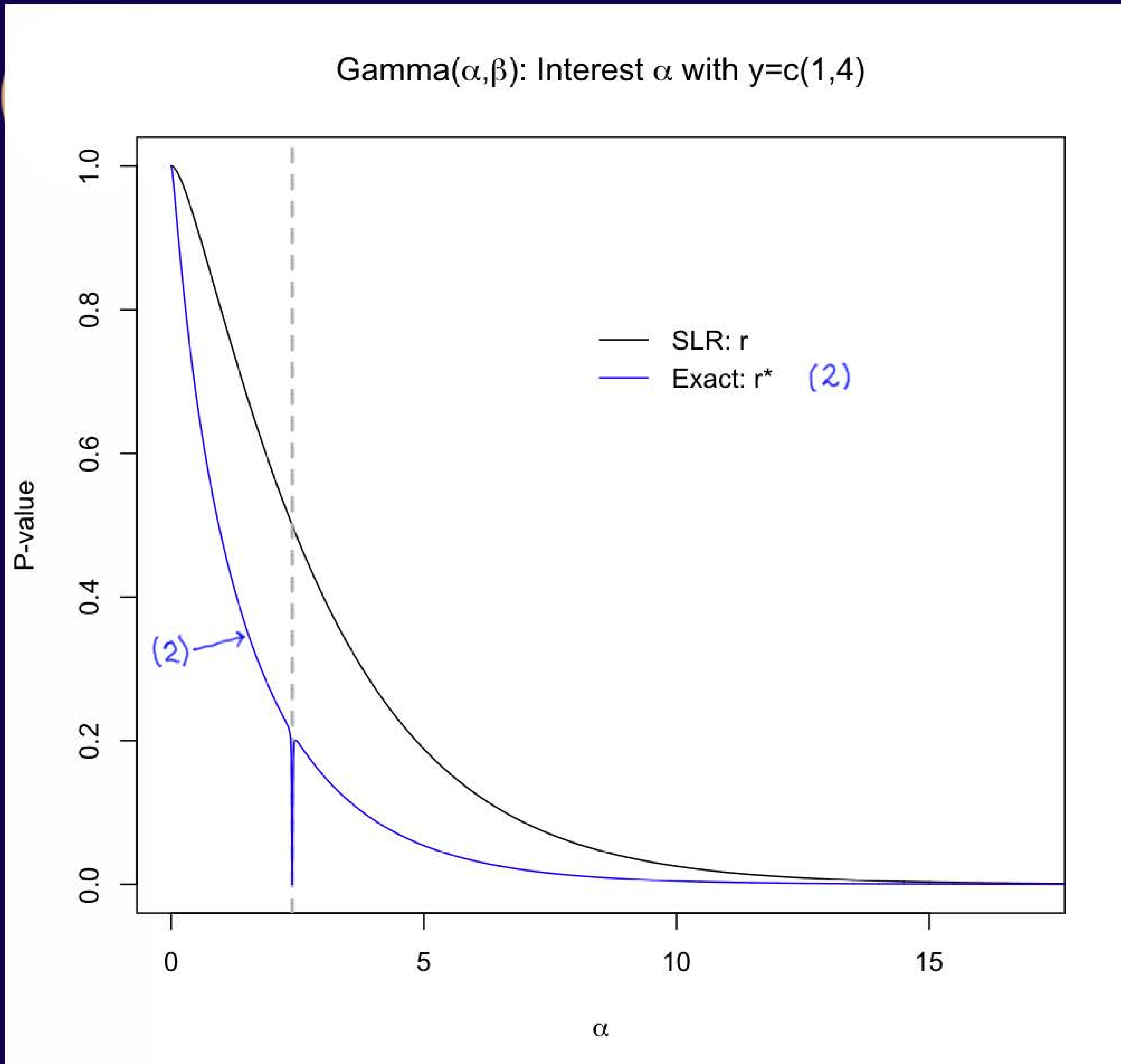
Example 2

Setup

- Model: $Y \sim Gamma(\alpha, \beta)$ where α is shape and β is rate
  pdf is $\frac{\beta^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\beta y}$
  Let $n = 2$ and data is $(y_1, y_2) = (1, 4)$

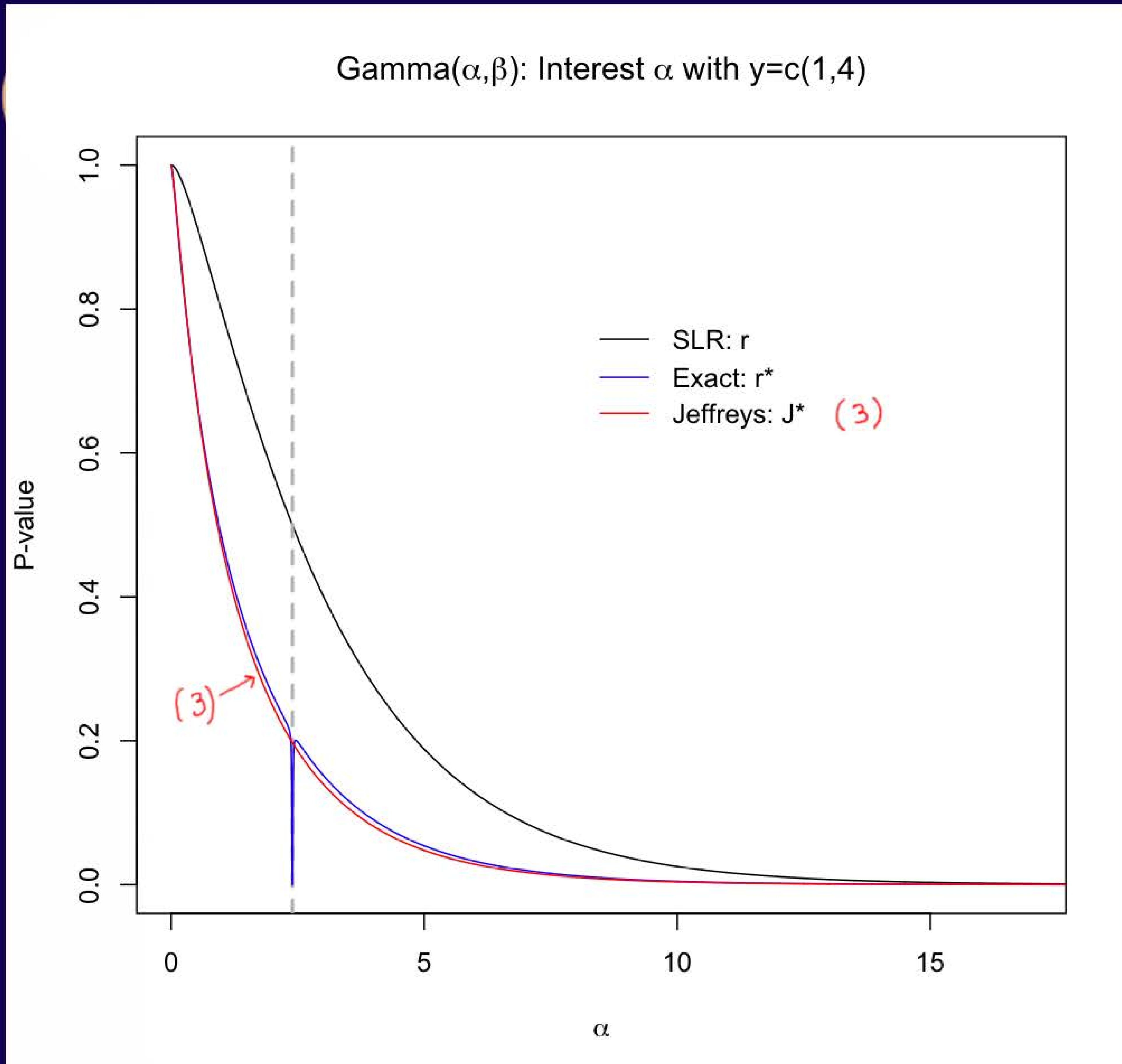
  Parameter of interest is α, β free nuisance


Gamma(α,β): Interest α with y=c(1,4)
SLR: r (1) $\bar{\Phi}(r)$
(1)
P-value
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
α

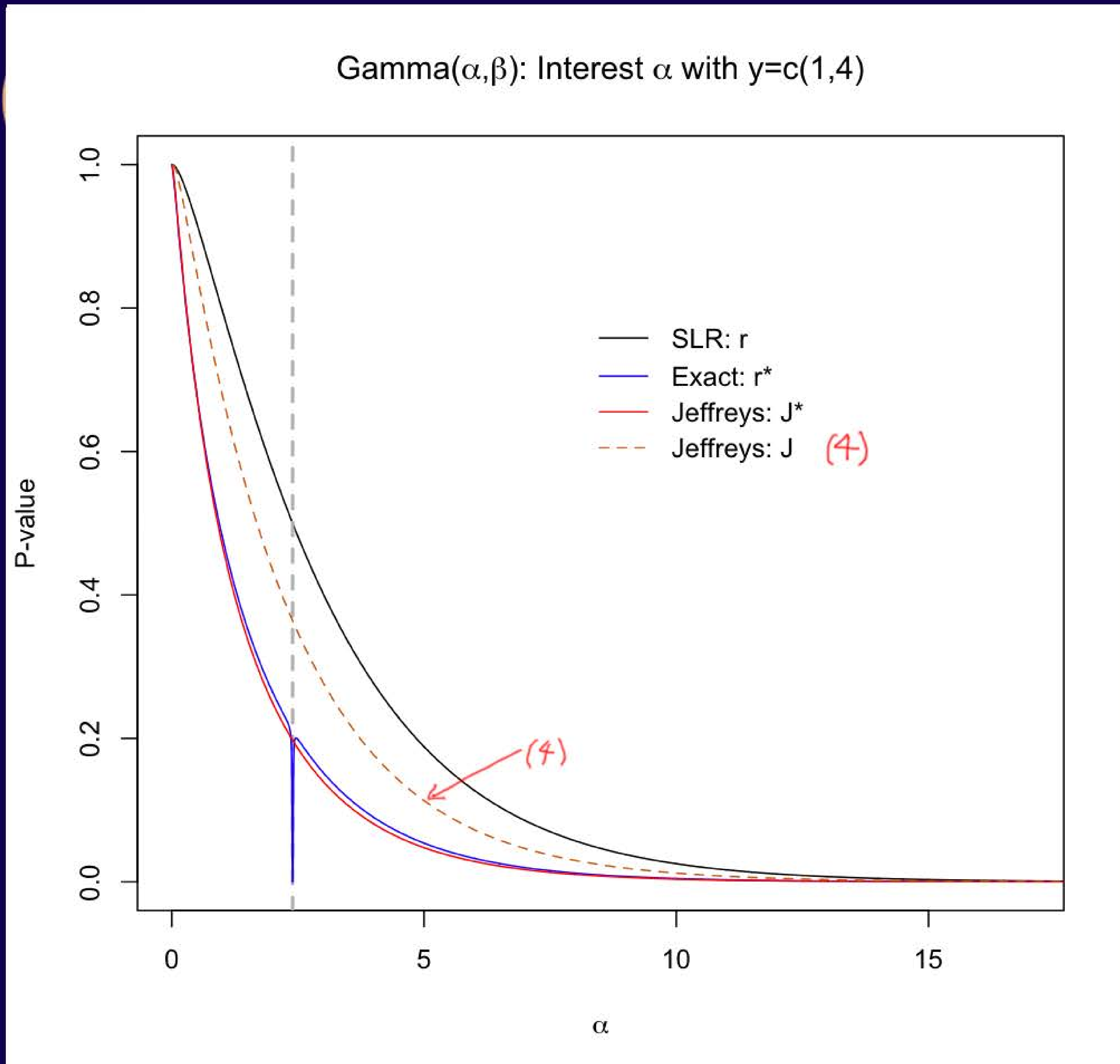


Gamma(α,β): Interest α with y=c(1,4)
P-value
α
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
SLR: r
Exact: r*
(2)
(2)

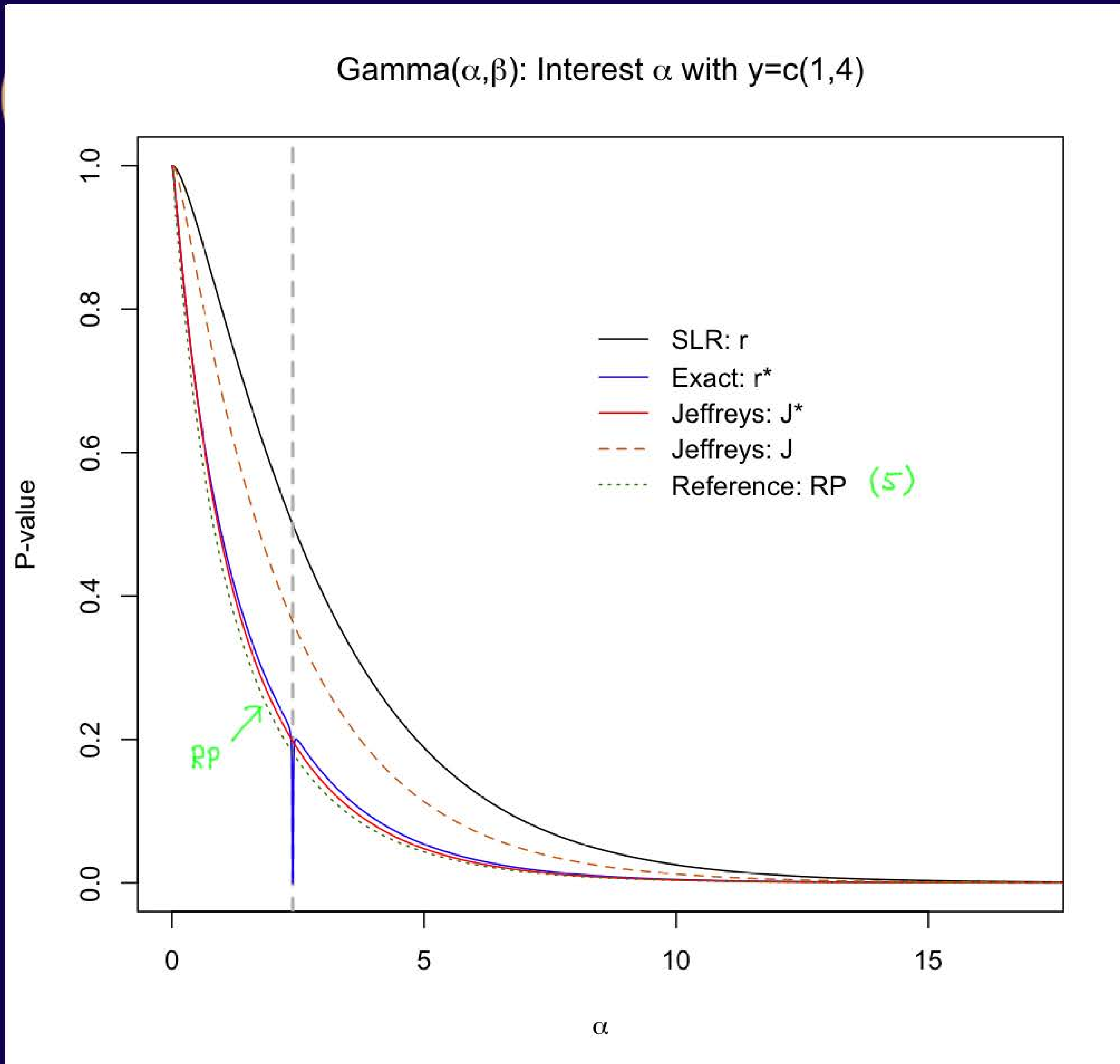
(2) $\Phi(r^*)$



(3) Profile Jeffreys



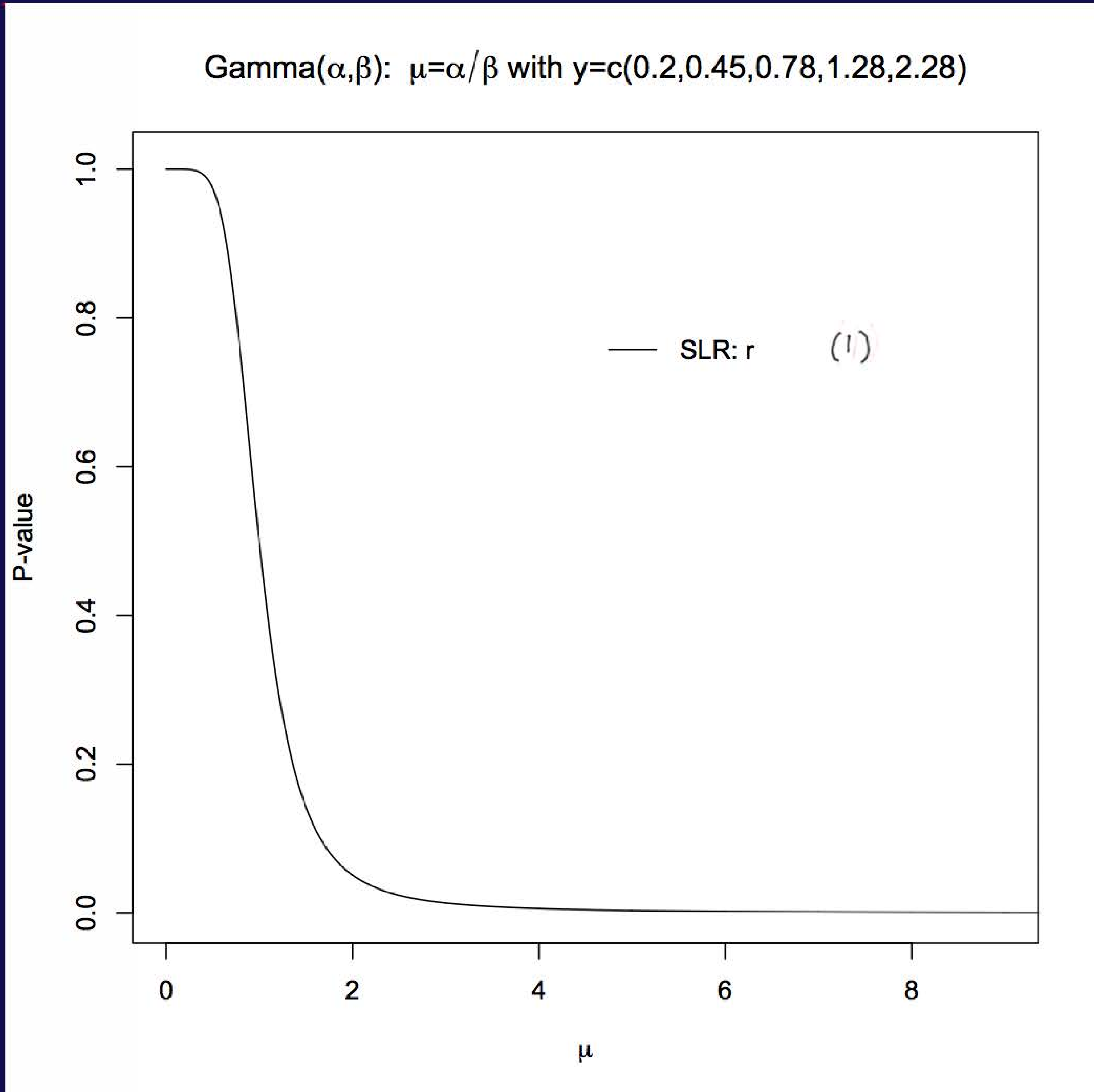
(4) Full Jeffreys



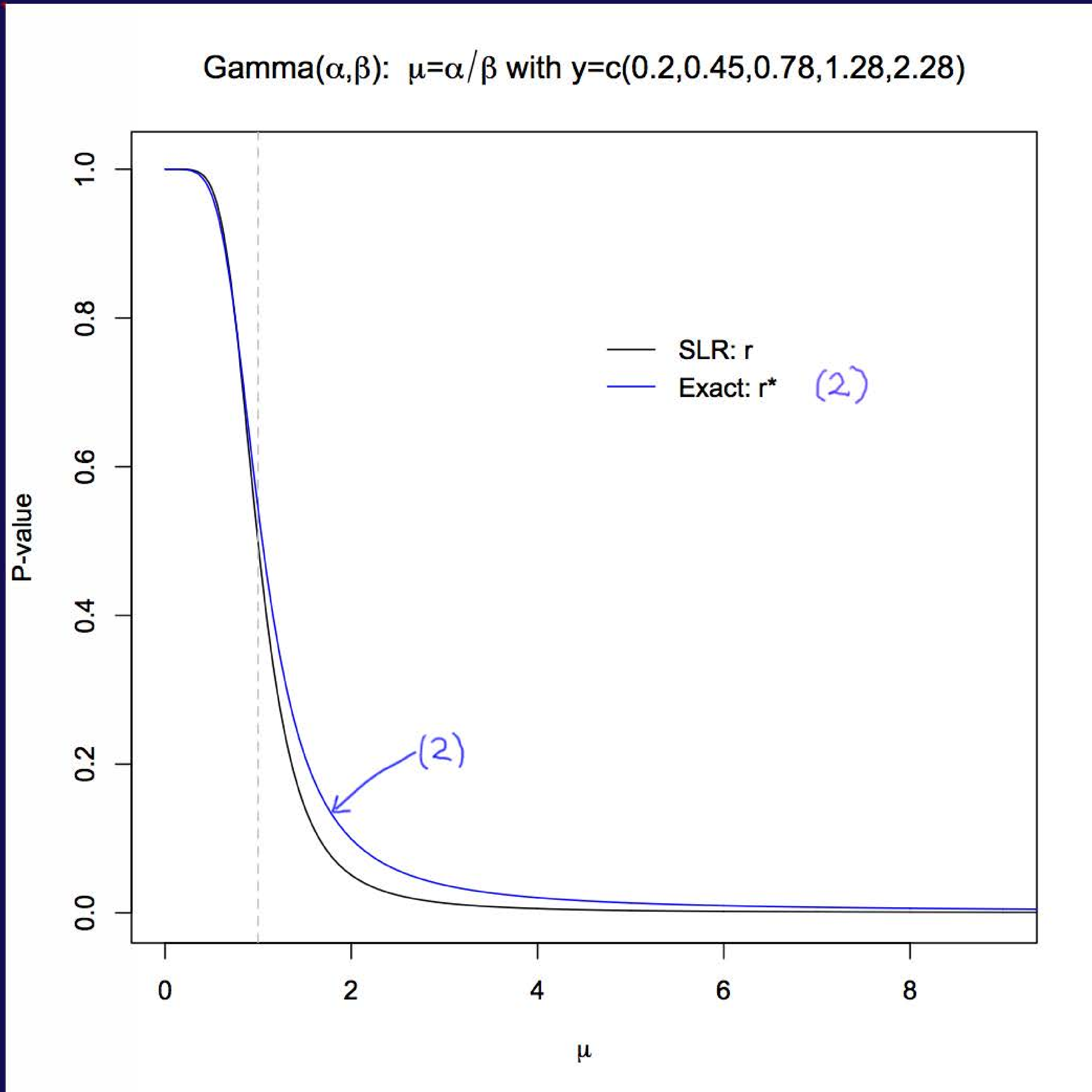
(5) Targeted Reference

Ex 3 Gamma(α, β) : Interest μ ... rotating

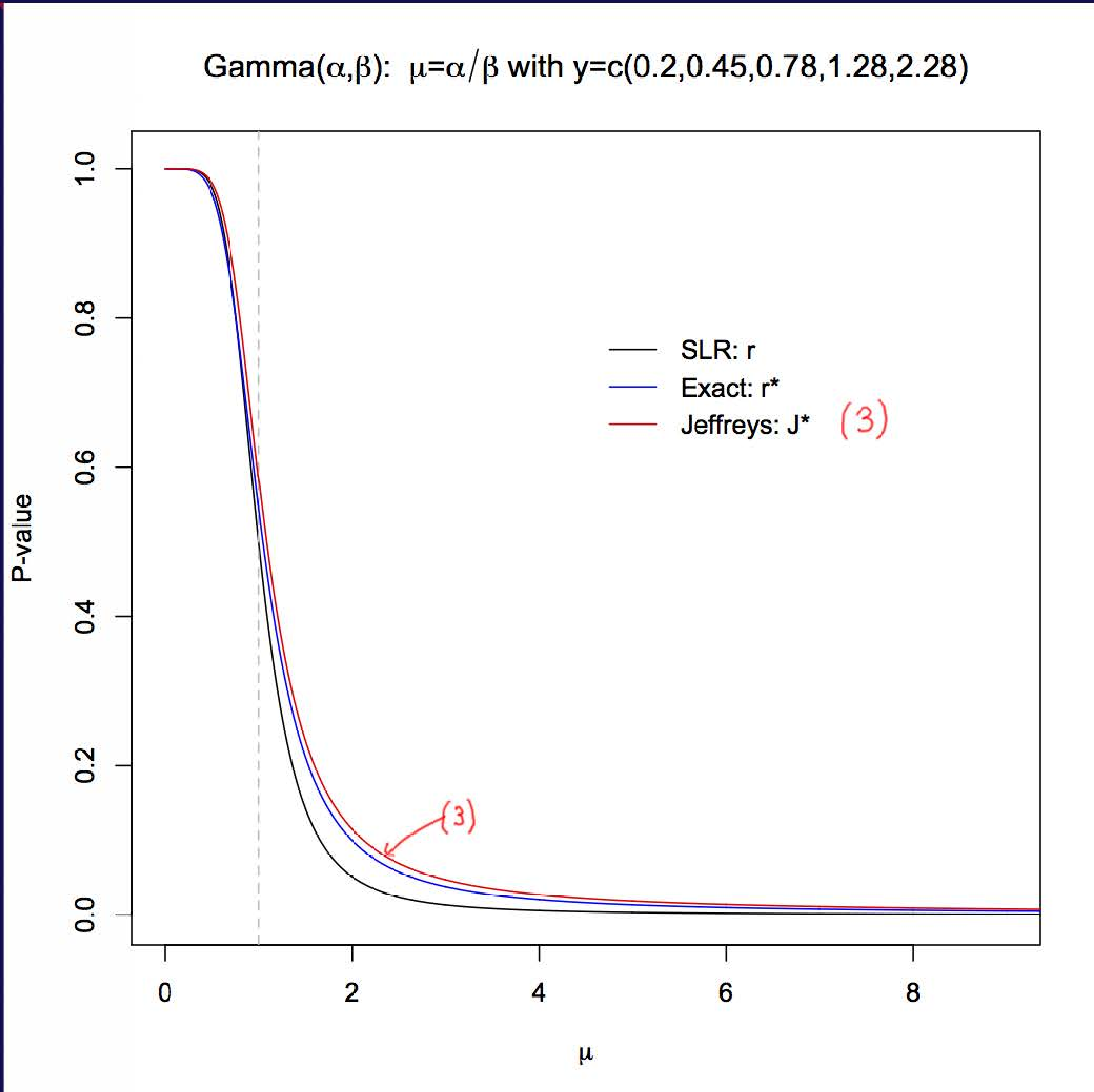
Gamma(α,β): $\mu=\alpha/\beta$ with y=c(0.2,0.45,0.78,1.28,2.28)


Gamma(α,β): μ=α/β with y=c(0.2,0.45,0.78,1.28,2.28)
P-value
1.0
0.8
0.6
0.4
0.2
0.0
SLR: r
(1)
0
2
4
6
8
μ

(1) $p(\mu) = \Phi(r)$


Gamma(α,β): μ=α/β with y=c(0.2,0.45,0.78,1.28,2.28)
SLR: r
Exact: r*
(2)
(2)
P-value
μ
0.0
0.2
0.4
0.6
0.8
1.0
0
2
4
6
8

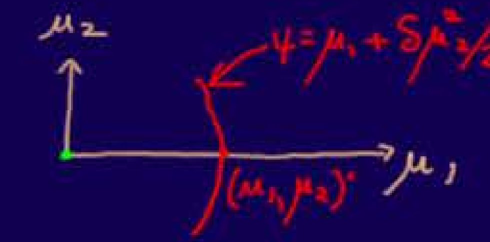
$p(\mu) = \Phi(r^*)$

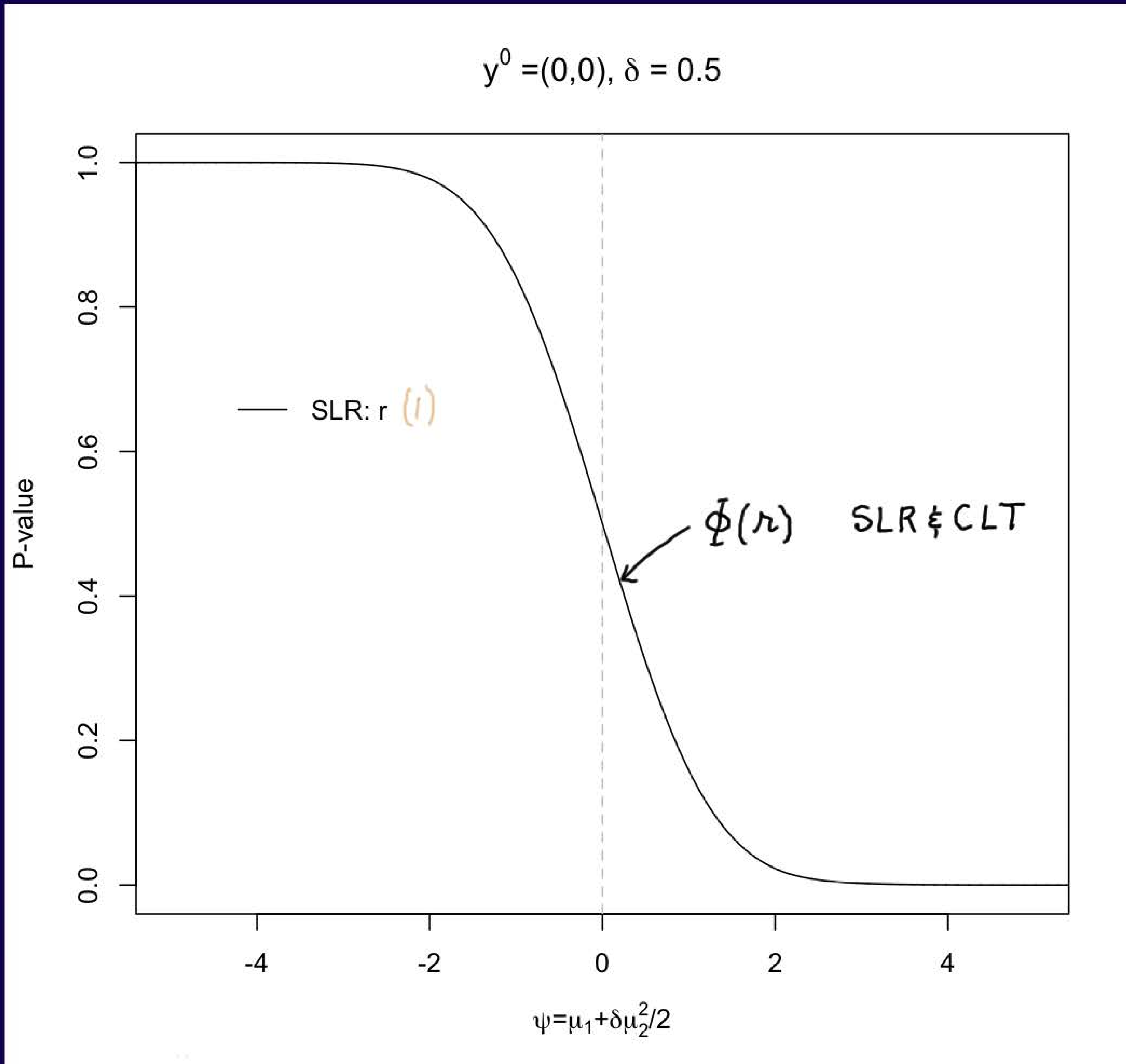

Gamma(α,β): μ=α/β with y=c(0.2,0.45,0.78,1.28,2.28)
P-value
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
μ
SLR: r
Exact: r*
Jeffreys: J*
(3)
(3)

p(μ) = Jeff*

Example 4 $N\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}; I\right)$ $\psi = \mu_1 + \delta\mu_2^2/2$

$y^0 = \begin{pmatrix}0\\ 0\end{pmatrix}$ Curved Interest (A story in itself!)



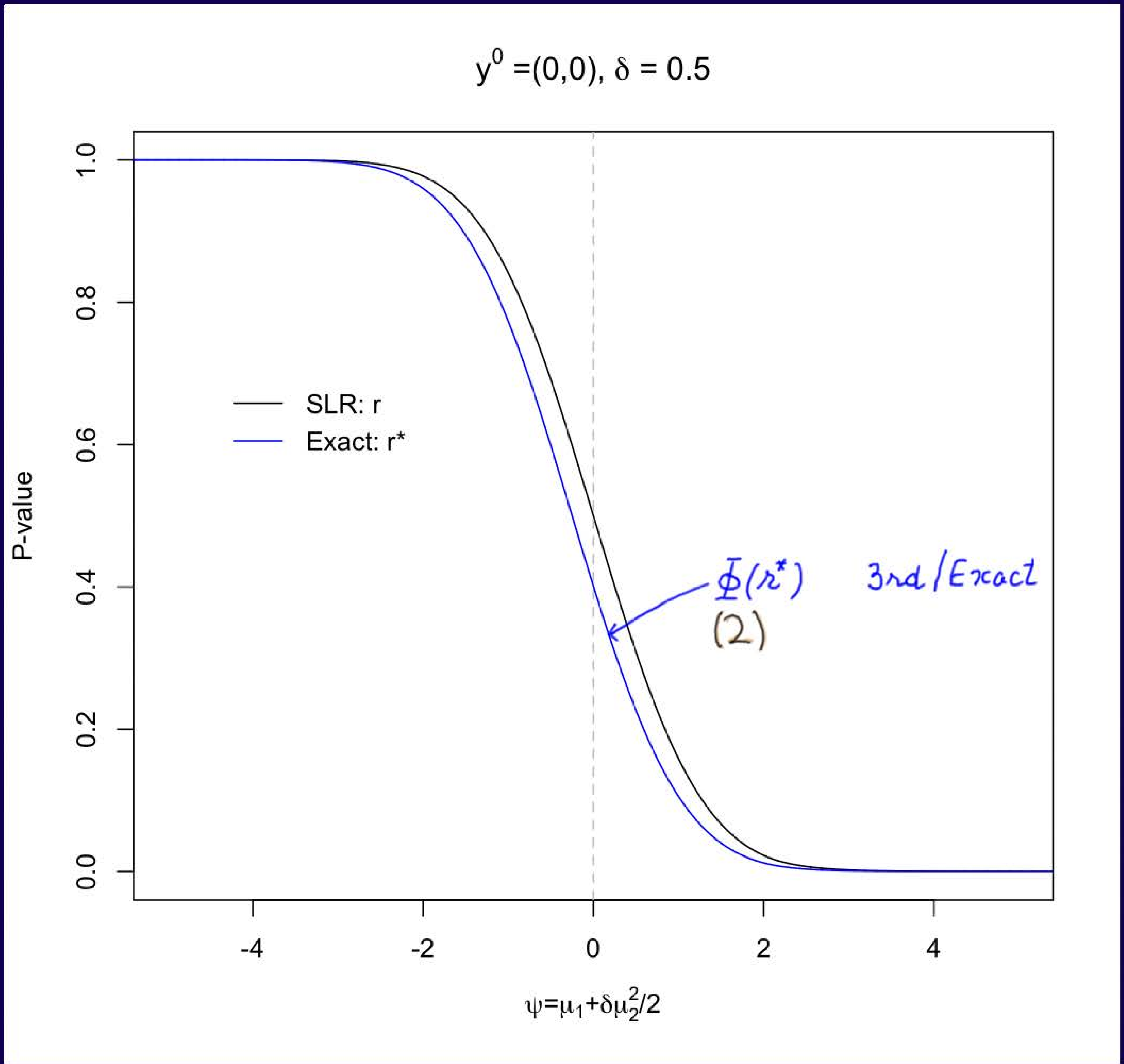


(1) SLR

$\Phi(r)$

Example 4 $N\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}; I\right)$ $\psi = \mu_1 + \delta\mu_2^2/2$

$y^\circ = \begin{pmatrix}0\\0\end{pmatrix}$



(1) SLR

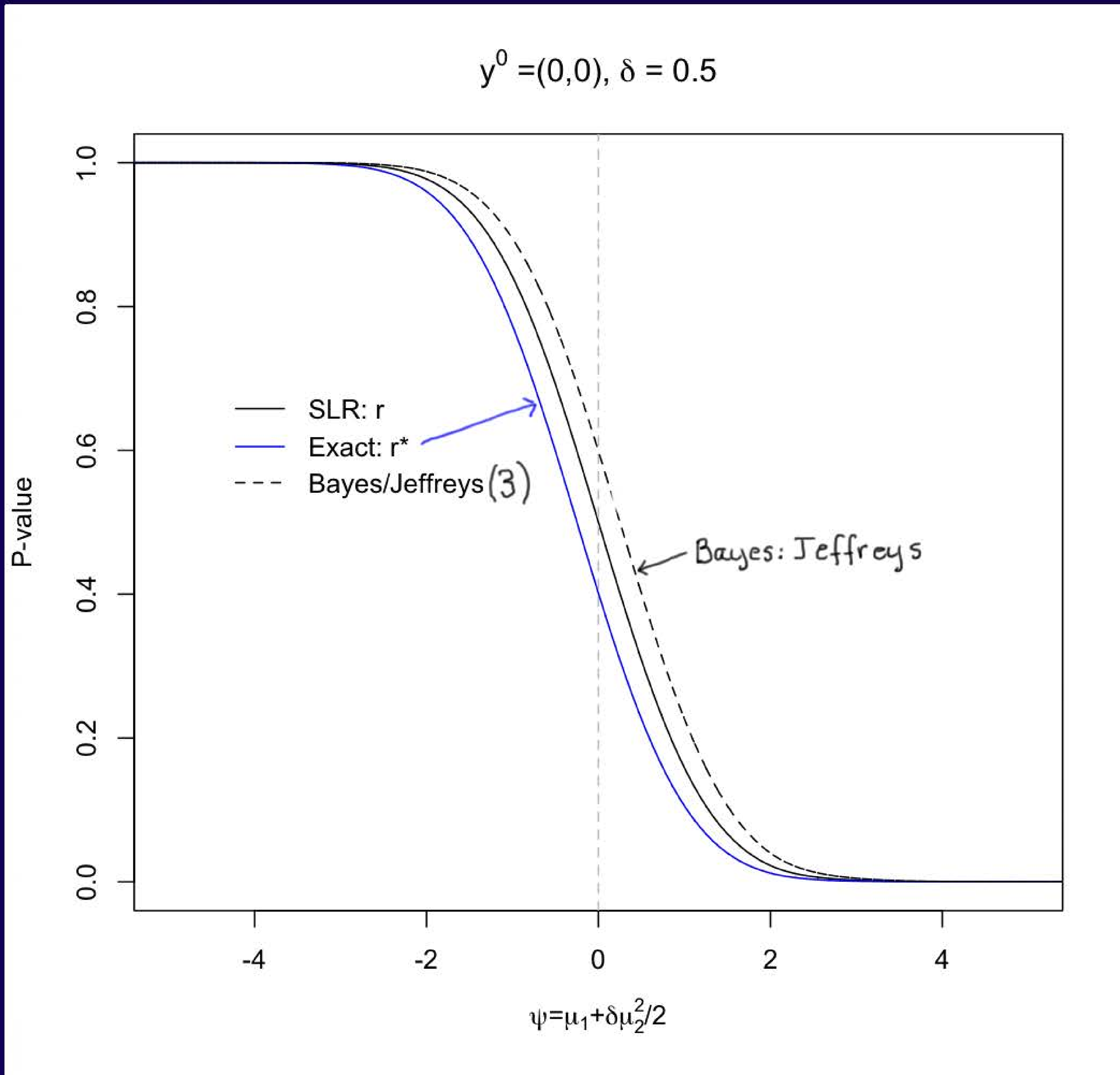
$\Phi(r)$

(2) Exact

$\Phi(r^*)$

Example 4 $N\left(\begin{smallmatrix}\mu_1\\ \mu_2\end{smallmatrix}; I\right)$ $\quad \psi=\mu_1+\delta\mu_2^2/2$

$y^0=\left(\begin{smallmatrix}0\\ 0\end{smallmatrix}\right)$



(1) SLR

$\Phi(r)$

(2) Exact

$\Phi(r^*)$

(3) Bayes

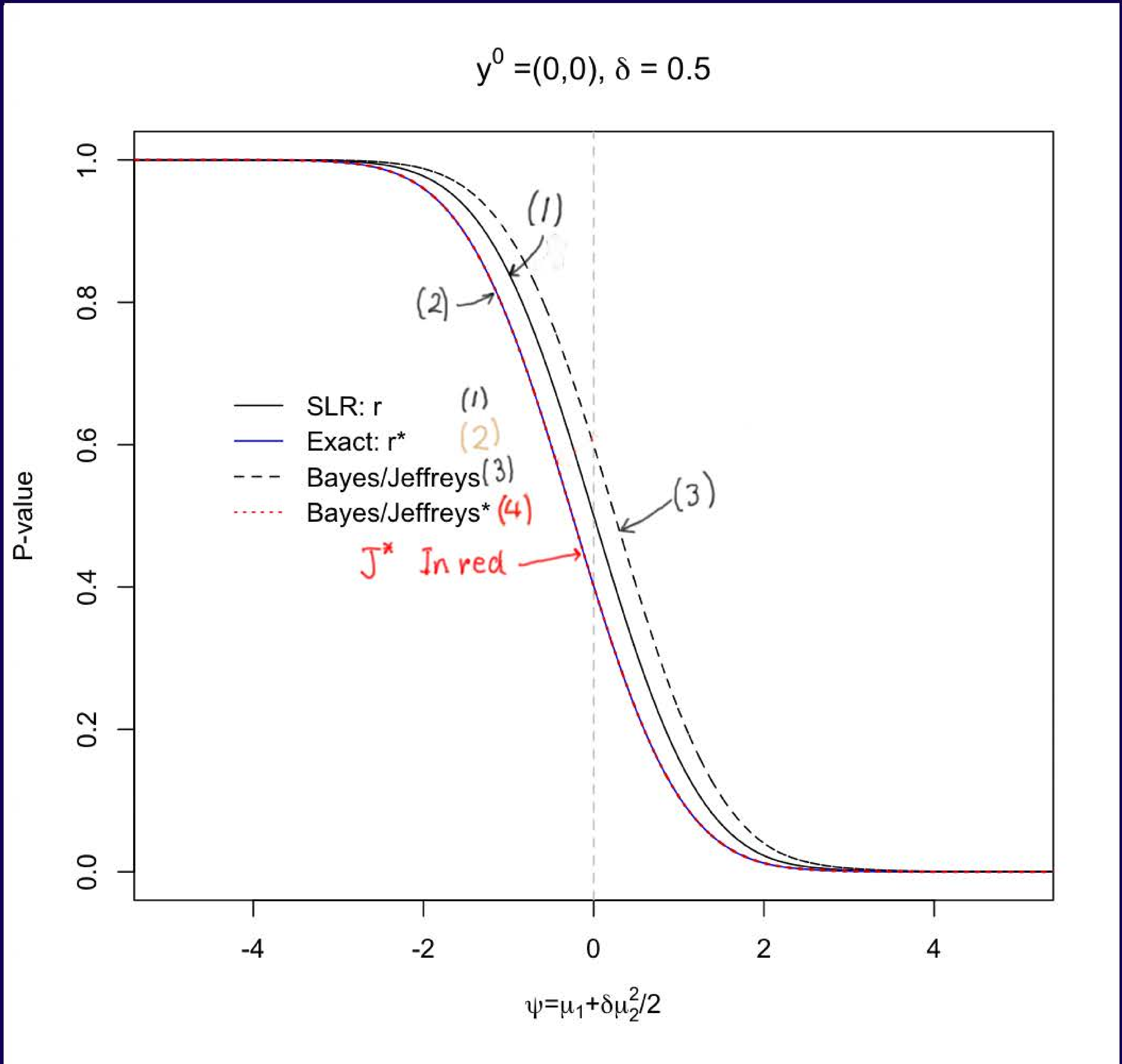
Plain Jeffreys

(Goes farther away from the Exact)

Worse than just L

Example 4 $N\left(\begin{smallmatrix}\mu_1\\ \mu_2\end{smallmatrix}; I\right)$ $\psi = \mu_1 + \delta\mu_2^2/2$

$y^0 = \left(\begin{smallmatrix}0\\0\end{smallmatrix}\right)$



(1) SLR
$\Phi(r)$

(2) Exact
$\Phi(r^*)$

(3) Bayes
Plain Jeffreys
(goes farther away from the Exact)

(4) J^*
Jeffreys but just on profile curve for interest α

J^* almost duplicates the exact

Discussion:

To get better Accuracy than: mle, score, or SLR $O(n^{-1/2})$

Bayes can:

)

Discussion:

To get better Accuracy than: mle, score, or SLR $O(n^{-1/2})$

Bayes can

1) Find what info is available $O(n^{-1})$ Recent lik. theory)

Discussion:

To get better Accuracy than: mle, score, or SLR $O(n^{-1/2})$

Bayes can

1) Find what info is available $O(n^{-1})$ Recent lik. theory)

2) Target on the particular parameter of interest

Discussion:

To get better Accuracy than: mle, score, or SLR $O(n^{-1/2})$

Bayes can

1) Find what info is available $O(n^{-1})$ Recent lik. theory)

2) Target on the particular parameter of interest

3) Use Jeffreys where info is ... on profile !

Discussion:

To get better Accuracy than: mle, score, or SLR $O(n^{-1/2})$

Bayes can

1) Find what info is available $O(n^{-1})$ Recent lik. theory)

2) Target on the particular parameter of interest

3) Use Jeffreys where info is ... on profile !

4) Calibrate if "non-linear"

Discussion:

To get better Accuracy than: mle, score, or SLR $O(n^{-1/2})$

Bayes can

1) Find what info is available $O(n^{-1})$ Recent lik. theory)

2) Target on the particular parameter of interest

3) Use Jeffreys where info is ... on profile!

4) Calibrate if "non-linear"

or go 3rd order and forget Bayes

How Bayes can deliver 2nd order Accuracy!

A long history
great collaborators: Nancy

A Wong	York
M Bédard	U de Montréal
W Lin	Toronto
A M Fraser	UBC
M J Fraser	Toronto

(Preliminary report)