

Another view of Composite Likelihood

D. A. S. Fraser

Statistical Sciences

U Toronto

2013 January 31

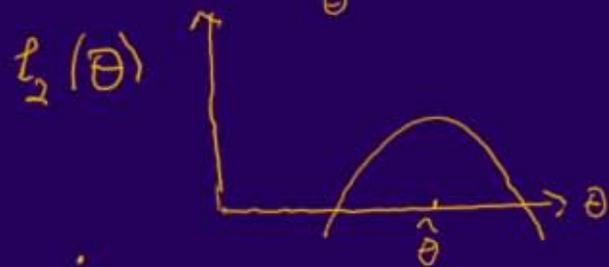
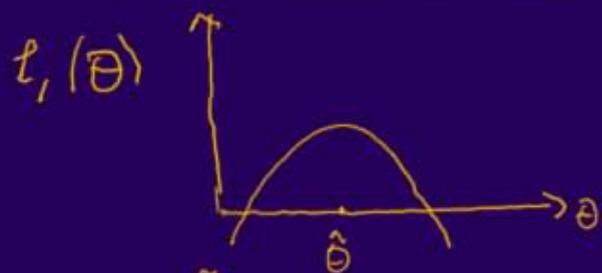
University Western Ontario

www.ustat.toronto.edu/dfraser/DWD2013_S
/UWO2013_L

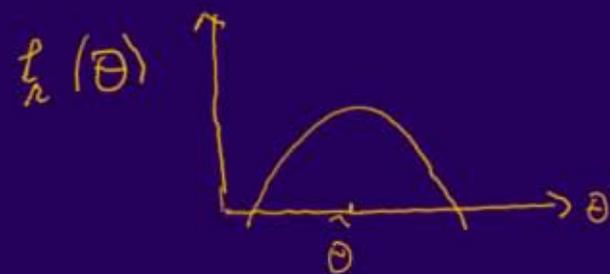
1	Combining likelihoods: Is there a problem?	2
2	Is there a role for asymptotics?	2
3	Combining first order	1
4	Constructed likelihood	1
5	1st-order log-L from data	1
6	Examples 1, 2, 3	3
7	Discussion	1
8	-Segmented Likelihood	1
9	-Vector parameter case	1
		<hr/>
		13

16 slides .. Segmented

1 Combining likelihoods: Is there a problem?



⋮



Have n observed log-likelihoods

- Want to combine them

- Add them up: $l(\theta) = l_1(\theta) + \dots + l_n(\theta)$

- dependence / overlap?

- and consistency of $\hat{\theta}$?

Lindsay (1988)

Cox & Reid (2004)

Varin (2004)

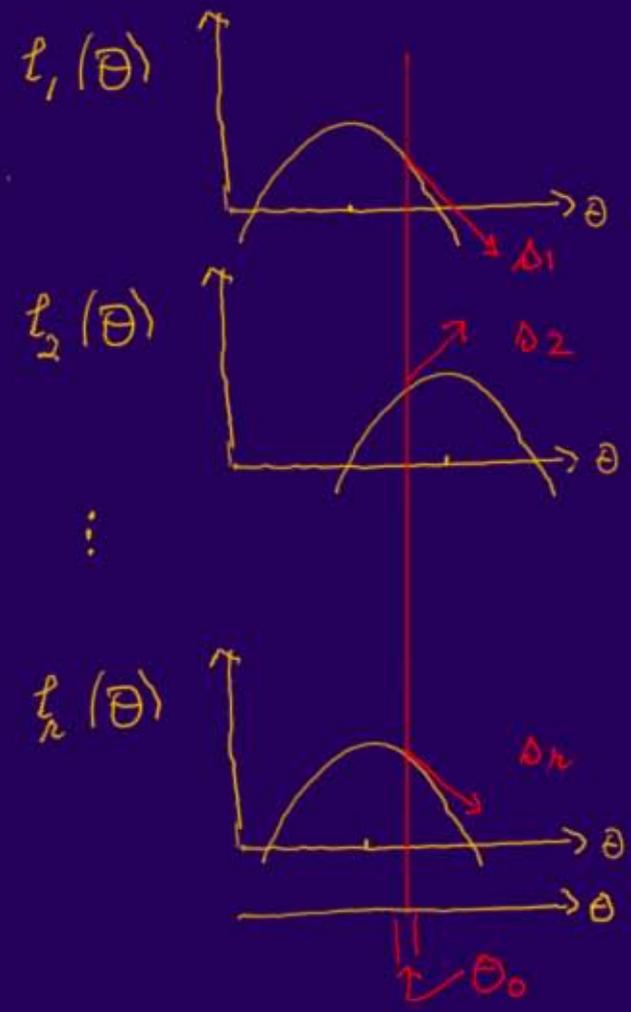
Widely; Applications; Theory

Why not get actual obs. likelihood?

⇒ available? accessible? feasible?

Often not!

2 A role for asymptotics? and least squares?



Asymptotics:

- Increasing data size n
- Separate dependence on n $O(n^{-1/2}), O(n^{-1})$ as in CLT proof
- Incredibly fruitful - Tests, p-values, confidence $xxx = 260$
- priors (agree with $\uparrow \dots$) $xxx = 265$

Here: ... Can we even get first order??

Take a "trial true" θ_0 {say from $\sum l_i(\theta)$ } and look $\theta_0 \pm$ 1st derivative ... i.e. 1st order

What do we have? Use $\theta \leftarrow \theta - \theta_0$ " " " "

$$l_1(\theta) = a + \delta_1 \theta + \dots$$

$$\vdots$$

$$l_n(\theta) = a + \delta_n \theta + \dots$$

Just scores:

and how do we combine?

$$l_1(\theta) = a + s_1\theta +$$

$$\vdots$$

$$l_n(\theta) = a + s_n\theta +$$

What distribution for s_1, \dots, s_n near $\theta = \theta_0$? Comes from context;
 1st order \Rightarrow Means, Variances

Need only mini-Bartlett's

$l(\theta; y) \quad s = l_{\theta}(\theta_0; y)$ $E(s; \theta_0) = 0$ $\frac{\partial}{\partial \theta} E(s; \theta) \Big|_{\theta_0} = i_{\theta\theta}(\theta_0)$

\swarrow scores at θ_0 $V(s; \theta_0) = l_{\theta\theta}(\theta_0)$ $E(s; \theta) = l_{\theta\theta}(\theta_0) \cdot (\theta - \theta_0)$

\nearrow call it θ

Thus... 1st order analysis:

$$\underline{s} = \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} \quad E(\underline{s}; \theta) = \underline{N}\theta = \begin{pmatrix} i_{11} & \dots & i_{1n} \\ \vdots & & \vdots \\ i_{n1} & \dots & i_{nn} \end{pmatrix} \theta$$

infos $N_{ii} = i_{ii}$

$$V(\underline{s}; \theta) = V = \begin{pmatrix} i_{11} & \dots & i_{1n} \\ \vdots & & \vdots \\ i_{n1} & \dots & i_{nn} \end{pmatrix}$$

cross infos $N_{ij} = i_{ij}$

where $N_{ii} = i_{11} = \text{var}(s_1; \theta_0)$ & needs only $f(x_1; \theta)$ model "easy"

$l_{12} = \text{cov}(s_1, s_2; \theta_0)$ & needs $f(x_1, x_2; \theta)$ model plus "calculations" ... later

3 Combining - first order

Variable s Mean = $n\theta$ Var = V n, V given $0 \leftarrow \theta_0$

Have: s° Use Θ for $\theta - \theta_0$... departure from Trial θ_0

Want to combine $l_i(\theta) = s_i^\circ \theta$ || $s \sim (E, \text{Var}) = (n\theta, V)$
and get best $l(\theta) = S\theta$ || s°

Least squares; unbiased estimation:

	Data	E	Var	Est $\hat{\theta}^\circ$	Var $\hat{\theta}$
GM	y°	$X\theta$	$\sigma^2 I$	$(X'X)^{-1} X'y^\circ$	$\sigma^2 (X'X)^{-1}$
GGM	y°	$X\theta$	Σ	$(X'\Sigma^{-1}X)^{-1} X'\Sigma^{-1}y^\circ$	$(X'\Sigma^{-1}X)^{-1}$
<u>Here</u>	s°	$n\theta$	V	$(n'V^{-1}n)^{-1} n'V^{-1}s^\circ$	$(n'V^{-1}n)^{-1}$ *

5 1st-order log-L from data ("2nd" unavailable!)

41

Given: Vector $\underline{\ell}(\theta)$ of log-likelihoods & Trial true θ_0

Calculate: $\underline{\rho} = \{ \ell_{i,\theta}(\theta_0), \dots, \ell_{n,\theta}(\theta_0) \}'$ score vector

$n = (n_{ii})$ info's

$V = (n_{ij})$ cross info's

Weight vector $n' V^{-1}$

\Rightarrow New log-Lik $\ell = n' V^{-1} \underline{\ell}(\theta)$

Information $n' V^{-1} n$

\Rightarrow New mle $\hat{\theta} = (n' V^{-1} n)^{-1} n' V^{-1} \underline{\rho}$

$= \theta_0 + \text{increment}$

log-L slopes at θ_0

Iterate if needed: θ_0, θ_1

6 Examples: $x_i \sim \phi(x_i - \theta)$ Normal($\theta, 1$)

Ex 1 X 's are also scores here

$$X_1 = x_1 + x_2$$

$$X_2 = x_1 + x_3$$

$$n = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \quad V = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

$$V^{-1} = \begin{pmatrix} 2/3 & -1/3 \\ -1/3 & 2/3 \end{pmatrix}$$

weights info

$$n' V^{-1} = \begin{pmatrix} 2/3 & 2/3 \end{pmatrix} \quad n' V^{-1} n = 8/3$$

The combined likelihoods:

add them

$$1) \quad l_{CL}(\theta) = l_1(\theta) + l_2(\theta) = -2\theta^2 + \theta(X_1 + X_2) \dots$$

Scale adjust

$$2) \quad l_{ACL}(\theta) = \frac{2}{5} \{ l_1(\theta) + l_2(\theta) \} = -\frac{4}{5}\theta^2 + \frac{2}{5}\theta(X_1 + X_2)$$

$$3) \quad l_{new}(\theta) = \frac{2}{3} l_1(\theta) + \frac{2}{3} l_2(\theta) = -\frac{4}{3}\theta^2 + \frac{2}{3}\theta(X_1 + X_2)$$

	$\hat{\theta}$	Inverse "info"	$\text{var} \hat{\theta}$
"info"			
4	$\frac{X_1 + X_2}{4}$ ✓	1/4 ✓	5/8 ✓
8/5	$\frac{X_1 + X_2}{4}$	5/8	5/8
8/3	$\frac{X_1 + X_2}{4}$	3/8	3/8

inverses

agree

Example ... symmetry ... ACL adjusts "info" but too much!

Ex 2

$$x_i \sim \phi(x_i - \theta) \quad \text{Normal}(\theta, 1)$$

asymmetry

$$y_1 = x_1$$

$$y_2 = x_1 + x_3$$

$$v = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad V = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

$$V^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$$

$$v' V^{-1} = (0, 1)$$

$$v' V^{-1} v = 2$$

Inverse "info"

var $\hat{\theta}$

add them

$$1) \ell_{CL}(\theta) = \ell_1(\theta) + \ell_2(\theta) = -\frac{3}{2}\theta^2 + \theta(y_1 + y_2) \dots$$

scale adjust

$$2) \ell_{ACL}(\theta) = \frac{3}{5} \{ \ell_1(\theta) + \ell_2(\theta) \} = -\frac{9}{10}\theta^2 + \frac{3}{5}\theta(y_1 + y_2)$$

only 2nd

$$3) \ell_{new}(\theta) = 0 + \ell_2(\theta) = -\theta^2 + \theta y_2$$

Corrects: "info" & mle

Asymmetry \Rightarrow both "info" & "estimate" corrections

$\hat{\theta}$	Inverse "info"	var $\hat{\theta}$
$\frac{y_1 + y_2}{3}$	$1/3$	$5/9$
$\frac{y_1 + y_2}{3}$	$5/9$ <small>Equal Too big</small>	$5/9$
$y_2/2$ <small>uses</small>	$1/2$ <small>corrected equal</small>	$1/2$

Exc 3

$$x_i \sim \phi(x_i - \theta) \quad \text{Normal}(\theta, 1)$$

62

$$y_1 = x_1 + x_2$$

$$y_2 = x_1 + x_3 + x_4 + x_5 + x_6$$

asymmetry

$$N = \begin{pmatrix} 2 \\ 5 \end{pmatrix} \quad V = \begin{pmatrix} 2 & 1 \\ 1 & 5 \end{pmatrix}$$

$$V^{-1} = \begin{pmatrix} 5/9 & -1/9 \\ -1/9 & 2/9 \end{pmatrix}$$

$$N' V^{-1} = \begin{pmatrix} 5/9 & 8/9 \end{pmatrix}$$

$$N' V^{-1} N = \frac{50}{9}$$

weights

info

$\hat{\theta}$

Inverse
"info"

$\text{var } \hat{\theta}$

$$\frac{y_1 + y_2}{7}$$

$$1/7$$

$$9/49$$

$$\frac{y_1 + y_2}{7}$$

$$9/49$$

equal
Too big

$$9/49$$

$$\frac{5y_1 + 8y_2}{50}$$

$$9/50$$

equal

$$9/50$$

add them

$$1) \ell_{CL}(\theta) = \ell_1(\theta) + \ell_2(\theta) = -\frac{7}{2} \theta^2 + \theta(y_1 + y_2) \dots$$

scale adjust

$$2) \ell_{ACL}(\theta) = \frac{9}{7} \{ \ell_1(\theta) + \ell_2(\theta) \} = -\frac{9}{2} \theta^2 + \frac{9}{7} \theta(y_1 + y_2)$$

$$3) \ell_{new}(\theta) = \frac{5}{9} \ell_1(\theta) + \frac{8}{9} \ell_2(\theta) = -\frac{25}{9} \theta^2 + \theta \left(\frac{5}{9} y_1 + \frac{8}{9} y_2 \right)$$

7 Discussion

66

If model symmetric under coord. permutations:

Indication: Scores symmetric

ACL: Estimate OK

Bartlett OK

Info under-states accuracy

If model not symmetric

Indication: CL and ACL may misrepresent

New: weighted likelihood OK

In general

get $l(\theta; X_1^o, \dots, X_n^o)$ $O(n^{1/2})$ & use only $f(X_i^o, X_j^o; \theta)$

Second order seems unavailable.

8 Segmented likelihood: getting $\hat{l}_{ij} = N_{ij}$

Ex 3 $y_1 = x_1 + x_2$ $x_i \sim \phi(x_i - \theta)$
 $y_2 = x_1 + x_3 + x_4 + x_5 + x_6$

Underlying randomness: has an 'intersection' & 'union'

$\tilde{y}_1 = x_2$ \tilde{y}_2 corresponds to "intersection" $\hat{l}_{12} = N_{12}$
 $\tilde{y}_2 = x_1$ $(\tilde{y}_1, \tilde{y}_2, \tilde{y}_3)$ corr. to "union" $\hat{l}_{(1,2)} = N_{11} + N_{22} - N_{12}$
 $\tilde{y}_3 = x_3 + x_4 + x_5 + x_6$

Use model $f(\tilde{y}_2; \theta)$

or model $f(\tilde{y}_1, \tilde{y}_2, \tilde{y}_3; \theta)$

to extract covariances $\hat{l}_{12} = N_{12}$

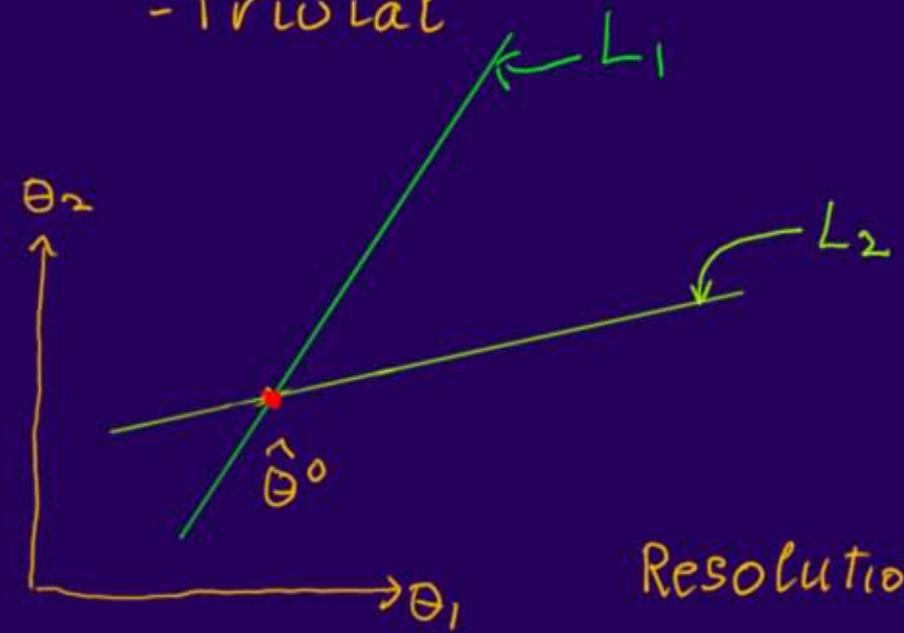
by an "information" calculation.

9 Vector $\theta = (\theta_1, \dots, \theta_p)$

- Intractable
- Trivial

Maybe $l_1(\theta)$ only about L_1 ?
 $l_2(\theta)$ " " L_2 ?

Should this affect combining ?



- or
- Multiple variance matrices
 - Trivial

Resolution:

- Do a combination for L_1
- Do a combination for L_2
- Do for each direction from data

It is well within the logic of the problem

- 1 Combining likelihoods: Is there a problem?
- 2 Is there a role for asymptotics?
- 3 Combining first order
- 4 Constructed likelihood
- 5 1st-order log-L from data
- 6 Examples 1, 2, 3
- 7 Discussion
 - 8 -Segmented Likelihood
 - 9 -Vector parameter case

Thank
you