Applied Asymptotics in R

Examples of the use of higher order likelihood inference

Alessandra R. Brazzale

Institute of Biomedical Engineering
Italian National Research Council, Padova

alessandra.brazzale@isib.cnr.it

The R User Conference 2006
Vienna, 15–17 June 2006

What is this talk about?

- an R package bundle

hoa

Higher Order (small sample) Asymptotics

$$n \longrightarrow \infty$$

for likelihood-based parametric inference

- . . . and where to read more about the subject

A toy example

- i.i.d. sample $y_1, \ldots, y_n$ from the Cauchy distribution

$$f(y_i;\theta) = \frac{1}{\pi\{1 + (y_i - \theta)^2\}}$$

- log likelihood function: $\ell(\theta; y) = -\sum_{i=1}^n \log\{1 + (y_i - \theta)^2\}$
- maximum likelihood estimator: $\hat\theta = \text{argmax}_\theta\, \ell(\theta; y)$

$n = 1$

$$\hat\theta = y$$

$$F(\hat\theta;\theta) = F(y;\theta) = \pi^{-1}\arctan(y - \theta)$$

Likelihood inference

confidence intervals and p-values are computed using

$$p(\theta;\hat\theta) = \Pr(\hat\Theta \le \hat\theta; \theta)$$

- exact: $p(\theta;\hat\theta) = \Pr(Y \le y; \theta)$
- approximate:

$$p(\theta;\hat\theta) = \Phi(\text{pivot}) + O_p(n^{-1/2})$$

  - Wald pivot: $w(\theta) = \sqrt{2}(y - \theta)$
  - likelihood root: $r(\theta) = \text{sign}(\hat\theta - \theta)\left[2\log\{1 + (y - \theta)^2\}\right]^{1/2}$
  - score pivot: $s(\theta) = \sqrt{2}(y - \theta)/\{1 + (y - \theta)^2\}$

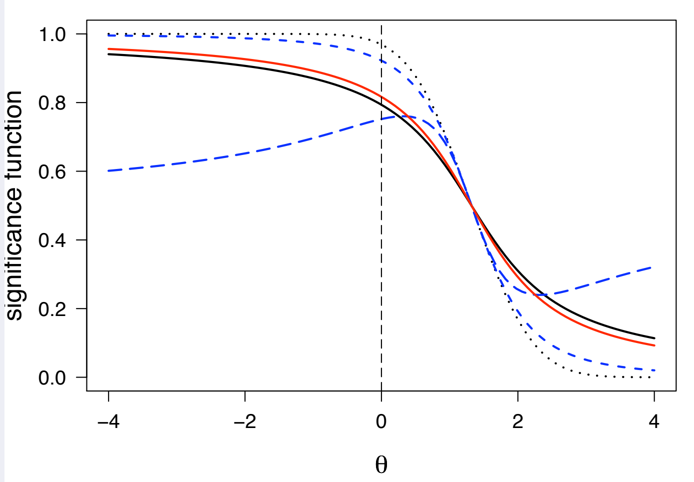
Can we do better?

$$p(\theta;\hat\theta) = \Phi(\text{pivot}) + O_p(n^{-3/2})$$

- modified likelihood root

$$r^*(\theta) = r(\theta) + \frac{1}{r(\theta)} \log \frac{s(\theta)}{r(\theta)}$$

$y = 1.32 \quad (n = 1)$



And what if $n > 1$?

There is no exact solution, but . . .

marg[hoa] package

```
> library( marg )
> set.seed( 321 )
> y <- rt( n = 15, df = 3 )
> y.rsm <- rsm( y ~ 1, family = student(3) )
> y.cond <- cond( y.rsm, offset = 1 )
> summary( y.cond, test = 0 )
```

p-values: 0.282 (Wald), 0.306 (r), 0.354 (r^*)

General theory

- $\theta = (\psi, \lambda)$, with scalar parameter of interest ψ
- significance function

$$p(\psi;\hat\psi) = \Pr(\hat\Psi \le \hat\psi; \psi)$$

- profile log likelihood: $\ell_p(\psi) = \ell(\psi, \hat\lambda_\psi; y)$
  - Wald statistic: $w(\theta) = j_p(\hat\psi)^{1/2}(\hat\psi - \psi)$
  - likelihood root: $r(\theta) = \text{sign}(\hat\psi - \psi)\left[2\{\ell_p(\hat\psi) - \ell_p(\psi)\}\right]^{1/2}$
  - score statistic: $s(\theta) = j_p(\hat\psi)^{-1/2} d\ell_p(\psi)/d\psi$

with $j_p(\psi) = -d^2\ell_p(\psi)/d\psi^2$

Higher order inference

Modified likelihood root

$$r^*(\psi) = r(\psi) + \frac{1}{r(\psi)} \log \frac{q(\psi)}{r(\psi)}$$

with $q(\psi)$ representing a suitable correction term

- $p(\psi; \hat{\psi}) = \Phi\{r^*(\psi)\} + O_p(n^{-3/2})$
- $r^*(\psi) = r(\psi) + r_{inf}(\psi) + r_{np}(\psi)$
  - $r_{inf}(\psi)$: information adjustment
  - $r_{np}(\psi)$: nuisance parameter adjustment

The hoa bundle

- `cond`: logistic regression

$$\Pr(Y_i = 1; \beta) = \frac{\exp(x_i^\top \beta)}{1 + \exp(x_i^\top \beta)}$$

- `marg`: linear nonnormal models

$$y_i = x_i^\top \beta + \sigma\varepsilon_i, \quad \varepsilon_i \sim f_0(\cdot)$$

- `nlreg`: nonlinear heteroscedastic regression

$$y_{ij} = \mu(x_i; \beta) + \omega(x_i; \beta, \rho)\varepsilon_{ij}, \quad \varepsilon_{ij} \sim N(0, 1)$$

- `csampling`: conditional sampling routines

airway data

```
> head(airway)

  response age sex lubricant duration type
1        0  48   1         0       45    0
2        0  48   1         0       15    0
3        1  39   0         1       40    0
4        1  59   1         0       83    1
5        1  24   1         1       90    1
6        1  55   1         1       25    1
```

Collet (1998)

airway data (2/3)

```
> airway.glm <- glm( formula(airway), family = binomial,
+                     data = airway )

> library( cond )

> airway.cond <- cond( airway.glm, offset = type1 )

> summary( airway.cond )
```

airway data (3/3)

```
Confidence intervals
--------------------
 level = 95 %
                                  lower  two-sided  upper
Wald pivot                        -3.486            0.2271
Wald pivot (cond. MLE)            -3.053            0.2655
Likelihood root                   -3.682            0.1542
Modified lik. root                -3.130            0.2558
Modified lik. (cont. corr.)       -3.592            0.5649

Diagnostics:
-----------
    INF       NP
0.05855 0.51426
```


calcium uptake data

```
> library( boot)

> head( calcium )

  time       cal
1 0.45   0.34170
2 0.45  -0.00438
3 0.45   0.82531
4 1.30   1.77967
5 1.30   0.95384
6 1.30   0.64080
```

Davison & Hinkley (1997, Example 7.7)

calcium uptake data (2/3)

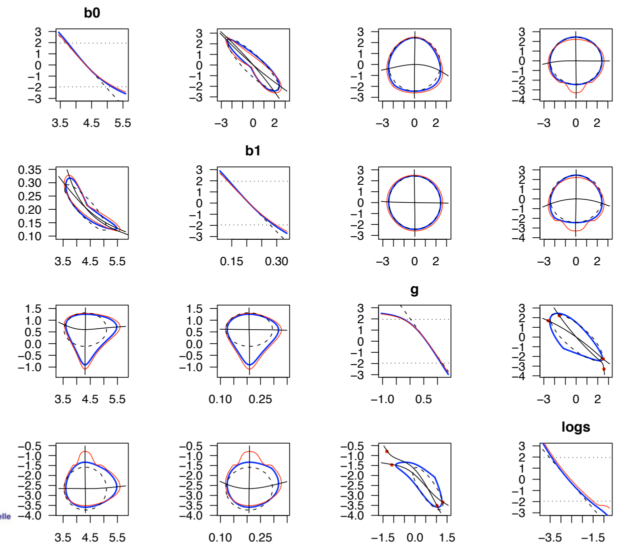
$$\mu(x_i;\beta) = \beta_0\{1 - \exp(-\beta_1 x_i)\}, \quad \omega^2(x_i;\gamma) = \sigma^2(1 + x_i^\gamma)$$

```
> library( nlreg )

> calcium.nl <-
+     nlreg( cal ~ b0 * (1 - exp(-b1 * time)),
+            weights = ~ 1 + time^g, data = calcium,
+            start = c(b0 = 4, b1 = 0.1, g = 0) )

> calcium.prof <- profile( calcium.nl )

> contour( calcium.prof, alpha = 0.05, lwd1 = 2,
+          lwd2 = 2 )
```

And if you wish to try more . . .

- Brazzale, A. R. (2005). hoa: An R package bundle for higher order likelihood inference. *Rnews*, Vol. 5/1, May 2005, pp. 20–27.

  R vignette in `hoa` v. 1.1-0

- Brazzale, A. R., Davison, A. C. and Reid, N. (2006). *Applied Asymptotics*. Cambridge University Press. (Forthcoming)

  `www.isib.cnr.it/~brazzale/CS`
  theory & implementation & examples and case studies

Credits

- Alessandra Salvan, Anthony C. Davison, Nancy Reid
- Ruggero Bellio
- Douglas M. Bates, Kurt Hornik, Torsten Hothorn

. . . and the useRs!