What is this talk about?

- an R package bundle

hoa

Higher Order (small sample) Asymptotics

\[ n \xrightarrow{\text{as}} \infty \]
for likelihood-based parametric inference

... and where to read more about the subject

Likelihood inference

confidence intervals and p-values are computed using

\[ p(\theta; \hat{\theta}) = \Pr(\hat{\theta} \leq \theta; y) \]

- exact: \[ p(\theta; \hat{\theta}) = \Pr( Y \leq y; \theta) \]
- 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:

\[ s(\theta) = \sqrt{2}(y - \theta) / \left( 1 + (y - \theta)^2 \right) \]

A toy example

i.i.d. sample \( y_1, \ldots, y_n \) from the Cauchy distribution

\[ f(y; \theta) = \frac{1}{\pi \left( 1 + (y - \theta)^2 \right)} \]

log likelihood function:

\[ \ell(\theta; y) = -\sum_{i=1}^n \log(1 + (y_i - \theta)^2) \]

maximum likelihood estimator:

\[ \hat{\theta} = \arg\max_{\theta} \ell(\theta; y) \]

\( n = 1 \)

\[ \hat{\theta} = y \]

\[ F(\hat{\theta}; \theta) = F(y; \theta) = \pi^{-1} \arctan(y - \theta) \]
Can we do better?

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

- modified likelihood root

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

And what if \( n > 1 \)?

There is no exact solution, but …

**marg** [hoa] package

```r
> 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 )
```

\[ p\text{-values: } 0.282 \text{(Wald), } 0.306 \text{ (r), } 0.354 \text{ (r*)} \]

**General theory**

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

\[ p(\psi; \hat{\psi}) = \Pr(\hat{\psi} \leq \hat{\psi}; \psi) \]

- profile log likelihood:

\[ \ell_p(\psi) = \ell(\psi, \hat{\lambda}; y) \]

  - Wald statistic:

\[ w(\theta) = \frac{d\ell_p(\hat{\psi})}{(\hat{\psi} - \psi)} \]

  - likelihood root:

\[ r(\theta) = \text{sign}(\hat{\psi} - \psi) \left( 2(\ell_p(\hat{\psi}) - \ell_p(\psi)) \right)^{1/2} \]

  - score statistic:

\[ s(\theta) = \frac{d\ell_p(\psi)}{d\psi} \]

with \( \ell_p(\psi) = -d\ell_p(\psi)/d\psi^2 \)
### Higher order inference

**Modified likelihood root**

\[ r^*(ψ) = r(ψ) + \frac{1}{r(ψ)} \log \frac{q(ψ)}{r(ψ)} \]

with \( q(ψ) \) representing a suitable correction term

- \( p(ψ; ̂ψ) = Φ\{r^*(ψ)\} + O_p(σ_n^{-3/2}) \)
- \( r^*(ψ) = r(ψ) + r_{inf}(ψ) + r_{np}(ψ) \)
  - \( r_{inf}(ψ) \): information adjustment
  - \( r_{np}(ψ) \): nuisance parameter adjustment

---

### The hoa bundle

- **cond**: logistic regression
  \[ Pr(Y_i = 1; β) = \frac{\exp(x_i^T β)}{1 + \exp(x_i^T β)} \]
- **marg**: linear nonnormal models
  \[ y_i = x_i^T β + σ_i ε_i, \quad ε_i \sim f_0(·) \]
- **nlreg**: nonlinear heteroscedastic regression
  \[ y_{ij} = μ(x_i; β) + ω(x_i; β, ρ)ε_{ij}, \quad ε_{ij} \sim N(0, 1) \]
- **csampling**: conditional sampling routines

---

### airway data

- **airway.data**

```r
> head(airway)
response age sex lubricant duration type
1  0 48 1 0  45 1
2  0 48 1 0  45 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)
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

Davison & Hinkley (1997, Example 7.7)
  
  R vignette in hoa v. 1.1-0

  
  www.isib.cnr.it/~brazzale/CS

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

... and the useRs!