STA 442F Review of Likelihood

Note Title 10/23/2007

Model $Y_1, \ldots, Y_n$ i.i.d. $f(y; \theta)$ $\theta \in \mathbb{R}^p$

Likelihood $L(\theta; \underline{y}) = L(\theta; y_1, \ldots, y_n)$

$$= \prod_{i=1}^{n} f(y_i; \theta)$$

Log-likelihood $\ell(\theta; \underline{y}) = \sum \ell(\theta; y_i)$

$$= \sum_{i=1}^{n} \log f(y_i; \theta)$$

Maximum likelihood estimate $\hat{\theta} = \arg\sup_{\theta} \ell(\theta)$

Score function $\ell'(\theta) = \ell'(\theta; y) = \dfrac{\partial \ell(\theta; \underline{y})}{\partial \theta}$ $p \times 1$ vector

- if $\ell(\theta)$ is smooth, 'well-behaved', then $\hat{\theta}$ is found by

$$\ell'(\hat{\theta}) = 0 \qquad \hat{\theta} = \hat{\theta}(y)$$

- this is usually a nonlinear equation

- observed information function $\quad j(\theta) = -l''(\theta; y)$

$p \times p$ matrix

- expected information function $\quad i(\theta) = E\, j(\theta)$

Example $\quad Y_i \sim \text{Gamma}(\beta, \mu)$

$$f(y_i; \beta, \mu) = \frac{1}{\Gamma(\beta)} \left(\frac{\beta}{\mu}\right)^{\beta} y_i^{\beta - 1} e^{-y_i \beta/\mu}$$

$l(\beta, \mu; y) =$

$\frac{\partial l}{\partial \beta}$

$\frac{\partial l}{\partial \mu}$

Inference based on the likelihood function

$$\frac{1}{\sqrt{n}}\, \ell'(\theta; y) \xrightarrow{d} N(0, i_1(\theta)) \qquad i_1(\theta) = \frac{i(\theta)}{n}$$

Leading to 3 approximations:

$$\ell'(\theta)\{j(\hat{\theta})\}^{-\frac{1}{2}} \;\dot{\sim}\; N(0, I)$$

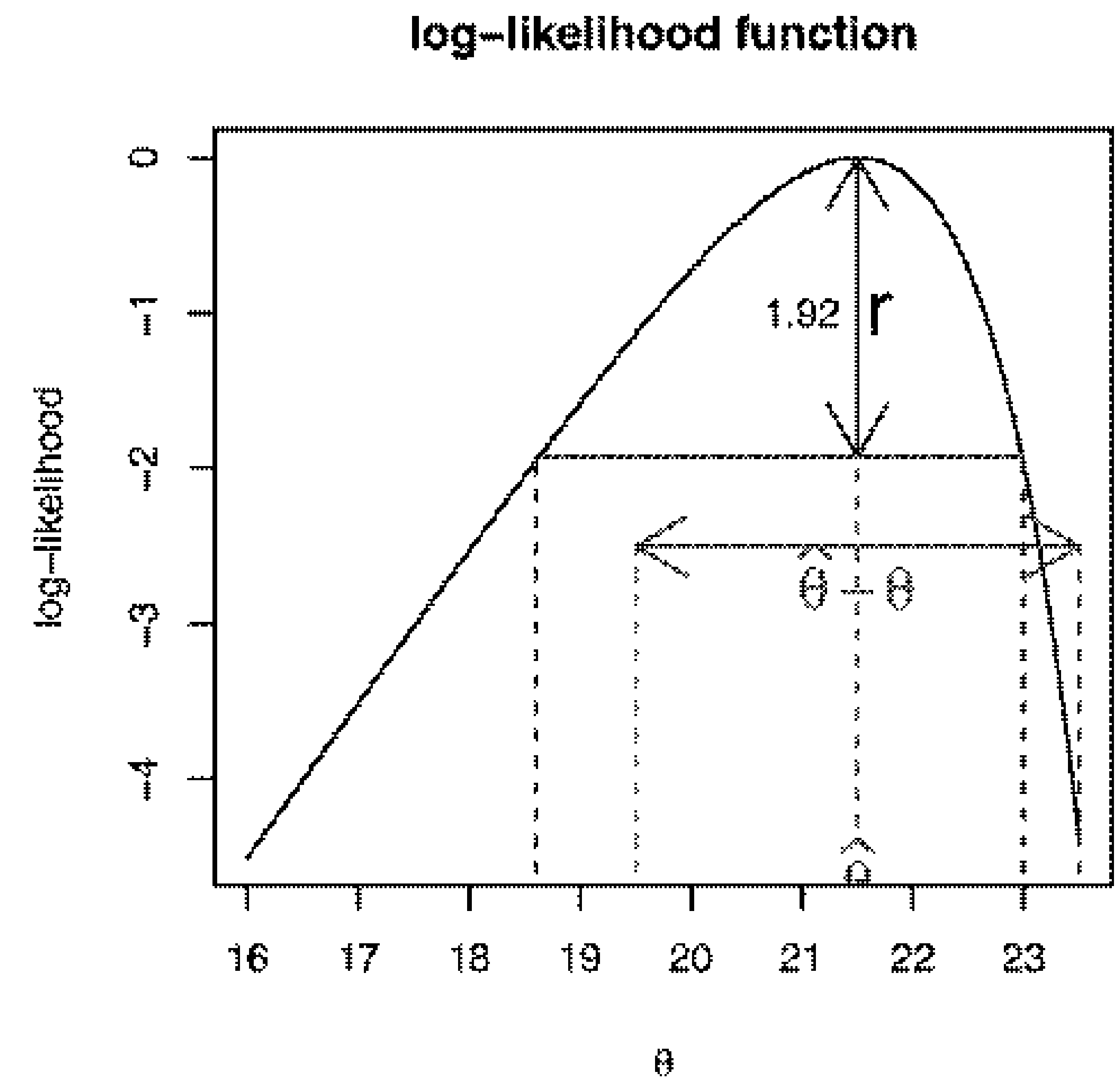
$$(\hat{\theta} - \theta)\{j(\hat{\theta})\}^{1/2} \;\dot{\sim}\; N(0, I)$$

$$2\{\ell(\hat{\theta}) - \ell(\theta)\} \;\dot{\sim}\; \chi^2_p$$

Special Case: $p = 1$

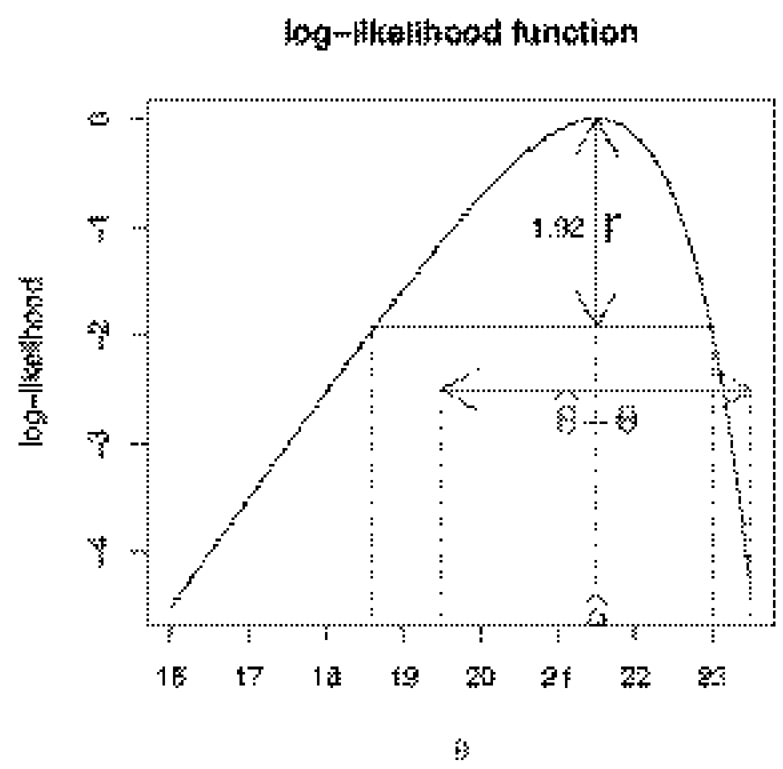
Likelihood inference

Parametric model: $f(y;\theta)$ e.g. $\exp\{-(y-\theta)-e^{-(y-\theta)}\}$



Likelihood inference

$\ell(\theta; y) = \log f(y;\theta) \qquad j(\hat\theta) = -\ell''(\hat\theta)$



standardized maximum likelihood est.: $q(\theta) = (\hat\theta - \theta)\{j(\hat\theta)\}^{1/2}$

likelihood root: $r(\theta) = \surd[2\{\ell(\hat\theta) - \ell(\theta)\}]$

standardized score function: $s(\theta) = \ell'(\theta)\{j(\hat\theta)\}^{-1/2}$

$$\xrightarrow{d} N(0,1)$$

If $\theta = \theta_0$, $\quad \hat\theta - \theta_0 \sim N(0, j(\hat\theta)^{-1})$

$$\frac{\hat\theta - \theta_0}{\sqrt{j(\hat\theta_0)^{-1}}} \sim N(0,1)$$

p-value

If $\quad \theta = \theta_0 \qquad 2\{\ell(\hat{\theta}) - \ell(\theta_0)\} \;\dot{\sim}\; \chi^2_1$

OR $\quad \hat{\theta} \pm z^{\alpha/2}\{j(\hat{\theta})\}^{-\frac{1}{2}}$

is a $1-\alpha$ CI for θ

OR $\quad \{\theta : 2\{\ell(\hat{\theta}) - \ell(\theta)\} \le 3.84\}$

is a 95% CI for θ

Whats about $p > 1$?

$$\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta} \;\dot{\sim}\; N_p\left(0,\; j^{-1}(\hat{\theta})\right)$$

- diagonal elements of $j^{-1}(\hat{\theta})$ estimate asy. variance of components of θ

$$\hat{\theta}_s \pm 1.96 \times \hat{s.e.}(\hat{\theta}_s) \quad \sim 95\% \text{ CI}$$

- this uses normal approx$^{=}$ to dist^{-} of $\hat{\underset{\sim}{\theta}}$

What is the likelihood ratio version of this?

$\hat{\theta}$ = mle = $\arg\sup_{\theta} \ell(\theta; y)$

$\hat{\hat{\theta}}$ = constrained mle = $\arg\sup_{\theta \in \Theta_0} \ell(\theta; y)$

1) e.g. $\frac{1}{\Gamma(\beta)} \left(\frac{\beta}{\mu}\right)^{\beta} y_i^{\beta-1} e^{-\beta y_i/\mu}$

$\theta = (\beta, \mu)$ $\hat{\theta} = (\hat{\beta}, \hat{\mu})$

- want to test $\beta = 1$ $\hat{\hat{\theta}} = (1, \hat{\hat{\mu}})$

or $\beta = \beta_0$ $\hat{\hat{\theta}} = (\beta_0, \hat{\hat{\mu}})$

2) eg $\theta = (\psi, \lambda)$, want inference on ψ

$\hat{\hat{\theta}} = (\psi, \hat{\lambda}_\psi)$ constrained mle

$2\{\ell(\hat{\psi}, \hat{\lambda}) - \ell(\psi, \hat{\lambda}_\psi)\} \xrightarrow{d} \chi^2_{\dim \psi}$