Likelihood

Arni Magnusson

United Nations University
Fisheries Training Programme

8–10 Dec 2014

Uncertainty

We have point estimates, excellent!

Harvest rate, reference points, . . .

But what about **uncertainty**?

Likelihood concepts

Relative probability

$$P(A_1) = 0.5 \qquad P(A_2) = 0.3 \qquad P(A_3) = 0.2$$

$$L(A_1) = 500 \qquad L(A_2) = 300 \qquad L(A_3) = 200$$

$$L(A_1) = 0.005 \quad L(A_2) = 0.003 \quad L(A_3) = 0.002$$

Likelihood concepts

Expresses how well the data **support** some parameter value or hypothesis

$$L(\theta|\text{data})$$

Like *RSS* but even more useful:

not only point estimate, but also **uncertainty**

Likelihood concepts

We can fit a model to many types of data at once and **combine** the likelihood components with simple multiplication

$$L = L_1 \times L_2 \times \cdots$$

Unified framework, for simple or complex models

Likelihood concepts

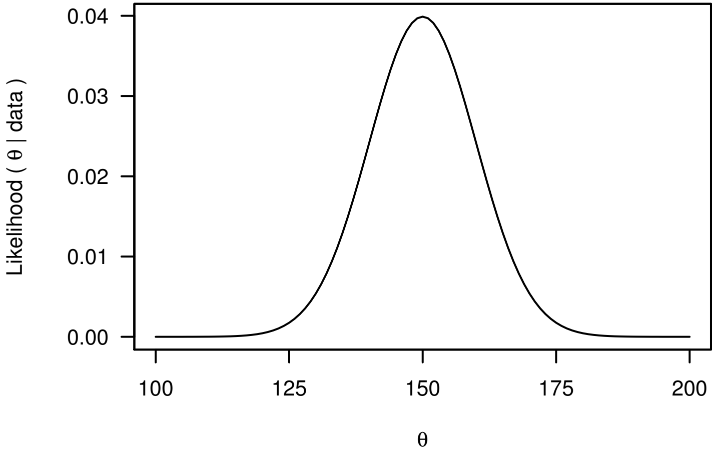
Choose between models with different number of parameters

$$2 \log \frac{L_1}{L_0} \sim \chi^2_{\Delta df}$$

$$\text{AIC} = -2 \log L + 2k$$

$$\text{BIC} = -2 \log L + \log(n) k$$

Maximum likelihood estimation



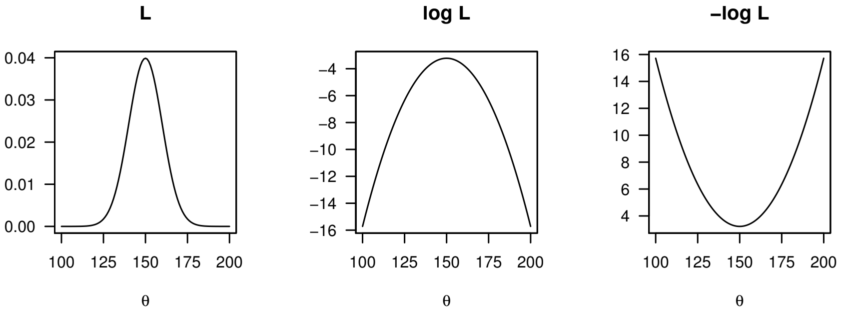
Log likelihood

Log transformation makes things easier

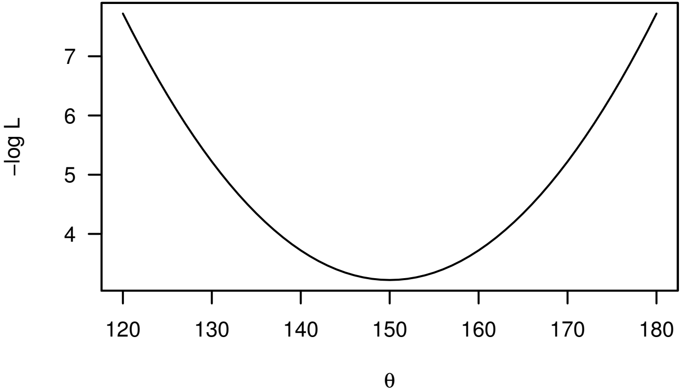
$$
\begin{aligned}
L(\theta|\text{data}) &= p(\text{data}|\theta) \\
& p(y_1, \ldots, y_n|\theta) \\
& p(y_1|\theta) \quad \times \quad \cdots \quad \times \quad p(y_n|\theta) \\
& \prod p(y_i|\theta)
\end{aligned}
$$

$$
\log L(\theta|\text{data}) = \sum \log p(y_i|\theta)
$$

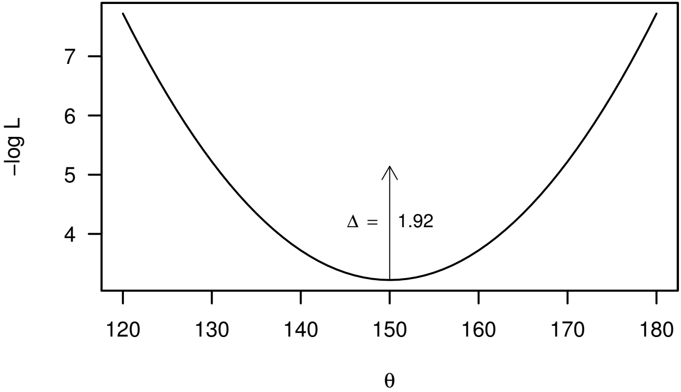
Log likelihood



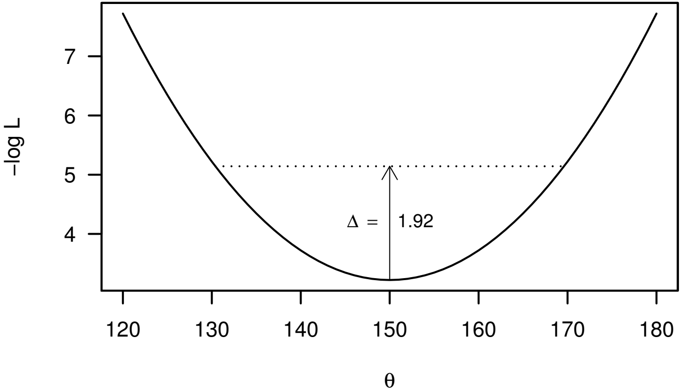
Confidence interval



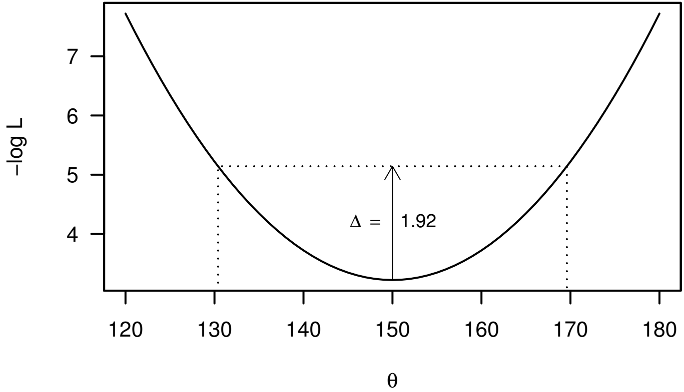
Confidence interval



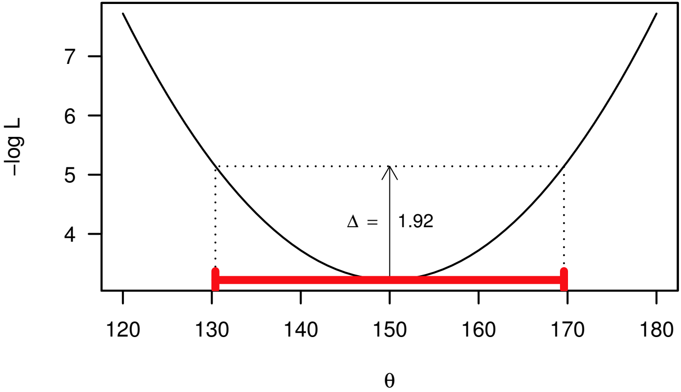
Confidence interval



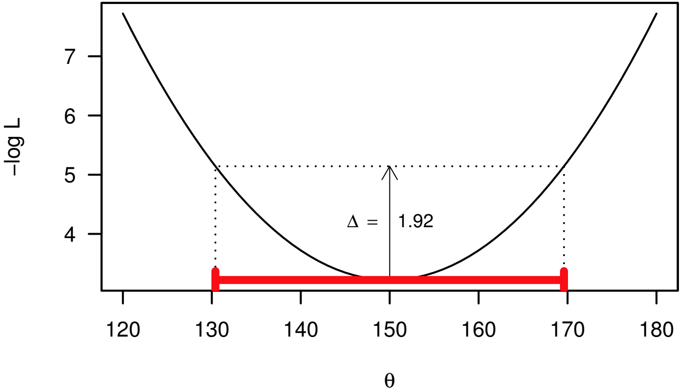
Confidence interval



Confidence interval



Confidence interval



$0.5\chi^2_{df=1} = 1.92$ for 95% confidence interval

Normal distribution

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\; e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$p(y_i|\theta) = \frac{1}{\sigma\sqrt{2\pi}}\; e^{-\frac{(y_i-\mu_i)^2}{2\sigma^2}}$$

$$L(\theta|y) = \prod\left(\frac{1}{\sigma\sqrt{2\pi}}\; e^{-\frac{(y_i-\mu_i)^2}{2\sigma^2}}\right)$$

$$-\log L = \left[0.5n\log(2\pi)\right] + n\log\sigma + \frac{\sum(y_i-\mu_i)^2}{2\sigma^2}$$

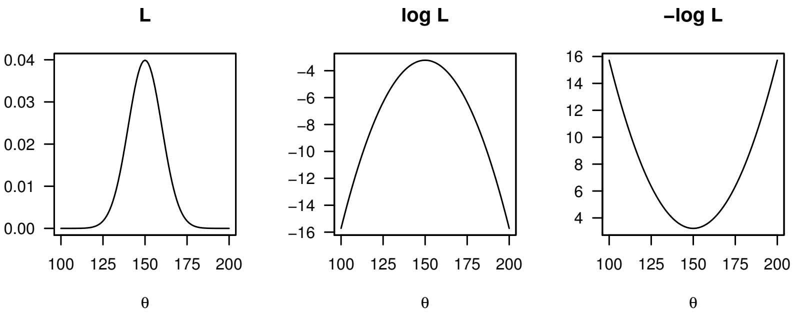
$$= \left[0.5n\log(2\pi)\right] + n\log\sigma + \frac{RSS}{2\sigma^2}$$

dnorm in R

```
L <- prod(dnorm(y, mu, sigma))
```

```
neglogL <- -sum(dnorm(y, mu, sigma, log=TRUE))
```

dnorm in R



`dnorm(theta, m=150, s=10)`

`dnorm(theta, m=150, s=10, log=T)`

`-dnorm(theta, m=150, s=10, log=T)`

Back to our Schaefer model

Use likelihood inference to evaluate uncertainty

Let's calculate 95% confidence interval for K

Fit model in spreadsheet

1. Get **data**
   ... and plot
2. Define **parameters**
   ... and consider transforming
3. Calculate **predictions**
   ... and fit by eye
4. Calculate **neglogL**
   ... and optimize

$t \quad B \quad C \quad I \quad Ifit \quad res2$

$B_{\text{init}} \quad r \quad K \quad q \quad \boldsymbol{\sigma}$

$$B_{t+1} = B_t + rB_t\left(1 - \frac{B_t}{K}\right) - C_t$$

$$\hat{I}_t = qB_t$$

$$-\log L =$$

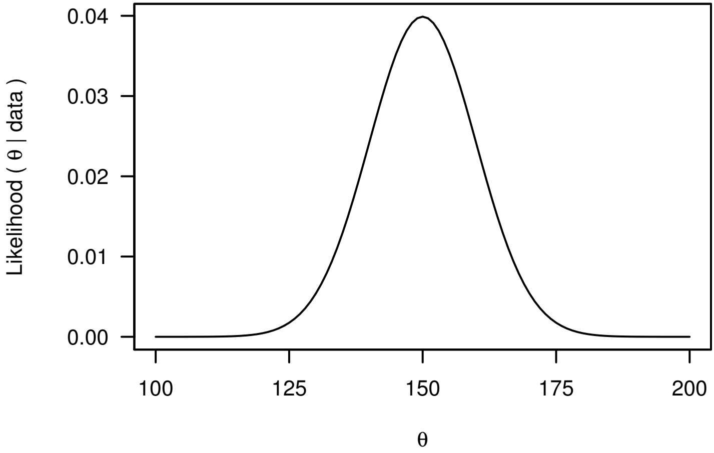
$$0.5n\log(2\pi) + n\log\sigma + \frac{RSS}{2\sigma^2}$$

Profile likelihood

1. Fix θ (a parameter of interest) at some value
2. Minimize $-\log L$ by estimating all other parameters
3. Save this value of $-\log L$

Repeat over a range of θ values

Profile likelihood



Discussion

- Difference between RSS and likelihood
- $B_{\text{init}} = K$?
- 95% confidence interval for MSY $rK/4$
- 95% confidence interval for B in final year

Discussion

- Difference between RSS and likelihood
- $B_{\text{init}} = K$?
- 95% confidence interval for MSY $rK/4$
- 95% confidence interval for B in final year

Discussion

- Difference between RSS and likelihood
- $B_{\text{init}} = K$?
- 95% confidence interval for MSY $\quad rK/4$
- 95% confidence interval for B in final year

Discussion

- Difference between RSS and likelihood
- $B_{\text{init}} = K$?
- 95% confidence interval for MSY $\quad rK/4$
- 95% confidence interval for B in final year