Likelihood II

Assessment workshop

Arni Magnusson

ICES, 18 Oct 2012

Outline

1. Likelihood - relative probability, support, combine data sources

2. Estimation - MLE, log likelihood, confidence interval

3. Normal distribution - $N(\mu, \sigma)$, dnorm

4. Profile likelihood - procedure, interpretation

Likelihood
Estimation
Normal
Profile

Relative probability
Support
Combine data sources
Model selection

Outline

1. Likelihood - relative probability, support, combine data sources

2. Estimation - MLE, log likelihood, confidence interval

3. Normal distribution - $N(\mu, \sigma)$, dnorm

4. Profile likelihood - procedure, interpretation

Likelihood
Estimation
Normal
Profile

Relative probability
Support
Combine data sources
Model selection

Likelihood concepts

Relative probability

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

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

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

Likelihood
Estimation
Normal
Profile

Relative probability
Support
Combine data sources
Model selection

Likelihood concepts

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

$$L\left(\theta\,|\,\text{data}\right)$$

Like *RSS* but even more useful:

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

Likelihood
Estimation
Normal
Profile

Relative probability
Support
Combine data sources
Model selection

Likelihood concepts

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

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

Unified framework, for simple or complex models

Likelihood
Estimation
Normal
Profile

Relative probability
Support
Combine data sources
Model selection

Likelihood concepts

Choose between models with different number of parameters

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

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

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

Likelihood
Estimation
Normal
Profile

Maximum likelihood
Log likelihood
Confidence interval

Outline

1. Likelihood - relative probability, support, combine data sources

2. Estimation - MLE, log likelihood, confidence interval

3. Normal distribution - $N(\mu, \sigma)$, dnorm

4. Profile likelihood - procedure, interpretation

Likelihood
Estimation
Normal
Profile

Maximum likelihood
Log likelihood
Confidence interval

Maximum likelihood estimation

Likelihood
Estimation
Normal
Profile

Maximum likelihood
Log likelihood
Confidence interval
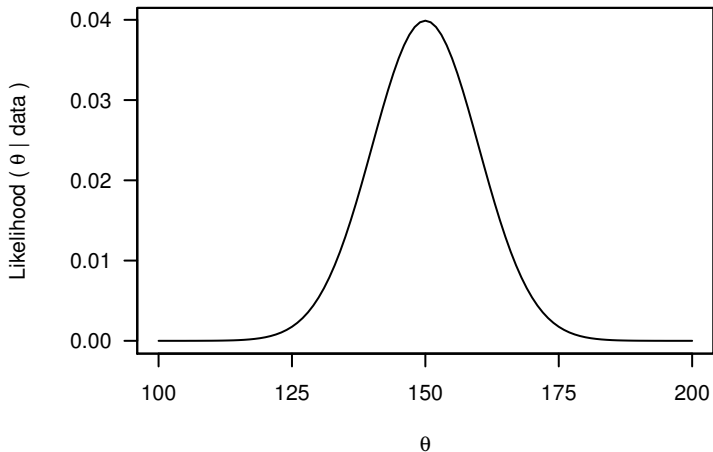
Log likelihood

Log transformation makes things easier

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

$$p(y_1, \ldots, y_n|\theta)$$
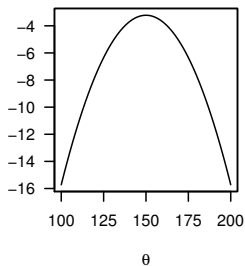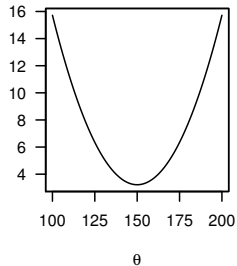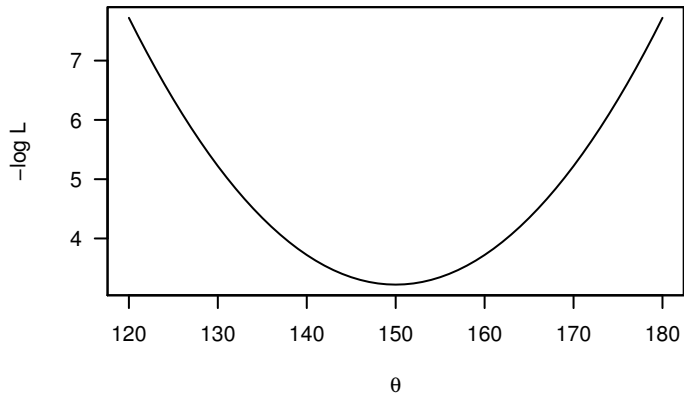
$$p(y_1|\theta) \times \cdots \times p(y_n|\theta)$$

$$\prod p(y_i|\theta)$$

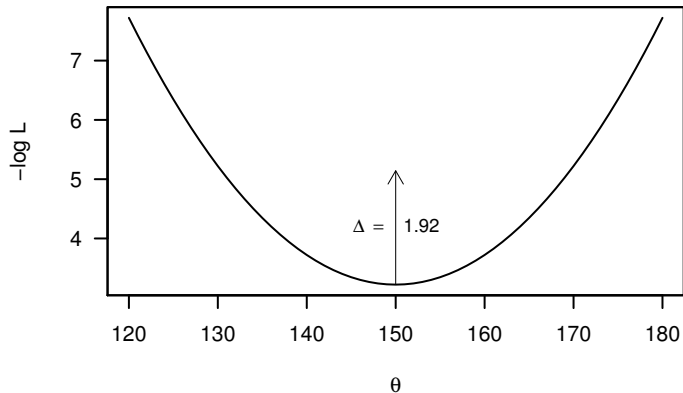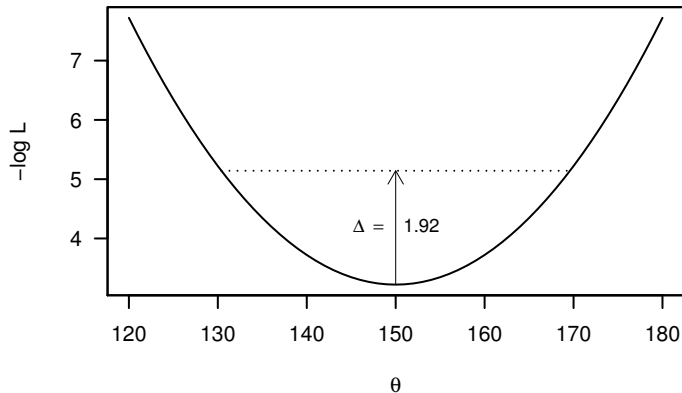$$\log L(\theta|\text{data}) = \sum \log p(y_i|\theta)$$

Likelihood
Estimation
Normal
Profile

Maximum likelihood
Log likelihood
Confidence interval

Log likelihood

Likelihood
Estimation
Normal
Profile

Maximum likelihood
Log likelihood
Confidence interval

Confidence interval

Likelihood
Estimation
Normal
Profile

Maximum likelihood
Log likelihood
Confidence interval

Confidence interval

Likelihood
Estimation
Normal
Profile

Maximum likelihood
Log likelihood
Confidence interval

Confidence interval

Likelihood
Estimation
Normal
Profile

Maximum likelihood
Log likelihood
Confidence interval

Confidence interval

Likelihood
Estimation
Normal
Profile

Maximum likelihood
Log likelihood
Confidence interval

Confidence interval

Likelihood
Estimation
Normal
Profile

Maximum likelihood
Log likelihood
Confidence interval

Confidence interval

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

Likelihood
Estimation
Normal
Profile

Normal distribution
dnorm

Outline

1. Likelihood - relative probability, support, combine data sources

2. Estimation - MLE, log likelihood, confidence interval

3. Normal distribution - $N(\mu, \sigma)$, dnorm

4. Profile likelihood - procedure, interpretation

Likelihood
Estimation
Normal
Profile

Normal distribution
dnorm

Normal distribution

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

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

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

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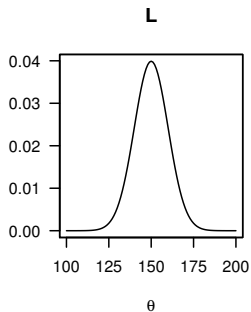
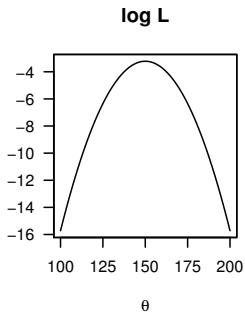$$= \left[0.5n\log(2\pi)\right] + n\log\sigma + \frac{RSS}{2\sigma^2}$$

Likelihood
Estimation
Normal
Profile

Normal distribution
dnorm

dnorm in R

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


neglogL <- -sum(dnorm(y, mu, sigma, log=TRUE))
```
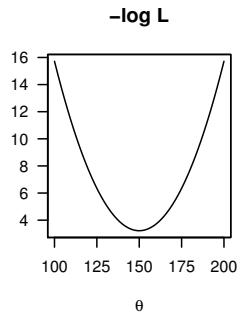
Likelihood
Estimation
Normal
Profile

Normal distribution
dnorm

dnorm in R

| **L** | **log L** | **–log L** |

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

Likelihood
Estimation
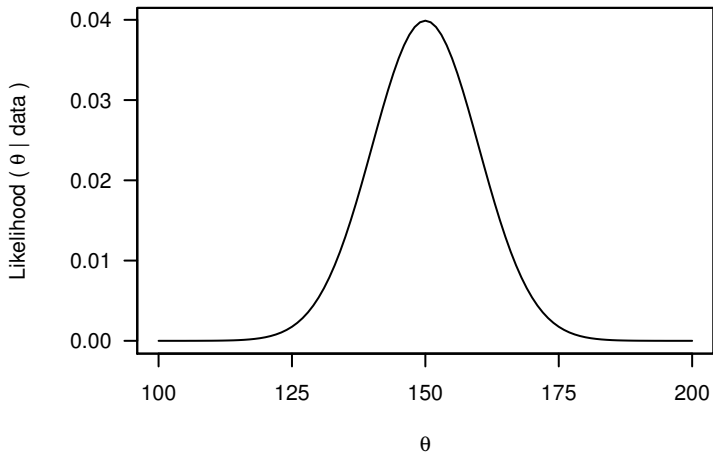Normal
Profile

Profile likelihood
Interpretation

Outline

1. Likelihood - relative probability, support, combine data sources

2. Estimation - MLE, log likelihood, confidence interval

3. Normal distribution - $N(\mu, \sigma)$, dnorm

4. Profile likelihood - procedure, interpretation

Likelihood
Estimation
Normal
Profile

Profile likelihood
Interpretation

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

Likelihood
Estimation
Normal
Profile

Profile likelihood
Interpretation

Profile likelihood

Likelihood
Estimation
Normal
Profile

Profile likelihood
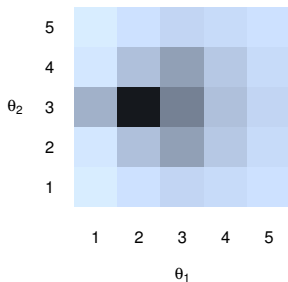Interpretation

Interpretation

Consider a 2-dimensional likelihood surface, describing the likelihood at different values of two parameters:

	$\theta_1 = 1$	$\theta_1 = 2$	$\theta_1 = 3$	$\theta_1 = 4$	$\theta_1 = 5$
$\theta_2 = 1$	5	15	25	20	15
$\theta_2 = 2$	10	40	60	35	20
$\theta_2 = 3$	50	100	75	40	25
$\theta_2 = 4$	10	40	60	35	20
$\theta_2 = 5$	5	15	25	20	15

Likelihood
Estimation
Normal
Profile

Profile likelihood
Interpretation

Interpretation

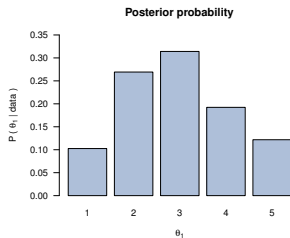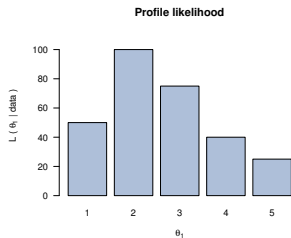Consider a 2-dimensional likelihood surface, describing the likelihood at different values of two parameters:

Likelihood
Estimation
Normal
Profile

Profile likelihood
Interpretation

Marginal distribution

In likelihood inference, we find the maximum likelihood for each value of θ_1 across all values of θ_2

In Bayesian inference, we integrate the likelihood over θ_2

Likelihood
Estimation
Normal
Profile

Profile likelihood
Interpretation

Interpretation

Would you place your bet on $\theta_1 = 2$ or $\theta_1 = 3$?

	$\theta_1 = 1$	$\theta_1 = 2$	$\theta_1 = 3$	$\theta_1 = 4$	$\theta_1 = 5$	sum
$\theta_2 = 1$	5	15	25	20	15	80
$\theta_2 = 2$	10	40	60	35	20	165
$\theta_2 = 3$	50	100	75	40	25	290
$\theta_2 = 4$	10	40	60	35	20	165
$\theta_2 = 5$	5	15	25	20	15	80
sum	80	210	245	150	95	780